Generative AI for cone-beam CT dose reduction and intensity correction in adaptive radiotherapy

Yan Chi Chan

München 2025

Aus der
Klinik und Poliklinik für Strahlentherapie und Radioonkologie
Klinikum der Ludwig–Maximilians–Universität München
Direktor: Prof. Dr. med. Claus Belka

Generative AI for cone-beam CT dose reduction and intensity correction in adaptive radiotherapy

Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften
an der Medizinischen Fakultät
der Ludwig–Maximilians–Universität München

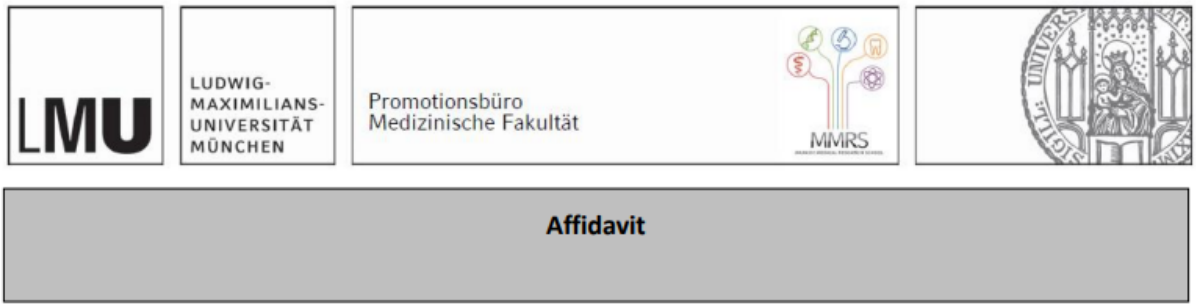
vorgelegt von
Yan Chi Chan
aus Hongkong

München 2025

Gedruckