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

M.Sc. Goran Stanić

born in: Novi Sad, Serbia

Oral examination: 17-10-2025

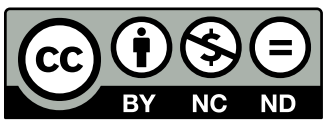
Deep Learning-Based Synthetic CT Images for Adaptive Radiotherapy

Referees: Prof. Dr. Oliver Jäkel

Prof. Dr. Mark Ladd

Mami

Nadam se da ćemo pronaći lek



Zusammenfassung

Die adaptive Strahlentherapie stützt sich auf tägliche Bildgebung, um anatomische Veränderungen während der Behandlung zu erfassen. Allerdings weisen aktuelle Verfahren wie die Kegelstrahl-Computertomographie (CBCT) und Anpassungstechniken unter Verwendung der deformierbaren Bildregistrierung (DIR) gelegentlich nicht die erforderliche Genauigkeit auf. Verbesserungen der CBCT bei der Erfassung und Rekonstruktion sind im Routineeinsatz nach wie vor begrenzt, während die DIR mit großen anatomischen Veränderungen zu kämpfen hat und keine robuste Qualitätssicherung bietet. Diese Mängel motivieren die Suche nach alternativen Anpassungsansätzen.

Als mögliche Lösung untersucht diese Arbeit synthetische CT-Bilder (synCT), die mit einem CycleGAN-Netzwerk erzeugt wurden. SynCTs zeigten im Vergleich zu CBCT eine verbesserte Bildqualität und eine bessere Übereinstimmung mit der täglichen Anatomie als DIR-basierte CT, was ihren Einsatz in klinisch relevanten Aufgaben unterstützt. Die Behandlungsplanung auf Basis von synCTs erzielte eine ähnliche Zielabdeckung und Schonung von Risikoorganen wie andere Anpassungsmethoden und ermöglichte das Training von Segmentierungsnetzwerken für Risikoorgane ohne hochwertige annotierte CBCT-Datensätze.

Diese Ergebnisse deuten darauf hin, dass Deep-Learning-basierte synCTs adaptive Strahlentherapie-Workflows stärken können, indem sie mehrere Einschränkungen bestehender Bildgebungsverfahren überwinden. Obwohl eine weitere Validierung an größeren und vielfältigeren Patientenkohorten erforderlich ist, ist synCT als klinisch nützliche Modalität vielversprechend und schließt die Lücke zwischen der aktuellen Praxis und dem Ziel einer zuverlässigen patientenspezifischen täglichen Behandlungsanpassung.

Abstract

Adaptive radiotherapy relies on daily imaging to capture anatomical changes during treatment, yet current modalities such as cone-beam CT (CBCT) and adaptation techniques using deformable image registration (DIR) occasionally lack the required accuracy. CBCT improvements in acquisition and reconstruction remain limited in routine use, while DIR struggles with large anatomical changes and lacks robust quality assurance. These shortcomings motivate the search for alternative adaptation approaches.

As a potential solution, this thesis investigates synthetic CT (synCT) images generated using a CycleGAN network. SynCTs demonstrated improved image quality compared with CBCT and closer agreement with daily anatomy than DIR-based CT, supporting their use in clinically relevant tasks. Treatment planning on synCTs achieved similar target coverage and sparing of organs at risk to other adaptation methods, and enabled training of organ at risk segmentation networks in the absence of high-quality annotated CBCT datasets.

These findings suggest that deep learning-based synCT images can strengthen adaptive radiotherapy workflows by overcoming several limitations of existing methods. While further validation on larger and more diverse patient cohorts is needed, synCT demonstrates potential as a clinically useful modality, bridging the gap between current practice and the goal of reliable patient-specific daily treatment adaptation.

Contents

Acronyms

ART adaptive radiotherapy

BEV beam's eye view

CBCT cone-beam computer tomography

CT computer tomography

DICE Sørensen-Dice coefficient

DIR deformable image registration

dirCT deformed CT

DVH dose-volume histogram

FOV field-of-view

GAN generative adversarial network

HU Hounsfield unit

IGRT image-guided radiotherapy

IMRT intensity-modulated radiotherapy

JSD Jensen-Shannon divergence

LINAC linear accelerator

LungCa lung cancer

MRI magnetic resonance imaging

MSE mean squared error

NSCLC non-small-cell lung cancer

NTCP normal tissue complication probability

PCa prostate cancer

pCT planning CT

PDD percentage depth dose

PET positron-emission tomography

PSNR peak signal-to-noise ratio

SCLC small-cell lung cancer

SDC Surface DICE coefficient

SSIM Structural similarity index measure

synCT synthetic CT

TCP tumor control probability

TPS treatment planning system

VOI volume of interest

Introduction

Cancer remains the second leading cause of death globally and continues to burden individuals, families and healthcare systems. While cancer incidence rates rise, improvements in treatment and early detection have stabilized or slightly reduced annual deaths. Among treatment modalities, radiotherapy plays a crucial role as clinicians use it to treat approximately 50% of all cancer patients. Radiotherapy has improved significantly over decades, with advancements that provide better tumor control and fewer side effects. Yet, it still faces fundamental precision challenges: it operates between the anatomical accuracy of surgery and the systemic reach of chemotherapy. The key issue involves balancing effective tumor targeting while sparing surrounding healthy tissue from radiation exposure. Achieving this precision relies heavily on accurate imaging of patient anatomy at the time of treatment. In particular, anatomical changes that occur between or during treatment sessions require methods to adapt treatment plans based on daily imaging. This need has led to the emergence of image-guided radiotherapy (IGRT), and more recently, adaptive radiotherapy (ART) approaches, which prove especially crucial in anatomically dynamic regions like the lungs and prostate.

Modern radiotherapy systems increasingly rely on daily imaging, such as cone-beam computer tomography (CBCT), to guide and adapt treatment. However, effectively integrating this daily imaging into treatment planning and dose delivery remains an open problem. Current research has focused on four main strategies for CBCT-based treatment adaptation: improving CBCT acquisition and reconstruction techniques, refining deformable image registration (DIR) algorithms, enhancing CBCT images using analytical methods, and synthesizing CT-equivalent images from CBCTs using neural networks. While researchers have extensively studied and implemented the first three approaches with varying degrees of success, the fourth option – image synthesis via deep learning – remains relatively underexplored, particularly regarding its clinical feasibility. It is still uncertain whether synthetic CT (synCT) images can fully replace planning CTs (pCTs) for critical tasks such as dose recalculation, treatment adaptation, or segmentation. This thesis seeks

to address that gap.

The present work aims to explore the clinical feasibility and limitations of synCT images in ART through exploratory, methodological, and comparative studies using lung and prostate cancer datasets. The limitations of CBCT and DIR-based CT (dirCT) methods in ART are evaluated, particularly regarding image quality and anatomical accuracy. A solution to these limitations is introduced through deep learning-based synCT images generated using the CycleGAN architecture. Finally, the work explores additional downstream clinical applications of synthetic images, such as direct dose calculation and automatic segmentation, as well as their integration into a modern treatment planning system (TPS).

In this thesis, retrospective radiotherapy data from open-access datasets and in-house clinical datasets was used, and two clinically commissioned treatment planning systems, Eclipse™ and Ethos™, were employed for treatment planning, optimization, and dose calculation, depending on the type of investigation. Given the nature of daily ART and available image modalities, ground truth anatomical verification is nearly impossible. Therefore, this research adopts a patient-specific approach with multiple complementary strategies that are accompanied by statistical analyses. Within this framework, each chapter presents a self-contained investigation with its own objectives and methods, contributing to a broader narrative. Collectively, this work critically examines the limitations of CBCT- and dirCT-based ART methods in dynamically changing anatomical sites, provides quantitative and qualitative insights into synCT image quality created using CycleGAN models, and presents comparative evaluations that demonstrate the advantages of ART over IGRT in general, as well as the benefit and promise of synCT-based ART in particular. In addition, a novel approach for training segmentation networks using synthetic images is introduced, leading to improved performance in downstream ART tasks such as automatic segmentation. Together, these findings highlight the potential of synthetic images to serve as a foundation for future ART workflows across a wide range of clinical applications.

Structure of the thesis

Chapter 1 provides background on cancer treatment, radiotherapy principles, CBCT technology with its advantages and drawbacks, and image-guided and adaptive radiotherapy with their relevance to lung and prostate cancers. *Chapter 2* establishes the groundwork for neural networks in general, and generative adversarial networks, specifically. It also covers image quality metrics and different evalua-

tion principles. Equipped with a strong deep learning foundation, we continue to *Chapter 3*, which presents a foundational comparison of CBCT and dirCT methods using lung cancer datasets, focusing on anatomical characteristics, and direct dose calculation performance. Its aim is to demonstrate a strong case for deep learning-based synthetic images. *Chapter 4* follows by detailing the development and evaluation of a 3DCycleGAN network for synthetic CT generation, using both lung and prostate datasets. *Chapter 5* applies the model generated images to a clinical task – evaluating synCT performance in treatment planning and daily dose recalculation – by comparing ART with IGRT strategies as well as different ART approaches with one another. Finally *Chapter 6* explores the use of synthetic CT in automatic segmentation tasks and introduces a new methodological framework for training segmentation networks. Chapters 3 to 6 follow a consistent structure with sections on Motivation, Methods, Results, Discussion, and Conclusion. While each chapter represents a standalone investigation, taken as a whole, the chapters chart a clear path from generating synCTs from daily CBCTs to demonstrating their clinical utility in ART.

1

Radiation Therapy Background

1.1 Cancer and radiation therapy

Cancer, also known as neoplasm or malignant tumor, is a group of diseases that occur when abnormal cells grow uncontrollably. Cancer can start in almost any organ or tissue of the body, it can go beyond its boundaries to invade neighboring parts of the body, or spread to other organs in a process called metastasizing [1]. With an estimated 20 million new yearly cases and 10 million deaths worldwide in 2022, it is the second leading cause of death after cardiovascular disease. According to recent data from The Global Cancer Observatory, the most commonly diagnosed cancers in both sexes combined are lung, breast, colorectum, prostate, stomach and liver cancers (Fig. 1.1). Notably, lung cancer remains the leading cause of cancer death overall, and the most commonly diagnosed cancer in men, with prostate cancer in the second place [2].

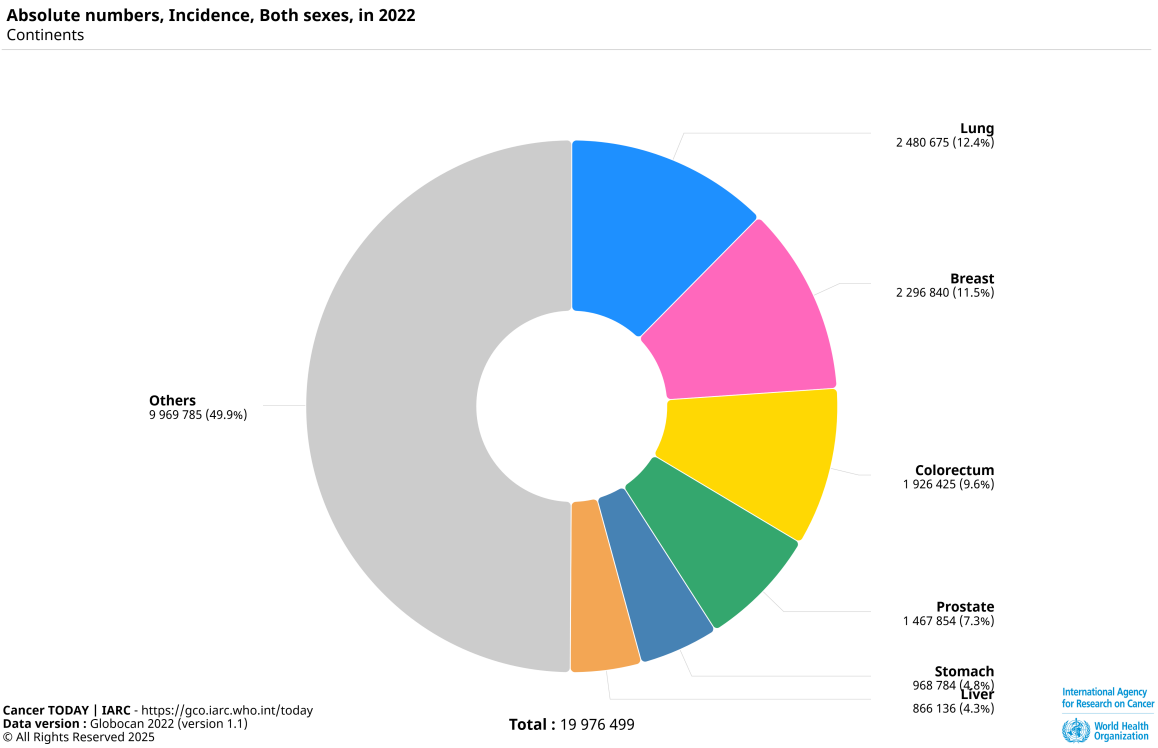


Figure 1.1. 2022 worldwide cancer incidence rates in percentages and absolute numbers for both sexes. Figure was created using an online tool from the International Agency for Research on Cancer [2].

Cancer treatment is inherently multimodal, involving a combination of surgery, chemotherapy, radiation therapy (radiotherapy), and more recently, targeted therapy and immunotherapy. Globally, it is estimated that up to 80% of patients will require surgery at some stage, and approximately 50% will receive radiotherapy [3]. Among the 50%, radiotherapy plays a crucial role in the management of both lung

and prostate cancers. It is the second most common treatment modality in prostate cancer after surgery [4] and it is used in both small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC) treatment depending on the progression of the disease [5–7].

The primary physical principle behind radiotherapy is ionizing radiation. Ionizing radiation is a type of radiation that has high enough energy to kick out one or more orbital electrons out of the shells of their atoms, leaving those atoms positively charged. In the field of radiotherapy ionizing radiation can be divided into photons (X-rays and γ-rays) and charged particles (electrons, protons, alpha particles, and heavy ions). Photons and charged particles have distinct ways of interacting with matter. Photons interact via the photoelectric effect, Compton (incoherent) scattering, pair production, and Rayleigh (coherent) scattering, while charged particles interact dominantly via Coulomb interactions, inelastic collisions with electrons, elastic collisions with nuclei, and radiative losses. Photons and charged particles will interact with matter in all of these ways with a certain probability that will depend on the energy of the particle and on the properties of the material. Regardless of the specific interaction mechanism, the fundamental result is the same: energy from the incident radiation is transferred to the absorbing medium. For protons and heavy ions, this energy transfer occurs continuously along their path as described by the Bethe-Bloch equation [8–10]

$$-\frac{dE}{dx} = Kz^2\frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ln\left(\frac{2m_ec^2\beta^2\gamma^2T_{\max}}{I^2}\right) - \beta^2 - \frac{\delta}{2} - \frac{C}{Z}\right], \tag{1.1}$$

where $K = 4\pi N_A r_e^2 m_e c^2 \approx 0.307\,\mathrm{MeV\,cm^2/g}$ is a constant depending on fundamental quantities, z is the charge number of the incident particle (e.g. $z = 1$ for protons, $z = 2$ for alpha particles), Z and A are the atomic number and atomic mass of the absorber (g/mol), $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$ are the particle velocity relative to the speed of light and the Lorentz factor, m_e is the electron rest mass, $T_{\max}$ is the maximum kinetic energy transferable to a free electron in a single collision, described by

$$T_{\max} = \frac{2m_ec^2\beta^2\gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2}, \tag{1.2}$$

I is the mean excitation (ionization) potential of the absorber, typically in the range of 10-100eV, depending on material, and δ and C are the density effect correction, accounting for medium polarization at high $\beta\gamma$, and the shell correction.

For photons, the energy transfer occurs through discrete interaction events that depend on the energy of the photon and the type of absorbing material. Each type

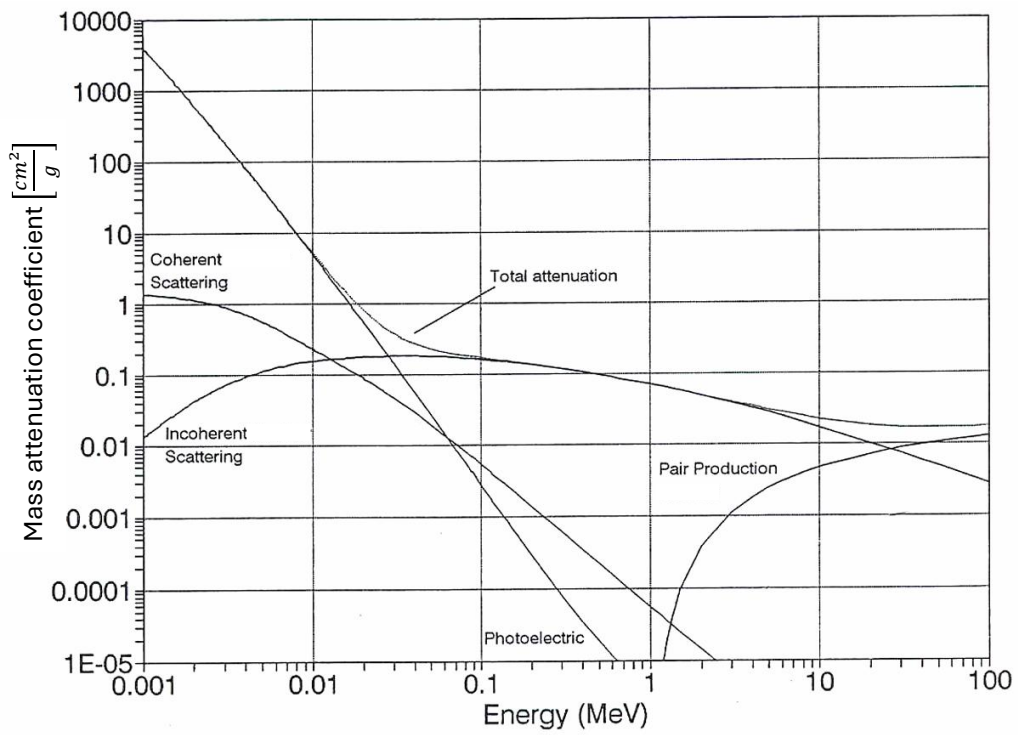


Figure 1.2. Photon mass attenuation coefficient μ/ρ for water as a function of photon energy. The graph displays the contributions from Coherent (Rayleigh) scattering, Incoherent (Compton) scattering, the Photoelectric effect, and Pair production, along with the total attenuation coefficient. Each interaction mechanism dominates at different energy ranges, illustrating the energy-dependent behavior of photon interactions with matter [11].

of interaction is characterized by an attenuation coefficient and the sum of all attenuation coefficients gives the total attenuation – a number describing how easily a material can be penetrated by a photon beam (Fig. 1.2).

For both photons and particles, the deposited energy is what makes radiation biologically effective and forms the basis for cancer treatment. The quantity that describes the energy per unit mass deposited by ionizing radiation in living matter is called *absorbed dose* and its SI unit of measure is gray (Gy) which is defined as J/kg. Similar to the absorbed dose, the *dose rate* describes the amount of absorbed dose per unit time and it is often indicated in μGy/h. Dose rate is used when describing deterministic radiation effects, while for stochastic effects, only the total absorbed dose matters. Even though the total absorbed dose determines the biological effects in tissue, radiotherapy also requires precise knowledge of how this dose is spatially distributed within the patient. This is often quantified using percentage depth dose (PDD) curves, which describe the percentage of deposited dose relative to maximum deposited dose as a function of depth in tissue. The curve characteristics vary depending on the particle type and its energy (Fig. 1.3).

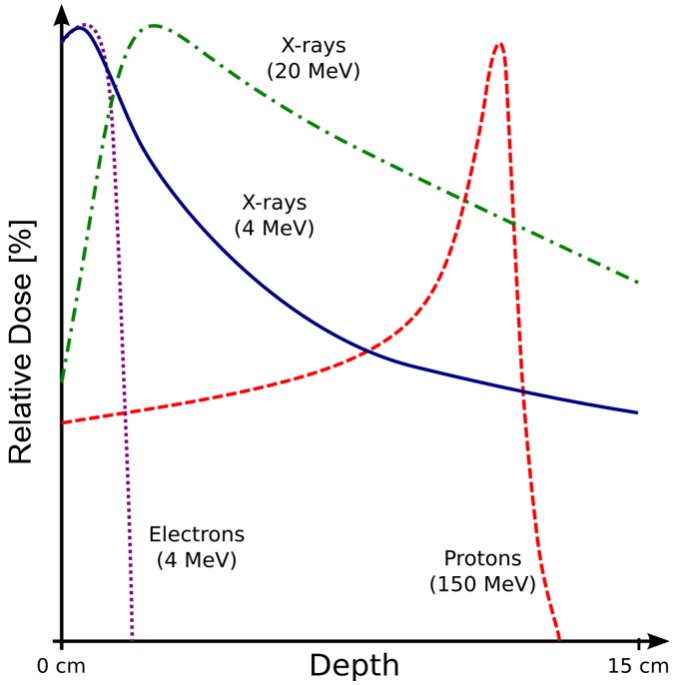


Figure 1.3. PDD curves showing relative dose as a function of depth in tissue for different radiation modalities: 4 MeV and 20 MeV X-rays (blue and green), 4 MeV electrons (purple), and 150 MeV protons (red). The photon curves exhibit a buildup region followed by exponential attenuation, with deeper penetration for higher energy X-rays. The electron beam shows a rapid dose fall-off after a shallow penetration depth. The proton curve highlights the characteristic Bragg peak, indicating maximal dose deposition at a specific depth. Adapted from [12, Chapter 29].

The biological effects caused by ionizing radiation occur through a cascade of three sequential phases: the physical phase where radiation particles interact with molecules in the cell causing ionization and free radical formation, the chemical phase where reactive radicals chemically modify biomolecules, and the biological phase involving enzymatic reactions and cellular regulation processes that determine cell fate [13, Chapter 1].

While radiation initially affects all components of the cell non-discriminately, the primary target of radiation is the DNA – the genetic material that carries crucial instructions about cell processes. Cell and DNA damage depends on the energy and type of radiation, and it can happen either directly or indirectly. Direct DNA damage is caused by charged particles such as protons and ions inducing single-, double- or clustered strand breaks. On the other hand, a more common, indirect DNA damage is caused by photons and electrons, where ionizing radiation produces highly reactive oxygen species (free radicals) that chemically modify the DNA structure. The cellular response to the radiation damage depends on

the cell's ability to repair DNA lesions at cell cycle checkpoints, with single-strand breaks being more easily repaired than double-strand breaks or clustered damage. Cells that cannot adequately repair radiation-induced damage undergo cell death either through early apoptotic pathways or late mitotic catastrophe after one or more failed division attempts [13, Chapter 3].

Cancer cells generally have a poorer DNA repair capacity than healthy cells due to defective DNA repair pathways, compromised cell cycle checkpoints, and many other disrupted mechanisms. This differential response between tumor and normal tissue provides the biological rationale for radiotherapy. Underlying this rationale are two characteristic sigmoid dose-response curves, one for tumor control probability (TCP) and the other for normal tissue complication probability (NTCP), which together define the therapeutic window (Fig. 1.4). The optimum choice of a treatment technique for a given tumor is such that it maximizes the TCP while minimizing the NTCP. For an effective radiotherapy treatment we are aiming at TCP $\geq$ 0.5 and NTCP $\leq$ 0.05. The further the NTCP curve is to the right of the TCP curve, the easier it is to achieve the radiotherapeutic goal, the larger is the therapeutic index and the less likely will it be that treatment causes complications. The therapeutic index generally refers to the ratio of the TCP and NTCP at a specified level of response (usually 0.05) for normal tissue (denoted as *max. tolerance* in Fig. 1.4), and it can be increased with chemotherapy and fractionation [14, Chapter 14].

Fractionation is one of the central principles of radiotherapy. It describes the process when treatment is delivered over a period of several weeks instead of a single session. The principle is effective due to 5 primary biological factors, the so-called 5 Rs of radiotherapy:

- Radiosensitivity - different cells have different sensitivity to radiation;
- Repair - cells can repair radiation damage. This is a complex process that involves repair of sublethal damage by a variety of repair enzymes and pathways. Cell repair is more effective in healthy cells than in tumors;
- Repopulation - cells repopulate (increase in numbers) while receiving fractionated doses of radiation. This is true both for healthy and tumor cells;
- Redistribution - populations of tumor cells are irradiated throughout the cell cycle phases, which increases the chance of killing them;
- Reoxygenation - during fractionated treatment, reoxygenation of hypoxic tumor cells occurs making them more radiosensitive to subsequent radiation exposure.

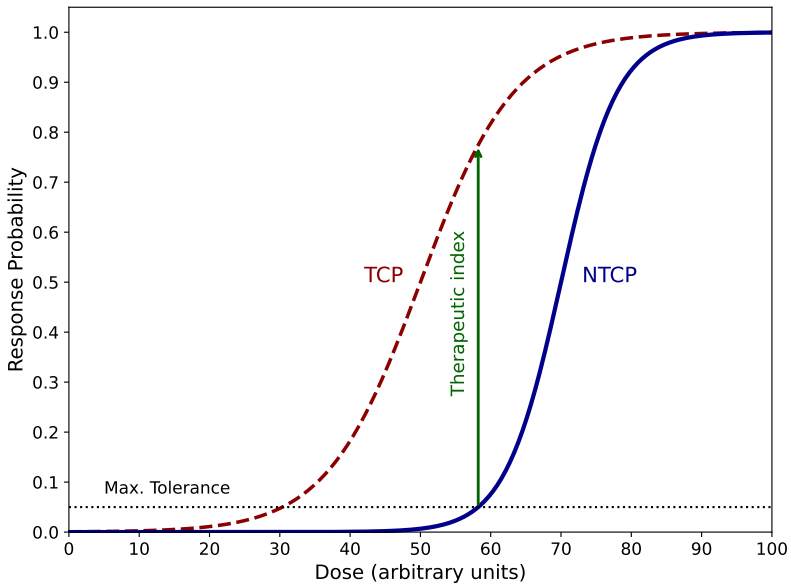


Figure 1.4. Illustration of the tumor control probability (TCP) curve (red), and normal tissue complication probability (NTCP) curve (blue), as a function of dose. The therapeutic index (green) represents the separation between effective tumor control and the onset of normal tissue complications, highlighting the optimal treatment window. The dashed horizontal line indicates the maximum tolerance threshold for normal tissue. Adapted from [15].

In current standard fractionation schemes, the dose is delivered to the patient 5 days per week for several weeks. This form of fractionation is also called normo-fractionation. The treatment can also be delivered to a patient multiple times during the day (hyper-fractionation) or only a couple of times but with a much larger dose per fraction (hypo-fractionation) [14, Chapter 14].

The most common form of radiation treatment is photon radiotherapy with 99% of patients receiving it worldwide. Due to its historical advantage and therefore wider availability, lower cost, as well as established effectiveness, it plays a central role in current clinical practice. It can be divided based on type of application into external beam radiotherapy (EBRT), brachytherapy, intravenous radiotherapy and intraoperative radiotherapy, where EBRT makes up the overwhelming majority of delivered treatments.

Today, photon EBRT is most commonly delivered using a medical linear accelerator (LINAC). Medical LINACs accelerate electrons to kinetic energies from 4 to 25 MeV using microwave radio-frequency (RF) fields. The electrons are accelerated following a straight path in accelerating waveguides. Once they reach the required energy they impinge on a target (often made out of tungsten), where high energy, mainly forward peaked photons are produced through Bremsstrahlung and are di-

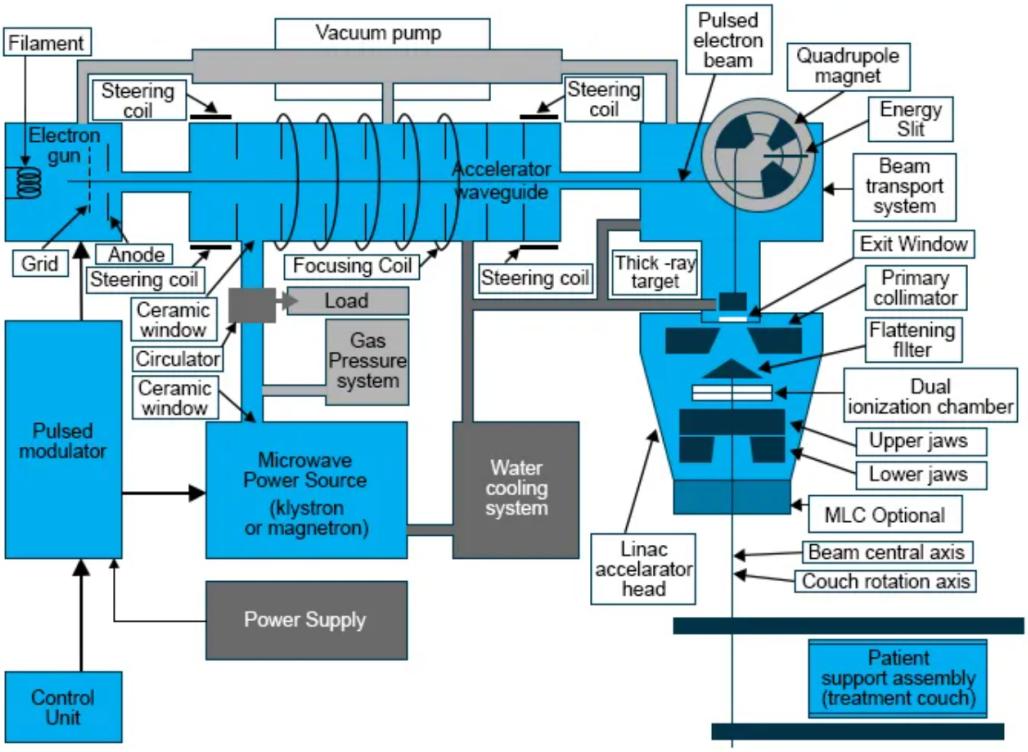


Figure 1.5. Schematic diagram of a medical linear accelerator (LINAC). EEWeb, accessed on June 24, 2025.

rected towards the patient using a series of collimators. A detailed schematic of a typical medical LINAC is illustrated in Fig. 1.5.

In order to deliver the appropriate dose from photon beams to the patient it is necessary to plan the treatment and in so doing, deliver high dose to the tumor and spare healthy organs surrounding it. A typical treatment planning process contains several necessary steps:

Initial imaging and diagnostics Before the treatment planning process begins it is important to have a complete understanding of the disease. This information is obtained using imaging techniques such as computer tomography (CT), magnetic resonance imaging (MRI), positron-emission tomography (PET and PET-CT), and through histopathological investigation of the tumor tissue.

Patient positioning and immobilization After initial imaging, the patient position on the treatment table is determined based on the location of the tumor and the anatomical site that will be irradiated. Also, depending on the required precision of the treatment delivery, patients will be fitted with immobilization devices such as vacuum cushions for thorax and abdomen irradiation, plastic masks and head rests for head and neck irradiation, or leg rests for pelvic region irradiation.

Simulation imaging The next step is to perform a simulation, or a planning CT (pCT). Unlike diagnostic CTs, which are optimized for anatomical visualization and diagnostic tasks, pCTs are specifically acquired with the patient in the treatment position using immobilization devices and protocols consistent with radiotherapy planning. pCTs are essential for accurate treatment planning and dose calculation, as they provide a reliable correlation between CT gray values, also known as Hounsfield unit (HU), and electron densities of tissues. It is also possible to perform a simulation MRI or PET-CT which will be close in time to the simulation CT and help with further steps in the treatment planning such as volume definition.

Volume definition Precise volume definition is key to any successful radiotherapy treatment. There are several volumes of interest that need to be defined during the treatment planning procedure [16, 17]:

- GTV (gross tumor volume) – Gross palpable or visible malignant growth determined using several imaging modalities as well as clinical examination and histopathological reports.
- CTV (clinical target volume) – Encompasses the GTV and nearby tissue that may harbor microscopic disease or is considered at risk, such as involved lymph nodes.
- ITV (internal target volume) – Includes the CTV along with a margin to account for internal motion (e.g., breathing, bladder filling), ensuring coverage in anatomical regions that are prone to movement during treatment.
- PTV (planning target volume) – The ITV with an additional margin to compensate for patient setup variability, machine precision, and treatment delivery uncertainties.
- Organs at risk – Normal tissues sensitive to radiation exposure that may be affected by the treatment plan, often requiring dose limitations or modified beam configurations to protect them.

The relationship between target volumes and organs at risk is visualized in Fig. 1.6.

Dose prescription and fractionation Based on the tumor grading and the complexity of the treatment, a radio-oncologist defines several dosimetric endpoints that describe requirements for the treatment dose. This means both how much dose

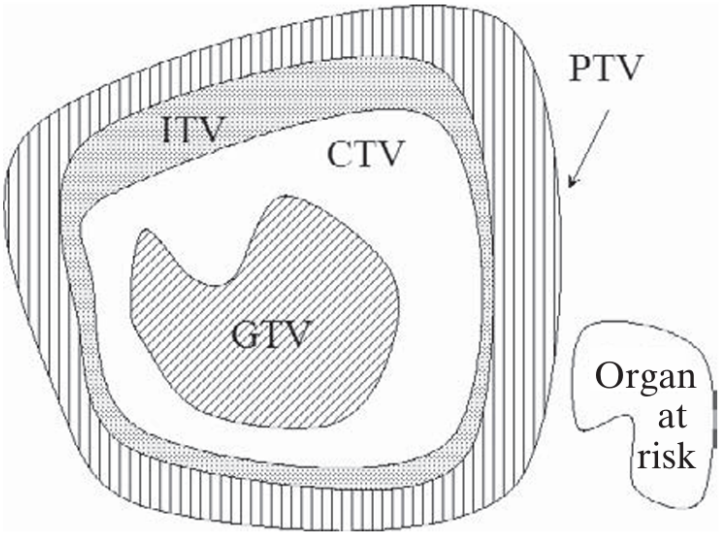


Figure 1.6. Volumes of interest – GTV, CTV, ITV, PTV, and organs at risk – as defined in ICRU Reports 50 and 62 [16, 17].

a tumor volume should receive, as well as which dose threshold is not allowed to be crossed when irradiating healthy organs at risk. The dosimetric endpoints for tumors are prescribed based on treatment outcome studies, whereas the limits on dose to organs at risk are based on toxicity studies. Some of the most common metrics that should be reported for proper comparison of outcome results are D_2, D_{95}, and D_{50} for target volumes and D_2, D_{mean}, and V_D for organs at risk [18]. These metrics are based on graphical representations that describe the amount of dose delivered to a volume of tissue, and they are called dose-volume histograms (DVHs) (see Fig. 1.7).

Treatment plan simulation – optimization – evaluation Following imaging, target and organ at risk definition, and dose prescription, it is time to simulate the treatment plan and configure the parameters necessary for patient irradiation. The treatment isocenter is first established, typically at the center of mass of the planning target volume, after which the beam geometry and delivery method (step-and-shoot or arc therapy) are defined. Finally, the treatment plan is optimized using internal algorithms of the treatment planning software. Now, a plan can be prepared for quality assurance by simulating beam's eye views (BEVs) for each beam for delivery tracking, after which the DVH and dosimetric endpoints can be evaluated to determine the quality of the plan. Ideally, an independent check should be performed by a second staff member and with a separate software.

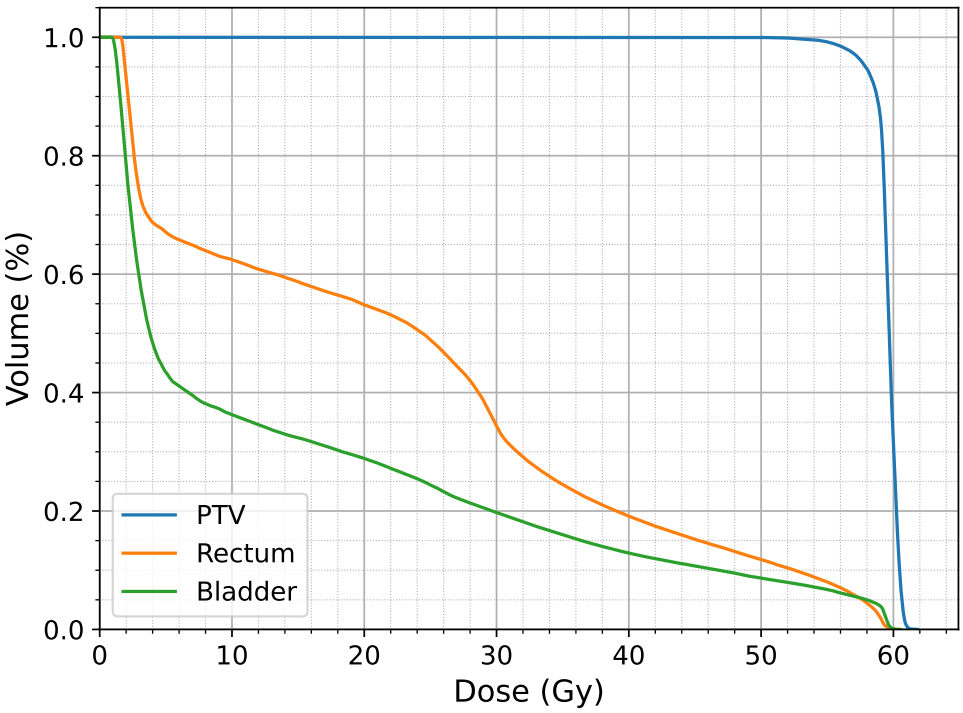


Figure 1.7. Example DVH of a PTV (blue) and two organs at risk, rectum (orange) and bladder (green).

Replicating simulation conditions on treatment day Since simulation, optimization and evaluation require significant time, treatment typically occurs on a separate day – often about a week after pCT acquisition. As a result, it is essential to reliably replicate the simulation conditions on the treatment day. This is accomplished using the same immobilization devices and, more importantly, through the use of integrated imaging, which enables accurate patient positioning at the correct treatment isocenter and allows comparing current BEVs with those generated during simulation.

Dose delivery Precise treatment delivery is crucial for adequate tumor coverage and OAR sparing. Any deviation between the planned and delivered dose can compromise treatment efficacy or increase the risk of toxicity. Therefore, it is essential to ensure that the actual dose delivered matches the planned distribution as closely as possible.

1.2 Image-guided radiotherapy (IGRT)

In the past, large PTV margins were employed to ensure tumor control despite geometric and setup uncertainties. However, this approach often led to unwanted

irradiation of healthy tissues [19, Chapter 1]. The landscape shifted in the 1990s with the introduction of intensity-modulated radiotherapy (IMRT), which significantly improved dose conformity. This advancement enabled dose escalation to the tumor while enhancing the sparing of OARs, thereby increasing the therapeutic ratio. The improved conformity, however, came at the cost of increased sensitivity to anatomical variations and setup errors due to the steep dose gradients it introduced. To fully exploit the benefits of IMRT while mitigating these risks, accurate and reproducible patient setup became essential [20, Chapter 2]. These anatomical variations and setup uncertainties, both inter- and intra-fractional, are statistically distributed, and arise from a variety of factors, including:

- *Variations in patient positioning*: Inter-fractional setup differences depend on the rigidity and consistency of immobilization devices or on patient movement during longer treatment sessions. They also include systematic components, e.g., patients may initially be tense during planning and early treatment sessions but gradually relax over time as they become more accustomed to the procedure.
- *Variable filling states of hollow organs*: These are typically inter-fractional and particularly relevant in pelvic and abdominal treatments, e.g. rectum, bladder, and bowels.
- *Tumor response to radiation*: Certain tumors, such as lymphomas or lung tumors, may experience substantial volume reduction during the course of treatment
- *Postoperative changes in adjuvant radiotherapy*: Resorption or seromas can lead to anatomical shifts.
- *Changes in patient weight*: Weight loss or weight gain during fractionated therapy lasting 5-7 weeks can alter body contours and internal anatomy, affecting dose distribution.
- *Respiratory motion*: Intra-fractional in nature and most pronounced near the diaphragm, meaning it significantly affects the position of lower lung tumors and liver lesions.
- *Physiological motions*: Swallowing, gastrointestinal peristalsis, and cardiac motion.

To account for geometric and anatomical uncertainties, safety margins, most commonly from the CTV to the PTV, are applied during treatment planning. These margins are typically derived from population level estimates, relying on general assumptions or averaged cohort data. However, patient-specific margins can often be significantly smaller, highlighting a critical opportunity for personalization in radiotherapy. This need for individualized adaptation forms the foundation of image-guided radiotherapy (IGRT) [21, Chapter 26], which emerged in response to the demands introduced by IMRT. By integrating 3D imaging into the treatment unit and into various stages of the radiotherapy workflow, IGRT enabled the detection and correction of setup errors and anatomical changes. Thereby, it reduced geometric uncertainties and enhanced the precision of dose delivery [19, Chapter 4]. In doing so, IGRT allowed for the safe reduction of safety margins, thereby improving normal tissue sparing without compromising tumor coverage.

1.2.1 Imaging in IGRT

IGRT relies on in-room orthogonal two-dimensional (2D), or three-dimensional (3D) imaging modalities to detect and quantify anatomical and setup uncertainties, enabling corrective actions before or during radiation therapy. These imaging systems include kilovoltage (kV)-based techniques (e.g., planar orthogonal imaging, kV cone-beam CT (CBCT), CT-on-rails), megavoltage (MV)-based imaging (e.g., portal imaging, MV-CT, MV-CBCT), magnetic resonance (MR)-based systems (e.g., MR-Linac, MRI-on-rails), as well as ultrasound methods [21, Chapter 26]. Among these, CBCT has become one of the most widely used modalities in clinical IGRT workflows, as the first on-board imaging device [22]. Its clinical adoption is largely driven by its ability to generate volumetric images directly on the treatment unit, providing 3D visualization of patient anatomy in the treatment position without the necessity to move the couch and the patient [23]. CBCT enables efficient assessment of patient alignment with equivalent spatial resolution as the pCT and discernible soft-tissue contrast, all at clinically acceptable imaging doses. The resulting volumetric datasets make CBCT a practical and effective tool for routine pre-treatment verification [24, Chapter II.6].

1.2.2 Cone-beam CT

The use of imaging in radiotherapy to verify treatment field placement dates back to the 1980s and 1990s, when early techniques employed portal images and diagnostic quality X-ray field verification devices. While these methods provided

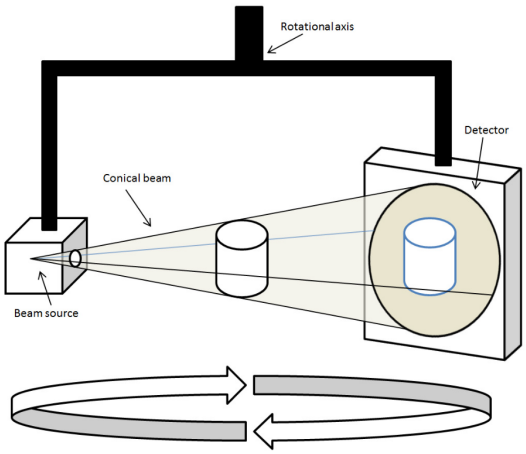


Figure 1.8. Geometrical principle of the cone-beam CT scanner [27].

limited geometric feedback, they laid the groundwork for more advanced imaging integration [25, 26]. A pivotal development occurred in 1999, when David Jaffray and colleagues introduced the first on-board system combining a kV X-ray source with a kV flat-panel imager mounted on a linear accelerator [22]. This innovation marked the beginning of kV CBCT in radiation therapy, providing volumetric imaging directly on the treatment unit.

CBCT imaging is achieved using a rotating gantry to which an X-ray source and flat-panel detector are fixed. A divergent, cone-shaped X-ray beam is directed through the center of the region of interest and captured on a two-dimensional area detector positioned opposite the source (Fig. 1.8). During a single gantry rotation – typically spanning 180 degrees or more – multiple sequential projection images are acquired over the field-of-view (FOV). This acquisition geometry contrasts with that of conventional fan-beam CT, which employs a narrow fan-shaped beam in a helical pattern to capture individual axial slices. In CBCT, a single rotational pass is sufficient for volumetric reconstruction, simplifying the mechanical requirements of the system and enabling integration with radiotherapeutic devices [23].

The widespread clinical use of CBCT has been facilitated by the development of high-resolution flat-panel detectors (FPDs), which typically rely on indirect conversion mechanisms. Similar to fan-beam CT technology, most CBCT systems employ a scintillator layer – commonly gadolinium oxysulfide (Gd_2O_2S:Tb) or caesium iodide (CsI:Tl) – to convert incident X-ray photons into visible light, which is then captured by an underlying photodiode array. Modern CsI-based scintillators are often preferred due to their columnar crystal structure, which minimizes lateral light spread and improves spatial resolution and dose efficiency. The photodiode arrays are embedded within a hydrogenated amorphous silicon (aSi:H) substrate,

with thin-film transistors integrated into the matrix to relay the stored charge as an image signal. This architecture enables high-resolution, real-time image acquisition with relatively low radiation doses compared to fan-beam CTs [28, 29].

The reconstruction of volumetric data from the acquired two-dimensional projections, known as cone-beam reconstruction, is a defining aspect of CBCTs. This process was pioneered by Feldkamp, Davis, and Kress, who expanded the fan-beam CT backprojection method to cone-beam geometry in the so-called FDK algorithm [30]. The FDK method remains a widely used approach for reconstructing CBCT volumes, offering an efficient and accurate way to convert cone-beam projection data into three-dimensional anatomical representations.

1.2.3 Registration in IGRT

In clinical IGRT, imaging alone is not sufficient to ensure accurate treatment delivery. To translate daily images into actionable information, it is necessary to correlate them with the reference images from the treatment planning phase. This process, known as *image registration*, enables geometric alignment of single- or multi-modality images (e.g., CT-CT, CT-MRI) to compare or integrate anatomical information across time points or imaging techniques [19, Chapter 5].

Image registration is generally formulated as an optimization problem, aiming to determine a transformation T that maps each point X in a reference (fixed) image I_a to the corresponding point Y in a target (floating) image I_b. Depending on the application, T may represent a rigid, affine, or non-rigid (deformable) transformation, and it can be performed in either 2D or 3D. In the context of IGRT, rigid registration is most commonly applied, as it assumes that the anatomy behaves as a single, inflexible body, allowing only six degrees of freedom (three translations and three rotations). Registration can be achieved manually, semi-automatically or automatically and the resulting transformation defines the spatial shift needed to align the patient on the treatment couch for accurate dose delivery.

1.2.4 Benefits of IGRT for prostate and lung cancer patients

On-board imaging, a major component in the IGRT process, has had a large impact on our understanding of the magnitude of the set-up error and the extent of motion during and between treatment fractions. It has for example given much more insight into the intrapatient, interpatient and intrafraction breathing motion variations and changes in the tumor position in lung and liver patients, as well as large positioning errors of the prostate. As a consequence, IGRT significantly reduces the

volume of irradiated healthy tissue and residual set-up error in prostate and lung cancer patients [31].

In prostate cancer, volumetric image guidance has enabled the evaluation of OAR geometry on a daily basis, allowing more accurate assessment of rectal and bladder filling. This is particularly important, as studies have shown reduced biochemical control in patients presenting with rectal distension in the pCT [32]. Furthermore, randomized trials have demonstrated that escalating radiation doses leads to improved biochemical control. While there is no conclusive evidence that IGRT alone improves long-term outcomes in prostate cancer, multiple studies suggest that its implementation is associated with reduced treatment-related toxicity [33].

In NSCLC, dose escalation has been shown to improve local control in both early and advanced stages. Image guidance has enabled more precise targeting, contributing to reduced rates of severe pneumonitis and improved overall survival. For early-stage NSCLC patients who are medically inoperable or refuse surgery, stereotactic body radiotherapy (SBRT) guided by volumetric imaging achieves survival outcomes comparable to those of surgical resection [33].

1.3 Adaptive radiation therapy (ART)

The next major step in the evolution of radiotherapy was introduced by Yan et al. [34], who reconceptualized treatment as a closed-loop feedback system in which the radiation plan is continuously updated throughout the treatment course in response to observed anatomical and biological changes. This paradigm, termed adaptive radiotherapy (ART), or more specifically, image-guided adaptive radiotherapy (IGART), represents a shift away from static, pre-planned treatment and toward a more responsive and individualized therapeutic approach.

A complete implementation of ART encompasses four essential components [35]:

- Treatment dose assessment – evaluating the delivered dose to the patient.
- Identifying and evaluating variation – detecting deviations from the planning baseline in anatomy or physiology.
- Decision-making – determining whether observed variations warrant modification of the treatment plan.

- Adaptive treatment modification – applying changes such as plan reoptimization, margin adjustment, or patient repositioning.

Depending on the timing of adaptation, ART can be classified into three operational modalities:

1. *Offline ART* involves acquiring periodic imaging during the treatment course. The images are then analyzed offline, allowing for manual or semi-automated segmentation and dose recalculation. Any necessary modifications are implemented in subsequent treatment sessions. This form of ART is clinically more accessible and resource-efficient, but lacks responsiveness to daily anatomical changes.

2. *Online ART* enables treatment adaptation immediately prior to dose delivery, while the patient remains on the treatment couch. This approach, however, requires efficient workflows and considerable automation. This includes fast segmentation, deformable image registration, dose computation, and plan reoptimization, which are made increasingly feasible through advances in computing power and AI-driven tools over the past decade [21, Chapter 26].

3. *Dynamic online ART* (or real-time ART) represents the most ambitious implementation. Here, treatment is adapted during actual dose delivery, accounting for intrafractional anatomical motion such as respiration or organ filling. Real-time imaging, processing, and control of beam modulation are required, making this form of ART technically demanding and currently limited to research or select clinical environments [36].

While IGRT corrects patient setup geometrically, ART goes further by adapting the dose distribution to account for changes in the patients anatomy and tumor response, thereby reducing the residual uncertainties that remain with IGRT. In lung and cancer patients these anatomical changes can occur at various time scales [37]:

- Seconds - cardiac and respiratory motion;
- Minutes - bladder filling and peristalsis;
- Days - prostate rotations, rectal filling and lung reventilation;
- Weeks - radiation induced weight loss and tumor regression.

More specifically, a differential motion of 2.5 mm was observed between the primary tumor and involved lymph nodes; a median increase in GTV of up to 35% was reported between simulation and the first treatment; and lung tumor regression during treatment ranged from 0.6% to 2.4% per day [38]. Changes in the atelectasis and pleural effusion are also regularly observed. When observing prostate cancer treatments with IGRT, most approaches are effective at correcting organ motion, however prostate rotations, which are primarily observed around the left-right axis with 5.1° systematic and 3.6° random error, remain uncorrected. In addition, radiation induced diarrhea causes rectal volume changes over the course of the treatment [37].

1.3.1 ART strategies

Over time, several practical approaches to ART have been developed to address anatomical changes during treatment. These methods differ in how and when the treatment plan is adapted ranging from preplanned adjustments based on early imaging to real-time changes made during a treatment session. Some strategies aim to capture average anatomical trends, while others respond to specific changes or follow a regular schedule. The following are commonly used approaches that illustrate the breadth of ART techniques [37]:

- Average anatomy model – similar to offline decision rules for setup error and organ motion, average anatomy model strategies rely on the quantification of anatomical changes over the first several fractions. An average anatomy model can be estimated by: deformable registration of the planning scan to the scans of the initial fractions; calculating the average deformation vector field and; deforming the planning scan and corresponding structures accordingly to obtain a synthetic scan representing the average anatomical configuration. A new treatment plan can subsequently be optimized on the average anatomy model. Optimal number of fractions ranges from about 5 to 12.

- Library of plans – different plans are typically made on a single CT scan with interpolated and possibly even extrapolated contours obtained from additional scans and using deformable registration. Alternatively, a library of plans can be made on different CT scans or using multiple CBCTs of the first few fractions.

- Triggered adaptation – most common form of offline adaptation, predating the term of adaptive radiotherapy. Triggered adaptation refers to the pro-

cess of adapting the treatment plan when a certain threshold is exceeded, e.g., when the patient experiences considerable anatomical changes such as weight loss. In most cases, the criteria for triggering an adaptation are qualitative, however they can also be quantitative by recalculating the treatment plan on daily images and accumulating the dose.

- Scheduled adaptation – scheduling the time points when the adaptive process will happen. Rescanning can occur once or several times (weekly).
- Online replanning – daily online replanning can account for systematic and random complex anatomical changes as well as time trends, since the treatment plan is reoptimized for every fraction. Therefore, it requires high image quality and both fast delineation and fast replanning within a single fraction, which has only been possible of late [39, 40].

1.3.2 Benefits of ART for prostate and lung cancer patients

ART has shown clear clinical benefits in prostate cancer treatment, with consistent improvements in precision, target coverage, and OAR sparing. Early clinical experience demonstrated that ART reduced systematic and random errors due to organ motion by half, enabling a 29% reduction in PTV volume and lowering the volume of rectal tissue receiving high doses, as well as an average dose reduction of 4.8 Gy to the anal wall [41]. ART also minimized the risk of geometric miss caused by rectal distension, leading to excellent biochemical control and very low rates of chronic gastrointestinal toxicity [42]. As techniques advanced, adaptive plans began to outperform scheduled ones in 78% of treatment fractions, offering better dosimetric quality while maintaining a practical workflow [43]. Clinical implementation further confirmed improvements in PTV coverage, with adaptive plans consistently maintaining dose constraints for the bladder and rectum in the majority of fractions [44]. Most recently, ART has been shown to significantly improve dose coverage and reduce variability in both target and OAR dosing, contributing to more consistent and reliable treatment delivery across fractions [45].

In addition to the benefits observed in prostate cancer, ART has also shown considerable promise in the treatment of lung cancer patients. Early work highlighted that incorporating soft tissue tumor matching and adaptive strategies could improve loco-regional control without increasing treatment-related toxicity [46]. Replanning based on anatomical changes during treatment, such as tumor shrinkage – though its reliability remains a subject of ongoing clinical debate, enabled better sparing of normal tissues like the lungs, heart, and spinal cord, while maintaining

the potential for dose escalation to the tumor [47]. CBCT imaging has been instrumental in identifying significant tumor volume changes during treatment, helping to select patients who may benefit from ART [48], and has been effectively used to reduce doses to OARs without compromising target coverage [49]. More recent studies have demonstrated that ART can reduce lung dose even without altering safety margins, thereby supporting either safer treatment or dose escalation strategies [50], while also significantly lowering the incidence of radiation pneumonitis and improving both progression-free and overall survival rate [51].

1.4 CBCT in ART

CBCT is currently the most widespread imaging technique used in the ART context. As such, the successful implementation of ART heavily relies on the quality of its images and processes that build on them such as [20, Chapter 2]:

- Optimized volumetric imaging protocols that ensure high-quality and timely anatomical data;
- Efficient CBCT segmentation techniques that allow for real-time or near real-time anatomical modeling;
- CBCT-based dose calculation and dose reconstruction, to evaluate deviations between planned and delivered dose distributions.

Despite its widespread adoption, CBCT has several intrinsic limitations that impact image quality and hence clinical utility in ART scenarios. The limitations of CBCT come from several factors occurring during image acquisition, image reconstruction, and some that are intrinsic to the CBCT technology. These limitations are called artifacts and they can be defined as visualized structures in the reconstructed data that is not present in the object under investigation [52].

1.4.1 Limitations of CBCT

Scatter radiation and image noise The cone-shaped beam employed in CBCT irradiates a large volume with each projection, resulting in a significant proportion of photons undergoing Compton scattering. The scattered radiation, which is detected omnidirectionally, contributes to image noise and does not accurately represent the attenuation profile along a defined beam path [53]. This effect is worsened by the large FOV, leading to a higher scatter-to-primary ratio than in fan-beam

CT systems [54]. Scatter is the principal source of image degradation in CBCT, producing artifacts such as cupping, inhomogeneous contrast, and noise enhancement. These artifacts compromise spatial uniformity and reduce the reliability of HU values, ultimately degrading soft tissue visibility and low-contrast resolution. Noise in CBCT images arises from both quantum (Poisson) and electronic (Gaussian) sources. Quantum noise is particularly elevated due to the use of low tube currents intended to reduce imaging dose. The resulting low photon count decreases the signal-to-noise ratio (SNR), which in turn causes greater variability in reconstructed attenuation values and HU numbers. Consequently, low-contrast resolution is impaired and diagnostic information may be lost [55].

Cone-beam artifacts The divergent nature of the cone-shaped X-ray beam can introduce additional artifacts, particularly at the periphery of the scanned volume. Detector pixels at the outer edges receive fewer projection data, leading to reduced sampling, image distortion, streaking, and increased peripheral noise. This phenomenon, known as the *cone-beam effect*, results from the non-uniform acquisition geometry and reduced attenuation sampling in peripheral detector rows [53].

Beam hardening effects As the polychromatic X-ray beam traverses the patient, lower-energy photons are preferentially absorbed, resulting in beam hardening. This leads to artifacts such as cupping, characterized by artificially low attenuation values in the center of homogeneous objects, and dark streaks or bands between high-density structures [54].

Detector and systemic limitations Flat-panel detectors used in CBCT systems are prone to non-linear responses and systemic artifacts that compromise their ability to accurately reflect X-ray attenuation. The heel effect and limitations in detector calibration further impair soft tissue contrast and spatial accuracy. Combined with the elevated scatter, these system-level characteristics restrict the utility of CBCTs in tasks requiring precise HU calibration [53,54,56].

Motion artifacts CBCT image acquisition involves longer gantry rotation times (5 s to 40 s) relative to diagnostic CT, increasing susceptibility to motion artifacts. Patient movement during scanning, especially of high-contrast structures such as bone or air cavities, results in streaking and image blurring, which can obscure critical anatomical boundaries [54,56].

1.4.2 Overcoming CBCT limitations

Each of the artifacts mentioned in the previous section have a potential to prevent the proper use of CBCT in ART whether it comes to semi-automatic and automatic segmentation tools, or dose calculation algorithms that require accurate HU distribution. Therefore, since the very beginning of clinical use of CBCTs for diagnostics and radiotherapy, there have been attempts to overcome the limitations arising from CBCT artifacts.

Hardware and software-based CBCT correction

Multiple strategies have been explored to reduce the impact of scattered radiation on image quality. These include physical methods such as enhanced collimation [57], the use of bowtie filters [58] and the implementation of antiscatter grids [59, 60]. Complementary efforts have focused on post-processing techniques aimed at correcting scatter-induced degradation, including convolution-based approaches [61], Monte Carlo simulations [62] and, more recently, deep learning algorithms [63]. Additional work has addressed other image quality issues, such as beam hardening [64] and motion artifacts [65]. With the help of these correction methods, the CBCT image quality has improved drastically, allowing advanced autosegmentation tools to be developed [39]. However, the distribution of HU values in the images is not yet at a point where it could reliably be used for direct dose calculation and treatment planning.

CBCT enhancement

Numerous enhancement techniques have been developed to exploit the geometric accuracy of CBCTs for ART. They aim to use the correct position and shape of organs in daily imaging while compensating for the poor image quality and unreliable HU values inherent to CBCT. Some of these techniques are:

- Electron density calibration – assigning electron densities based on an HU calibration curve, that can be population-based or patient-specific, and voxel-based or image slice-based [66–70].
- Bulk density override – assigning fixed densities to anatomical regions (e.g., air, water, bone). Density override techniques partially circumvent the noise and artifacts in CBCT images and produce better agreement with pCT-based dose calculations compared to population-based conversion curves [71–73].

- Histogram matching – adjusting the CBCT so that its histogram aligns with that of a pCT. It improves the HU value distribution, particularly at tissue interfaces such as soft tissue and air [74,75].

Deformable image registration (DIR)

An alternative method to navigate the problems of using CBCT images for dose calculation, and the one that will be discussed and investigated in more detail in this thesis, is deformable image registration (DIR). DIR algorithms have been developed as a method to approximate electron density maps which reflect the daily patient anatomy [76–78]. In contrast to rigid registration, they are based on establishing a spatial correspondence between different image acquisitions (pCT – source image, and CBCT – target image) by using non-linear dense transformations, or a spatially varying deformation model [79]. DIR can therefore account for local anatomical changes such as tumor shrinkage, organ motion, or weight loss, and thus enable non-uniform, region-specific adjustments (Fig. 1.9). As a result, we can create a synthetic CT (synCT) image with anatomical characteristics of a CBCT and HU value distribution and image quality of a pCT.

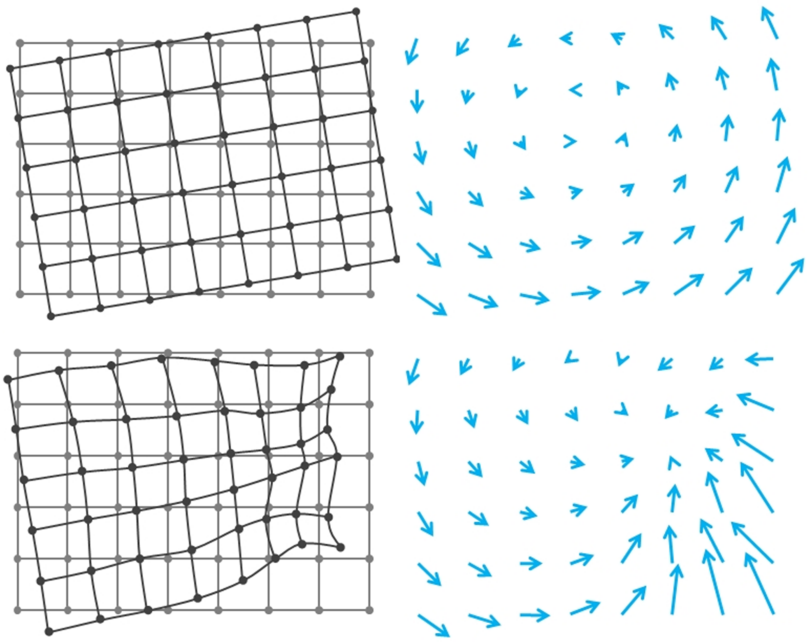


Figure 1.9. Graphical representation of rigid (top) and deformable (bottom) image registration, with the resulting deformation vector fields on the right.

Existing DIR methods can be classified into parametric or model-based, and non-parametric. Parametric methods include B-spline, thin-plate spline, and linear elastic finite element, while non-parametric methods include optical flow, viscous

fluid, etc. [80]. Most existing algorithms rely primarily on image intensity similarities and do not incorporate the physical properties of tissues. This can result in anatomically implausible outcomes, such as non-physiological deformations (e.g., bone warping). To address this, regularization techniques and the inclusion of prior anatomical or physiological knowledge are increasingly being used to constrain the deformation field and improve realism in the registration process [81].

Voxel-to-voxel correspondence provided by deformable registration has wide applications in ART. It is not merely used to create synCTs, but also for improved tumor target definition, image enhancement, propagation of anatomical contours across image sets, and calculation of accumulated dose in deforming organs [20, Chapter 2].

While widely implemented, DIR can be unreliable in cases of significant anatomical changes (e.g., organ filling, air pockets, atelectasis, tumor shrinkage), it is sensitive to the variable quality of the target image, and it lacks robust quality assurance methods to ensure the accuracy and clinical acceptability of the resulting deformed image [66,80,82–84]. Given these limitations, recent advances have turned towards alternative strategies that do not rely on deformation-based mapping.

Neural networks for synthetic CT generation

Lately, generative artificial intelligence (AI) models have opened an avenue for transforming lower-quality CBCT images into high-quality synCT images in a process called image-to-image translation. Such models preserve the daily anatomy while approximating the electron density characteristics of the pCT. Among a wide variety of neural network architectures [85–90], CycleGAN-based models [91] have shown promising results, with the added advantage of enabling training on unpaired images, reducing training instability, and mitigating generative hallucinations. They have also demonstrated performance that matches, and in some cases surpasses, CBCT enhancement and DIR-based techniques [92–96]. The deep learning basis of these models will be further explored in the following chapter and the utility of synCT images generated with image-to-image translation networks will be thoroughly discussed in the rest of the thesis.

2

Deep Learning Background

2.1 Motivation

Given the increasing role of AI in medical imaging, deep learning has become increasingly important in modern radiotherapy research. As highlighted in the previous chapter, generating synCTs through image-to-image (I2I) translation is one such application where deep learning plays a central role. To fully understand the mechanisms behind synCT generation and the design decisions behind these models, a foundational understanding of deep learning is essential. This chapter introduces the key concepts and architectures underpinning I2I translation, beginning with artificial neural networks, and progressing through convolutional neural networks, generative adversarial networks, and finally, the CycleGAN – an architecture at the center point of the thesis.

2.2 Artifical neural network

All I2I networks are artificial neural networks (ANNs), a class of machine learning models loosely inspired by the structure and function of the human brain [97, 98]. ANNs consist of layers of interconnected computational units called *neurons*. Typically, an ANN receives a vectorized input that is passed into the input layer. The information is then propagated through subsequent *hidden layers* through a series of mathematical operations. These operations occur at each neuron and usually involve a linear combination of the input values and the internal parameters of the neuron, known as *weights* and *biases*. This process can be expressed as

$$y = f(x) = \sum_{i=1}^{n} w_i x_i + b, \tag{2.1}$$

where f represents the neuron operation, x is the vectorized input, x_i are individual input features, w_i are the corresponding weights, and b is the bias term (see Fig. 2.1). The output of the neuron y is then passed through a non-linear activation function, such as Rectified Linear Unit (ReLU) or sigmoid (Fig. 2.2), allowing the network to model complex, non-linear relationships. The processed information flows through subsequent layers toward the final layer, forming a fully connected neural network (FCNN), an ANN in its simplest form (Fig. 2.3). At the final layer the outputs of the network can be read out and interpreted. This directional flow of data from the input to the otput layer is referred to as the *forward pass*.

To learn a task, an ANN must be trained on a dataset consisting of correlated input-output pairs. After each forward pass, the output of the network is compared to the expected output, also known as the *ground truth*, using a *loss function* that

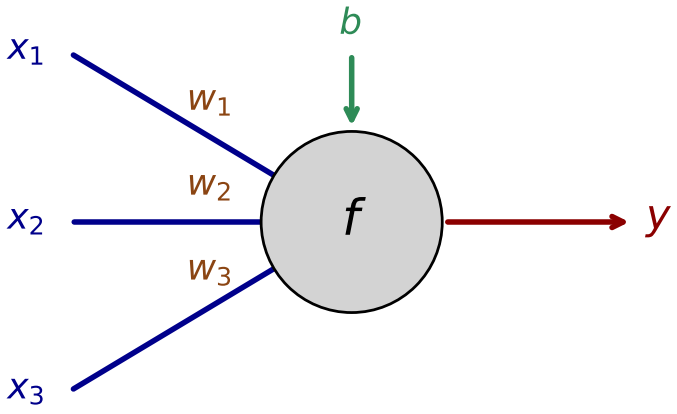


Figure 2.1. Schematic representation of a neuron in an ANN showing three input features x_1, x_2, and x_3, three weights w_1, w_2 and w_3, a bias term b, and an output y.

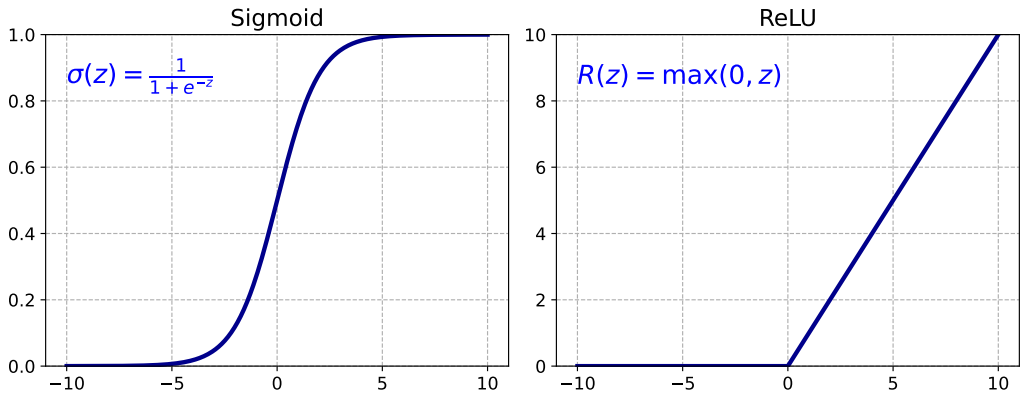


Figure 2.2. Visualization of two commonly used activation functions in artificial neural networks. The sigmoid function (left) forces input values into the range (0, 1), enabling smooth, non-linear transformations. The ReLU function (right) outputs zero for negative inputs and a linear response for positive inputs.

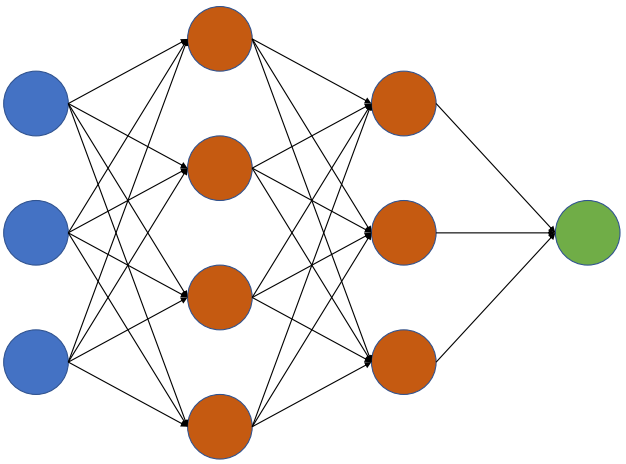


Figure 2.3. Schematic representation of a fully connected neural network and the forward pass, with input (blue), hidden layers (orange), and the output (green).

quantifies the error. This error is then propagated backward through the network to receive gradients of the weights in a process known as *backpropagation*. The network then adjusts its weights and biases using optimization algorithms such as *stochastic gradient descent*, gradually improving its performance over time.

2.3 Convolutional neural network

In ANNs designed to process image data, one of the most widely used architectures is the convolutional neural network (CNN). While CNNs share many core principles with simpler FCNNs, such as learning through layers of weighted connections and backpropagation, they are optimized to handle the spatial structure of visual data [99].

CNNs typically consist of repeated sequences of convolution, non-linear activation, and pooling layers, followed by one or more fully connected layers that map the learned features to the output. This layered structure allows CNNs to learn increasingly abstract and complex representations as the depth of the network increases, forming the foundation for tasks like image classification.

Convolutional layers are the core components of CNNs, responsible for extracting spatial features from image data. They apply small, learnable matrices called kernels, or filters, that slide across the input image, performing an element-wise multiplication with the pixel values they cover. The results are then summed to produce a single value in the output feature map (see Fig. 2.4). This operation, known as convolution, enables the network to detect meaningful visual patterns such as edges, textures, and shapes.

Following each convolutional layer, a pooling layer is used to reduce the spatial dimensions of the feature maps. Pooling operations, such as max pooling or average pooling, summarize small regions of the input by selecting either the maximum or average value of the feature map, respectively. This reduces the number of parameters, improves computational efficiency, and helps make the network more robust to small translations or distortions in the input [98].

Next, a non-linear activation function, most commonly ReLU, is applied which introduces non-linearity into the model.

Finally, the resulting feature maps are flattened into a one-dimensional vector and passed through one or more fully connected layers, similar to those used in standard ANNs. These dense layers combine the learned features to perform the final task, e.g., classification (Fig. 2.6).

By capturing local patterns and maintaining the spatial structure of the input

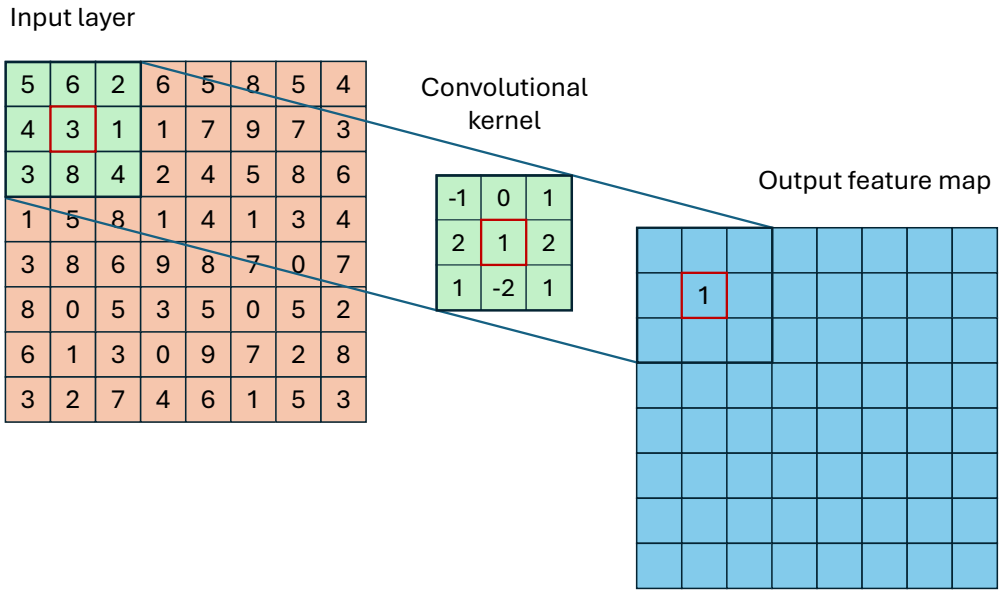


Figure 2.4. Visualization of a convolution operation. A 3×3 kernel is applied to a 3×3 region of the input (source layer) by performing element-wise multiplication and summing the results. The output is written to the corresponding position in the destination layer. This example illustrates how local features are extracted through convolution.

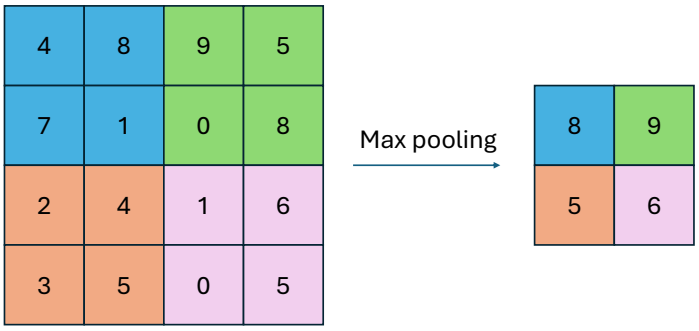


Figure 2.5. Example of a max pooling operation. The maximum value is selected from each 2×2 region, reducing spatial dimensions while preserving important features.

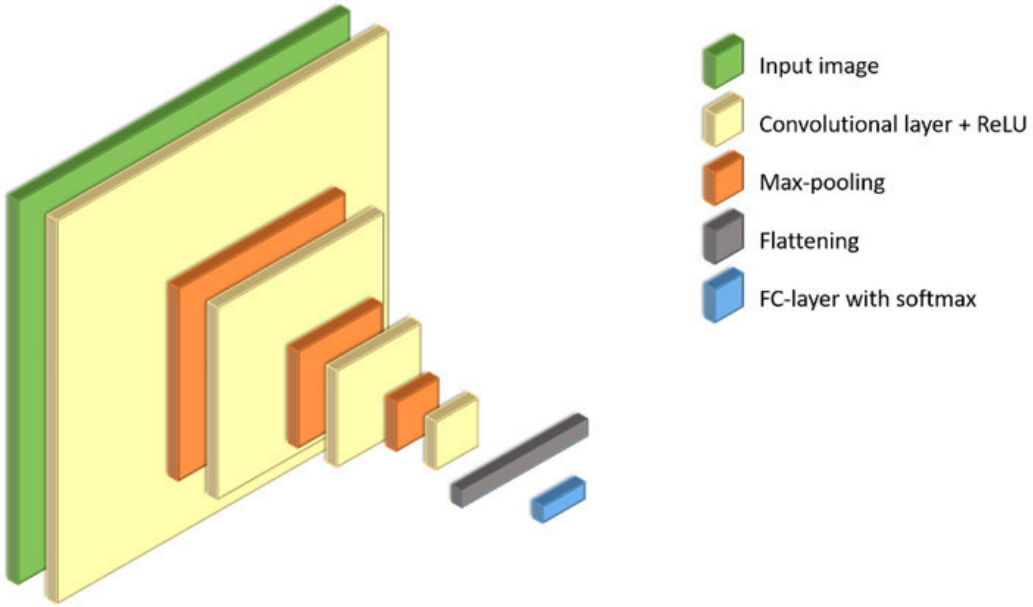


Figure 2.6. Overview of a convolutional neural network (CNN). The input image is processed through convolutional, pooling, and activation layers to extract features, followed by fully connected layers for prediction [100].

data, CNNs have proven highly effective across a wide range of computer vision tasks. However, as deeper architectures with a larger number of layers were explored to improve performance, a problem of vanishing gradients appeared, where gradients become too small during backpropagation to effectively update earlier layers. This issue hindered the training of very deep networks and limited their practical utility. To address this, Residual Networks (ResNets) were introduced [101], proposing the use of residual connections, or skip connections, which allow gradients to bypass one or more layers and propagate directly to earlier parts of the network (Fig. 2.7).

These connections prevent the vanishing gradient issue, enabling the stable training of much deeper CNNs and improving performance across a variety of vision tasks. As a result, ResNet-style architectures have become foundational components in many modern CNN-based systems, particularly in deep generative models and I2I translation architectures.

2.4 Autoencoder architecture

A widely adopted ANN architecture for I2I tasks is the autoencoder. In this setup, the encoder is typically implemented using a down-sampling CNN, which progressively reduces the spatial dimensions of the input image while increasing the

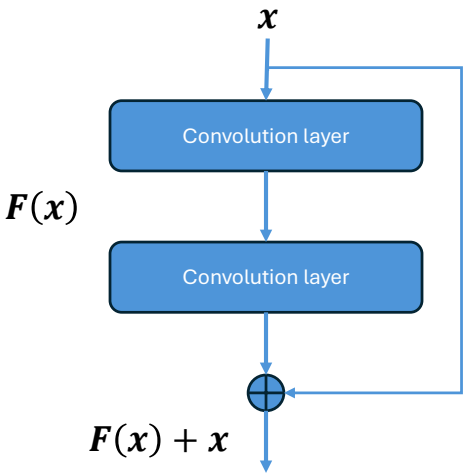


Figure 2.7. Residual layer with a skip connection, allowing input to bypass one or more convolutional layers. This helps keep gradients from vanishing and enables the training of deeper networks.

number of feature channels. The down-sampling process helps the network focus on high-level semantic features by compressing the image into a compact latent representation [102]. The decoder, on the other hand, is an up-sampling CNN that performs the reverse operation. It gradually reconstructs the spatial resolution of the image from the latent space. The up-sampling process aims to restore the original image size and structure, ultimately generating an output image from the encoded features [103]. Autoencoders are used in tasks such as image denoising, segmentation, and synthesis, where the spatial structure and content of the input must be transformed or regenerated in a meaningful way.

Among the most powerful autoencoder variants is the U-Net, which extends the basic autoencoder structure with skip connections between encoder and decoder layers [104]. These connections allow low-level spatial features from the encoding path to be reused in the decoding path (see Fig. 2.8). This leads to more spatially coherent outputs which is particularly important in tasks such as biomedical segmentation. While U-Net uses skip connections to preserve spatial detail, residual connections, as used in ResNets, are another technique that improves information flow in deep networks. They are now commonly integrated into autoencoder architectures for image synthesis tasks.

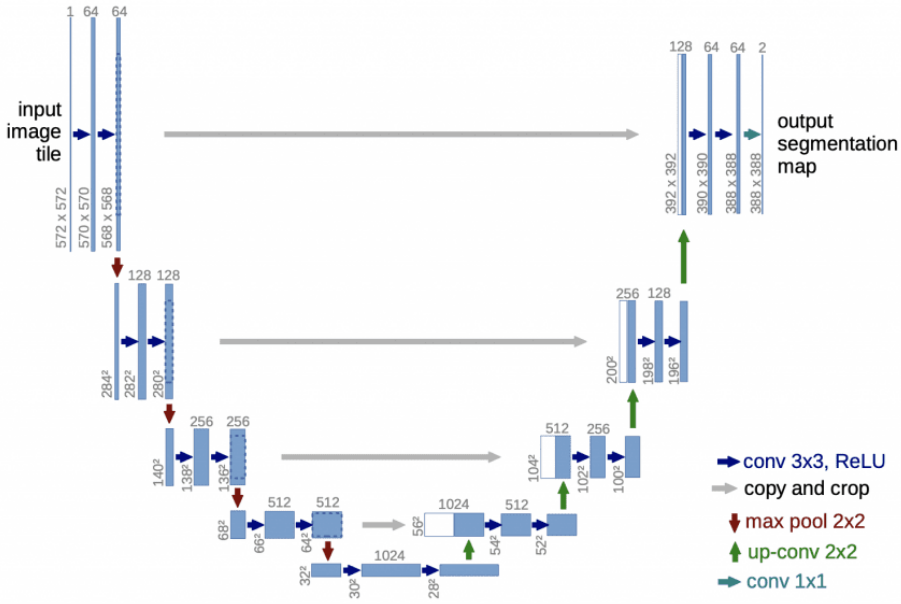


Figure 2.8. U-Net architecture for image segmentation, featuring a symmetric encoder-decoder structure (blue blocks) with skip connections that transfer spatial information from down-sampling to up-sampling paths (gray arrows) [104].

2.5 Generative adversarial network

The first major breakthrough in the area of I2I translation was introduced in the form of a *generative adversarial network (GAN)* [105]. The model is conceptualized with having two competing neural networks, a *generator* G, and a *discriminator* D. In the context of this thesis, the generator is a deep autoencoder CNN (e.g., U-Net or ResNet), and the discriminator is a deep CNN similar to the one depicted in Fig. 2.6. In the original formulation, the goal of the generator G is to learn a distribution $p_g(x)$ that closely approximates the true data distribution $p_{\text{data}}(x)$, starting from a random noise input $z \sim p_z(z)$. The discriminator D, on the other hand, aims to distinguish whether a given sample comes from the real data distribution p_x or from the generated distribution p_g. During training, the discriminator is optimized to *maximize* the probability of correctly identifying real versus generated samples, 1 versus 0, respectively. Formally, D is trained to maximize the following objective

$$\mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))]. \tag{2.2}$$

At the same time, the generator is trained to *minimize* this objective. In practice, this is often done by minimizing the second term alone

$$\mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))], \tag{2.3}$$

which encourages G to produce outputs that D cannot distinguish from real data. This relationship can be seen as D and G playing a two-player minimax game

which is described by

$$\min_G \max_D \left\{ V(D,G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))] \right\}, \quad (2.4)$$

where $V(D,G)$ is the loss function [105].

At the start of training, the generator produces synthetic data that does not resemble real examples. Although this might seem like an easy classification task for the discriminator, it too is untrained at this stage and struggles to distinguish real from fake data. As training progresses and both networks are exposed to more examples, the discriminator becomes better at telling real and fake apart, while the generator simultaneously improves at creating samples that more closely match the real distribution.

This interaction can be thought of as a competitive game, like a forger and a detective: in the beginning, the forgeries are poor and easily spotted, but the detective is inexperienced. Over time, as the detective becomes better at spotting fakes, the forger also improves their craft in an attempt to fool the detective. The key to successful training is maintaining a balance, if one becomes too powerful too quickly, the other stops learning effectively. The goal is for both to improve in sync, pushing each other toward better performance.

The setup described in this section is not directly sufficient for I2I tasks as the images are generated from random noise samples. However, it forms the foundation for more advanced architectures such as conditional GANs (cGANs) [106], where the generator is conditioned on an input image to produce a corresponding output. These conditional variants are essential for I2I tasks, where the goal is to translate an image from one domain (e.g., CBCT) to another (e.g., pCT).

2.6 CycleGAN

One of the main challenges in training GAN-based networks for medical image translation is the scarcity of paired data. To create synCT images from daily CBCTs we would need to construct a dataset of aligned pCT and CBCT image pairs. That would require either perfectly consistent anatomy across acquisitions or the application of DIR to spatially align the images. As discussed in Section 1.4.2, DIR has its own limitations and uncertainties. Therefore, we require a model capable of learning from unpaired data drawn from two different domains.

A notable solution to this problem is the CycleGAN architecture [91], which enables training on unpaired data by using a cycle-consistency constraint. CycleGAN consists of two generator-discriminator pairs: one generator G_{A2B} ($G_{A2B} : A \rightarrow B$)

translates images from domain A (e.g., CBCT) to domain B (e.g., pCT), while the other generator G_{B2A} ($G_{B2A} : B \rightarrow A$) performs the reverse translation. Correspondingly, two discriminators D_A and D_B evaluate the realism of generated images in each domain.

The overall training objective includes three loss components:

- Adversarial loss $\mathcal{L}_{GAN}$ – defined as $V(G, D)$ in Eq. (2.4), is used to match the distribution of generated images to the distribution of the images in the target domain.

- Cycle-consistency loss $\mathcal{L}_{cyc}$ – newly introduced term that guarantees the mapping from an input image to a desired output. Therefore, translated images $G_{A2B}(I_a)$ and $G_{B2A}(I_b)$ are fed back into the opposite generators forming a forward and a backward cycle. The output of such an image translation cycle should bring I_a back to the original image $I_a \rightarrow G_{A2B}(I_a) \rightarrow G_{B2A}(G_{A2B}(I_a)) \approx I_a$, and vice versa for I_b. In the final form $\mathcal{L}_{\text{cyc}}$ is written as

$$\begin{aligned}\mathcal{L}_{\text{cyc}}(G_{A2B}, G_{B2A}) = \mathbb{E}_{I_a \sim p_{\text{data}}(I_a)} \left[|G_{A2B}(G_{B2A}(I_a)) - I_a|\right] \\ + \mathbb{E}_{I_b \sim p_{\text{data}}(I_b)} \left[|G_{B2A}(G_{A2B}(I_b)) - I_b|\right].\end{aligned} \tag{2.5}$$

- Identity loss $\mathcal{L}_{identity}$ – encourages each generator to preserve the image if it is already from the target domain, e.g., generator G_{A2B} should perform an identity operation if an image from domain B is used as input

$$\begin{aligned}\mathcal{L}_{\text{identity}}(G_{A2B}, G_{B2A}) = \mathbb{E}_{I_a \sim p_{\text{data}}(I_a)} \left[|G_{A2B}(I_b) - I_b|\right] \\ + \mathbb{E}_{I_b \sim p_{\text{data}}(I_b)} \left[|G_{B2A}(I_a) - I_a|\right].\end{aligned} \tag{2.6}$$

The final form of the loss function can be written as

$$\begin{aligned}\mathcal{L}(G_{A2B}, G_{B2A}, D_A, D_B) = \mathcal{L}_{GAN}(G_{A2B}, D_B) + \mathcal{L}_{GAN}(G_{B2A}, D_A) \\ + \lambda_1 \mathcal{L}_{\text{cyc}}(G_{A2B}, G_{B2A}) \\ + \lambda_2 \mathcal{L}_{\text{identity}}(G_{A2B}, G_{B2A}).\end{aligned} \tag{2.7}$$

Here, λ_1 and λ_2 are weighting factors that control the importance of the cycle-consistency and identity losses, respectively.

The schematic representation of the network architecture is depicted in Fig. 2.9, and the exact configuration of the network that was used in the thesis will be described in more detail in Chapter 4 together with the dataset used for training and the training pipeline.

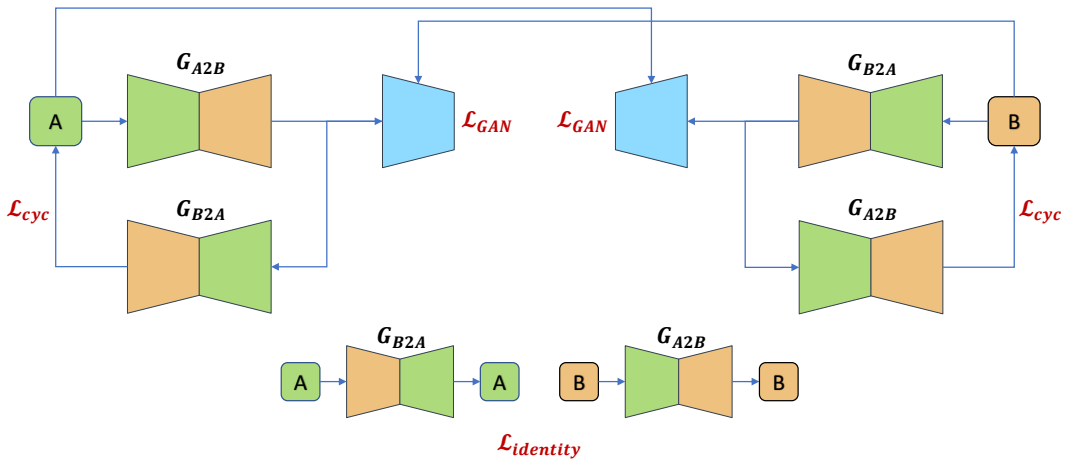


Figure 2.9. Schematic representation of the CycleGAN architecture where, thanks to a cycle-consistency constraint, two generators and two discriminators are used to translate images between two domains without needing paired data. Same colors (orange and green) represent same domains. Discriminators are visualized in blue.

2.7 Image quality evaluation and metrics

Evaluating the performance of a neural network trained on unpaired data can be challenging due to the absence of a clear ground truth. Acquisition-wise, pCT and CBCT images are often days apart and registering them deformably has its limitations (Section 1.4.2).

To overcome this, we rely on a combination of quantitative metrics and qualitative visual inspection. Each type of metric has its strengths and limitations, so using a mix of them often leads to a more reliable conclusion.

2.7.1 Pixel-wise metrics

These metrics compare two images by measuring the differences between their corresponding pixel values. They are straightforward but can be misleading if the images are misaligned – a common case in medical images when we want to compare scans of the same patient on two different days.

Mean absolute error (MAE) MAE measures the average absolute difference between corresponding pixel values in two images I and $\hat{I}$:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |I_i - \hat{I}_i| \,. \tag{2.8}$$

Mean squared error (MSE) MSE calculates the average of squared differences between pixels in two images I and $\hat{I}$, emphasizing larger errors:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(I_i - \hat{I}_i)^2 \,. \tag{2.9}$$

Peak signal-to-noise ratio (PSNR) Derived from MSE, PSNR measures the quality of the image in terms of the relationship between signal and noise. It is defined by

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}}\right), \tag{2.10}$$

where MAX is the maximum possible pixel value.

2.7.2 Information-based metrics

These metrics assess similarity based on the distribution of intensity values, making them more robust to small misalignment like shifts or rotations.

Normalized mutual information (NMI) NMI measures the amount of shared information between two images. It is often used in image registration and it is described by

$$\text{NMI}(X, Y) = \frac{H(X) + H(Y)}{H(X, Y)}, \tag{2.11}$$

where $H(X)$ and $H(Y)$ are marginal entropies of images X and Y and $H(X, Y)$ is the joint entropy. Higher values indicate more similarity.

Normalized cross-correlation (NCC) NCC evaluates how well one image correlates with another, normalized by their variances. It is given by

$$\text{NCC}(X, Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \cdot \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}, \tag{2.12}$$

where $\bar{X}$ and $\bar{Y}$ are mean pixel intensities of images X and Y. A value close to 1 indicates high similarity.

Jensen-Shannon divergence (JSD) A symmetric version of the Kullback-Leibler (KL) divergence, JSD measures the similarity between two probability distributions. It is given by

$$\text{JSD}(P \parallel Q) = \frac{1}{2}D_{\text{KL}}(P \parallel M) + \frac{1}{2}D_{\text{KL}}(Q \parallel M), \tag{2.13}$$

where P and Q are distributions of the two domains, $M = \frac{1}{2}(P + Q)$ and $D_{\mathrm{KL}}(P \parallel Q) = \sum_i P(i) \log\left(\frac{P(i)}{Q(i)}\right)$ is the KL divergence. Metric values are bounded by 0 and 1 and lower values suggest closer similarity.

2.7.3 Anatomical consistency metrics

When anatomical similarity is important, such as preserving organ or body contour shapes, we can use several distance metrics suited for the task. For these metrics we need to have tumor and OAR segmentations at our disposal.

Hausdorff distance (HD) HD measures how far two subsets of a metric space are from each other, or in our case, the maximum distance between the edge points of two contours [107]. It is defined as

$$\mathrm{HD} = \max\left\{\sup_{x \in X} d(x, Y), \sup_{y \in Y} d(X, y)\right\}, \tag{2.14}$$

where sup represents the supremum operator, inf the infimum operator, and where $d(a, B) := \inf_{b \in B} d(a, b)$ quantifies the distance from a point $a \in X$ to the subset $B \subseteq X$. A smaller Hausdorff distance means better overlap and anatomical consistency. In the context of this thesis, a 95th percentile was chosen for the evaluation of HD (HD_{95}).

Sørensen-Dice coefficient (DICE) DICE measures the overlap between two discrete sets, in our case, the volumetric overlap between two contours [108, 109]. It is given as follows

$$\mathrm{DICE} = \frac{2|X \cap Y|}{|X| + |Y|}, \tag{2.15}$$

where X and Y are contours of the same structure from two different images. Values closer to 1 indicate better overlap and values closer to 0 worse overlap.

Surface DICE coefficient (SDC) SDC measures the overlap between two surfaces within a certain tolerance margin [110]. It is defined as

$$\mathrm{SDC(X,Y)} = \frac{|S_X \cap B_Y^{(\tau)}| + |S_Y \cap B_X^{(\tau)}|}{|S_X| + |S_Y|}, \tag{2.16}$$

where $B_i^{(t)}$ is the border region of the surface S_i at a given tolerance τ, and S_i is the boundary of the segmentation mask. In this thesis, a tolerance margin of 2 mm was chosen based on the clinical practice in photon radiation therapy to intervene when

deviations are in the order of 2 mm or larger [111]. Metric values are bounded by 0 and 1 and higher value of SDC indicates better surface-wise anatomical matching between structures.

2.7.4 Structural or perceptual metrics

Structural similarity index measure (SSIM) Perceptual metric that quantifies image similarity by comparing structural information, contrast, and luminance between two images. Unlike pixel-wise metrics, SSIM models the way humans perceive image quality. It is described by

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{2.17}$$

where μ_x and μ_y are mean intensities of images X and Y, σ_x^2 and σ_y^2 are variances of X and Y, σ_{xy} is the covariance, and C_1 and C_2 are empirical constants that stabilize the division when denominators are small. In medical imaging, a high SSIM score suggests that the synthetic image preserves important structural details relevant for clinical interpretation. The values are bounded between 0 and 1.

2.7.5 Population-based metrics

Some metrics like Inception Score (IS) [112] and Fréchet Inception Distance (FID) [113] are widely used in computer vision for comparing image distributions. However, they require large sample sizes and are not well suited for smaller medical datasets, so they are mentioned for completeness but are not included in the evaluation of images in this thesis.

2.7.6 Qualitative evaluation

In addition to quantitative metrics, visual inspection remains a crucial part of image quality assessment, especially in the medical domain. One common approach is to plot **difference maps** between the synthetic and reference images to highlight regions of change or error. **Histogram overlays** of HU distributions can reveal whether tissue intensities are being realistically reproduced. **Line profiles** drawn across specific anatomical regions can further show how well structural edges and intensity gradients are preserved. Finally, **direct visual assessment** by medical professionals or researchers can help judge the degree of **artifact reduction**, **anatomical consistency**, and **preservation of clinical detail**, all of which are essential for real-world applicability.

3

Limitations of Dose Calculation on DIR-Based CT and CBCT

3.1 Motivation

Chapter 1 outlined several challenges that limit reliable use of CBCT or DIR-based CT for treatment planning and dose calculation. DIR methods often fail to capture large anatomical changes, and clinicians usually cannot detect or investigate these failures because robust quality assurance techniques do not exist in the registration domain. At the same time, even the most advanced CBCT enhancement and correction techniques still struggle to overcome increased scatter, noise, beam hardening, and motion artifacts.* These limitations can lead to suboptimal dose delivery, potentially compromising treatment outcomes. Manual intervention and replanning, the current fallback strategies, delay treatment and place additional burden on clinical resources.

Deep learning-based synthetic CT offers a third pathway that could overcome both limitations by producing high-quality images suitable for dose calculation and adaptation. This chapter experimentally demonstrates the limitations of CBCT- and DIR-based dose calculation and motivates integration of synthetic imaging into ART workflows. Weekly treatment fractions from lung cancer cases are analyzed, and direct dose calculations on both CBCT and deformed CT images are performed. These experiments highlight time-dependent anatomical changes, their impact on dose distributions across modalities, and the potential of synthetic CT to better represent anatomical variations typically seen in adaptive lung radiotherapy.

3.2 Methods

3.2.1 Dataset

Longitudinal radiotherapy datasets from 16 consecutive lung cancer (LungCa) patients treated at the University Clinic Heidelberg (UKHD, Heidelberg, Germany) and the German Cancer Research Center (DKFZ, Heidelberg, Germany) within the MARS trial [40] were collected for this investigation. All patients agreed to pseudonymized analyses and anonymized publication. Data from the site contained both SCLC and NSCLC cases and it was collected from patients who had undergone treatment on the Ethos™ device (Varian, a Siemens Healthineers company, Palo Alto, CA, USA) at the DKFZ. Fractionation schemes varied among patients, in-

*The datasets used in this study were acquired prior to the introduction of Hypersight™, the current best-in-class CBCT imager by Varian, a Siemens Healthineers company. Images obtained with this new detector may offer significant improvements and hold promise for advancing ART practices.

Table 3.1. Gender representation, disease stage, and age of LungCa patients used in the study.

Site	Gender (Male/Female)	Disease stage	Age (Mean, range)
LungCa	6/10	T1-T4	70, 50-82

cluding 15 fractions (3 Gy/fraction), 25 fractions (1.8 Gy/fraction), 30 fractions (2 Gy/fraction) and 33 fractions (2 Gy/fraction). An overview of characteristics of LungCa patients included in the study is shown in Table 3.1.

One fraction from each week of treatment was selected and two types of image scans were collected from each fraction for the investigation:

- CBCT – acquired with the Ethos™ on-board kV CBCT imaging device with 125 kVp tube voltage, 2 mm slice thickness, 0.96 mm pixel spacing, and 18 to 20cm FOV size in the cranio-caudal direction.
- Deformed CT (dirCT) – created using the Velocity™ deformable image registration algorithm (Varian, a Siemens Healthineers company, Palo Alto, CA, USA) by registering the pCT to the daily CBCT. The software uses an elastic deformation model, with Mattes mutual information as the cost function [114], and a B-spline deformation model parametrized by a grid of control points [115].

An example of a CBCT and a dirCT from one patient on the same treatment day can be seen in Fig. 3.1. In addition, for each collected fraction, a plan that was used in the clinic to deliver the treatment to the patient was extracted from the Ethos™ TPS version 1.2 (Varian, a Siemens Healthineers Company, Palo Alto, CA, USA).

Due to a major study constraint, the number of patients was reduced from 16 to 9. Specifically, CBCT images were required to have a non-truncated FOV along all beam paths within each treatment plan. Otherwise, comparisons between recalculated doses on CBCT and dirCT would be rendered invalid.

3.2.2 Study design

Dose recalculation was performed using a clinically commissioned TPS Eclipse™ version 16.0 (Varian, a Siemens Healthineers Company, Palo Alto, CA, USA). For each patient and treatment fraction, the CBCT, dirCT, and the delivered plan were imported into Eclipse™. Body contours were initially automatically generated using Eclipse internal algorithms, then manually reviewed and corrected as needed.

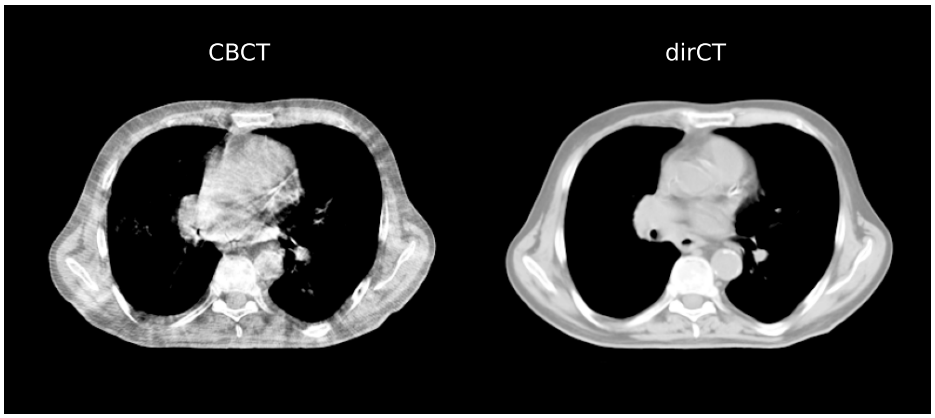


Figure 3.1. CBCT (left) and dirCT (right) of a LungCa patient on the same treatment day.

When exporting CBCT and dirCT images from the Ethos™ TPS, both remained in the same frame of reference, enabling the same volumes of interest (VOIs) to be assigned to both images. VOIs were generated using the Ethos™ autosegmentation tool on the CBCT and subsequently reviewed and manually corrected by clinical staff as part of the standard clinical workflow. This, however, did not imply perfect alignment between CBCT contours and visible anatomy on the dirCT, but it did enable the use of the same clinical workflow applied in routine Ethos™ operations. The structures analyzed in this study included the CTV, PTV, both lungs, esophagus, spinal canal, heart and stomach.

Dose calculations were performed using the Acuros XB (AXB) algorithm, integrated within the Eclipse™ TPS, with user-defined HU-ρ_e and HU-ρ_m calibration curves. Separate curves were used for CBCT and dirCT, both of which were derived by scanning a 32 cm diameter Gammex™ RMI-467 phantom (GammexCo, Bad Münstereifel, Germany) shown in Fig. 3.2. The CBCT calibration curve was based on scans from the Ethos™ on-board CBCT imager, while the dirCT calibration curve was created using a Siemens SOMATOM® Definition Flash scanner at DKFZ. For each investigated fraction, the dose was recalculated using the corresponding delivered plan. In this analysis, each fraction was treated as a surrogate for the entire treatment course, meaning that comparisons of dose metrics were made in terms of total accumulated dose – as these metrics are conventionally defined – rather than per-fraction dose.

Figure 3.2. Gammex^{TM} RMI-467 phantom (32 cm diameter) used for calibration of the lung region.

3.2.3 Evaluation

Two primary aspects were evaluated in this study: time-dependent changes in lung volume across image types and the variation of dose metrics in the target and OARs over time.

Lung volume investigation

Lungs were selected as the primary organ of interest due to their high susceptibility to anatomical changes between treatment fractions [116]. Phenomena such as atelectasis, reventilation, and tumor regression can significantly affect lung morphology, thereby influencing the shape of surrounding tissues [117]. These changes, in turn, alter the environment along the beam path and impact dose deposition within the target volume [118].

Manually delineated lung contours were available for CBCT images; however, corresponding segmentations were not present for dirCTs. To address this, lung contours on the dirCTs were generated using TotalSegmentator v2.0.5 (TS) [119], a deep learning-based segmentation tool. TS has demonstrated high accuracy and robustness across a range of datasets [119, 120], which made it a reliable choice for this task. With lung segmentations available for both CBCT and dirCT, the week-to-

week change of total lung volume was plotted for each modality, along with their inter-modality differences. To minimize potential bias introduced by differing segmentation approaches – manual for CBCT and automated for dirCT – volume differences were baseline-corrected using the first treatment fraction for each patient. This normalization enabled the analysis to focus on relative anatomical changes over time, rather than absolute differences between modalities.

Dosimetric investigation

The second part of the evaluation assessed the temporal behavior of clinically relevant dose metrics across both image types. Standard metrics commonly used in LungCa treatment evaluation were selected, including $D_{0.03cc}$ for the esophagus, V_{20} for both lungs, and D_{95} for the CTV. For each patient, the metrics were extracted from CBCT- and dirCT-based dose distributions across all available weekly fractions. Line plots were generated to visualize the time-dependent change of each dose metric and to assess how values behaved between the two modalities.

3.3 Results

Lung volume investigation

Fig. 3.3 illustrates the changes in lung volume across all nine patients in the study, with treatment weeks on the x-axis and lung volume in cm^3 on the y-axis. The plots highlight the dynamic nature of lung morphology from week to week and, more importantly, the volume differences between CBCT (blue line) and dirCT (orange line). For some patients, the volume difference between the two modalities fluctuates over time, while for others it remains relatively stable.

To more clearly assess week-to-week changes, baseline-corrected volume differences were plotted in Fig. 3.4, where the y-axis now represents the difference in lung volume between CBCT and dirCT. These plots show that, with the exception of Patients 3 and 5, all patients exhibit persistent and fluctuating discrepancies in lung volume between the two modalities. While these differences do not directly imply variations in dose distributions or dose metrics, they do indicate how the underlying anatomy can vary both across patients, and fraction-to-fraction for a given patient.

Figure 3.3. Weekly lung volume for all nine patients, shown for CBCT (blue) and dirCT (orange). The plots illustrate changes in time in lung anatomy throughout the treatment course and highlight differences in volume between two image types.

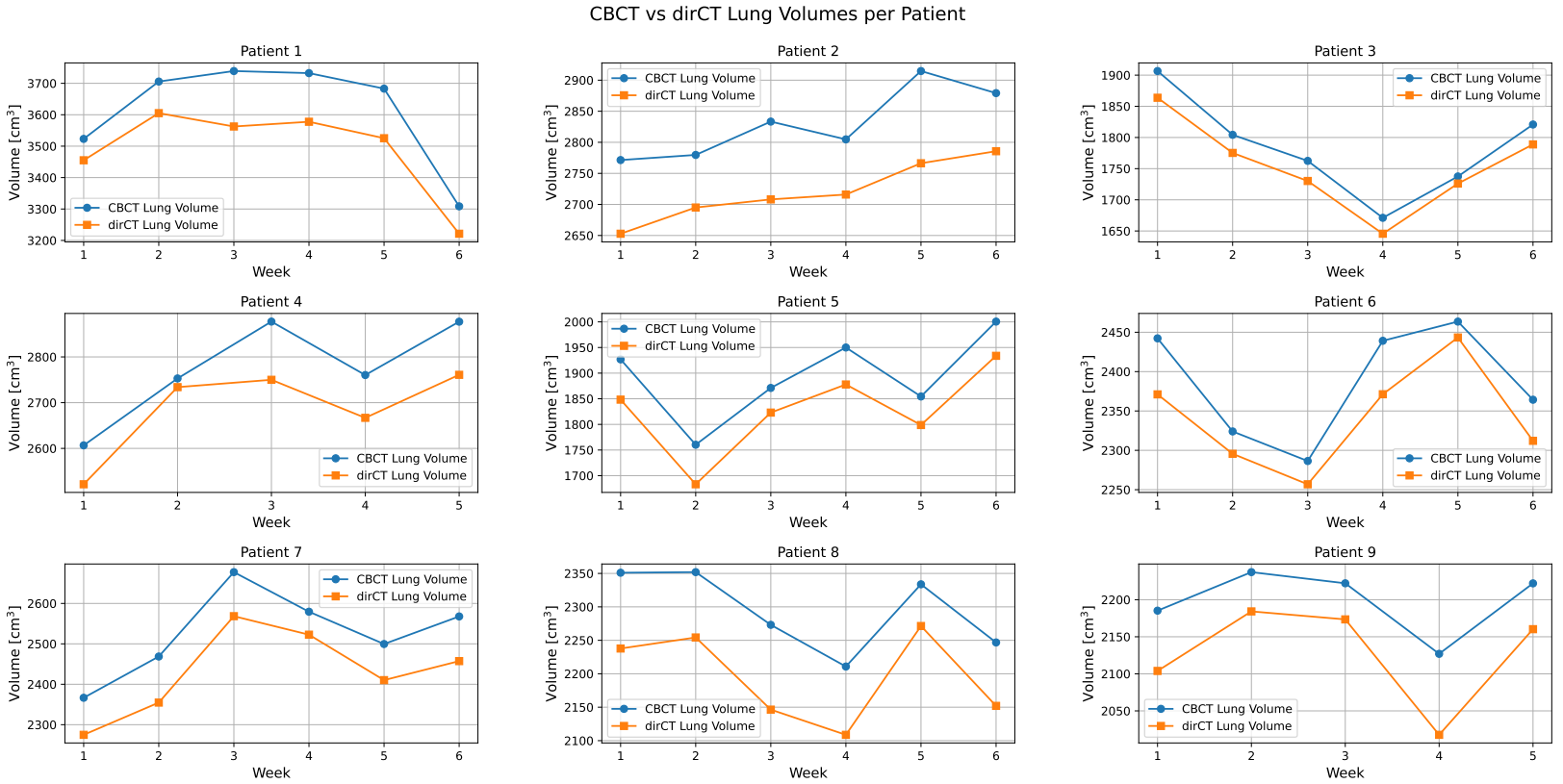
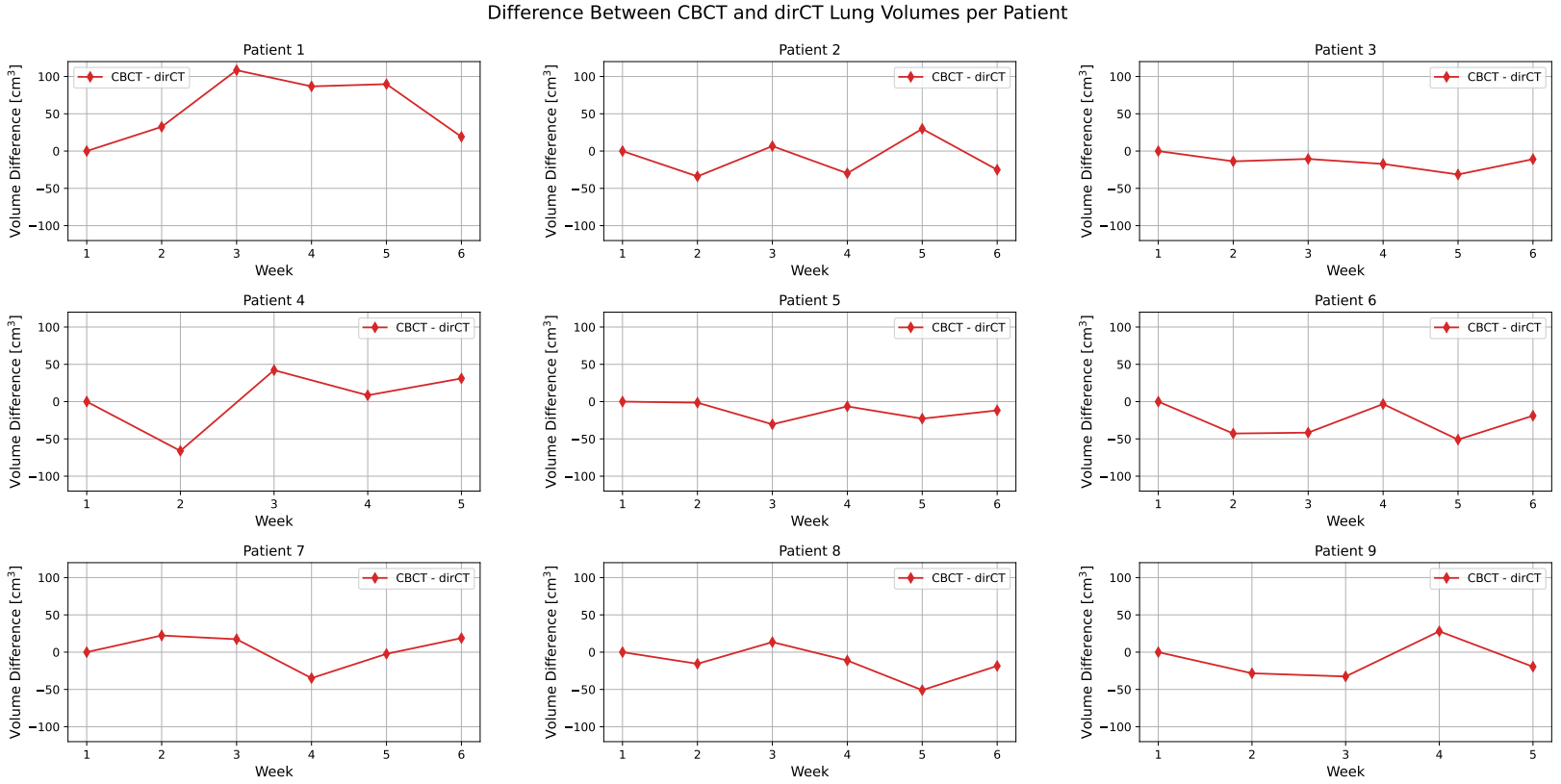


Figure 3.4. Baseline-corrected lung volume differences between CBCT and dirCT for each patient over the treatment course. The plots highlight the relative fluctuation in volume between image types.



Dosimetric investigation

To evaluate the dosimetric impact of anatomical differences, Figs. 3.5 to 3.7 were generated to visualize the weekly trend of key dose metrics across patients. Similar to the lung volume findings, some patients demonstrated consistent behavior between CBCT- and dirCT-based dose calculations, while others presented with unpredictable behavior:

- Fig. 3.5 shows the change of the V_{20} metric for both lungs. Patients 2, 3, 7, 8 and 9 exhibit a consistent offset between the CBCT and the dirCT curve. Patient 1 shows a more significant change towards the end of the treatment with the two curves swapping places. Finally, Patient 5 shows an exact match between the two curves, while Patients 4 and 6 exhibit minor variations in differences between dirCT and CBCT depending on the week.

- In Fig. 3.6, $D_{0.03cc}$ for the esophagus is presented. Patients 2, 3, 7, and 8 show agreement between the two image types with a constant offset in dose. The time-dependent change of the $D_{0.03cc}$ metric for other patients display irregular patterns of behavior – curves are crossing, converging, diverging, and changing places, which indicates that the investigated region and the corresponding dose metric are very sensitive to changes in the anatomy and image quality.

- Finally, in Fig. 3.7, the D_{95} metric for the CTV is visualized. Here, Patients 4, 5, and 6 (except in week three), as well as Patients 7 and 9, show comparable dose values when accounting for the scale of the y-axis. The rest present with irregular patterns similar to Fig. 3.6.

Figure 3.5. Week-to-week progression of the V_{20} dose metric for both lungs across all nine patients. Blue and orange lines represent values derived from CBCT and dirCT, respectively. The plots demonstrate consistent differences in lung dose estimates between the two modalities over time.

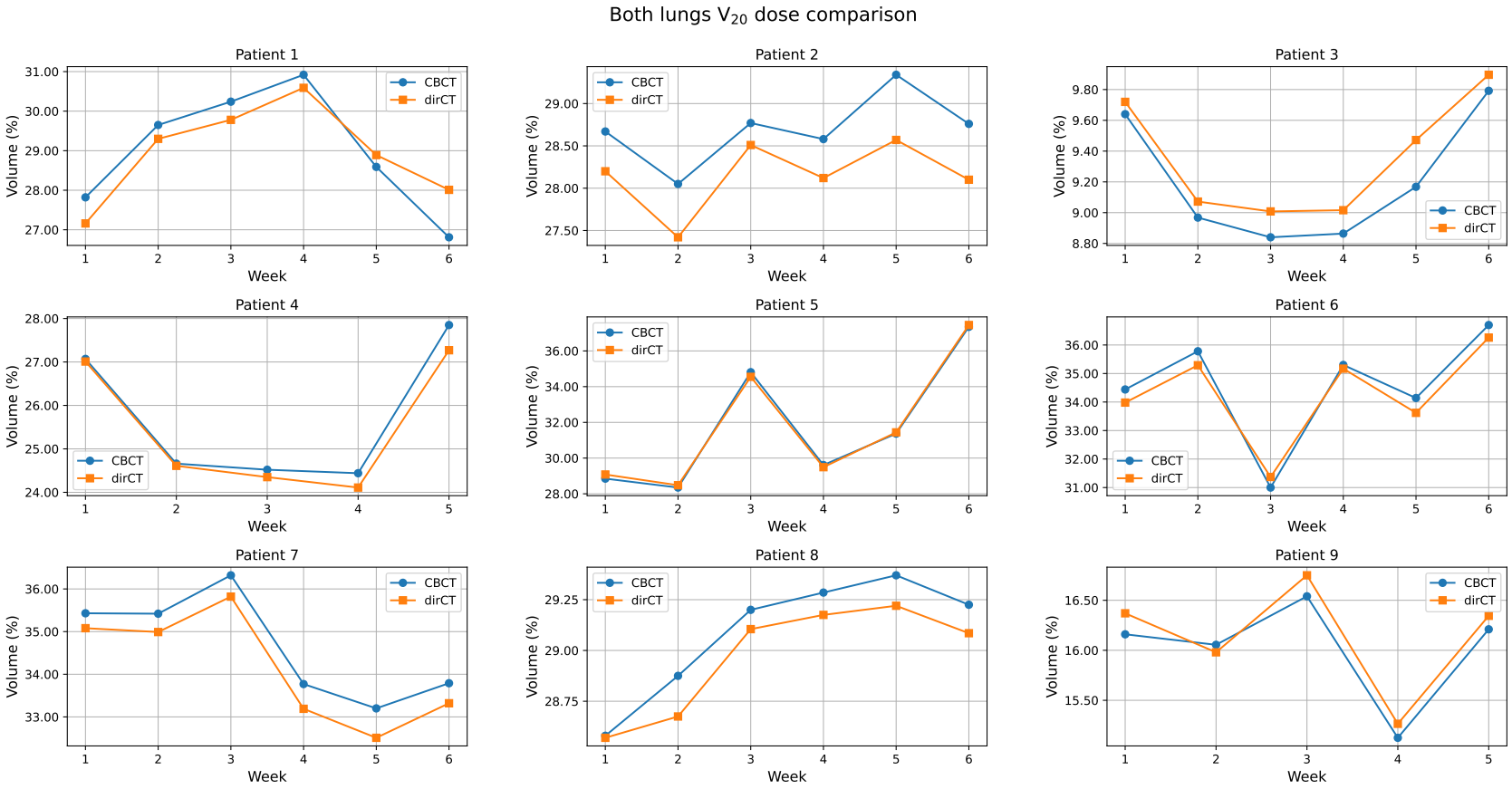


Figure 3.6. Weekly change of the esophagus $D_{0.03cc}$ dose metric for all nine patients. CBCT-based values are shown in blue and dirCT-based values in orange. The plots demonstrate modality-dependent variations and their progression across the treatment course in high-dose exposure to the esophagus.

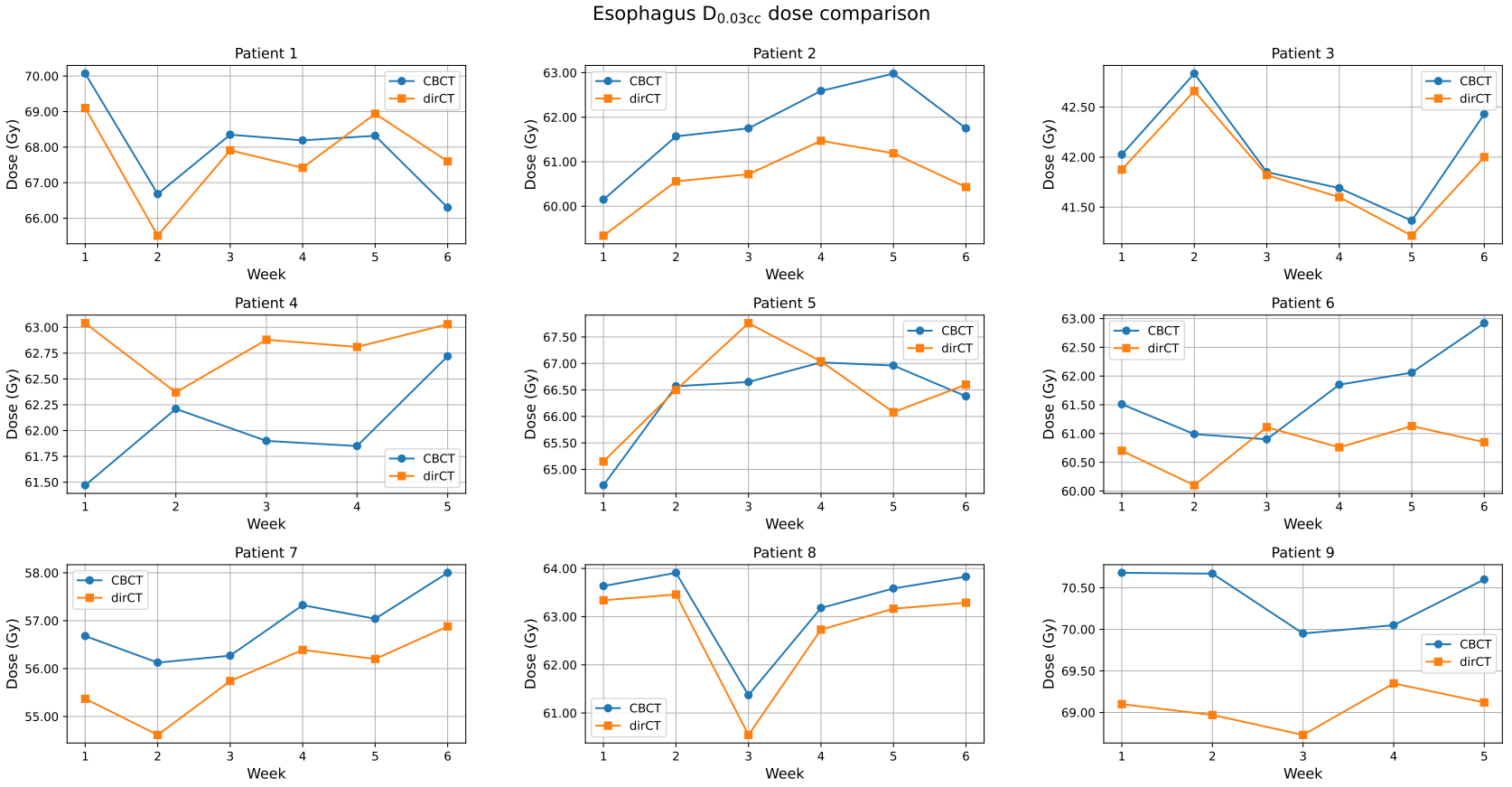
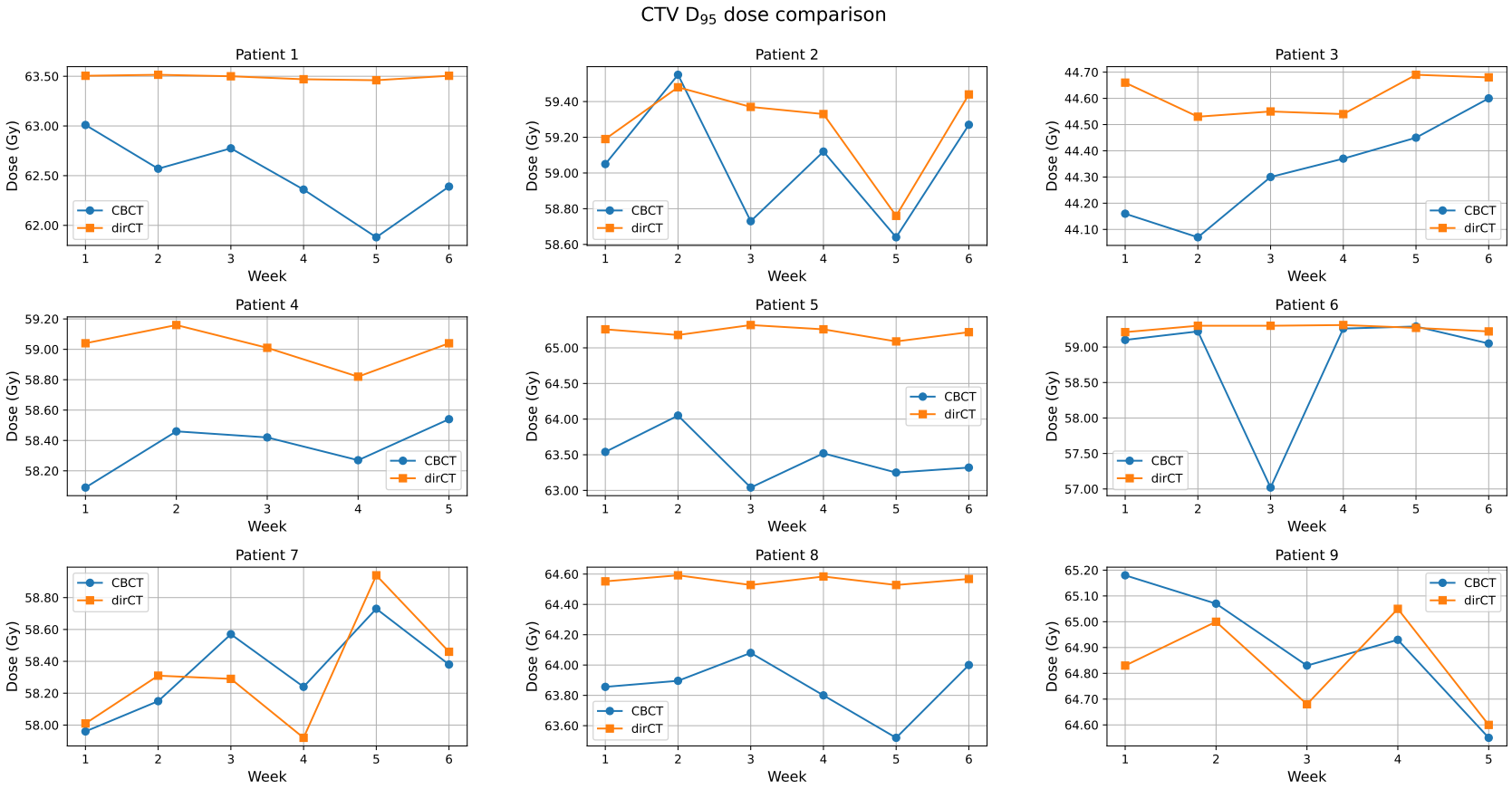


Figure 3.7. Weekly changes in CTV D_{95} values for all nine patients, comparing CBCT (blue) and dirCT (orange) dose calculations. The plots reveal variations in target coverage across the treatment course and differences in dose estimation between two image types.



3.4 Discussion

The primary objective of this chapter was to investigate anatomical fluctuations between CBCT and dirCT images and to assess how differences between these images influence dose distributions throughout the course of treatment. The results demonstrated that nearly all patients exhibited fluctuating differences in lung volume between the two image sets. However, these findings should be interpreted with caution, as lung contours were generated using different segmentation approaches: CBCT images were initially segmented automatically and then corrected by a clinician, while dirCT images were segmented using the TS algorithm without manual refinement. A more balanced comparison could be achieved by having the same clinician manually delineate lung contours on the dirCT images, thereby minimizing interobserver variability and eliminating bias introduced by differing segmentation techniques.

The dosimetric analysis revealed varying outcomes depending on the VOI. For the lungs, although the delivered dose changed from week to week, the difference between CBCT- and dirCT-based dose estimates remained relatively stable throughout the treatment course. One possible interpretation of these findings is that the observed fluctuations in lung volume between CBCT and dirCT are relatively small compared to the overall lung size. For example, in the most pronounced case (Patient 1), a volume difference of approximately 150 cm^3 corresponds to only about 4% of the total lung volume, suggesting that its direct impact on beam path and dose deposition may be limited. Instead, systematic shifts in the V_{20} dose metric, and its occasional variation, are more likely driven by two factors: first, global shifts of HU values in CBCT caused by noise and motion artifacts (which will be examined in greater detail in later chapters), and second, the spatial location of the anatomical change. Since V_{20} quantifies the proportion of lung receiving 20 Gy, changes occurring in regions outside beam paths or in low-dose areas would have minimal influence on the metric. This stands in contrast to Figs. 3.6 and 3.7. The global shift in HU values and the increased noise and motion artifacts were still present in CBCT images, however this time, the fluctuating week-to-week differences in delivered dose were much more pronounced. This is likely due to the increased sensitivity of these areas to small anatomical shifts, given their proximity to the high-dose regions and steep dose gradients. Therefore, even small anatomical irregularities between the dirCT and CBCT arising from the inaccuracies of the DIR algorithm could amplify the divergence between the planned and delivered dose.

It is also important to consider the known limitations of CBCT image quality in the thoracic region, where motion artifacts are dominant. These artifacts can significantly affect the accuracy of dose delivery both to the tumor and the surrounding healthy tissue. Therefore, even with accurate anatomical delineations, the reliability of CBCT-based dose calculation may remain limited in regions affected by respiratory motion. Future work investigating this topic could benefit from incorporating state-of-the-art CBCT technologies such as Varian Hypersight™ [121], and introducing novel AI-based DIR methods [122]. While such improvements could mitigate some of the observed challenges in this study specifically, and in ART in general, the current findings ultimately suggest that deep learning-based synCT may provide a more robust and reliable foundation for adaptive workflows.

3.5 Conclusion

The investigation presented in this chapter highlights not only how anatomical differences and dose metrics evolve throughout the treatment course, but also how the two imaging modalities on which ART workflows rely – CBCT and dirCT – can be unreliable for clinical decision-making. As a result, for some patients, the planned dose may closely match the delivered dose, while for others, significant discrepancies may occur. Plans optimized on daily contours and dirCT images, an approach employed by leading treatment planning systems such as Ethos™, may generate plans that differ from the actual delivered dose. In these cases, prescriptions and dose constraints would appear satisfied computationally, but not necessarily in the patient.

These findings point to the potential value of synthetic CT images generated through deep learning methods. Such images offer the dual advantage of accurate daily anatomy and image quality comparable to that of the pCT. The development and training of neural networks for synthetic image generation, as well as their application across various clinical tasks will be explored in Chapters 4 to 6.

4

CycleGAN-Based Synthetic CT Generation

4.1 Motivation

Chapter 3 addressed the limitations of correction and registration methods outlined in Chapter 1, and quantified the magnitude of anatomical differences and time-dependent changes in clinically relevant dose metrics. These limitations highlight the need for alternative imaging strategies that can provide both anatomical accuracy and dosimetric reliability through good image quality. In this chapter, a CycleGAN architecture for generating synthetic CT images from CBCTs was implemented, inspired by the potential of deep learning-based image synthesis to act as a reliable modality in ART workflows.

4.2 Methods

4.2.1 Network architecture

A 3D extension of the original CycleGAN architecture [91] was developed to enable image-to-image translation using volumetric data. In this 3DCycleGAN framework, the generator and discriminator networks are designed to process 3D image patches, allowing direct translation between volumetric images. As discussed in Section 2.6, the training of a CycleGAN involves the construction of its two key components: a generator and a discriminator.

Generator The generator was ResNet-based [101] with 3 downsampling convolutional blocks (convolution – instance normalization – ReLU non-linear activation), 6 residual blocks (convolutional blocks with skip connections as defined in Section 2.3), and 3 upsampling convolutional blocks. The first convolutional layer contained 64 filters which were doubled with every convolutional block.

Discriminator The discriminator architecture was based on a 3-layer Markovian patch classifier, which penalizes structures at the scale of image patches, not complete images. It models the image as a Markov random field, assuming independence between pixels separated by more than a patch diameter [91]. Similar to the generator, the first convolutional layer contained 64 filters and instance normalization was used in individual layers.

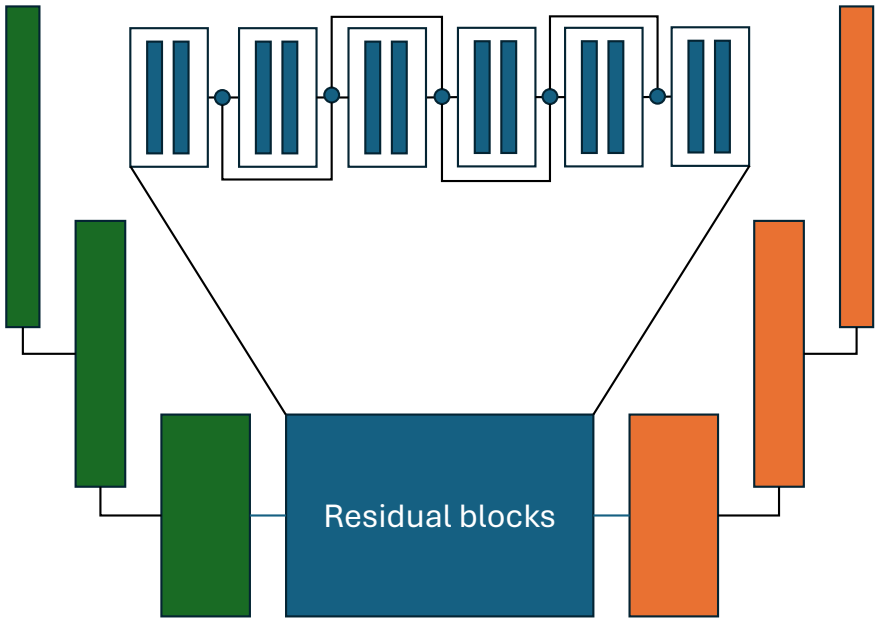


Figure 4.1. Schematic representation of the ResNet-based generator used in the thesis, with 3 downsampling convolutional blocks (green), two upsampling convolutional blocks (orange), and 6 residual blocks (blue).

4.2.2 Dataset

In addition to the LungCa dataset described in Section 3.2.1, data from another 27 prostate cancer (PCa) patients within the MARS trial [40], who had undergone treatment on the Ethos™ at the DKFZ, were used for this study. All PCa patients agreed to pseudonymized analyses and anonymized publication. All but two PCa cases were treated in a normo-fractionated manner with 3Gy/fraction for 20 fractions - one patient received 2.25Gy/fraction for 34 fractions and one was treated in a hypo-fractionated manner with 7.5Gy/fraction for 5 fractions. An overview of characteristics of PCa patients is shown in Table 4.1.

Table 4.1. Gender representation, disease stage, and age of PCa patients used in the study.

Site	Gender (Male/Female)	Disease stage	Age (Mean, range)
PCa	27/0	T1-T2	72, 52-86

Three types of images scans were collected from each patient (LungCa and PCa) for the purpose of network training and evaluation:

- pCT - acquired with a Siemens SOMATOM® Definition Flash (Siemens Healthineers, Erlangen, Germany), with 120 kVp tube voltage, 3 mm slice thickness, and 1.27 mm pixel spacing.

- CBCT - acquired with the Ethos™ on-board kV CBCT imaging device with 125 kVp tube voltage, 2 mm slice thickness, 0.96 mm pixel spacing, 18 to 20cm FOV size in the cranio-caudal direction.
- dirCT - created using Velocity™ deformable image registration algorithm by registering the pCT to the daily CBCT. Slice thickness and pixel spacing same as in the pCT.

Preprocessing pCT and CBCT images were split into training/validation/test set (PCa: 21/1/5 and LungCa: 12/1/3). As a form of data augmentation in the training phase, multiple fraction CBCT images were selected at evenly spaced treatment time points, three from each PCa patient and six from each LungCa patient. All pCT images were rigidly registered to CBCT images using SimpleITK [123] and SimpleElastix [124] and cropped to the extent of the CBCT FOV. All content outside the patient contour was removed from images using TS v2.0.5 and its body delineation tool. Finally, CT-values in all images were clipped to [-1000, 2000] HU and normalized to [-1.0, 1.0].

4.2.3 Training and inference

Loss functions, training procedure and model selection Three loss functions were used for network training – adversarial loss, cycle consistency loss, and identity loss [91]. To prevent discriminator overfitting two methods were employed: one-sided label smoothing, and addition of random noise to the generated image before forwarding it to the discriminator [125, 126]. Lambda parameters for forward cycle loss, backward cycle loss, and identity loss were 10, 10 and 0.1. *Two separate networks were trained, one for each anatomical region*. They were trained in a supervised manner using the Adam optimizer with betas $= (0.5, 0.999)$ for 200 epochs with an initial learning rate $\mathrm{lr}_{\mathrm{init}} = 0.00005$ for the first 100 epochs and $\mathrm{lr} = 0.5 \cdot \mathrm{lr}_{\mathrm{init}}$ for the second 100 epochs. Due to memory limitations, training was performed patch-wise using five image patches from each image pair with patch size (256, 256, 32) in the left-right (x), anterior-posterior (y), and cranio-caudal (z) direction. To enhance data variability, patches were randomly re-sampled at the beginning of each epoch. Four metrics were monitored on the validation set: MSE, PSNR, SSIM, and JSD, as defined in Section 2.7. Adversarial loss, cycle consistency loss and identity loss were used for monitoring training stability and convergence (Fig. 4.2). Checkpoints were saved every 10 epochs, and the model used for inference was selected based on the best performance in at least three of the four metrics,

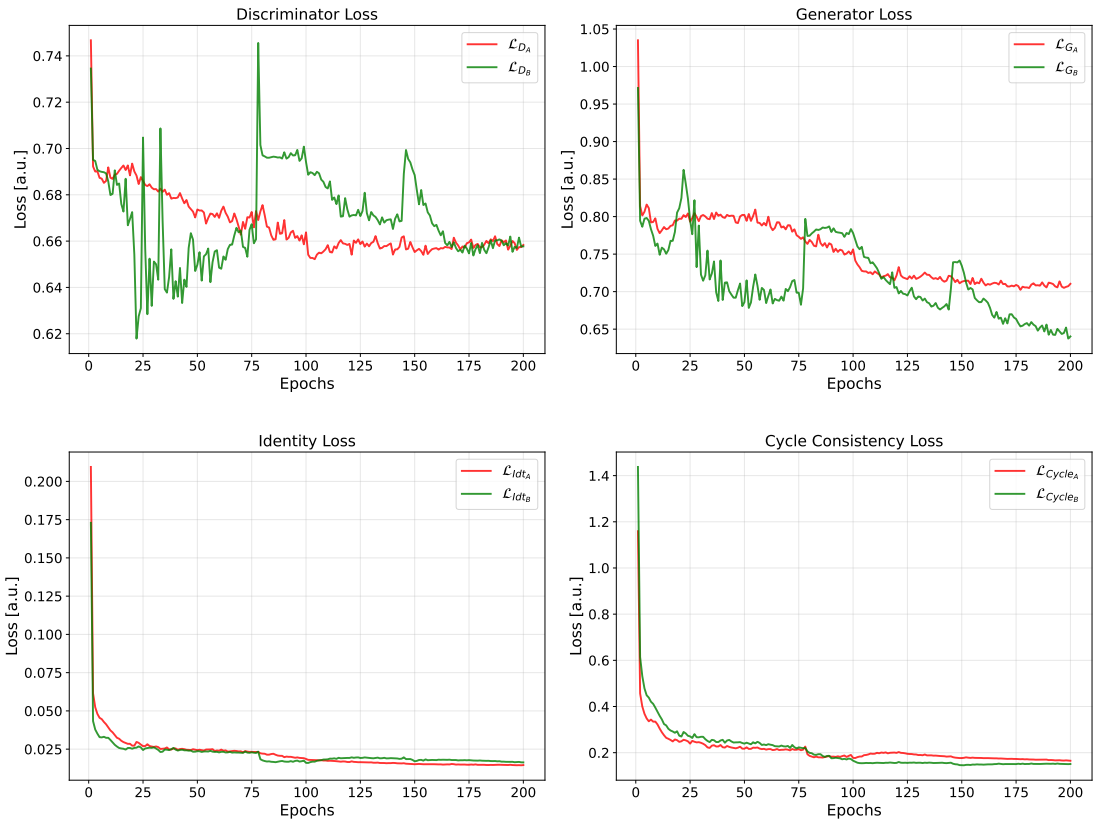


Figure 4.2. Loss function monitoring during training. Top: discriminator and generator loss; bottom: identity and cycle consistency loss – two characteristic losses used in CycleGAN training.

computed over the entire image volume.

Inference Daily CBCTs from five independent PCa patients and three independent LungCa patients were used as inputs for network inference. The criteria for test patient inclusion was that the patient body had to be fully covered by the CBCT FOV in the x and y direction due to its application in downstream dose calculation tasks in Chapter 5. As the network was trained on randomly sampled image patches, inference was also performed in a patch-wise manner. Overlapping 3D patches were extracted using a stride of 64 pixels in the x and y direction and 8 slices in the z direction. The final output volume was reconstructed by averaging voxel intensities in overlapping regions, which was followed by rescaling the intensity values to the range of [-1000, 2000] HU.

Hardware All training and inference was executed on a computer with an AMD Ryzen 9 5900X Processor, 128 GB RAM, with an NVIDIA RTX A6000, and 48 GB VRAM, using Python Version 3.9.13, PyTorch 2.2.0 and CUDA Version 11.8.

4.3 Results

4.3.1 Qualitative evaluation

Qualitative evaluation was performed by observing the resulting images, the perceived anatomical consistency and the reduction of imaging artifacts. Resulting synCT images for three LungCa and three PCa patients are visualized in Fig. 4.3 and Fig. 4.4 with pCT, CBCT and synCT shown next to each other. The arrows in the images point to anatomical features which are preserved during CBCT-synCT translation, which cannot be found in the pCT.

PCa In Fig. 4.3, we can see that, in all visualized cases, the difference in rectal filling between pCT and CBCT is notable, and more importantly anatomically preserved between CBCT and synCT. Further qualitative evaluation reveals a reduction in beam hardening artifacts in the region around the prostate and seminal vesicles, which lies between dense bony structures such as femoral heads and pelvic bones (in the left-right direction). Additionally, the noise level in the synCT appears visibly lower compared to the CBCT.

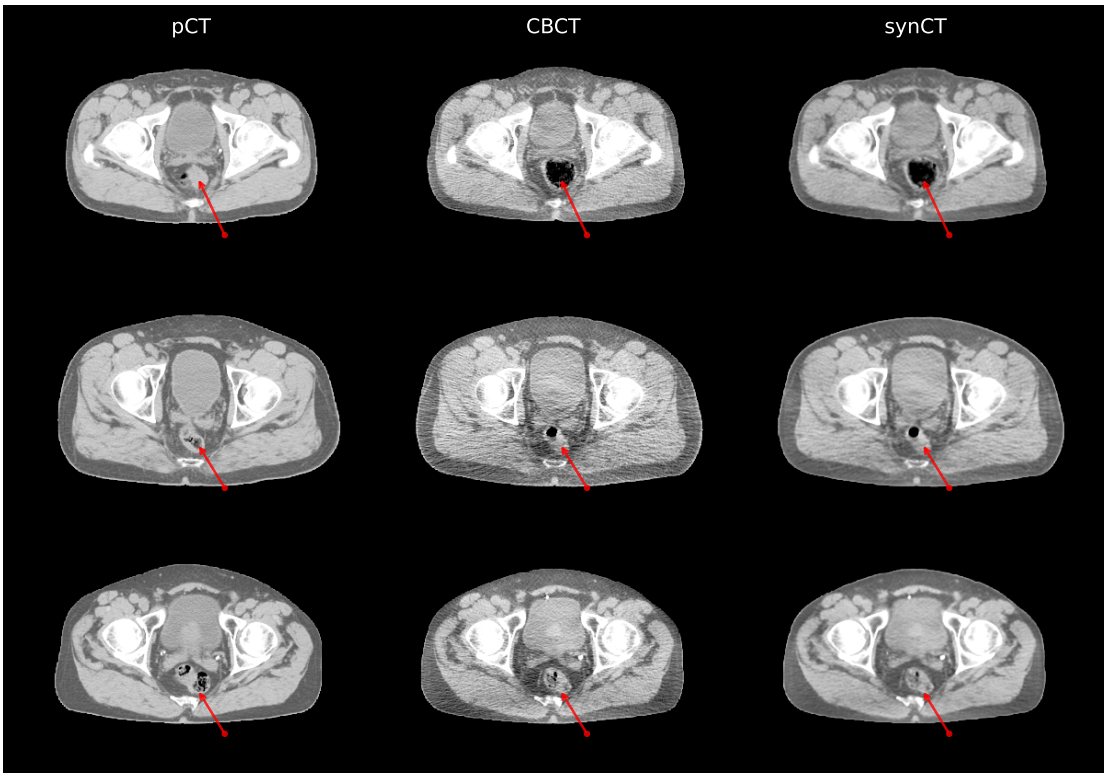


Figure 4.3. Comparison of CT scans (pCT, CBCT, synCT) from three PCa patients. Red arrows point to anatomical features which are preserved during CBCT-synCT translation.

LungCa Fig. 4.4 presents a more nuanced case. In the top row, the example shows tumor volume changes from pCT to CBCT that are consistently preserved in the synCT. The bottom row illustrates changes in lung volume, specifically the re-expansion of a previously collapsed lung. The CBCT lung images exhibit noticeably lower quality, with pronounced motion artifacts caused by patient's breathing. While synCT demonstrates a reduction in these motion artifacts, the most severe ones remain partially visible. As with the PCa cases, image noise is also visibly reduced in the synCT.

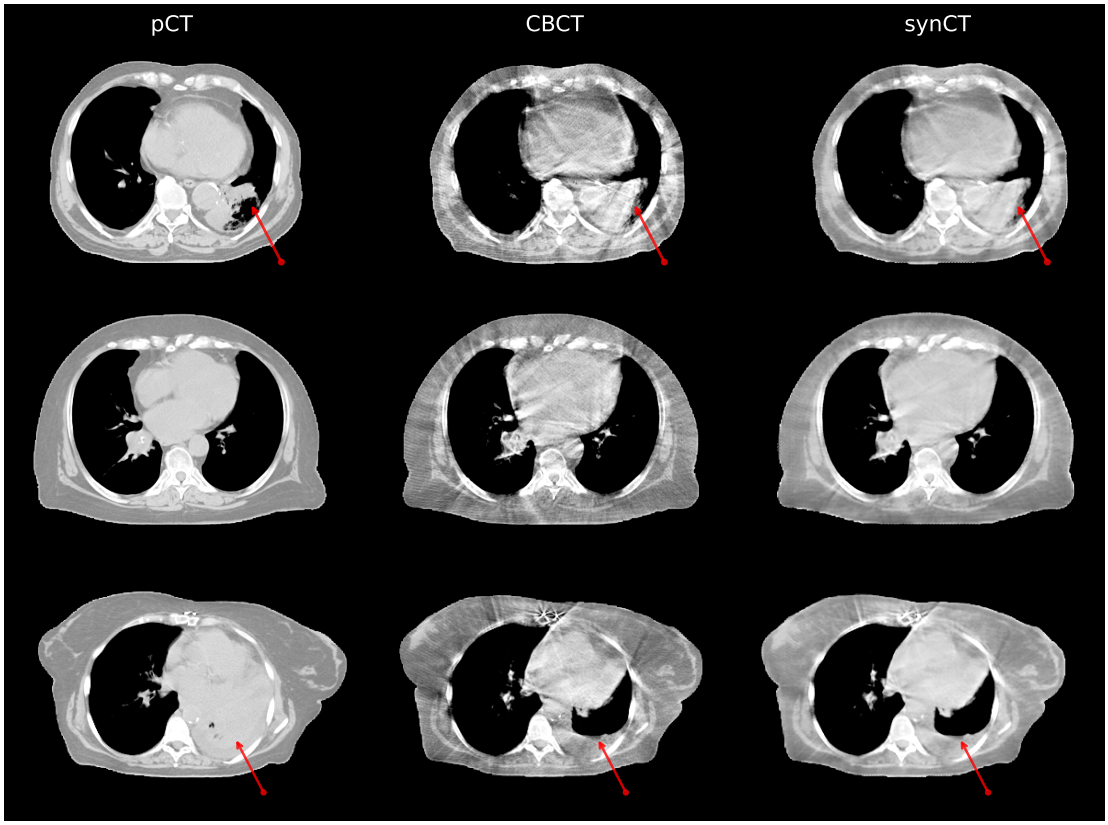


Figure 4.4. Comparison of CT scans (pCT, CBCT, synCT) from three LungCa patients. Red arrows point to anatomical features which are preserved during CBCT-synCT translation.

4.3.2 Quantitative evaluation

Following the qualitative evaluation, MAE, SSIM, JSD and DSC were calculated to quantitatively show the similarity between the pCT and synCT in terms of image quality (Table 4.2), and the similarity between the CBCT and synCT in terms of patient anatomy (Table 4.3). The image quality metrics, on average, show an improvement of the synCT image quality compared to CBCT and the structure similarity metric shows that the body contour visible on the synCT image is more

Table 4.2. Image quality metrics (SSIM and JSD) for CBCT-pCT and synCT-pCT comparison.

	MAE ↓		SSIM ↑		JSD ↓	
	CBCT	synCT	CBCT	synCT	CBCT	synCT
PCa	28.5 ± 4.1	28.6 ± 6.4	0.90 ± 0.02	**0.93 ± 0.01**	0.19 ± 0.03	**0.16 ± 0.04**
LungCa	46.2 ± 5.0	**42.8 ± 4.4**	0.87 ± 0.02	**0.89 ± 0.03**	0.29 ± 0.03	**0.20 ± 0.01**
All	35.2 ± 9.6	**33.9 ± 9.0**	0.88 ± 0.03	**0.91 ± 0.03**	0.23 ± 0.05	**0.18 ± 0.04**

Table 4.3. Contour similarity metric (SDC) for pCT-CBCT and synCT-CBCT comparison.

	SDC ↑	
	pCT	synCT
PCa	0.983 ± 0.002	**0.997 ± 0.001**
LungCa	0.978 ± 0.007	**0.995 ± 0.001**
All	0.981 ± 0.005	**0.996 ± 0.001**

similar to the one on the CBCT than on the pCT – DSC is consistently above 99%. Same results hold for both disease sites.

In addition to image quality metrics, HU histogram analysis was conducted for the three image types, pCT, CBCT, and synCT, revealing a consistent improvement in HU distribution in synCT compared to CBCT. Representative comparisons are shown in Fig. 4.5 and Fig. 4.6 for one PCa and one LungCa case, respectively. Across all cases, the synCT histograms exhibited a narrower profile within the dominant soft tissue range and demonstrated a double-peak structure that more closely resembled that of the pCT. These findings are supported by values of JSD (Table 4.2) which is a histogram-based metric. The histogram analysis also supports the qualitative findings shown in Figs. 4.3 and 4.4, where the network was unable to fully correct severe artifacts in LungCa cases. Conversely, the relatively high quality of the original CBCT images in PCa cases enabled the network to generate synCT images that closely matched the appearance of the pCT.

Finally, a line profile of HU values was analyzed to assess image consistency. To minimize bias from potential misalignments between pCT and CBCT images, the analysis was performed using dirCTs, as defined in Sections 3.2.1 and 4.2.2. One PCa and one LungCa case were selected, and HU line profiles were extracted for CBCT, synCT, and dirCT images.

For the PCa case (Fig. 4.7), the HU profiles show generally good agreement among all three image types. However, occasional deviations are observed in the CBCT values, which diverge from the overall trend. In contrast, the profiles for

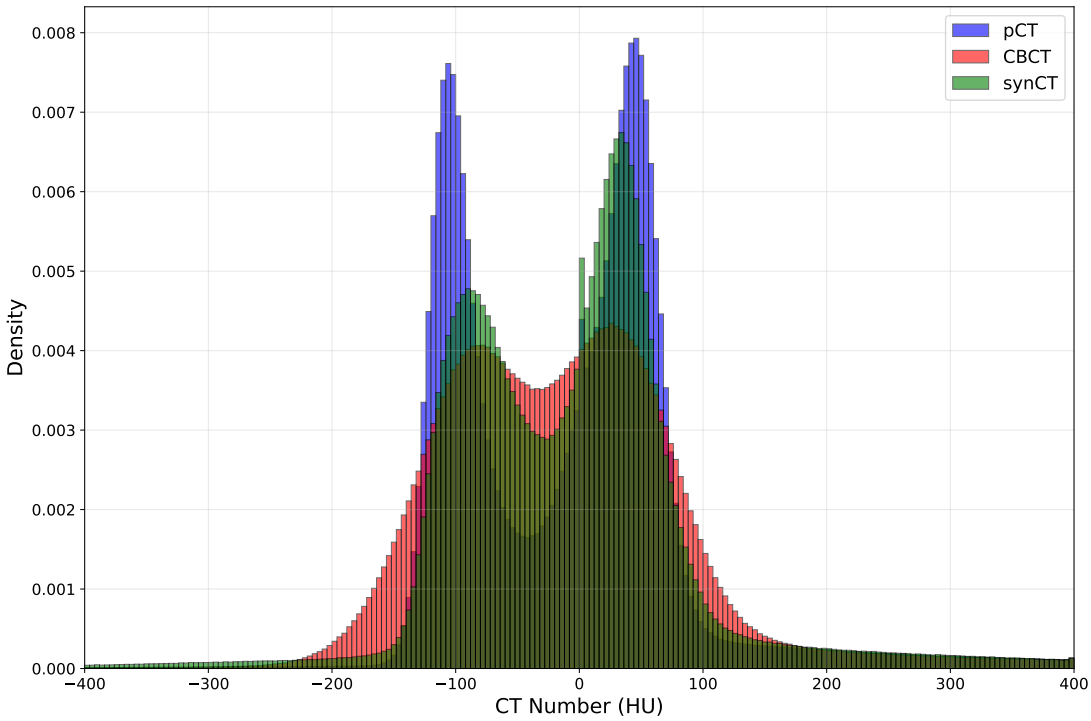


Figure 4.5. Histogram comparison of HU distributions of one PCa patient. Three CT image variants, pCT (blue), CBCT (red) and synCT (green), are shown for the most dominant non-background HU value range [-400, 400].

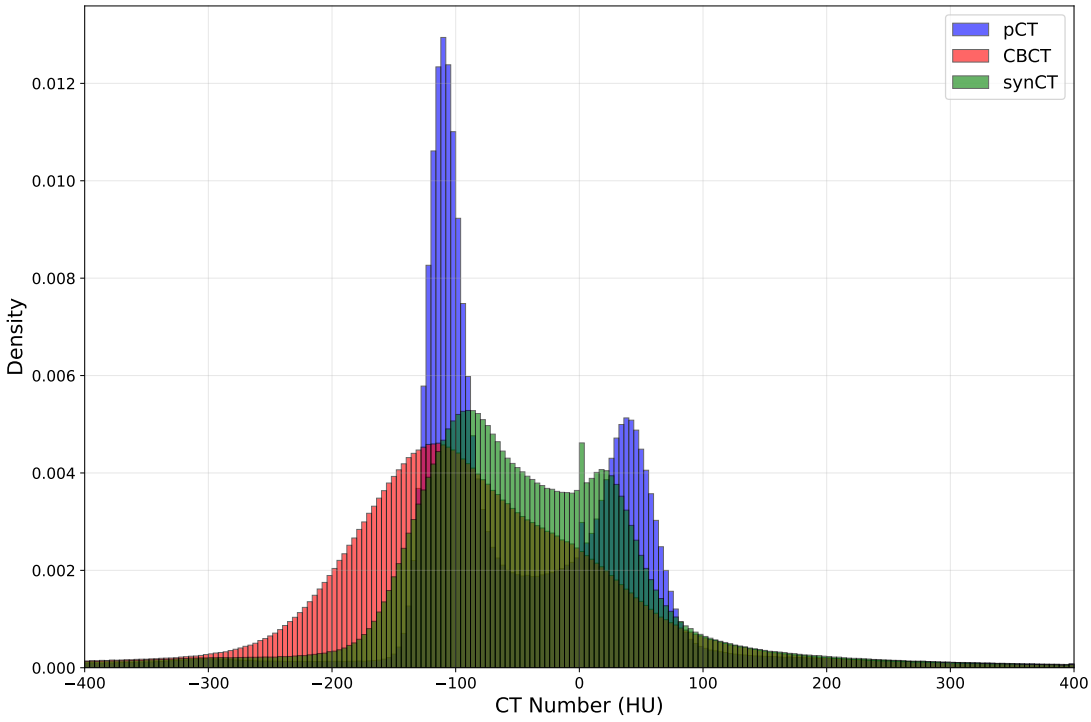


Figure 4.6. Histogram comparison of HU distributions of one LungCa patient. Three CT image variants, pCT (blue), CBCT (red) and synCT (green), are shown for the most dominant non-background HU value range [-400, 400].

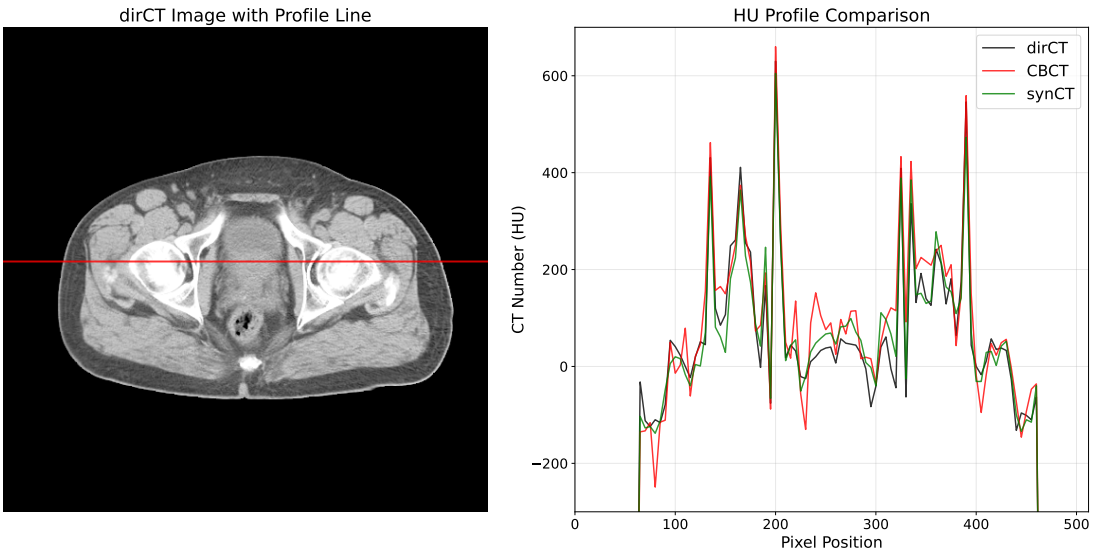


Figure 4.7. Line profiles of HU values for dirCT (black), CBCT (red), and synCT (green) in a representative PCa case. The profiles were extracted along the same line (left figure, red) for all three image types.

dirCT and synCT (black and green lines, respectively) exhibit more consistent behavior and better alignment with each other.

As shown in Fig. 4.8, the dirCT image displays a slightly altered body contour and a minor tilt of the pelvic bone compared to the pCT. Nonetheless, large-scale anatomical deformations, such as the rectum, remain unchanged and show no significant deformation across the two image types, whereas in CBCT and synCT the change of the rectal filling is noticeable (see Fig. 4.3 middle row).

The HU profile of a LungCa case (Fig. 4.9) reveals a similar yet more intricate pattern compared to the previous example. Once again, the synCT profile more closely follows the profile of the dirCT. This is especially visible in the valley be-

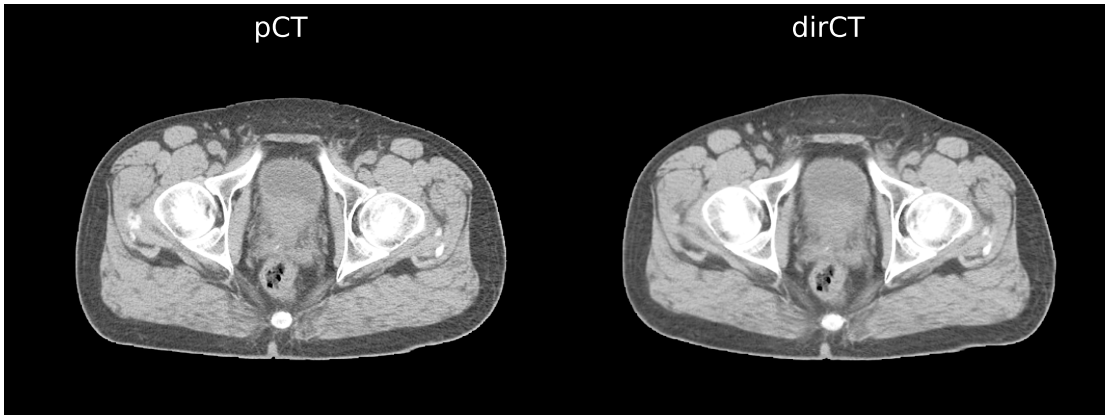


Figure 4.8. Comparison of dirCT and pCT images for the PCa patient used in the HU profile evaluation.

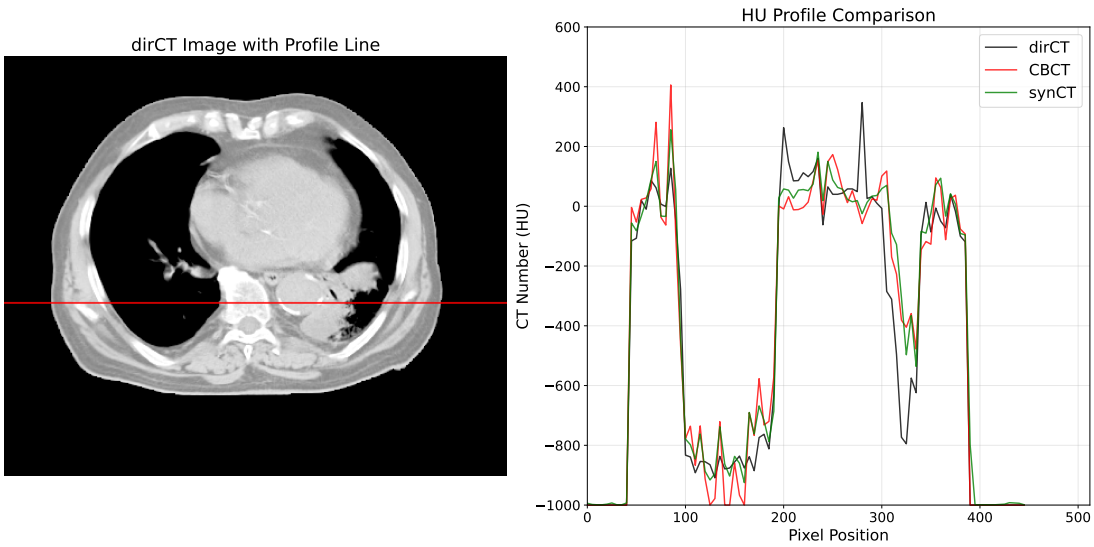


Figure 4.9. Line profiles of HU values for dirCT (black), CBCT (red), and synCT (green) in a representative PCa case. The profiles were extracted along the same line (left figure, red) for all three image types.

tween pixel positions 100 and 200. The erratic behavior visible in the CBCT profile (red) due to strong motion artifacts is smoothed out in the synCT image which brings it closer to the dirCT profile. Beyond illustrating the distribution of HU values along the selected profile line, this example also more clearly reflects key anatomical differences between image types, as highlighted earlier in Fig. 4.4. For instance, the first peak corresponds to a section of the profile that passes close to the rib in the dirCT image, slightly grazing its edge. However, in both the CBCT and synCT, this anatomical detail appears more pronounced. This is due to the fact that the position of the rib in the CBCT and synCT is lower, which means that the profile line cuts directly through it. In the region corresponding to the spine (pixel position 200-250), a noticeable difference can be observed in the location of the dense outer bone in both the CBCT and the synCT, which stands in contrast to the same region in dirCT. Around pixel position 280, a sharp peak appears due to calcification in the aorta, which is present in all three images but slightly shifted upwards in the CBCT and synCT. Finally, the region between voxels 300 and 350 clearly reflects the tumor size differences pointed out in Fig. 4.4, top row. Looking at Fig. 4.10 one can see that the act of deformation managed to slightly move the tumor to the right but did not fill in the region that is visible in the CBCT.

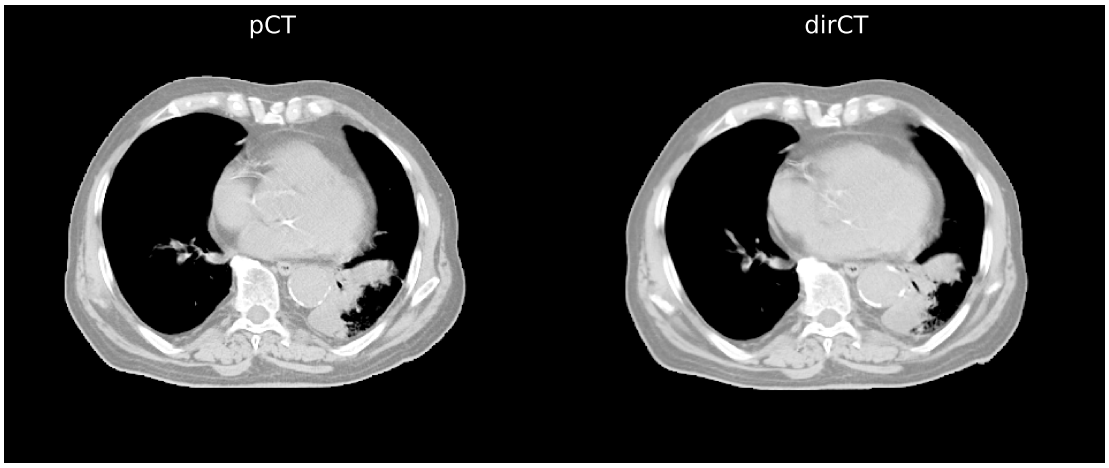


Figure 4.10. Comparison of dirCT and pCT for the LungCa patient used in the HU profile evaluation.

4.4 Discussion

A key component of this study was the comparison of image quality across the evaluated image variants. This assessment included multiple classes of quantitative evaluation metrics, histogram comparisons, and HU line profile analyses. synCTs demonstrated image quality more similar to pCTs than to CBCTs across all tested metrics, which aligns with previously published studies on the topic [93, 127, 128]. While there was some overlap in the metrics when the standard deviation was taken into account, the per-patient analysis revealed that synCTs consistently provided a more accurate representation of daily anatomy than CBCTs. This suggests that the observed improvements are both statistically and clinically robust, and not limited to a subset of individual patients.

Notably, synCTs demonstrated greater anatomical consistency with the daily anatomy compared to pCTs. This was especially evident in the accuracy of the outer body contour, which showed closer alignment with the anatomy visible in the CBCTs, as shown by the surface-based metric. This finding contrasts with the results of [129], who reported notable discrepancies in synCT accuracy, particularly near the treatment couch on both lateral sides and in the anterior abdominal region. In the present study, such discrepancies were not observed.

In terms of motion artifacts, the findings are consistent with those of [130]. In LungCa cases, where motion artifacts tend to be more pronounced, the CycleGAN approach was not able to eliminate all of them. However, it successfully reduced the intensity of more severe artifacts and effectively removed the less prominent ones.

4.5 Conclusion

This chapter demonstrated that synCT images generated using a 3DCycleGAN model not only offer improved image quality over CBCTs but also exhibit greater anatomical consistency with the daily anatomy compared to pCTs and dirCTs. In the tested metrics, synCTs more closely resembled reference images in terms of HU distribution and preserved critical anatomical structures visible in CBCT which was supported by both visual assessment and SDC evaluation of body contours. These findings confirm and expand upon prior work, while also highlighting limitations, such as residual motion artifacts in lung cancer cases, that call for further optimization.

However, improved image quality does not guarantee clinical usefulness. For a synCT to be viable in ART workflows, it must support downstream treatment planning tasks with comparable or superior performance to existing imaging methods. In other words, the true measure of success is whether this improved visual and quantitative fidelity translates into more accurate dose calculations and enhanced segmentation performance. The next two chapters will address this directly. Chapter 5 evaluates the use of synCTs for direct dose calculation, comparing dosimetric accuracy across different image types. Chapter 6 then investigates use of synthetic images in automatic segmentation tasks. Together, these task-based evaluations aim to assess the extent to which synCT images can support downstream components of adaptive treatment planning beyond image quality alone.

5

Synthetic CT for Treatment Planning and Direct Dose Calculation

5.1 Motivation

Accurate dose calculation is essential for ART, where treatment plans must adapt to anatomical changes over the course of treatment. This chapter presents a comparative evaluation of three image types used in ART workflows: CBCT, deformed pCT via DIR (dirCT), and synCT generated using a CycleGAN model (more details in Chapter 4). Ideally, CBCT would provide image quality and HU accuracy sufficient to support all stages of treatment planning directly, from segmentation to dose calculation. In such a scenario, a dedicated electron density calibration curve could be used to convert HU values into electron densities for accurate dose calculation, similar to pCT. However, as discussed in previous chapters, current CBCT systems often suffer from low soft tissue contrast and HU inconsistency, limiting their direct use in planning. In this study, the best-in-class CBCT images are used to assess their potential under optimal conditions. A more established clinical approach is to use DIR to deform the pCT to match the daily anatomy visible in CBCT. This method enables dose calculation based on the updated anatomy, but its accuracy can be unreliable in the presence of large anatomical deformations between the two image modalities. A third and increasingly explored option is the use of deep learning-based synCT generation, where models like CycleGAN translate CBCT images into synCTs that aim to preserve daily anatomy while improving HU consistency. As outlined in Chapter 4, synCT offers a promising alternative to DIR by preserving daily anatomical detail while achieving HU accuracy sufficient for clinically acceptable dose calculations. This chapter evaluates the clinical trade-offs among these three approaches across two anatomically and clinically distinct tumor sites – prostate and lung cancer – with the goal of identifying the most appropriate ART strategy for different treatment scenarios.

5.2 Methods

5.2.1 Retrospective treatment planning study design

To study the effect three adaptive workflows (CBCT-, dirCT- and synCT-based) have on daily delivered dose a research version of the Ethos™ TPS v1.2 was used. In addition to the adaptive workflows with CBCT, dirCT and synCT images, the IGRT workflow on the pCT was calculated as a reference to contrast the overall effect of adaptation schemes, independent of image choice. Both the ART and the IGRT workflows were investigated on eight patients (PCa – 5, LungCa – 3). The

Table 5.1. Tabular representation of common target and OAR volumes in PCa patients and LungCa patients.

	PCa	NSCLC
Targets	CTV	CTV
	PTV	PTV
OAR	Rectum	Lung (left and right)
	Bladder	Esophagus
	Seminal vesicles	Spinal canal
	Sigmoid colon	Heart
	Penile bulb	Stomach
	Femur head (left and right)	

choice of patients was based on the test dataset used for network evaluation in Chapter 4.

For the ART branch of the treatment planning study, available images included, the CBCT from the first treatment fraction and two additional CTs derived from this CBCT: a dirCT generated using the Ethos™ registration algorithm as described in Chapter 3, and a synCT produced using the 3DCycleGAN model described in Chapter 4. A structure set (OARs and target volumes) from the first treatment session was used for all ART cases. On the other hand, the IGRT branch of the study was based on the initial pCT. Therefore, the structure set from the reference pCT plan was used. A list of common structures in both branches of the study can be seen in Table 5.1.

All images were resampled to match the voxel spacing of the CBCT, cropped to the CBCT FOV, registered to a common frame of reference, and imported into the research TPS for further analysis. All content outside the body contour of the patient was removed from images using TS v2.0.5 and its body delineation tool. This removal, which includes the treatment couch, did not affect dose calculations, as the algorithm automatically inserts the correct couch outline and composition. Treatment intents and dose prescriptions, also called treatment directives, were imported into the TPS and used for plan optimization. The treatment isocenter was positioned at the center of the bounding box encompassing the union of all targets; if this union exceeded 26 cm along the longitudinal axis of the treatment unit, a second isocenter was added.

Dose calculation was carried out using an Ethos™ dose calculation algorithm AXB, with user-defined HU-ρ_e and HU-ρ_m calibration curves. A single calibration curve was used for pCT, dirCT and synCT. It was obtained by scanning a Gam-

Figure 5.1. Circular acrylic holder (16 cm diameter) with Gammex™ RMI-467 inserts used for calibration of the pelvic region.

mex™ RMI-467 phantom (Fig. 3.2) using a Siemens SOMATOM® Definition Flash at the DKFZ. Two other calibration curves were used for dose calculation on the CBCT – one for the PCa and one for the LungCa region. These were created using the Ethos™ on-board CBCT imaging device with phantoms of two different sizes to mimic the pelvic and the lung region - Gammex™ RMI-467 (32 cm diameter) for LungCa and a circular acrylic holder with Gammex™ RMI-467 inserts (16 cm diameter). The small phantom is shown in Fig. 5.1.

One primary intent (reference) plan was generated for each image set (pCT, CBCT, dirCT, and synCT) per patient. After evaluating two plans with 9 and 12 equidistant beams, the one best fulfilling clinical constraints was exported. Although the number of beams in the selected plan could differ between image sets (e.g., 9 beams for pCT and 12 for synCT), comparisons were always made using the same plan and beam configuration, recalculated on other image modalities, without re-optimization. The order how dose was recalculated on different images using reference plans is shown in Fig. 5.2.

5.2.2 Evaluation

Dosimetric evaluation

Dosimetric analysis was performed to compare three ART strategies: synCT-based, dirCT-based, and CBCT-based. In addition, the IGRT workflow on the pCT was used as the reference to isolate and contrast the effects of the adaptation schemes independently of imaging modality differences. The evaluation focused on tumor coverage and OAR sparing, assessed per patient and based on clinical prescriptions, most of which were aligned with ICRU Report 83 [18]. Some institution-

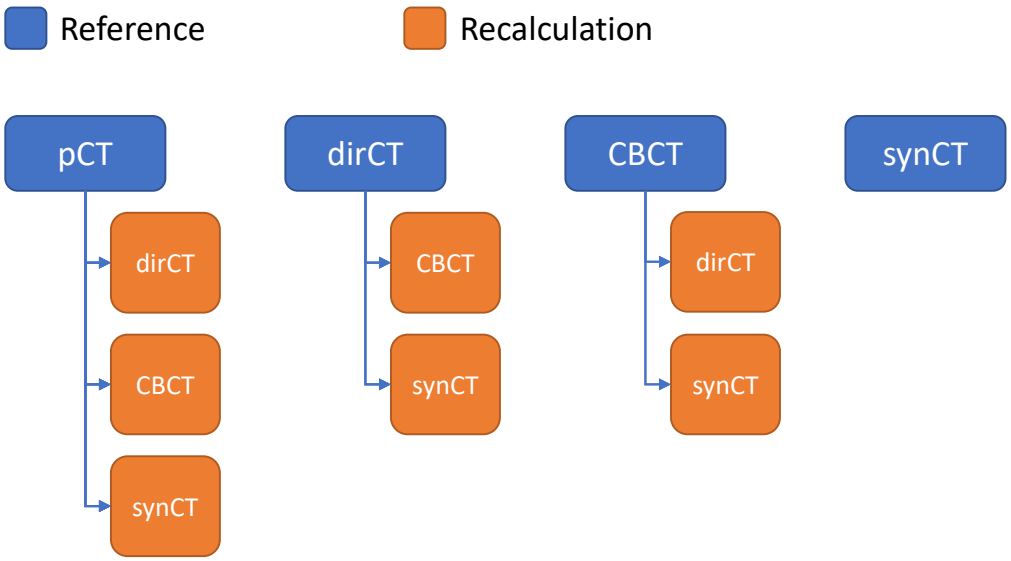


Figure 5.2. Overview of reference plans and their corresponding recalculations on alternative imaging modalities. Each reference plan, originally generated on a specific image set (pCT, CBCT, dirCT, or synCT), was recalculated on different image sets to evaluate dose differences across IGRT and ART workflows.

specific adaptations were applied:

- Tumor coverage was quantified using $V_{95\%\text{Dose}}$ instead of D_{95}, allowing for normalized comparisons across variable prescription doses.
- For certain OARs – specifically the esophagus, stomach, spinal canal, and rectum – $D_{0.03\text{cc}}$ was used as a surrogate for D_{max}.

All analyzed structures are listed in Table 5.1. DVHs were computed using an in-house developed Python script. They were then used to extract relevant dose metrics for the analysis. To support visual inspection, voxel-wise dose and image difference maps were generated and displayed slice-wise, always comparing each recalculated plan to its corresponding reference plan (see Fig. 5.2).

In the end, relative dose differences were plotted for all metrics and all combinations of reference-recalculated plans, according to Fig. 5.2. Most relevant metrics in PCa and LungCa treatment were highlighted.

Correlation study

To assess the relationship between geometric and dosimetric consistency across imaging modalities, the Pearson correlation coefficient was calculated between SDC and relative dose differences. The SDC was computed for the body contour with a

2 mm tolerance, restricted to slices within ±5 of those containing the PTV. Relative dose difference δ_{dose} was defined as the mean absolute percentage difference between the recalculated and reference dose distributions within the same slice range of the body contour

$$\delta_{dose} = \frac{1}{N_s N_v} \sum_{m=1}^{N_s} \sum_{n=1}^{N_v} \frac{\left| s_{m,n}^{recalc} - s_{m,n}^{ref} \right|}{s_{m,n}^{ref}} \cdot 100 \,, \tag{5.1}$$

where s_{recalc} and s_{ref} are the recalculated and reference dose, N_s is the number of slices, and N_v the number of voxels that received dose. Reference and recalculated plans were defined as described in Fig. 5.2. For each patient and each pair of reference and recalculated plans, a single SDC value was obtained together with its corresponding relative dose difference. These pairs of values formed the dataset used for the correlation analysis. Correlation plots were created for LungCa cases, PCa cases, and a combination of both. All computations were performed using in-house Python scripts.

5.3 Results

5.3.1 Dosimetric evaluation

The results in this section will be organized by adaptation type – IGRT or different ART methods – and each adaptation type will cover PCa and LungCa cases separately.

IGRT

Initial comparisons were performed using the IGRT workflow applied to the pCT as a baseline reference. Representative axial slices are shown for three PCa and three LungCa cases, depicting dose differences between recalculated plans using CBCT, dirCT and synCT images, each compared to the reference pCT plan (Figs. 5.3 and 5.4). Dose differences between image variants varied across patients.

PCa In Patient 1, large discrepancies were observed around the body outline, which appeared consistent across all images. Within the body, dose distributions on the CBCT and synCT were more similar to each other than to the dirCT. No major anatomical differences were present, except in the rectum, which, in the observed slice, was empty in the CBCT and synCT but completely filled in the dirCT. Patient 2 also showed similar discrepancies at the body outline for all three images,

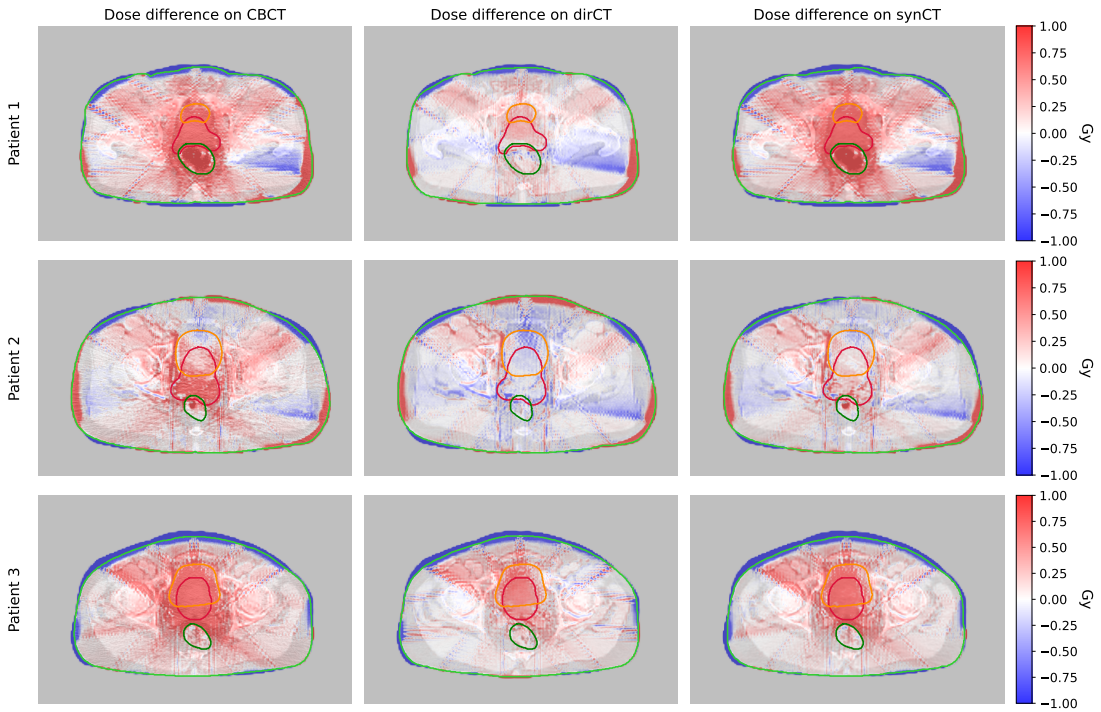


Figure 5.3. Example dose difference maps for three PCa cases. Structures are colored according to the following rule: Body contour – light green, rectum – dark green, bladder – orange, and PTV – red.

with one visible difference in the anterior region of the body contour in the dirCT. Rectal filling differed slightly between dirCT and the other two images. In this case, the synCT dose distribution aligned more closely with the dirCT, while the CBCT showed larger deviations. A noticeable difference between synCT and dirCT appeared in the posterior part of the PTV, in the region of seminal vesicles. Patient 3 presented a scenario similar to Patient 1, however this time the dose difference on the synCT was leaning more towards the one in the dirCT. Dose differences at the body outline were consistent across all images for Patient 3, and no significant anatomical differences were noted.

LungCa In the lung cohort, Patient 1 showed increasing dose differences moving from dirCT to synCT and then to CBCT. A reduction in dose differences was observed in the right lung (image left) in the synCT compared to the CBCT, though not to the extent seen in the dirCT. Dose differences along the body outline were consistent across all three images. Within the PTV, both the dose distribution and the underlying anatomy were more similar between the CBCT and synCT than between each of them and the dirCT. Patient 2 exhibited lower overall dose differences compared to other LungCa patients. The body outline showed consistent

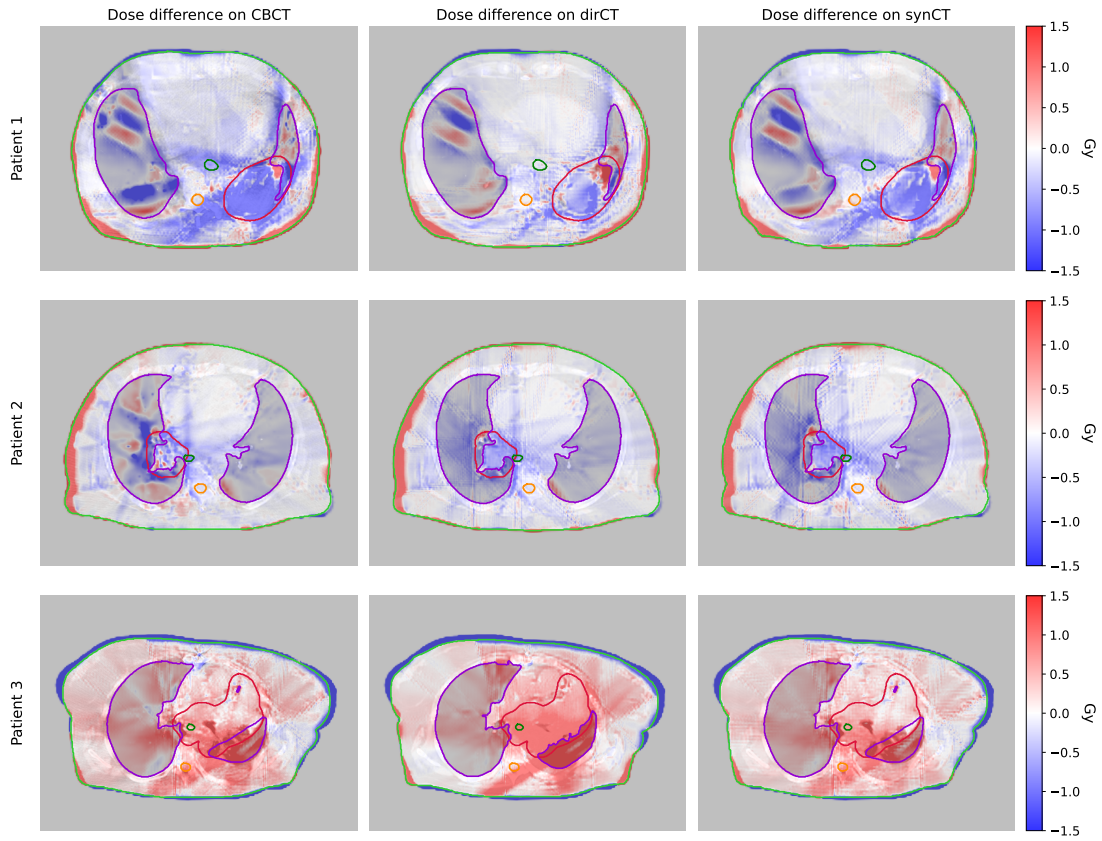


Figure 5.4. Example dose difference maps for three LungCa cases. Structures are colored according to the following rule: Body contour – light green, lungs – purple, esophagus – dark green, spinal canal – orange, and PTV – red.

discrepancies across all image types, while a larger deviation was observed in the right lung on the CBCT image relative to the dirCT and synCT. For Patient 3, CBCT and synCT dose difference maps were similar, with an additional reduction in differences in the right lung (image left) on the synCT difference map. In contrast to previous cases, dirCT showed larger discrepancies, particularly in the high-dose region around the PTV and in the left lung (image right).

To facilitate a more detailed analysis, the region around the esophagus was isolated in the LungCa case, and the region around the rectum was isolated in the PCa case (Fig. 5.5). Dose distributions from all four modalities (pCT, CBCT, dirCT and synCT) are displayed on a single slice, illustrating dose conformity, as well as the spatial relationship between the esophagus and the high-dose region in the LungCa case, and between the rectum and the high-dose region in the PCa case. Changes in size and position of OARs show a consequential change in the amount of dose deposited to them. The increase in delivered dose for the two cases visualized in Fig. 5.5 is supported by a quantitative dose comparisons for selected OARs in Table 5.2. The table includes the $D_{0.03cc}$ for the esophagus, spinal canal, stom-

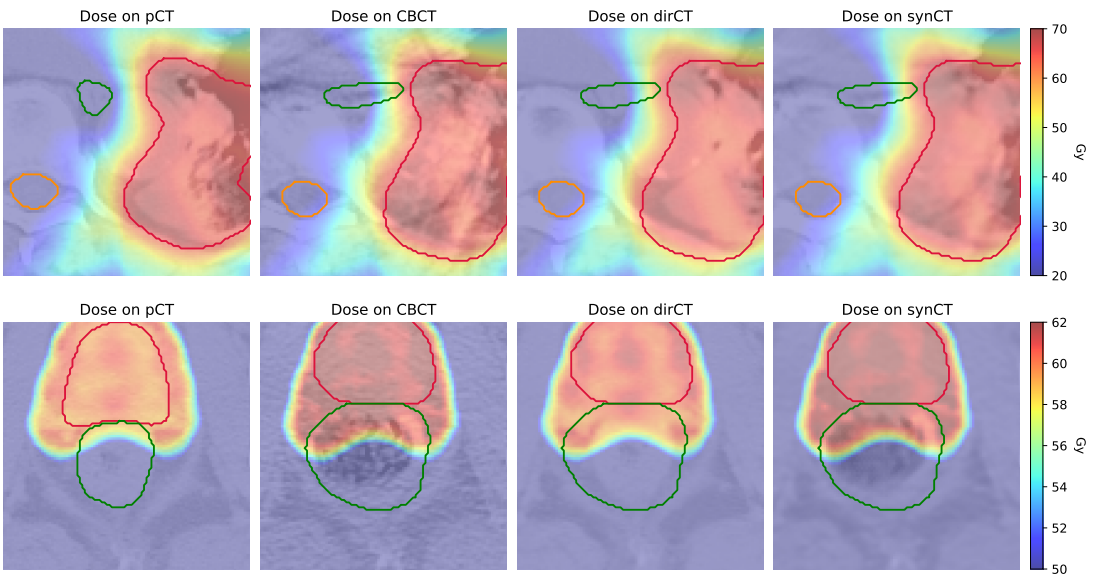


Figure 5.5. Esophagus (top) and rectum (buttom) overdosing in the IGRT scenario. Representation of contours by color is the same as in Fig. 5.3 and Fig. 5.4.

Table 5.2. Quantitative dose comparison for selected OARs of one LungCa and one PCa case highlighting the excess in delivered dose to sensitive tissue.

		LungCa			PCa	
		Esophagus	Spinal canal	Stomach/bowel	Rectum	
		D0.03cc	D0.03cc	D0.03cc	D0.03cc	V40
Reference	pCT	22.15	21.48	20.93	59.9	19.13
Recalculation	CBCT	59.1	27.39	34.72	63.03	30.34
	dirCT	58.65	27.77	35.05	61.26	29.54
	synCT	59.1	27.79	35.42	63	30.32

ach/bowel, and $D_{0.03cc}$ and V_{40} for the rectum, highlighting changes introduced by the IGRT workflow on different images. It should be noted that the changes in dose are generally consistent across all three image sets with an exception of $D_{0.03cc}$ of the rectum where the dose increase for dirCT is comparably smaller than for CBCT and synCT.

ART: dirCT optimized plan

PCa Dose difference maps generated by optimizing plans on the dirCT and recalculating them on CBCT and synCT images, showed consistent patterns across all three patients (Fig. 5.6). Differences along the body outline between CBCT and synCT were comparable for each case. In Patient 1, the overall dose distributions in both difference maps appeared similar, with more pronounced differences con-

fined to the high-dose region and the rectum. These differences are attributable to the rectal filling, which was present in the dirCT but absent in both CBCT and synCT. Patient 2 showed more substantial discrepancies between CBCT and dirCT, particularly within the PTV. The synCT dose distribution, by contrast, more closely matched the dose distribution in the dirCT. However, some heterogeneity was visible in the posterior PTV where the seminal vesicles are located. Patient 3 followed a similar trend, with the synCT dose distribution aligning more closely with the dirCT than with the CBCT, especially within the high-dose region.

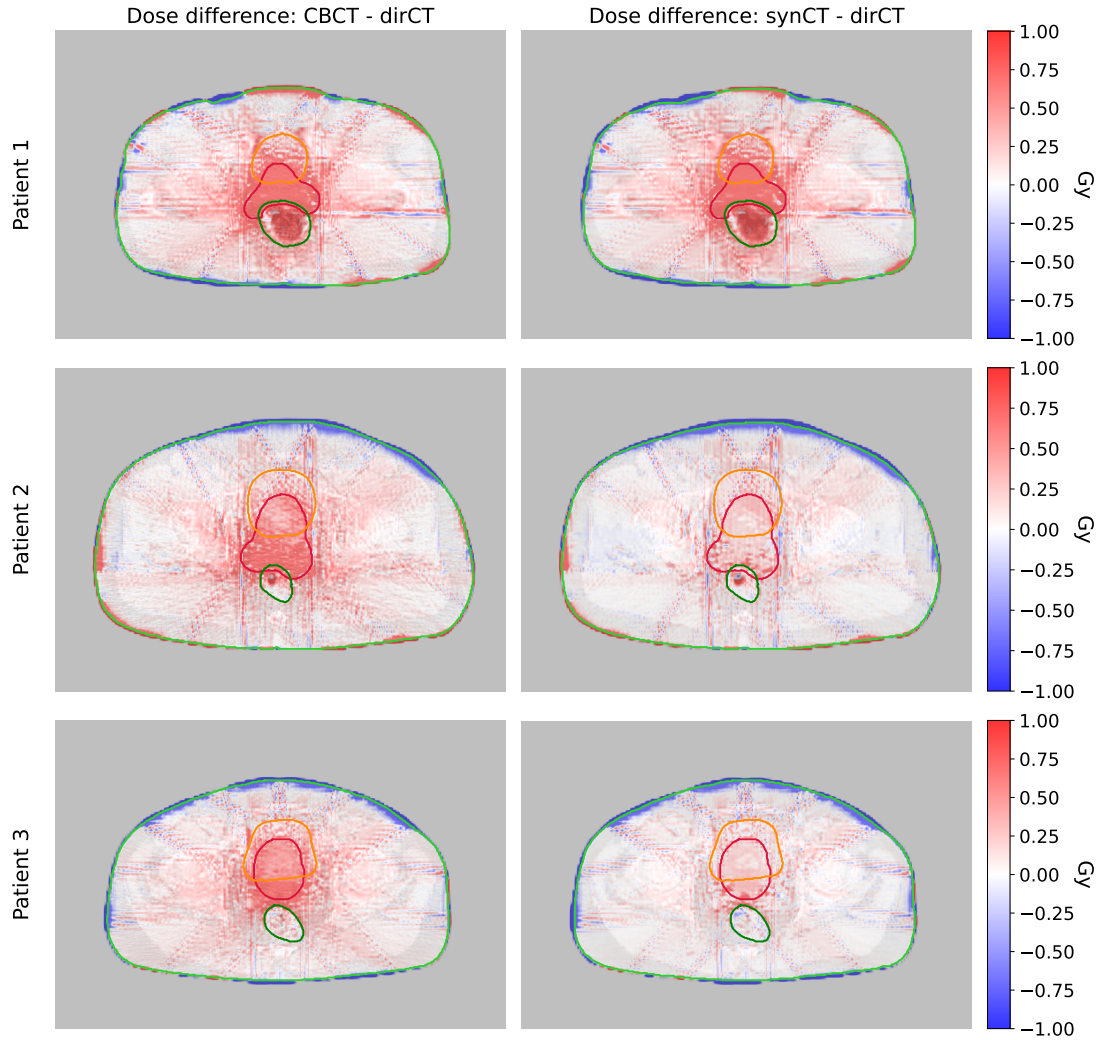


Figure 5.6. Dose differences for three PCa cases for a plan optimized on the dirCT. Representation of contours by color is the same as in Fig. 5.3.

LungCa In the lung cohort, dose difference maps showed overall larger discrepancies compared to the prostate case, reflecting increased sensitivity to anatomical variation. Although the general trends remained similar to those seen in the IGRT

scenario, the scale of the differences was smaller (visible dose difference range [-1, 1], compared to [-1.5, 1.5] in IGRT).

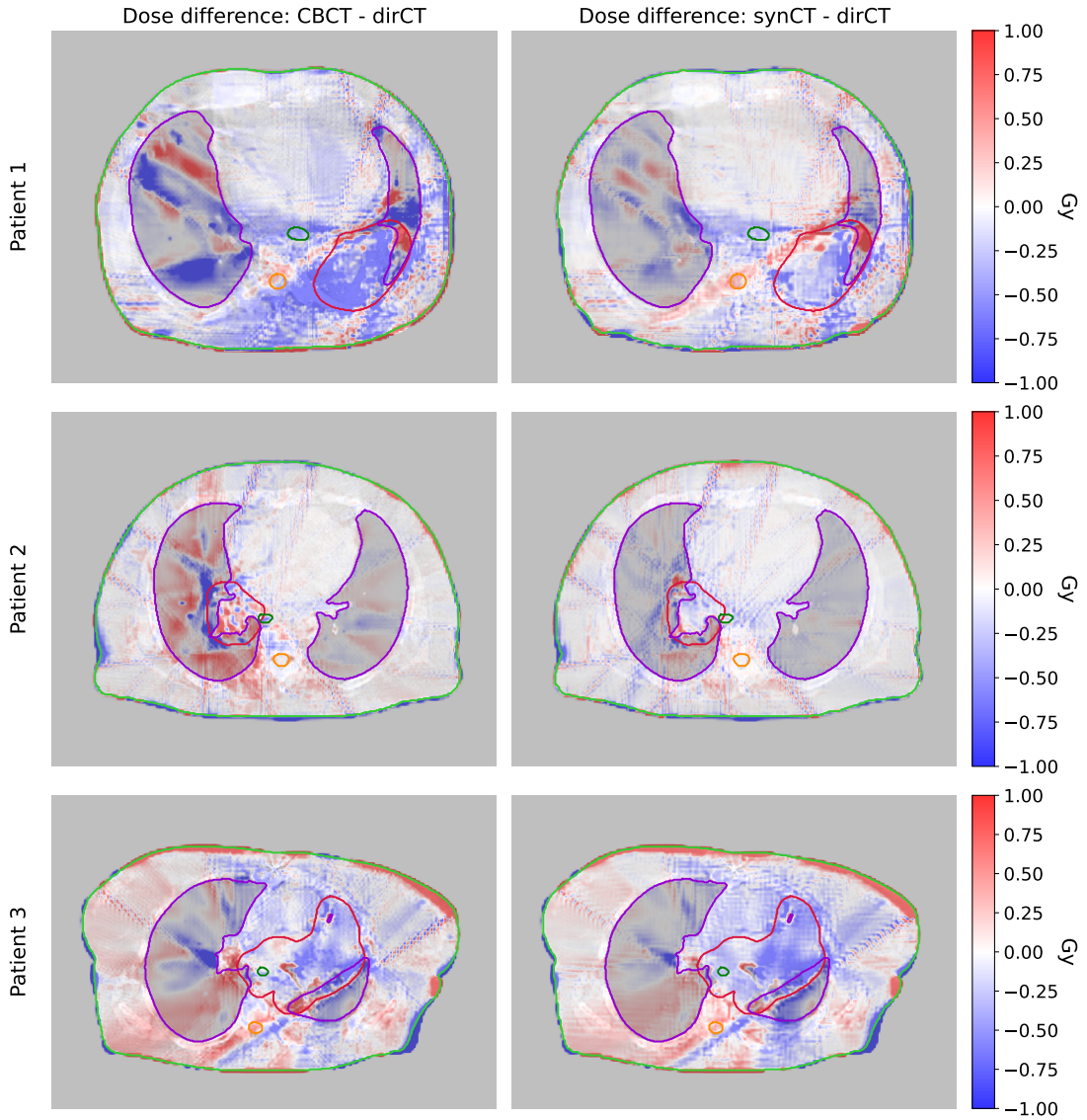


Figure 5.7. Dose differences for three LungCa cases for a plan optimized on the dirCT. Representation of contours by color is the same as in Fig. 5.4.

In Patient 1, the synCT difference map showed a reduction in dose deviations within both lungs and the PTV when compared to the CBCT. However, one consistent feature was observed in both maps: a region of elevated dose in the anterior part of the PTV at its interface with the lung. This corresponds to an anatomical feature present in the dirCT but not in the CBCT or synCT. Patient 2 displayed overall smaller dose differences than Patient 1, except in the lungs and PTV. The CBCT difference map revealed more distinct hot and cold spots in these regions, while the synCT map appeared more uniform. Patient 3 exhibited the largest differences along the body outline. The CBCT map showed more pronounced dose

variations in the right lung (image left) and at the boundary between the left lung (image right) and the PTV, with widespread discrepancies across the entire image.

ART: CBCT optimized plan

PCa For all three prostate patients, dirCT-CBCT dose difference maps showed patterns similar to those observed in the dirCT-optimized plan, but with an inverted sign, which was expected due to the change in reference plan. In the synCT-CBCT difference maps, the most prominent differences appeared in the PTV and bladder regions across all patients, with Patient 1 showing the smallest differences. A notable observation was the appearance of dose discrepancies around the body contour in the dirCT difference maps, which were largely absent in the corresponding synCT difference maps. This contrast highlights the alignment of synCT and CBCT in terms of external body anatomy, which was not the case for dirCT.

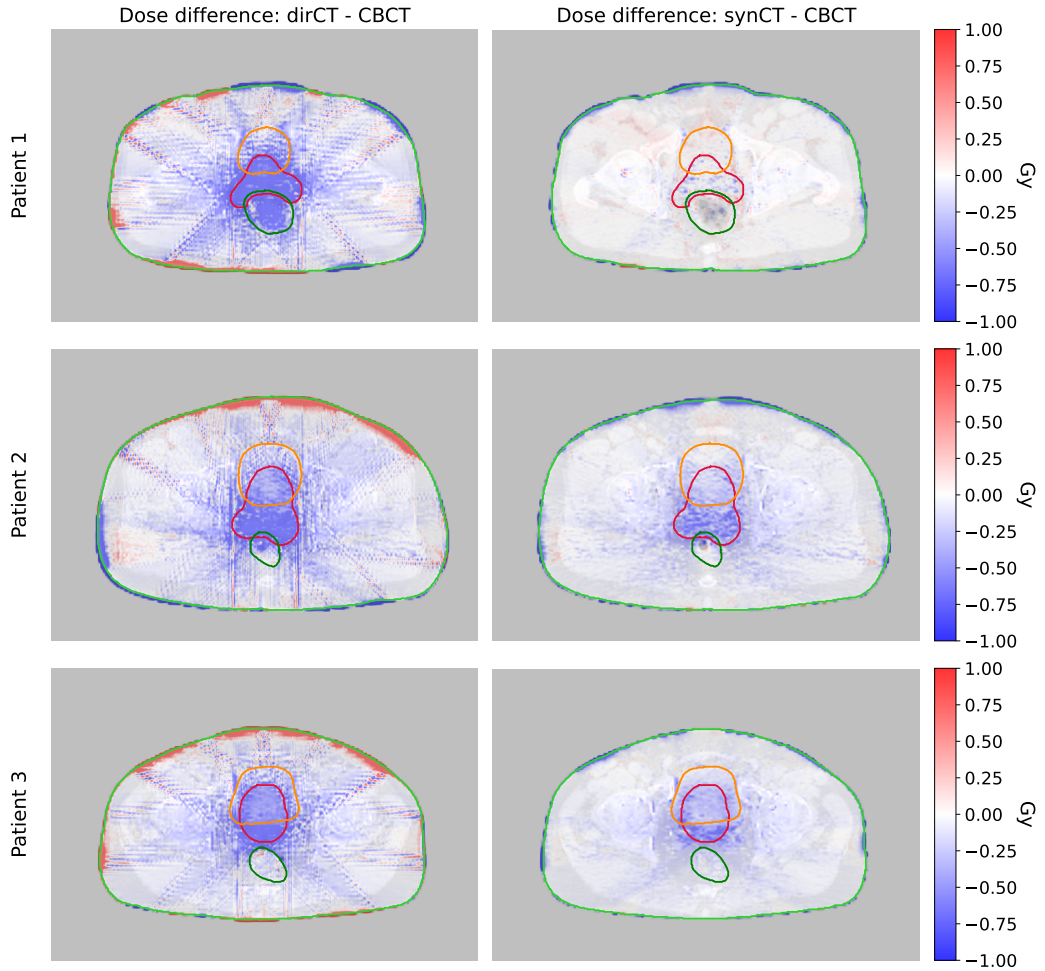


Figure 5.8. Dose differences for three PCa cases for a plan optimized on the CBCT. Representation of contours by color is the same as in Fig. 5.3.

LungCa Dose difference maps for the dirCT-CBCT combination closely mirrored the CBCT-dirCT results, with differences primarily reversed in sign, as expected due to the change in the reference plan.

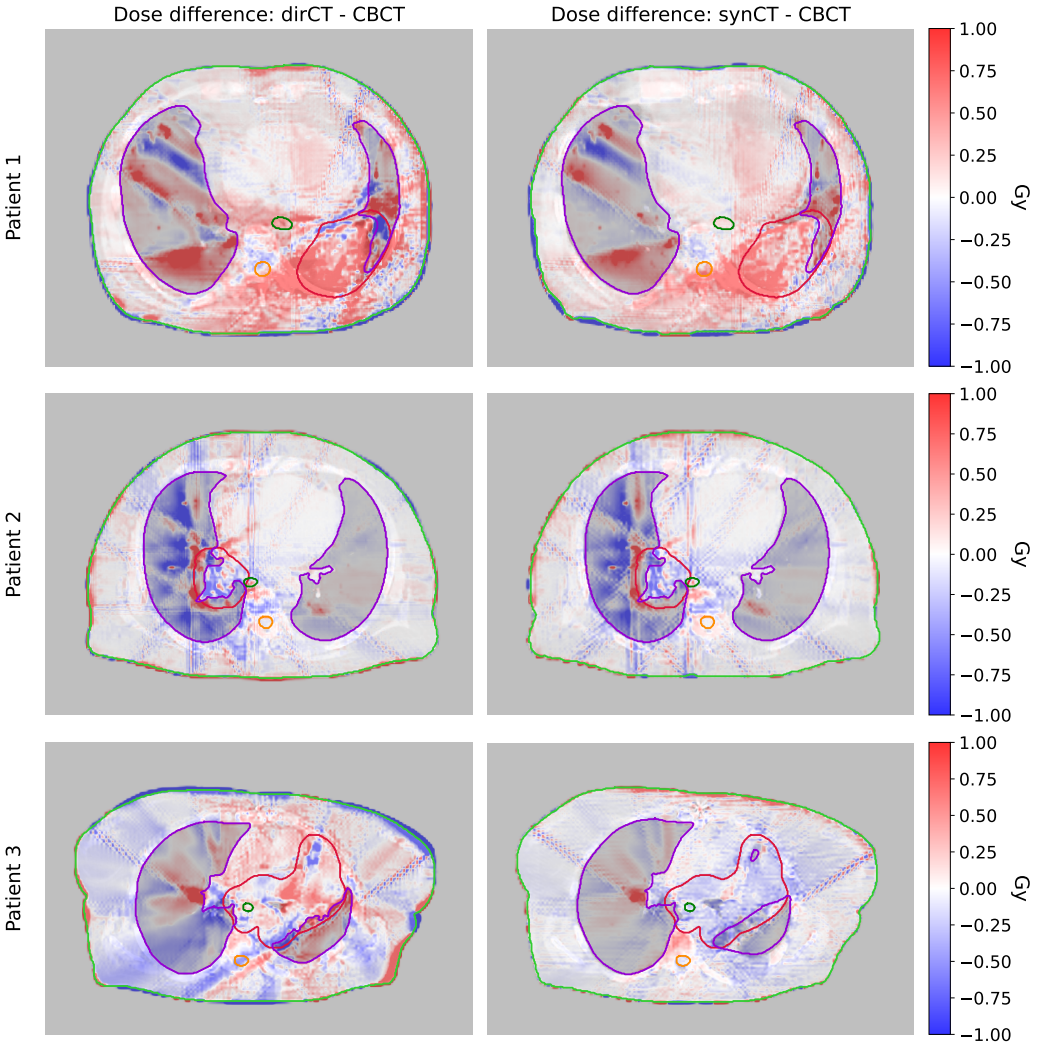


Figure 5.9. Dose differences for three LungCa cases for a plan optimized on the CBCT. Representation of contours by color is the same as in Fig. 5.4.

In Patient 1, motion-related artifacts in the lungs were similarly visible in both difference maps, indicating that the dirCT and synCT images were more comparable in image quality than the CBCT. A difference emerged in the anterior region of the PTV at its interface with the lung, where the synCT and CBCT appeared more anatomically aligned. Patient 2 showed highly similar difference maps, reflecting minimal anatomical variation between image sets. In this case, the benefit of synCT over CBCT was evident through reduced motion artifacts, particularly in the lungs. Patient 3 exhibited larger discrepancies in the dirCT-CBCT map, primarily due to anatomical mismatches, especially along the body outline. While some

differences persisted in the synCT-CBCT map, they once again suggested improved image quality relative to CBCT.

ART: synCT optimized plan

So far, results covered the IGRT method, and two ART methods, dirCT-based and CBCT-based. For all optimized plans, the effect on synCT images was shown as well. For the sake of completeness, the plan was optimized on the synCT to show that the optimization algorithm managed to find a clinically acceptable treatment plan, but recalculations were not performed on different modalities. As we saw in previous paragraphs, dose differences between two image types are complementary but have an opposite sign depending if the plan was optimized on the first or on the second image type. Therefore, the aim of creating a treatment plan and optimizing it on synCT was to show that all prescriptions and dose constraints could be satisfied as with any other image type. The plan optimization resulted in clinically acceptable plans for all PCa and LungCa cases without exception.

Overview of adaptation methods

Finally, Fig. 5.10 consolidates dose metrics across all cases, treatment-relevant structures and adaptation methods. Each data point represents a per-patient metric, grouped by adaptation type. Key metrics, such as D_{95} for the CTV, $D_{0.03cc}$ for OARs like the esophagus, spinal canal, stomach/bowel and rectum are highlighted to indicate trends across the complete patient cohort. The figure enables direct comparison of tumor coverage and OAR sparing performance between the pCT-based (IGRT), dirCT-based, and CBCT-based strategies and it clearly shows the benefit of plan adaptation. Looking at adapted plans, CBCT- and dirCT-based plans seem to have similar distributions of relative dose differences in the complete patient cohort, however there is a difference between the two distributions and it is statistically significant both when one looks at all evaluated dose metrics and only the key highlighted metrics (see Appendix B).

5.3.2 Correlation study

Following the dosimetric analysis, the correlation between dose differences and body outline similarity was evaluated. After calculating the SDC and the relative dose difference for pairs of data points as described in Section 5.2.2 and Table 5.2, a Pearson correlation coefficient was calculated for LungCa and PCa cases, and a

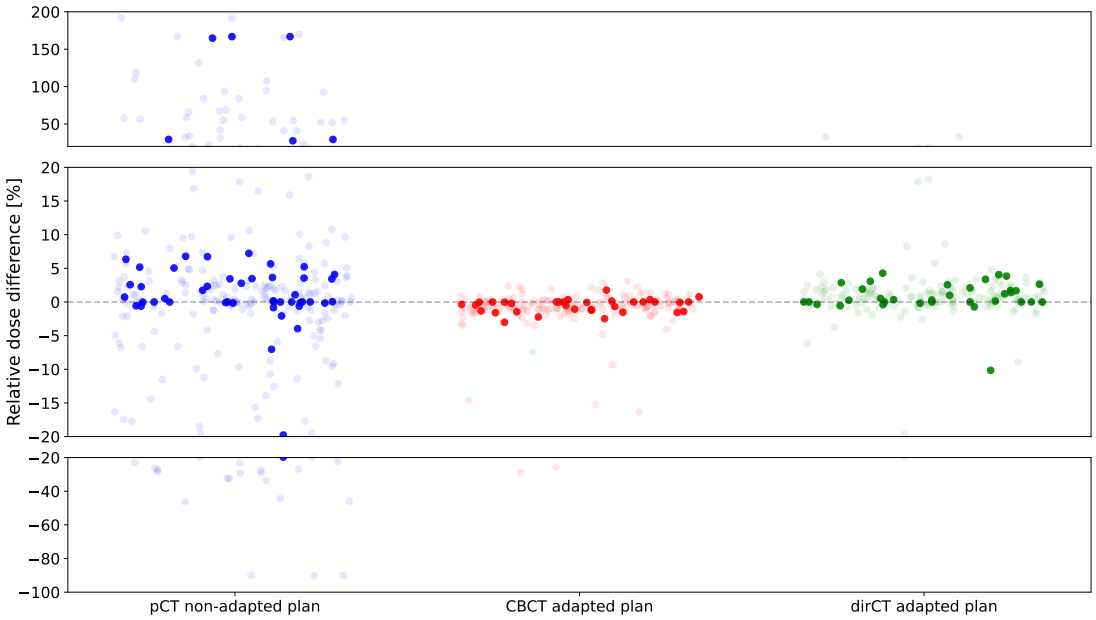


Figure 5.10. Consolidated dose metrics across all patients, structures, and adaptation methods. Each point represents a per-patient value, grouped by adaptation type. Key metrics include CTV D_{95} and $D_{0.03cc}$ for OARs such as the esophagus, spinal canal, stomach/bowel, and rectum.

combination of both (Table 5.3). Scatter plots of calculated metrics and a linear least squares fit are presented in Fig. 5.11, and Figs. A.1 and A.2 in Appendix A.

Pearson correlation coefficients and scatter plots indicate a strong negative correlation between the difference in dose and the difference in patient body outline. The larger the DSC, the larger the overlap between body contours, and the smaller the dose difference.

Table 5.3. Pearson correlation coefficient between SDC and the relative dose difference for LungCa, PCa, and all cases together.

	LungCa cases	PCa cases	All cases
Pearson corr. coef.	-0.94	-0.80	-0.83

5.4 Discussion

In this study, an IGRT strategy and three ART strategies were compared – dirCT-based, CBCT-based, and synCT-based – across two distinct disease sites: prostate and lung cancer. This evaluation is clinically relevant, as it provides insights into the ongoing shift from IGRT to ART workflows, and offers practical guidance to

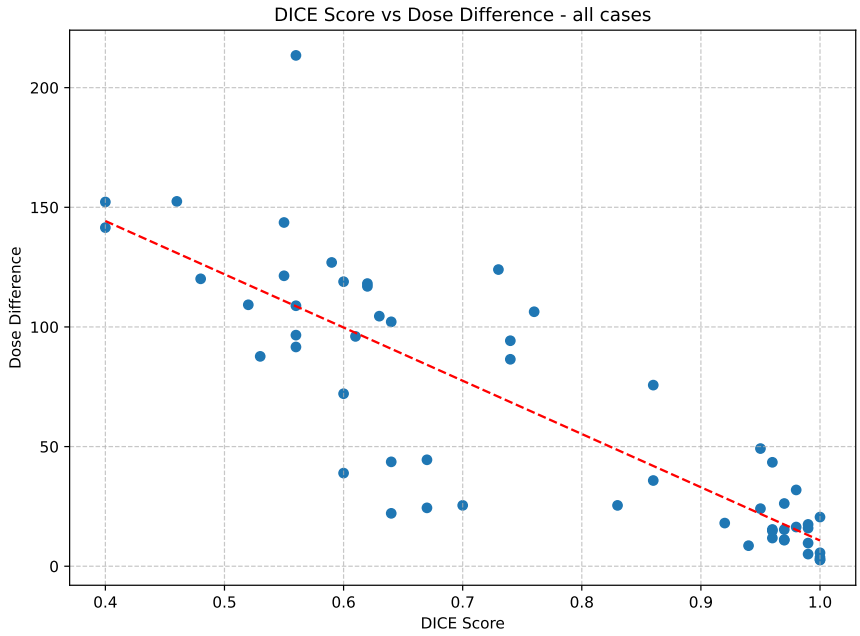


Figure 5.11. Scatter plot showing the relationship between body outline similarity (SDC) and relative dose difference for all investigated cases. Each point represents a recalculation pair as detailed in Fig. 5.2, with a linear least squares fit used to visually assess the correlation.

clinicians and medical physicists navigating a growing number of ART implementation pathways. This is the first study to import externally generated AI-based synthetic CTs into the Ethos™ TPS, and to benchmark their use within Ethos dose calculation and optimization framework against both conventional IGRT and deformable strategies in a controlled, dosimetrically grounded analysis.

Dosimetric evaluation: IGRT To fully explore the dosimetric impact of each imaging strategy, the IGRT workflow was analyzed first by recalculating the reference (pCT-based) plan on CBCT, dirCT, and synCT datasets. This analysis followed the first column of the graph in Fig. 5.2, offering insights into how different image types affect dose when the underlying plan is held constant. The recalculation revealed large deviations between dose distributions on the pCT and the recalculated dose on other imaging modalities. This discrepancy was especially pronounced along the outer body contour, where differences in body shape led to changes in beam attenuation from multiple directions which can influence the accumulated dose in the body to both the tumor and adjacent healthy tissue. The Pearson correlation analysis confirmed a strong association between body outline mismatches and observed dose deviations. In LungCa cases, dose differences of similar magnitude

appeared in the lung parenchyma which was often traceable to rib misalignment, as well as in the high-dose region surrounding the PTV. These differences are likely due to anatomical discrepancies between the pCT and daily images (dirCT, CBCT, synCT), as well as intrathoracic motion. In PCa cases, dose deviations emerged primarily in and around the rectum for CBCT- and synCT-based recalculations, but were less evident in the dirCT-based recalculated dose. This may be due to the inability of the deformable registration algorithm to accurately capture large internal deformations in the rectal region, particularly those caused by variations in rectal filling and gas bubble distribution. Another important observation from the IGRT-based analysis is the potential for significant OAR overdosing. As shown in Figure 5.5 and Table 5.2, recalculating the reference plan on daily images revealed that substantial increases in dose to critical structures can occur, particularly when anatomical misalignment is pronounced. This is in line with previous findings in both PCa [45] and LungCa irradiation [48, 51]. While not present in all patients, these deviations occurred in all LungCa cases and 2 out of 5 PCa cases which puts into question the reliability of IGRT-based workflows in certain clinical scenarios.

Dosimetric evaluation: ART Adapting the plan on dirCT and recalculating it on synCT and CBCT images improved tumor coverage across all cases compared to IGRT. While recalculations on CBCT still showed some underdosage in the PTV, the overall improvement highlights the value of adapting the treatment plan on accurate body contours and daily structures, even if the underlying anatomy in the dirCT is not fully accurate due to registration limitations. Looking at the LungCa cases in the dirCT adapted plan, dose difference maps revealed residual hot and cold spots in CBCT and synCT recalculations, reflecting anatomical differences to the reference image near the tumor and lung regions. Dose deviations in the lung parenchyma were reduced when recalculating on synCT versus CBCT, suggesting that the 3DCycleGAN-based synCT mitigates motion artifacts in CBCT images, which are common for thorax imaging. When adapting directly on CBCT, differences emerged between sites. In LungCa cases, the CBCT-adapted plan recalculated on synCT and dirCT showed a closer match between two dose difference maps, again pointing to synCTs corrective effect. In PCa, CBCT and synCT were more similar in image quality and anatomy, resulting in similar dose distributions and less pronounced dose differences. However, in both anatomical regions, CBCT-dirCT difference maps showed similar discrepancies along the body outline and with certain OARs, reflecting the limitations of deformable registration in handling both the internal organ and skin deformation.

Overview of adaptation methods

Finally, the graph with accumulated data points from all relevant metrics (Fig. 5.10) confirms two things: first, it highlights the benefit of ART irrespective of the workflow, and second, it shows that there is a difference between dirCT- and CBCT-optimized plans. Considerations like motion artifacts on CBCTs or registration errors in dirCTs need to be taken with care when choosing the ART workflow.

Clinical implications of synCT in ART The results of this study support the clinical value of using synCTs for ART. Compared to CBCT, synCT offers improved image quality, reduced artifacts, and more consistent HU distributions, which are critical for accurate dose calculation. By more accurately reflecting daily patient anatomy, synCTs strengthen the connection between planned and delivered dose, leading to safer and more effective treatments. For patients, this translates to better tumor coverage and more reliable sparing of organs at risk, particularly in anatomically dynamic regions like the thorax and pelvis. However, synCTs are not without limitations. Their performance depends on the quality and diversity of the training data used to generate them. Severe artifacts in the original CBCT may not be fully corrected, and out-of-distribution cases can result in inaccuracies in HU values or anatomical representation. This limits the generalizability of current models across institutions, scanners, or patient populations without retraining or extensive validation. Future work should focus on expanding training datasets and improving model robustness to ensure consistent performance in diverse settings.

5.5 Conclusion

In this study, three ART strategies were compared across PCa and LungCa cases: CBCT-, dirCT-, and synCT-based. The results show that all three approaches are feasible for daily treatment adaptation and that adaptation consistently improves dose accuracy compared to IGRT alone.

Among them, synCTs provided better image quality than CBCT and more accurate anatomical representation than dirCT, resulting in more reliable dose distributions, particularly in areas with heterogeneous anatomy. dirCT-based planning showed benefits from using daily contours, but its accuracy was limited in regions with large deformations, such as the rectum, or lung parenchyma. These findings suggest that the choice of ART method can affect treatment quality, and that synCTs offer a promising option for centers looking to implement daily adaptation.

6

Synthetic Images for Autosegmentation

This chapter is based on the paper titled “Cross-Modality Supervised Prostate Segmentation on CBCT for Adaptive Radiotherapy” by Kovács and Stanić, et al., accepted for publication and in press for the MICCAI 2025 Proceedings. Most of the results and examples discussed here are derived from that publication.

6.1 Motivation

In Chapter 5, we have seen how synCT images generated using deep learning methods can be used for treatment planning and dose calculation in ART workflows. This chapter will focus on the use of synthetic images and deep learning algorithms in the context of automatic segmentation tasks in PCa.

Accurate organ delineation is crucial for PCa radiotherapy. As discussed in Section 1.3, continuous anatomical changes require frequent dose adaptations to precisely target the prostate while minimizing radiation exposure to surrounding healthy tissue [131, 132]. In Section 1.4.1 we saw that CBCT is widely used for treatment adaptation, but its low image quality and suboptimal soft tissue contrast make organ delineation time-consuming and error-prone [133, 134]. Furthermore, soft tissue deformations that occur during the lengthy contouring process can render delineations outdated before treatment begins [135]. All of this makes high-quality annotations sparse and hinders the adoption of AI-driven segmentation in clinical workflows.

Existing automated solutions fall short in addressing these challenges, e.g., DIR algorithms, which transfer segmentations from pCT to CBCT, can be unreliable in cases of strong local soft tissue deformations [136, 137]. However, state-of-the-art CNN models, particularly the nnU-Net [138], have demonstrated strong performance in medical image segmentation, but training a reliable model requires large, well-annotated datasets, which is an obstacle for CBCT due to the lack of high-quality ground truth (GT). Monte Carlo simulations can bridge the pCT and CBCT domains for CNN training [139], but they are complex, device-specific, and prone to modeling errors. Another approach involves using CycleGANs to generate synthetic CTs or MRIs from CBCTs to aid in automatic contouring of OAR [140, 141].

This study proposes a cross-modality supervised domain translation framework for prostate segmentation in CBCT, addressing the challenges of low image quality and the scarcity of annotations. Instead of enhancing the quality of CBCT to the pCT level, we generate synthetic CBCT (synCBCT) from pCT using a CycleGAN-based translation pCT $\mapsto$ CBCT. This enables training segmentation models on synCBCT while leveraging high-quality pCT-derived annotations. Furthermore, anatomy-aware augmentation [142] enhances the robustness to soft tissue deformations, ensuring reliable segmentation in real-world clinical scenarios.

This is the first implementation of a cross-modality supervised learning method for CBCT segmentation that effectively bridges the domain gap without requiring manual CBCT annotations. By combining generative modeling with robust aug-

mentation strategies, this method facilitates clinically viable AI-driven segmentation for ART, reducing the annotation burden while improving treatment precision.

6.2 Methods

The key innovation of this study is to provide high-quality pCT-derived GT for training segmentation networks on CBCT images, as outlined in Fig. 6.1. To achieve this, the pCT $\mapsto$ CBCT branch of a 3DCycleGAN was used to convert high-quality pCT images into synthetic CBCT (see Fig. 2.9 in Section 2.6).

This fully learnable strategy avoids the pitfalls of simulation- and deformation-based methods. The nnU-Net was then trained using synCBCT images generated from pCT while transferring the original pCT segmentations as GT. This approach allowed the network to leverage high-quality pCT annotations while operating in the CBCT domain. As a result, this method bridged the domain gap without compromising segmentation quality.

6.2.1 Characteristics of datasets

The study included 56 images from six institutes as detailed in the following. All images used in the study were resampled with an in-plane spacing of 0.875 mm and a slice thickness of 2.5 mm. TS v2.0.5 was used to extract the body contour and remove the imaging table or the treatment table. OARs included in the study were the femur, bladder, rectum, penile bulb, prostate, and seminal vesicles.

SPARK & Gold Atlas pCT images from two external open access datasets, SPARK [143] and Gold Atlas [144] were used. 2 out of 5 treatment centers from the SPARK dataset were included in the study, with 7 patients (4+3). The rest were excluded due to the study criterion that the prostate contours should exist and that they are separated from the seminal vesicles. All 3 centers in the Gold Atlas dataset were represented by 19 patients (8+7+4). According to the original Gold Atlas study, the contours associated with the pCT were first delineated on MRIs, and the pCTs were subsequently deformably registered to the same MRIs allowing them to inherit the contours. On the other hand, the SPARK dataset contained organ structures that were contoured directly on the pCT. The distance between the prostate and the penile bulb was measured to assess compliance with the ESTRO ACROP guidelines [133], which specify a minimum separation of 1 cm between these structures. The mean distance between the prostate and the penile bulb was

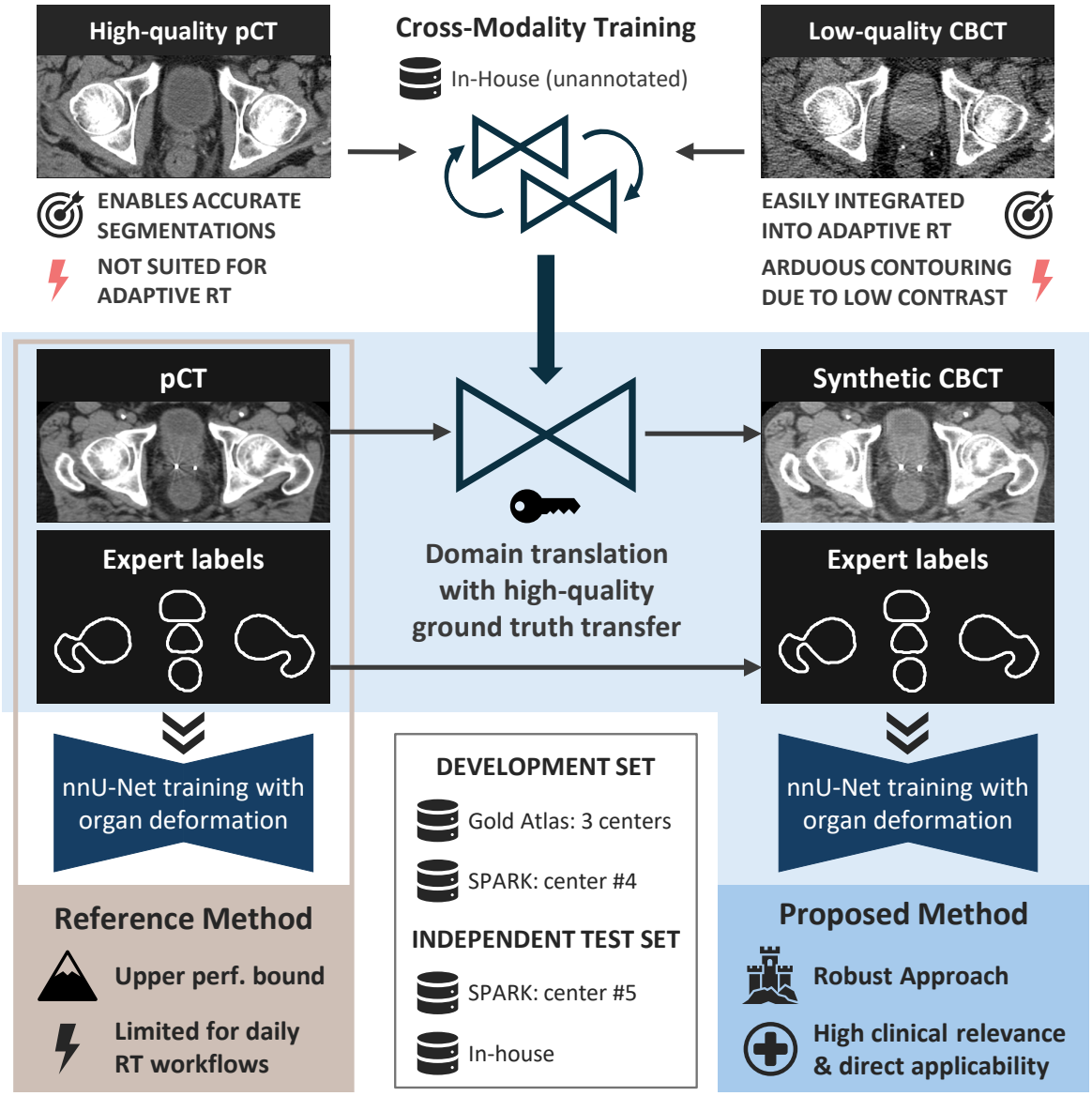


Figure 6.1. The workflow of the proposed framework that aims to address annotation scarcity on low-quality CBCT images. By combining generative neural networks with robust augmentation strategies, the method facilitates clinically viable AI-driven segmentation for ART.

12.0 ± 4.8 cm, with 8 patients (42.1%) failing to meet the minimum separation requirement.

Unannotated in-house pCT and CBCT 27 PCa patients, as described in Section 4.2.2, were chosen for the study. A pCT and three equally spaced fraction CBCT images from each patient were used for 3DCycleGAN training.

Annotated in-house CBCT CBCTs of three separate patients treated with Ethos™ at DKFZ were contoured by a medical student following the ESTRO delineation guidelines. The standard procedure developed with the guidance of a senior radiation oncologist was used to ensure proper interpretation and application.

6.2.2 Experimental setting

1. How does CBCT with pCT-derived GT compare to high-quality pCT in the context of model training? Two segmentation models were trained using the same pCT-derived GT: one on the pCT images and another on the synCBCT images generated via the pCT $\mapsto$ CBCT branch of the 3DCycleGAN. This ensured that the only difference in training was the information content within the image domain. The model trained on pCT served as an upper performance bound, while the model trained on low-quality synCBCT with pCT-derived GT aimed to match its performance. Additionally, model robustness was increased by simulating soft tissue deformations during model training [142] and its parameters were optimized during pCT training. An independent test set from a single center in the SPARK dataset was used for the evaluation. Original pCT images were used for evaluation of the model trained on pCTs and the corresponding synCBCT images for the model trained on synCBCTs.

2. How does the proposed method perform on clinical data? Segmentation models were evaluated on three increasingly difficult cases from the independent annotated in-house CBCT dataset, with distinct anatomical properties.

- Slim patient with moderate beam hardening artifacts
- Patient with prostate growth into the bladder
- Obese patient with severe beam hardening $\Rightarrow$ barely visible organ contours

This dataset served as a clinical validation and also helped determine whether training directly on CBCT is necessary or if pCT-trained models suffice.

6.2.3 Evaluation

The performance of the 3DCycleGAN network was quantitatively evaluated using SSIM, MSE, and JSD, as defined in Section 2.7. As demonstrated in Chapter 4, the generated images were closer to the target modality than to the input. Qualitatively, the generated images were visually inspected, confirming their anatomical fidelity. To evaluate the segmentation model and assess label similarity, three metrics were used: DICE, surface dice (SDC) and Hausdorff distance (HD), with evaluation parameters defined in Section 2.7. The metrics were calculated for three structures: prostate, rectum, and bladder. To ensure clinical relevance, the evaluation was limited to image slices within the typical extent of prostate PTV in radiotherapy, which is 7-10 mm in the inferior and posterior directions [134]. Due to truncated rectum segmentations in the development dataset, the inferior limit was set to the maximum extent of the prostate contour. Adherence to the ESTRO guidelines was also evaluated regarding the distance from the prostate to the penile bulb.

6.2.4 Model training of the proposed strategy

Domain-translation network training The network that was trained on the PCa dataset for the purposes of various investigations in the thesis (Chapters 4 and 5) was used for this study as well. No changes were made to the model weights. The only notable exception was inverting the direction of inference from CBCT $\mapsto$ synCT to pCT $\mapsto$ synCBCT. This meant using the generator G_{B2A} instead of G_{A2B} as depicted in Fig. 2.9 in Section 2.6.

Segmentation network training The 29 patients were stratified by medical center and available GT structures into a **training set** (4 from ■ SPARK center #4, 19 from all 3 ■ GOLD ATLAS centers) and a **test set** (3 from ■ SPARK center #5 and 3 from ■ the annotated in-house dataset). The test set included the prostate, bladder, rectum, and penile bulb, while the training set also included the femur and seminal vesicles. Images were globally z-score normalized using nnU-Net preprocessing. 3D nnU-Net models were trained in 5-fold cross-validation (5fCV), adopting the anatomy-informed augmentation [142] to enhance robustness against organ deformations. Hyperparameters for the deformations were optimized based on 5fCV results. In the final model, either rectum or bladder deformations were applied with a 7.5 % probability each, using a Gaussian kernel $\sigma = 8$ and deformation amplitude sampled uniformly in $C = [-75, 75]$. An example rectum deformation can be seen in Fig. 6.2. Postprocessing was done by the nnU-Net pipeline. The code for

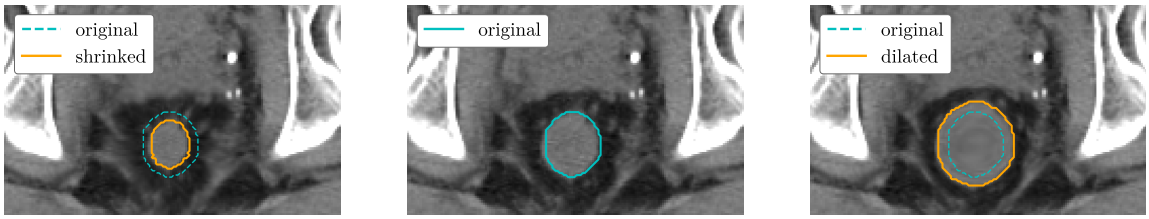


Figure 6.2. Example of the anatomy-informed augmentation used during network training. Left figure shows the original rectum contour in blue and shrunk rectum in orange, center figure shows the original rectum contour, and the right figure shows the dilated rectum in orange, with the original contour for comparison.

both model training pipelines is publicly available at {https://github.com} /DKFZ-OpenMedPhys/3DcycleGAN & /MIC-DKFZ/anatomy_informed_DA.

6.3 Results

The first experiment, which as a goal had incorporating rectal and bladder deformations during model training showed slightly improved or unchanged 5fCV segmentation performance, see Table 6.1. Then, after adopting the augmentation pipeline the models trained on pCT and synCBCT with pCT-derived GT were compared (first research question in Section 6.2.2). The results in Table 6.2 show comparable performance, even with slight HD improvements for the proposed synCBCT based method. The proposed method was just 0.01 below the upper bound in all DICE and SDC metrics, except for the prostate SDC, where a drop of 4 percent points was noted.

To address the clinical need and viability of the proposed method (second research question in Section 6.2.2), the performance of models trained on synCBCT with pCT-derived GT was compared against those trained on planning CT on an ■ annotated in-house dataset. The synCBCT-trained models demonstrated significant improvements in prostate and rectum segmentation across all evaluation metrics, while pCT-trained models showed superior bladder segmentation (Table 6.3). Observing the results on a case-by-case basis on three patients described in Section 6.2.2 showed variable results (see Table 6.4) which will be further discussed in Section 6.4.

The prostate-penile bulb distance for the synCBCT-trained model was 10.0 ± 4.7 mm for the ■ SPARK center #5, and 10.0 ± 2.5 mm for the ■ annotated in-house dataset.

Table 6.1. Model training on pCT with and without simulating organ deformations.

Training with	Prostate			Rectum			Bladder		
organ deform.	HD [mm]	SDC	DICE	HD [mm]	SDC	DICE	HD [mm]	SDC	DICE
No	5.0	0.63	0.83	7.3	0.65	0.79	4.7	0.80	0.88
	±1.2	±0.12	±0.06	±2.8	±0.15	±0.08	±3.5	±0.10	±0.06
Yes	5.0	**0.64**	**0.83**	**6.9**	**0.66**	0.79	**4.6**	**0.81**	**0.89**
	±1.2	±**0.11**	±**0.05**	±**2.5**	±**0.15**	±0.08	±**3.2**	±**0.10**	±**0.06**

Table 6.2. Segmentation results and differences on the independent ■ SPARK center #5 between the pCT-trained model (upper bound) and the synCBCT-trained model with pCT-derived GT (proposed).

Training	Prostate			Rectum			Bladder		
Modality	HD [mm]	SDC	DICE	HD [mm]	SDC	DICE	HD [mm]	SDC	DICE
pCT	4.2	**0.63**	**0.84**	3.3	**0.90**	0.89	4.2	**0.83**	**0.89**
(upper bound)	±0.6	±**0.09**	±**0.01**	±1.0	±**0.06**	±0.01	±2.4	±**0.10**	±**0.06**
synCBCT	**4.1**	0.59	0.83	**3.0**	0.89	0.89	**4.2**	0.82	0.88
(proposed)	±**0.4**	±0.09	±0.02	±**1.1**	±0.09	±0.02	±**2.3**	±0.11	±0.07
Difference	0.1	0.04	0.01	0.3	0.01	0.00	0.0	0.01	0.01

Table 6.3. Segmentation results on the ■ annotated in-house dataset (CBCT) for the pCT-trained model and the synCBCT-trained model with pCT-derived GT (proposed).

Training	Prostate			Rectum			Bladder		
Modality	HD [mm]	SDC	DICE	HD [mm]	SDC	DICE	HD [mm]	SDC	DICE
pCT	8.8 ±3.6	0.43 ±0.23	0.75 ±0.11	7.7 ±6.4	0.76 ±0.14	0.83 ±0.05	**2.8 ±1.2**	**0.91 ±0.06**	**0.94 ±0.03**
synCBCT (proposed)	**7.2 ±2.4**	**0.50 ±0.23**	**0.78 ±0.10**	**7.5 ±4.3**	**0.79 ±0.04**	**0.83 ±0.01**	4.4 ±3.2	0.86 ±0.04	0.87 ±0.05

Table 6.4. Segmentation results of the ■ annotated in-house dataset case-by-case analysis for a synCBCT-trained model.

In-house	Prostate			Rectum			Bladder		
patient	HD [mm]	SDC	DICE	HD [mm]	SDC	DICE	HD [mm]	SDC	DICE
#1	6.6	0.62	0.83	12.5	0.75	0.83	2.5	0.91	0.93
#2	5.0	0.65	0.85	5.1	0.78	0.82	8.3	0.83	0.86
#3	9.8	0.23	0.66	5.0	0.83	0.85	2.5	0.85	0.83

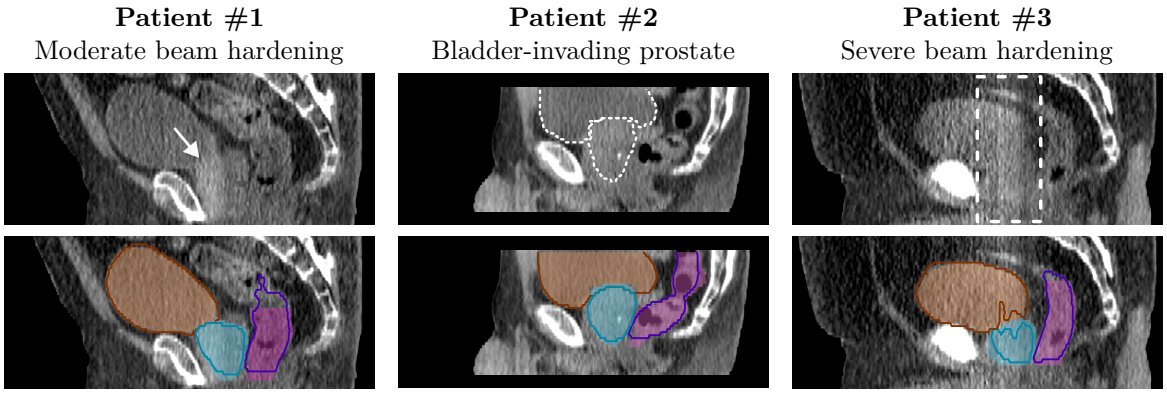


Figure 6.3. Visualization of segmentation results for three patients from the annotated in-house dataset. The top row highlights case characteristics (white arrow and white dashed lines) presented in the title of each column. The bottom row shows predicted contours (solid colored lines) overlaid on the ground truth (shaded regions): bladder (orange), prostate (light blue), and rectum (purple).

6.4 Discussion

The integration of AI-based segmentation into ART workflows involves several technical challenges, especially related to the quality and availability of training data in CBCT imaging. While many existing methods depend on manual annotations of low-quality CBCT scans, this work investigates an alternative approach based on domain adaptation. The discussion that follows focuses on several key aspects: the influence of synthetic training data on segmentation performance, the effect of anatomical deformation on model robustness, and the potential clinical implications of the proposed framework.

Robust performance approaching the upper bound - research question 1

The segmentation model trained on synCBCT paired with pCT-derived GT achieved results comparable to those of a model trained on high-quality pCT, showing no significant drop in performance and in some cases even yielding competitive outcomes. This indicates that the model was able to extract sufficient features to compensate for the reduced image quality of CBCT. Although both pCT and CBCT rely on physics of X-ray attenuation to generate an image, they differ in acquisition parameters, scatter levels, and reconstruction techniques. These differences, along with typically lower contrast-to-noise ratio and greater susceptibility to scatter artifacts of CBCTs, especially in the pelvic region, pose a challenge for soft-tissue segmentation [145, 146]. The strong performance of the model despite these fac-

tors may be attributed to the use of high-quality training labels. Furthermore, accounting for organ deformation, particularly bladder and rectal volume changes, contributed to improved segmentation accuracy.

Promising clinical feasibility - research question 2

Model performance The prostate segmentation results on the annotated in-house CBCT dataset indicate that models trained directly on target domain data may offer advantages over those trained solely on pCT. While the evaluated cases were considered challenging, the findings should be viewed as preliminary due to the limited sample size. Performance across OARs was mixed: rectum segmentation showed signs of improvement, whereas bladder segmentation was broadly consistent with results from the SPARK dataset. A case-by-case analysis showed that the model managed moderate beam hardening artifacts and instances of prostate growth invading the bladder reasonably well but struggled in cases with severe imaging artifacts. This is consistent with expectations, since such complex scenarios were likely underrepresented in the development set. Overall, the observed trends are encouraging and highlight the potential value of further studies using larger and more diverse training datasets.

Clinical implementation Beyond segmentation accuracy, inference speed and seamless integration into clinical workflows are essential for practical deployment. In the proposed framework, only the segmentation network is executed at inference time, with processing times under 30 s per CBCT on an NVIDIA RTX 2080 GPU (<10.7 GB memory usage). This is significantly faster than pipelines involving sequential synthesis and segmentation steps. The rapid inference supports the time-sensitive nature of prostate radiotherapy, where minimizing on-table delays is critical to avoid anatomical changes that could compromise treatment accuracy.

Compliance with ESTRO ACROP guidelines Segmentation models predicted prostate and penile bulb structures with inter-organ distances that generally adhered to ESTRO ACROP guidelines across the SPARK and in-house datasets. However, a few individual cases did not meet the minimum required separation. This suggests that the models were able to learn spatial relationships between structures to the extent supported by the training data. Notably, approximately 40% of patients in the development dataset did not satisfy the minimum separation requirement, which likely limited the models ability to consistently predict guideline-compliant distances. As such, training on fully ESTRO ACROP-compliant datasets

would likely improve the alignment with guideline-based expectations in evaluation scenarios.

6.5 Conclusion

This work aimed to support the development of clinically applicable, AI-driven segmentation of the prostate and OARs for PCa ART, by addressing the challenge of annotation scarcity in low-quality CBCT images – a major barrier to routine clinical implementation. To this end, a domain adaptation framework was proposed that leverages synthetic CBCT images paired with high-quality pCT-derived annotations. By explicitly incorporating organ deformations into the training process, the method was designed to improve the model robustness to daily anatomical variations, which are common in prostate radiotherapy workflows.

The results demonstrated that this approach can achieve segmentation performance comparable to models trained directly on high-quality images, even when evaluated on a set of challenging, real-world clinical cases. These findings support the feasibility of deploying AI-based segmentation tools in ART settings, offering a practical path toward reducing manual workload while maintaining high precision in treatment adaptation.

Summary and Outlook

With cancer cases continuing to increase worldwide and the need for more precise treatment options becoming ever more urgent, ART has emerged as a promising radiotherapy treatment avenue by tailoring treatment plans to daily anatomical changes. In this context, the present work explores the feasibility of integrating synCT images, generated with a 3DCycleGAN neural network, into clinical ART workflows. The approach is evaluated against existing CBCT- and dirCT-based methods in lung and prostate cancer cases, while also extending the discussion to downstream applications, such as automatic organ segmentation, where synthetic images may provide additional benefits.

Initial investigations in Chapter 3 examined whether CBCT and dirCT images provide sufficient accuracy for daily dose calculations in adaptive workflows. The analyses revealed notable changes in lung volume over the course of treatment in CBCT and dirCT images in most patients, reflecting limitations of the DIR algorithm used to deform pCTs to daily scans. Dosimetric evaluations indicated a systematic shift in delivered dose, likely due to HU inaccuracies in CBCT, as well as time-dependent differences in the magnitude of this shift. These results suggest that individual anatomical changes, such as tumor shrinkage or organ motion can influence dose accuracy in unpredictable ways. Taken together, the findings emphasize the need for patient-specific scrutiny of CBCT- and dirCT-based adaptation methods. Future work could investigate advanced CBCT technologies such as Varian Hypersight™ and implement novel AI-based DIR methods to address these limitations. Nonetheless, the findings of this study indicate that deep learning-based synCTs may offer a more robust foundation for adaptive workflows.

Building on this motivation, the next stage of the work, presented in Chapter 4, examined whether deep learning-generated synCTs could address both the image quality issues of CBCT and the anatomical inconsistencies of dirCT. Two 3DCycleGAN models were trained to generate synCTs, one for the prostate and one for the lung region. The models were then evaluated using quantitative metrics, histogram comparisons, and HU profiles. Across all measures, synCTs aligned more

closely with pCTs than with CBCTs, and patient-specific analyses confirmed that these improvements were consistent across cases. The anatomical fidelity of the outer body contour was particularly improved – a feature critical for accurate dose delivery. Although motion artifacts persisted in some lung cases, their severity was reduced relative to CBCT and scatter-related distortions were substantially mitigated. With larger and more diverse training datasets, these residual issues can be further minimized. In general, the findings suggest that synCT offers a more faithful representation of daily anatomy than existing methods, providing more precise HU distributions and more accurate patient geometry. At the same time, image quality alone is not sufficient to establish clinical usefulness. The crucial test lies in whether these improvements translate into concrete benefits in treatment planning and segmentation. Subsequent studies, therefore, turned to downstream tasks to evaluate the use of synCT for treatment adaptation, direct dose calculation and automatic segmentation, in order to more comprehensively assess its potential contribution to adaptive workflows.

The treatment adaptation scenario was investigated using the dosimetric analysis performed in Chapter 5 which highlighted both the limitations of IGRT workflows and the potential of adaptive strategies based on synthetic images. When reference treatment plans based on pCT were recalculated on daily CBCT, dirCT, and synCT datasets, large deviations in dose distributions were observed, particularly along the outer body contour and in regions affected by motion or organ deformation. These discrepancies highlight how strongly the delivered dose depends on anatomical differences between planning and daily images, raising concerns about the reliability of IGRT in certain clinical scenarios. By contrast, adapting plans on daily images improved tumor coverage and organ sparing across cases. While residual inconsistencies linked to registration errors in dirCTs or motion artifacts in CBCTs remained, recalculations on synCT consistently reduced these effects. In lung cases, synCT proved especially valuable in mitigating CBCT motion artifacts, whereas in prostate cases, it provided HU distributions comparable to pCTs. Together, the results support synCT as a more reliable basis for adaptive workflows, combining improved image quality with clinically relevant dosimetric accuracy.

The use of synthetic images in segmentation tasks was finally investigated in Chapter 6. More specifically, the study addressed the integration of AI-based segmentation into ART workflows, with a focus on overcoming the limitations of network training on low-quality CBCT data. By adopting a domain adaptation strategy, segmentation models trained with synthetic CBCT images paired with high-quality labels achieved performance comparable to models trained on pCTs and,

in some cases, even matched their accuracy. This indicates that synthetic training data can compensate for the reduced image fidelity of CBCT and that accounting for organ deformation improves robustness, particularly in the pelvic region where bladder and rectal changes are common. Although the small cohort size makes these findings preliminary, strong performance in difficult cases highlights the potential of this approach. Beyond accuracy, inference speed proved clinically viable, with segmentations produced in less then 30 s per CBCT, supporting time-sensitive workflows in prostate treatment. Overall, the study suggests that synthetic data can strengthen the reliability and efficiency of automatic segmentation, reducing the reliance on manual annotations of poor-quality CBCTs and paving the way for more streamlined ART workflows.

Taken together, the studies presented in this thesis demonstrate that synCT images generated with deep learning have the potential to address some of the persistent challenges of ART. By improving image quality, reducing artifacts, and providing a closer match to daily anatomy, synCT is able to support more accurate dose calculations and other important tasks in ART such as OAR autosegmentation. At the same time, the findings of this work emphasize that image synthesis cannot be regarded as a universal solution. Its clinical utility will depend on the availability of sufficiently large and diverse training datasets, rigorous validation across patient populations, and careful integration into established clinical workflows. Future research should therefore focus on expanding training cohorts and assessing the generalizability of models across institutions and imaging systems. With continued development, synthetic images have strong potential to become an integral component of ART workflows, bringing the field closer to the objective of accurate, patient-specific treatment adaptation in routine clinical practice.

Appendix A

Correlation Analysis

The correlation between the dose difference and the similarity in the body outline was evaluated in Section 5.2.2. Scatter plots of data points used to calculate the correlation and to plot a linear least squares fit are presented in Figs. A.1 and A.2, for the PCa and LungCa cases, respectively.

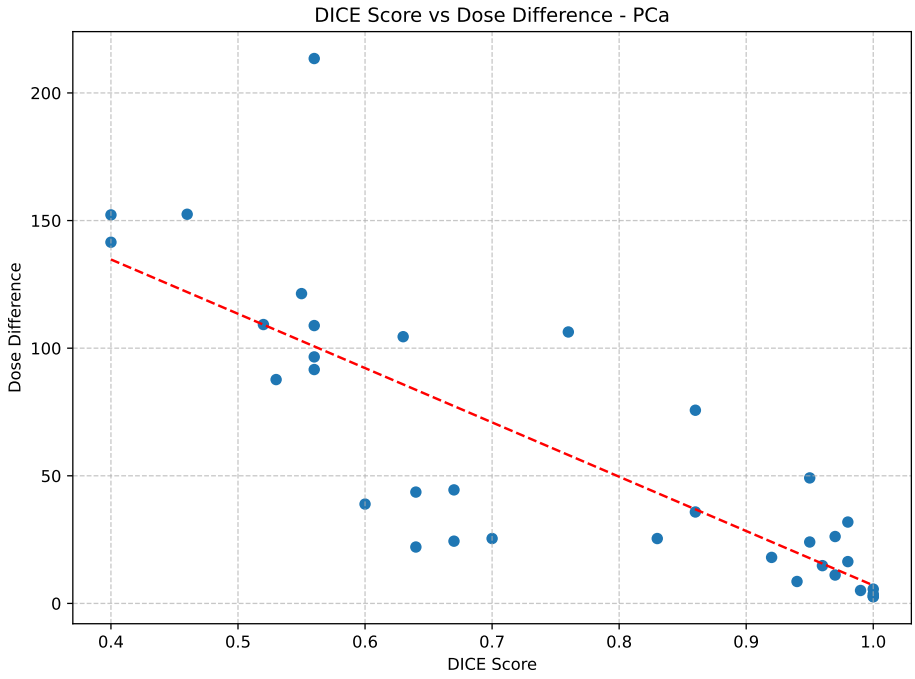


Figure A.1. Scatter plot showing the relationship between body outline similarity (SDC) and relative dose difference for PCa cases. Each point represents a recalculation pair as detailed in Fig. 5.2, with a linear least squares fit used to visually assess the correlation.

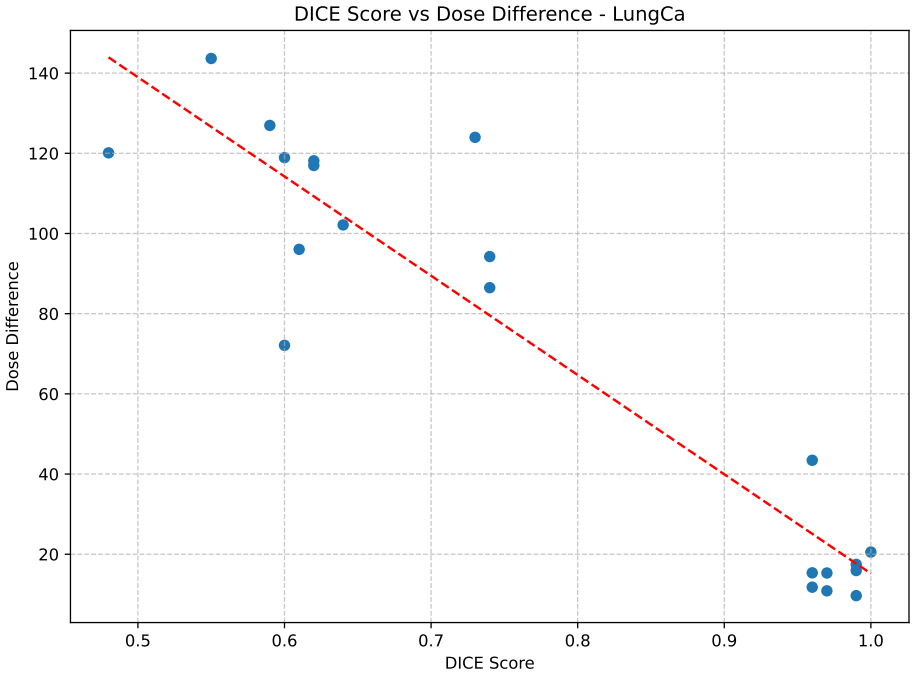


Figure A.2. Scatter plot showing the relationship between body outline similarity (SDC) and relative dose difference for LungCa cases. Each point represents a recalculation pair as detailed in Fig. 5.2, with a linear least squares fit used to visually assess the correlation.

Appendix B

Statistical Analysis

In Fig. 5.10, Chapter 5, consolidated dose metrics are presented for all patients, treatment-relevant structures, and adaptation strategies. At first glance, the distributions for CBCT- and dirCT-based plans appear very similar; however, a closer examination reveals that the relative dose difference values between the two methods are statistically distinct. A series of paired t-tests were performed both on the full set of extracted metrics and on a subset of key clinically relevant parameters. In both cases, the resulting p-values were below the 0.05 threshold, indicating statistically significant differences between CBCT- and dirCT-based recalculations. The corresponding *Cohen's d* effect size measure suggested that the difference was of medium magnitude for the key parameters and small for the complete metric set. Boxplot representations of these comparisons are provided in Figs. B.1 and B.2, illustrating the degree of overlap and separation between distributions.

An additional observation concerns the shift in mean relative dose difference between the two aggregated graphs. When considering selected key metrics, the spread is narrower, whereas focusing on all metrics results in a more pronounced shift of the mean values. This can be attributed to the larger number of outliers in the complete dirCT dataset, which exert smaller influence when the analysis is restricted to a smaller subset of metrics. These findings indicate that while the average performance of dirCT and CBCT may appear comparable, the presence of outliers in the dirCT-based recalculations may undermine its reliability for certain structures or clinical scenarios.

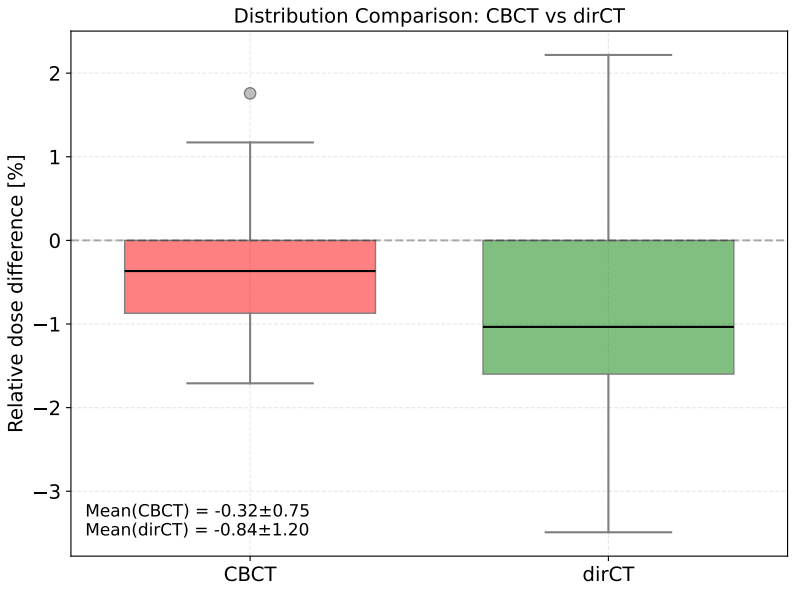


Figure B.1. Box-plots show the difference in relative dose difference distributions between key dose metrics such as $V_{95\%\text{Dose}}$ for the CTV, and $D_{0.03cc}$ for OARs like the esophagus, spinal canal, stomach/bowel and rectum. The relative dose difference was calculated between the reference plan and recalculated plans as per Fig. 5.2.

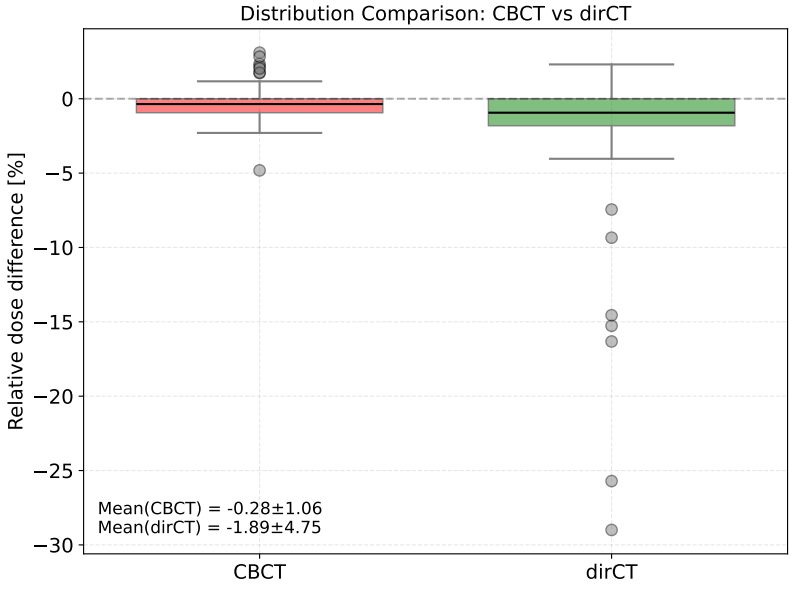


Figure B.2. Box-plots show the difference in relative dose difference distributions between all evaluated dose metrics. The relative dose difference was calculated between the reference plan and recalculated plans as per Fig. 5.2.

Bibliography

[1] World Health Organization (WHO). Cancer. https://www.who.int/news-room/fact-sheets/detail/cancer [Access Date: June 2, 2025].

[2] World Health Organization (WHO), International Agency for Research on Cancer. The Global Cancer Observatory - Cancer Today. https://gco.iarc.who.int/today/en/dataviz/ [Access Date: June 2, 2025].

[3] Barton, M. B., Jacob, S., Shafiq, J., et al. Estimating the demand for radiotherapy from the evidence: A review of changes from 2003 to 2012. *Radiother. Oncol.*, 112(1):140–144, July 2014. Epub 2014 May 12.

[4] for Cancer Research, T. I. Cancer. https://www.icr.ac.uk/about-us/icr-news/detail/our-research-impact-prostate-cancer-radiotherapy [Access Date: July 2, 2025].

[5] Rodríguez De Dios, N., Navarro-Martin, A., Cigarral, C., et al. GOECP/SEOR radiotheraphy guidelines for non-small-cell lung cancer. *World Journal of Clinical Oncology*, 13(4):237–266, April 2022.

[6] Glatzer, M., Schmid, S., Radovic, M., Früh, M., and Putora, P. M. The role of radiation therapy in the management of small cell lung cancer. *Breathe (Sheffield, England)*, 13(4):e87–e94, December 2017.

[7] Couñago, F., de la Pinta, C., Gonzalo, S., et al. GOECP/SEOR radiotherapy guidelines for small-cell lung cancer. *World Journal of Clinical Oncology*, 12(3):115–143, March 2021.

[8] Bethe, H. Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie. *Annalen der Physik*, 397(3):325–400, 1930.

[9] Bloch, F. Zur Bremsung rasch bewegter Teilchen beim Durchgang durch Materie. *Annalen der Physik*, 408(3):285–320, 1933.

[10] Eidelman, S., Hayes, K., Olive, K., et al. Review of particle physics. *Physics Letters B*, 592(1-4):1–1109, 2004.

[11] United State Nuclear Regulatory Commission. Interaction of photons with matter. `https://www.nrc.gov/docs/ml1122/ML11229A667.pdf` [Access Date: July 6, 2025], 2011.

[12] Gidley, P. W. and DeMonte, F., editors. *Temporal Bone Cancer*. Springer Cham, first edition, 2018.

[13] Joiner, M. C., van der Kogel, and J., A., editors. *Basic Clinical Radiobiology*. CRC Press, Boca Raton, fifth edition, 2018.

[14] Podgorsak, E., International Atomic Energy Agency, V. A., of Medical Physicists/Canadian College of Physicists in Medicine (Canada), C. O., et al. *Radiation oncology physics: A handbook for teachers and students*. IAEA, July 2005.

[15] El Naqa, I., Pater, P., and Seuntjens, J. Monte Carlo role in radiobiological modelling of radiotherapy outcomes. *Physics in medicine and biology*, 57:R75–97, 05 2012.

[16] Landberg, T., Chavaudra, J., Dobbs, J., et al. ICRU Report 50: Prescribing, Recording, and Reporting Photon Beam Therapy. *Reports of the International Commission on Radiation Units and Measurements*, os-26(1):67–70, 1993.

[17] Landberg, T., Chavaudra, J., Dobbs, J., et al. ICRU Report 62: Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU 50). *Reports of the International Commission on Radiation Units and Measurements*, os-32(1):48–51, 1999.

[18] on Radiation Units, I. C. and Measurements. ICRU Report 83: Prescribing, Recording, and Reporting Photon-Beam Intensity-Modulated Radiation Therapy (IMRT): Contents. *Journal of the ICRU*, 10(1):1–3, 2010.

[19] Maqbool, M., editor. *An Introduction to Medical Physics*. Biological and Medical Physics, Biomedical Engineering. Springer Cham, first edition, 2017.

[20] Timmerman, R. and Xing, L. *Image-Guided and Adaptive Radiation Therapy*. Wolters Kluwer Health, 2012.

[21] Schlegel, W., Karger, C., and Jäkel, O., editors. *Medizinische Physik*. Springer eBook Collection. Springer Spektrum, Berlin, 2018.

[22] Jaffray, D. A., Drake, D. G., Moreau, M., Martinez, A. A., and Wong, J. W. A radiographic and tomographic imaging system integrated into a medical linear accelerator for localization of bone and soft-tissue targets. *Int. J. Radiat. Oncol. Biol. Phys.*, 45(3):773–789, 1999.

[23] Jaffray, D. A., Siewerdsen, J. H., Wong, J. W., and Martinez, A. A. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 53(5):1337–1349, 2002.

[24] Bortfeld, T., Schmidt-Ullrich, R., Neve, W., and Wazer, D. *Image-Guided IMRT*. Springer Berlin Heidelberg, 2006.

[25] Biggs, P. J., Goitein, M., and Russell, M. D. A diagnostic X ray field verification device for a 10 MV linear accelerator. *Int. J. Radiat. Oncol. Biol. Phys.*, 11(3):635–643, 1985.

[26] Sephton, R. and Hagekyriakou, J. A diagnostic-quality electronic portal imaging system. *Radiotherapy and Oncology*, 35(3):240–247, 1995.

[27] Scholten, J., Kos, A., Richardson, M., and Campion, K. Cone Beam Computed Tomography: Technology review, dose, and utility considerations and regulatory bodies. *Journal of Contemporary Chiropractic*, 6(1):92–99, may 25 2023.

[28] Miracle, A. and Mukherji, S. Conebeam CT of the Head and Neck, Part 1: Physical principles. *American Journal of Neuroradiology*, 30(6):1088–1095, 2009.

[29] Pauwels, R., Araki, K., Siewerdsen, J. H., and Thongvigitmanee, S. S. Technical aspects of dental CBCT: State of the art. *Dentomaxillofacial Radiology*, 44(1):20140224, 10 2014.

[30] Feldkamp, L. A., Davis, L. C., and Kress, J. W. Practical Cone-beam algorithm. *J. Opt. Soc. Am. A*, 1(6):612–619, Jun 1984.

[31] Dawson, L. A. and Jaffray, D. A. Advances in Image-guided Radiation Therapy. *Journal of Clinical Oncology*, 25(8):938–946, 2007. PMID: 17350942.

[32] de Crevoisier, R., Tucker, S. L., Dong, L., et al. Increased risk of biochemical and local failure in patients with distended rectum on the planning CT for prostate cancer radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 62(4):965–973, July 2005.

[33] Boda-Heggemann, J., Lohr, F., Wenz, F., Flentje, M., and Guckenberger, M. kV Cone-Beam CT-based IGRT: A Clinical Review. *Strahlentherapie und Onkologie*, 187(5):284–291, 2011.

[34] Yan, D., Vicini, F., Wong, J., and Martinez, A. Adaptive Radiation Therapy. *Phys. Med. Biol.*, 42(1):123–132, 1997.

[35] Yan, D. Adaptive radiotherapy: merging principle into clinical practice. *Seminars in radiation oncology*, 20 2:79–83, 2010.

[36] Benitez, C. M., Chuong, M. D., Künzel, L. A., and Thorwarth, D. MRI-Guided Adaptive Radiation Therapy. *Seminars in Radiation Oncology*, 34(1):84–91, 2024. MR-guided Radiation Therapy.

[37] Sonke, J. J., Aznar, M., and Rasch, C. Adaptive Radiotherapy for Anatomical Changes. *Semin Radiat Oncol*, 29(3):245–257, 2019.

[38] Sonke, J.-J. and Belderbos, J. Adaptive radiotherapy for lung cancer. *Seminars in Radiation Oncology*, 20(2):94–106, April 2010.

[39] Archambault, Y., Boylan, C., Bullock, D., et al. Making on-line adaptive radiotherapy possible using artificial intelligence and machine learning for efficient daily re-planning. *Med Phys Int J*, 8(2), 2020.

[40] Kim, J.-Y., Tawk, B., Knoll, M., et al. Clinical Workflow of Cone Beam Computer Tomography-Based Daily Online Adaptive Radiotherapy with Offline Magnetic Resonance Guidance: The Modular Adaptive Radiotherapy System (MARS). *Cancers*, 16(6), 2024.

[41] Nijkamp, J., Pos, F. J., Nuver, T. T., et al. Adaptive Radiotherapy for Prostate Cancer Using Kilovoltage Cone-Beam Computed Tomography: First Clinical Results. *Int. J. Radiat. Oncol. Biol. Phys.*, 70:75–82, 1 2008.

[42] Park, S. S., Yan, D., McGrath, S., et al. Adaptive image-guided radiotherapy (IGRT) eliminates the risk of biochemical failure caused by the bias of rectal distension in prostate cancer treatment planning: clinical evidence. *International Journal of Radiation Oncology, Biology, Physics*, 83(3):947–952, July 2012. Epub 2011 Dec 28.

[43] Byrne, M., Archibald-Heeren, B., Hu, Y., et al. Varian ethos online adaptive radiotherapy for prostate cancer: Early results of contouring accuracy, treatment plan quality, and treatment time. *Journal of Applied Clinical Medical Physics*, 23(1):e13479, 2022.

[44] Zwart, L. G., Ong, F., ten Asbroek, L. A., et al. Cone-beam computed tomography-guided online adaptive radiotherapy is feasible for prostate cancer patients. *Physics and Imaging in Radiation Oncology*, 22:98–103, 2022.

[45] Fischer, J., Fischer, L. A., Bensberg, J., et al. CBCT-based online adaptive radiotherapy of the prostate bed: first clinical experience and comparison to nonadaptive conventional IGRT. *Strahlentherapie und Onkologie*, November 2024.

[46] Tvilum, M., Khalil, A. A., Møller, D. S., Hoffmann, L., and Knap, M. M. Clinical outcome of image-guided adaptive radiotherapy in the treatment of lung cancer patients. *Acta Oncologica*, 54(9):1430–1437, 2015. Epub 2015 Jul 24.

[47] Kataria, T., Gupta, D., Bisht, S. S., et al. Adaptive radiotherapy in lung cancer: dosimetric benefits and clinical outcome. *The British Journal of Radiology*, 87(1038):20130643, 2014. Epub 2014 Mar 17.

[48] Duffton, A., Harrow, S., Lamb, C., and McJury, M. An assessment of cone beam CT in the adaptive radiotherapy planning process for non-small-cell lung cancer patients. *British Journal of Radiology*, 89(1062):20150492, 04 2016.

[49] Berkovic, P., Paelinck, L., Gulyban, A., et al. Adaptive radiotherapy for locally advanced non-small cell lung cancer: dosimetric gain and treatment outcome prediction. *Acta Oncologica*, 56(11):1656–1659, 2017.

[50] Hoppen, L., Sarria, G. R., Kwok, C. S., et al. Dosimetric benefits of adaptive radiation therapy for patients with stage III non-small cell lung cancer. *Radiation Oncology*, 18(1):34, 2023.

[51] Møller, D. S., Lutz, C. M., Khalil, A. A., et al. Survival benefits for non-small cell lung cancer patients treated with adaptive radiotherapy. *Radiotherapy and Oncology*, 168:234–240, 2022.

[52] Pahadia, M., Katkar, R., and Geha, H. Effect of a Motion Artifact Correction System on Cone-Beam Computed Tomography Image Characteristics. *Cureus*, 15(2):e35016, February 2023.

[53] Scarfe, W. C. and Farman, A. G. What is cone-beam CT and how does it work? *Dental Clinics of North America*, 52(4):707–730, October 2008.

[54] Srinivasan, K., Mohammadi, M., and Shepherd, J. Applications of linac-mounted kilovoltage Cone-beam Computed Tomography in modern radiation therapy: A review. *Polish Journal of Radiology*, 79:181–193, 2014.

[55] Steinke, M. F. and Bezak, E. Technological approaches to in-room CBCT imaging. *Australasian Physics & Engineering Sciences in Medicine*, 31(3):167–179, 2008.

[56] Siewerdsen, J. H. and Jaffray, D. A. Cone-beam computed tomography with a flat-panel imager: Magnitude and effects of x-ray scatter. *Medical Physics*, 28(2):220–231, 2001.

[57] Endo, M., Tsunoo, T., Nakamori, N., and Yoshida, K. Effect of scattered radiation on image noise in cone beam CT. *Medical Physics*, 28(4):469–474, 2001.

[58] Graham, S. A., Moseley, D. J., Siewerdsen, J. H., and Jaffray, D. A. Compensators for dose and scatter management in cone-beam computed tomography. *Medical Physics*, 34(7):2691–2703, 2007.

[59] Siewerdsen, J. H., Moseley, D. J., Bakhtiar, B., Richard, S., and Jaffray, D. A. The influence of antiscatter grids on soft-tissue detectability in cone-beam computed tomography with flat-panel detectors. *Medical Physics*, 31(12):3506–3520, 2004.

[60] Alexeev, T., Kavanagh, B., Miften, M., and Altunbas, C. Two-dimensional antiscatter grid: A novel scatter rejection device for Cone-beam computed tomography. *Medical Physics*, 45(2):529–534, 2018.

[61] Star-Lack, J., Sun, M., Kaestner, A., et al. Efficient scatter correction using asymmetric kernels. In Samei, E. and Hsieh, J., editors, *Medical Imaging 2009: Physics of Medical Imaging*, volume 7258, page 72581Z. International Society for Optics and Photonics, SPIE, 2009.

[62] Watson, P. G., Mainegra-Hing, E., Tomic, N., and Seuntjens, J. Implementation of an efficient Monte Carlo calculation for CBCT scatter correction: phantom study. *Journal of Applied Clinical Medical Physics*, 16(4):216–227, 2015.

[63] Maier, J., Sawall, S., Knaup, M., and KachelrieSS, M. Deep Scatter Estimation (DSE): Accurate Real-Time Scatter Estimation for X-Ray CT Using a Deep Convolutional Neural Network. *Journal of Nondestructive Evaluation*, 37(3):57, 2018.

[64] Cai, M., Byrne, M., Archibald-Heeren, B., et al. Decoupling of bowtie and object effects for beam hardening and scatter artefact reduction in iterative cone-beam CT. *Physical and Engineering Sciences in Medicine*, 43(4):1161–1170, December 2020.

[65] Santaella, G. M., Wenzel, A., Haiter-Neto, F., Rosalen, P. L., and Spin-Neto, R. Impact of movement and motion-artefact correction on image quality and interpretability in CBCT units with aligned and lateral-offset detectors. *Dentomaxillofacial Radiology*, 49(1):20190240, 09 2019.

[66] Ma, C., Cao, J., Yin, Y., and Zhu, J. Radiotherapy dose calculation on KV cone-beam CT image for lung tumor using the CIRS calibration. *Thoracic Cancer*, 5:68–73, 1 2014.

[67] Elstrøm, U. V., Olsen, S. R. K., Muren, L. P., Petersen, J. B. B., and Grau, C. The impact of CBCT reconstruction and calibration for radiotherapy planning in the head and neck region a phantom study. *Acta Oncologica*, 53:1114–1124, 8 2014.

[68] Kurz, C., Dedes, G., Resch, A., et al. Comparing cone-beam CT intensity correction methods for dose recalculation in adaptive intensity-modulated photon and proton therapy for head and neck cancer. *Acta Oncologica*, 54:1651–1657, 10 2015.

[69] MacFarlane, M., Wong, D., Hoover, D. A., et al. Patient-specific calibration of cone-beam computed tomography data sets for radiotherapy dose calculations and treatment plan assessment. *Journal of Applied Clinical Medical Physics*, 19:249–257, 3 2018.

[70] Shinde, P., Jadhav, A., Shankar, V., et al. Evaluation of kV-CBCT based 3D dose calculation accuracy and its validation using delivery fluence derived dose metrics in Head and Neck Cancer. *Physica Medica*, 96:32–45, 4 2022.

[71] Fotina, I., Hopfgartner, J., Stock, M., et al. Feasibility of CBCT-based dose calculation: Comparative analysis of HU adjustment techniques. *Radiotherapy and Oncology*, 104:249–256, 8 2012.

[72] van Zijtveld, M., Dirkx, M., and Heijmen, B. Correction of conebeam CT values using a planning CT for derivation of the dose of the day. *Radiotherapy and Oncology*, 85:195–200, 11 2007.

[73] Dunlop, A., McQuaid, D., Nill, S., et al. Comparison of CT number calibration techniques for CBCT-based dose calculation. *Strahlentherapie und Onkologie*, 191:970–978, 12 2015.

[74] Kidar, H. S. and Azizi, H. Enhancement of Hounsfield unit distribution in cone-beam CT images for adaptive radiation therapy: Evaluation of a hybrid correction approach. *Physica Medica*, 69:269–274, 1 2020.

[75] Gong, H., Liu, B., Zhang, G., et al. Evaluation of Dose Calculation Based on Cone-Beam CT Using Different Measuring Correction Methods for Head and Neck Cancer Patients. *Technology in Cancer Research and Treatment*, 22:153303382211483, 1 2023.

[76] Brock, K. K., Mutic, S., McNutt, T. R., Li, H., and Kessler, M. L. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. *Medical Physics*, 44:e43–e76, 7 2017.

[77] Oh, S. and Kim, S. Deformable image registration in radiation therapy. *Radiation Oncology Journal*, 35:101–111, 6 2017.

[78] Iliadou, V., Economopoulos, T. L., Karaiskos, P., et al. Deformable image registration to assist clinical decision for radiotherapy treatment adaptation for head and neck cancer patients. *Biomedical Physics and Engineering Express*, 7:055012, 9 2021.

[79] Sotiras, A., Davatzikos, C., and Paragios, N. Deformable medical image registration: a survey. *IEEE Transactions on Medical Imaging*, 32(7):1153–1190, July 2013. Epub 2013 May 31.

[80] Varadhan, R., Karangelis, G., Krishnan, K., and Hui, S. A framework for deformable image registration validation in radiotherapy clinical applications. *Journal of Applied Clinical Medical Physics*, 14:192–213, 1 2013.

[81] Bauer, C. J., Teske, H., Walter, A., et al. Biofidelic image registration for head and neck region utilizing an in-silico articulated skeleton as a transformation model. *Physics in Medicine and Biology*, 68(9):095006, apr 2023.

[82] Rigaud, B., Simon, A., Castelli, J., et al. Evaluation of Deformable Image Registration Methods for Dose Monitoring in Head and Neck Radiotherapy. *BioMed Research International*, 2015:1–16, 2015.

[83] Barber, J., Yuen, J., Jameson, M., et al. Deforming to Best Practice: Key considerations for deformable image registration in radiotherapy. *Journal of Medical Radiation Sciences*, 67:318–332, 12 2020.

[84] Paganelli, C., Meschini, G., Molinelli, S., Riboldi, M., and Baroni, G. Patientspecific validation of deformable image registration in radiation therapy: Overview and caveats. *Medical Physics*, 45, 10 2018.

[85] Dahiya, N., Alam, S. R., Zhang, P., et al. Multitask 3D CBCT-to-CT translation and organs-at-risk segmentation using physics-based data augmentation. *Medical Physics*, 48(9):5130–5141, 2021.

[86] Zhang, Y., Yue, N., Su, M.-Y., et al. Improving CBCT quality to CT level using deep learning with generative adversarial network. *Medical Physics*, 48(6):2816–2826, 2021.

[87] Suwanraksa, C., Bridhikitti, J., Liamsuwan, T., and Chaichulee, S. CBCT-to-CT Translation Using Registration-Based Generative Adversarial Networks in Patients with Head and Neck Cancer. *Cancers*, 15(7), 2023.

[88] Wynne, J. F., Lei, Y., Pan, S., et al. Rapid unpaired CBCT-based synthetic CT for CBCT-guided adaptive radiotherapy. *Journal of Applied Clinical Medical Physics*, 24(10):e14064, 2023.

[89] Peng, J., Qiu, R. L. J., Wynne, J. F., et al. CBCT-Based synthetic CT image generation using conditional denoising diffusion probabilistic model. *Medical Physics*, 51(3):1847–1859, 2024.

[90] Hu, Y., Zhou, H., Cao, N., Li, C., and Hu, C. Synthetic CT generation based on CBCT using improved vision transformer CycleGAN. *Scientific Reports*, 14(1):11455, 2024.

[91] Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

[92] Kida, S., Kaji, S., Nawa, K., et al. Visual enhancement of Cone-beam CT by use of CycleGAN. *Medical Physics*, 47(3):998–1010, 2020.

[93] Deng, L., Zhang, M., Wang, J., Huang, S., and Yang, X. Improving cone-beam CT quality using a cycle-residual connection with a dilated convolution-consistent generative adversarial network. *Phys. Med. Biol.*, 67(14), 2022.

[94] Deng, L., Ji, Y., Huang, S., Yang, X., and Wang, J. Synthetic CT generation from CBCT using double-chain-CycleGAN. *Computers in Biology and Medicine*, 161:106889, 2023.

[95] Chan, Y., Li, M., Parodi, K., et al. Feasibility of CycleGAN enhanced low dose CBCT imaging for prostate radiotherapy dose calculation. *Phys. Med. Biol.*, 68(10), 2023.

[96] Rusanov, B., Hassan, G. M., Reynolds, M., et al. Transformer CycleGAN with uncertainty estimation for CBCT based synthetic CT in adaptive radiotherapy. *Physics in Medicine and Biology*, 69(3):035014, jan 2024.

[97] Fukushima, K. Cognitron: A self-organizing multilayered neural network. *Biological Cybernetics*, 20(3):121–136, 1975.

[98] LeCun, Y., Boser, B., Denker, J., et al. Handwritten Digit Recognition with a Back-Propagation Network. In Touretzky, D., editor, *Advances in Neural Information Processing Systems*, volume 2. Morgan-Kaufmann, 1989.

[99] LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. *Nature*, 521(7553):436–444, 2015.

[100] Politz, F., Kazimi, B., and Sester, M. Classification of Laser Scanning Data Using Deep Learning. In *Photogrammetrie–Fernerkundung–Geoinformatik–Kartographie–2018*, volume 27, pages 597–610. Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., 03 2018.

[101] He, K., Zhang, X., Ren, S., and Sun, J. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[102] Hinton, G. E. and Salakhutdinov, R. R. Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(5786):504–507, 2006.

[103] Noh, H., Hong, S., and Han, B. Learning Deconvolution Network for Semantic Segmentation, 2015. arXiv:1505.04366 [cs.CV].

[104] Ronneberger, O., Fischer, P., and Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[105] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. Generative Adversarial Networks, 2014. arXiv:1406.2661 [stat.ML].

[106] Mirza, M. and Osindero, S. Conditional Generative Adversarial Nets, 2014. arXiv:1411.1784 [cs.LG].

[107] Rogers, C. A. *Hausdorff Measures*. Cambridge University Press, 1998.

[108] Sorensen, T. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biologiske Skrifter*, 5:1–34, 1948.

[109] Dice, L. R. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[110] Nikolov, S., Blackwell, S., Zverovitch, A., et al. Clinically applicable segmentation of head and neck anatomy for radiotherapy: deep learning algorithm development and validation study. *Journal of Medical Internet Research*, 23(7):e26151, 2021.

[111] Wagenaar, D., Kierkels, R. G. J., van der Schaaf, A., et al. Head and neck IMPT probabilistic dose accumulation: Feasibility of a 2 mm setup uncertainty setting. *Radiotherapy and Oncology*, 154:45–52, January 2021.

[112] Salimans, T., Goodfellow, I., Zaremba, W., et al. Improved techniques for training gans. In Lee, D., Sugiyama, M., Luxburg, U., Guyon, I., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016.

[113] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. GANs trained by a two time-scale update rule converge to a local nash equilibrium. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 66296640, Red Hook, NY, USA, 2017. Curran Associates Inc.

[114] Pluim, J., Maintz, J., and Viergever, M. Mutual-Information-Based Registration of Medical Images: A Survey. *Medical Imaging, IEEE Transactions on*, 22:986 – 1004, 09 2003.

[115] Rueckert, D., Aljabar, P., Heckemann, R. A., Hajnal, J. V., and Hammers, A. Diffeomorphic Registration Using B-Splines. In Larsen, R., Nielsen, M., and Sporring, J., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, pages 702–709, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[116] van Zwienen, M., van Beek, S., Belderbos, J., et al. Anatomical changes during radiotherapy of lung cancer patients. *International Journal of Radiation Oncology, Biology, Physics*, 72(1):S111, 2008.

[117] Tennyson, N., Weiss, E., Sleeman, W., et al. Effect of variations in atelectasis on tumor displacement during radiation therapy for locally advanced lung cancer. *Advances in Radiation Oncology*, 2(1):19–26, 2017.

[118] Kavanaugh, J., Hugo, G., Robinson, C. G., and Roach, M. C. Anatomical adaptation–early clinical evidence of benefit and future needs in lung cancer. *Seminars in Radiation Oncology*, 29(3):274–283, 2019.

[119] Wasserthal, J., Breit, H.-C., Meyer, M. T., et al. TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. *Radiology: Artificial Intelligence*, 5(5):e230024, 2023.

[120] Ding, M., Maspero, M., Littooij, A. S., et al. Deep learning-based auto-contouring of organs/structures-at-risk for pediatric upper abdominal radiotherapy. *Radiotherapy and Oncology*, 208:110914, 2025.

[121] Robar, J. L., Cherpak, A., MacDonald, R. L., et al. Novel technology allowing cone beam computed tomography in 6 seconds: A patient study of comparative image quality. *Practical Radiation Oncology*, 14(3):277–286, 2024.

[122] Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J., and Dalca, A. V. Voxelmorph: A learning framework for deformable medical image registration. *IEEE Transactions on Medical Imaging*, 38(8):1788–1800, 2019.

[123] Lowekamp, B. C., Chen, D. T., Ibanez, L., and Blezek, D. The Design of SimpleITK. *Frontiers in Neuroinformatics*, 7, 2013.

[124] Marstal, K., Berendsen, F., Staring, M., and Klein, S. SimpleElastix: A user-friendly, multi-lingual library for medical image registration. In *International Workshop on Biomedical Image Registration (WBIR)*, Las Vegas, Nevada, USA, 2016.

[125] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, 2016.

[126] Sønderby, C. K., Caballero, J., Theis, L., Shi, W., and Huszár, F. Amortised MAP Inference for Image Super-resolution. *CoRR*, abs/1610.04490, 2016.

[127] Harms, J., Lei, Y., Wang, T., et al. Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. *Medical Physics*, 46(9):3998–4009, 2019.

[128] Maspero, M., Houweling, A. C., Savenije, M. H. F., et al. A single neural network for cone-beam computed tomography-based radiotherapy of head-and-neck, lung and breast cancer. *Physics and Imaging in Radiation Oncology*, 14:24–31, April 2020.

[129] Kurz, C., Maspero, M., Savenije, M. H. F., et al. CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation. *Physics in Medicine and Biology*, 64(22):225004, 2019.

[130] Gao, L., Xie, K., Wu, X., et al. Generating synthetic CT from low-dose cone-beam CT by using generative adversarial networks for adaptive radiotherapy. *Radiation Oncology*, 16(1):202, October 2021.

[131] Ghilezan, M., Yan, D., and Martinez, A. Adaptive radiation therapy for prostate cancer. In *Seminars in radiation oncology*, volume 20, pages 130–137. Elsevier, 2010.

[132] Peng, C., Ahunbay, E., Chen, G., et al. Characterizing interfraction variations and their dosimetric effects in prostate cancer radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 79(3):909–914, 2011.

[133] Salembier, C., Villeirs, G., De Bari, B., et al. ESTRO ACROP consensus guideline on CT-and MRI-based target volume delineation for primary radiation therapy of localized prostate cancer. *Radiotherapy and Oncology*, 127(1):49–61, 2018.

[134] Li, G., Li, Y., Wang, J., et al. Guidelines for radiotherapy of prostate cancer (2020 edition). *Precision Radiation Oncology*, 5(3):160–182, 2021.

[135] Voyant, C., Biffi, K., Leschi, D., Briancon, J., and Lantieri, C. Dosimetric uncertainties related to the elasticity of bladder and rectal walls: Adenocarcinoma of the prostate. *Cancer/Radiothérapie*, 15(4):270–278, 2011.

[136] Acosta, O., Dowling, J., Drean, G., et al. Multi-atlas-based segmentation of pelvic structures from CT scans for planning in prostate cancer radiotherapy. In *Abdomen and thoracic imaging: an engineering & clinical perspective*, pages 623–656. Springer, 2013.

[137] Thor, M., Petersen, J. B., Bentzen, L., Høyer, M., and Muren, L. P. Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer. *Acta Oncologica*, 50(6):918–925, 2011.

[138] Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., and Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2):203–211, 2021.

[139] Abbani, N., Baudier, T., Rit, S., et al. Deep learning-based segmentation in prostate radiation therapy using Monte Carlo simulated cone-beam computed tomography. *Medical Physics*, 49(11):6930–6944, 2022.

[140] Fu, Y., Lei, Y., Wang, T., et al. Pelvic multi-organ segmentation on cone-beam CT for prostate adaptive radiotherapy. *Medical physics*, 47(8):3415–3422, 2020.

[141] Dai, X., Lei, Y., Wynne, J., et al. Synthetic CT-aided multiorgan segmentation for CBCT-guided adaptive pancreatic radiotherapy. *Medical physics*, 48(11):7063–7073, 2021.

[142] Kovacs, B., Netzer, N., Baumgartner, M., et al. Anatomy-informed data augmentation for enhanced prostate cancer detection. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 531–540. Springer, 2023.

[143] Keall, P., Nguyen, D. T., O'Brien, R., et al. Real-time image guided ablative prostate cancer radiation therapy: results from the TROG 15.01 SPARK trial. *Int. J. Radiat. Oncol. Biol. Phys.*, 107(3):530–538, 2020.

[144] Nyholm, T., Svensson, S., Andersson, S., et al. MR and CT data with multiobserver delineations of organs in the pelvic areaPart of the Gold Atlas project. *Medical physics*, 45(3):1295–1300, 2018.

[145] Siewerdsen, J. H. and Jaffray, D. A. Cone-beam computed tomography with a flat-panel imager: magnitude and effects of x-ray scatter. *Medical Physics*, 28(2):220–231, 2 2001.

[146] Lechuga, L. and Weidlich, G. Cone Beam CT vs. Fan Beam CT: A Comparison of Image Quality and Dose Delivered Between Two Differing CT Imaging Modalities. *Cureus*, 8(9):e778, 9 2016.

Documentation on the Use of AI-Tools

The AI tools used during the writing of this thesis were responsibly utilized in accordance with the *Deutsche Forschungsgemeinschaft (DFG)* guidelines.

All scientific ideas, concepts, interpretations, and conclusions were developed independently and without the influence of AI-based tools. All methods and results are based on the independent work of the author using conventional scientific methods. The AI-based tool was used exclusively for linguistic revision, specifically for improving the grammar of already self-written text. The tools that were used were ChatGPT v4.0 by OpenAI and Claude v3.5-Sonnet by Anthropic.

Objectives of Using AI-based Tools

- Grammatical recommendations for self-written text
- Grammatical review of the author's own translation of the English Abstract into German

Mode of Use of AI-based Tools

The workflow consisted of the following steps:

1. Manual insertion of exclusively self-written paragraphs into the AI model with a request for grammatical suggestions
2. Critical review of the AI suggestions by the author
3. Implementation of suitable suggestions without altering the scientific content or meaning. No text passage was directly or entirely adopted from an AI tool.

Acknowledgments

I am deeply grateful to my supervisors at the German Cancer Research Center (DKFZ) Prof. Dr. Oliver Jäkel, Dr. Kristina Giske and Dr. Niklas Wahl. Without you this thesis would have remained a dream. Thank you for guiding me and teaching me how to do proper science.

Many thanks to Prof. Dr. Mark Ladd as second referee, and to JProf. Dr. Loredana Gastaldo and Prof. Dr. Jörg Jäckel for their role in my examination.

This work would not have been possible without the financial support from the Varian Research Grant and my industry supervisor Dr. Pascal Paysan, whose flexibility and openness to new ideas helped shape this thesis. I also have to extend my gratitude to the Varian branch in Baden, Switzerland, and the people who made me feel a part of the crew during my two-week research stay.

My sincere appreciation extends to the clinical radiation therapy team at the DKFZ lead by PD Dr. med. Fabian Weykamp. The work you are doing is sacred and every day you remind us what our work is all about.

Special thanks go to the medical physics team at the DKFZ, Dr. Peter Häring, Clemens Lang and Mona Lifferth. Your help with the treatment planning systems, patient data and dose recalculation was invaluable and I cannot thank you enough.

I am especially grateful to my two amazing research groups, Computational Patient Models (E0401) and Radiotherapy Optimization (E0404). I am very happy I met so many wonderful people that made my stay at the DKFZ feel like I belonged.

I must not forget our departmental IT team consisting of Christoph Döttling and Kai-Ernst Wilberg, our saviors in time of need and shoulders to cry on when a piece of equipment breaks.

I am grateful to my sister Marina and my dad Milan for cheering me on from afar, and to my late mom Ruža for urging me to pursue my passion and return to the field of medical physics.

Finally, I wish to thank my partner in crime and my infinite well of support, Ajla Nurkanović. Your love and faith in me kept me afloat when the academic and personal seas got rough. Thank you for being my anchor in a foreign land and for not allowing me to overwork – your travel bug made us fully enjoy the time when we were off duty. The pages in this book would have been empty without you.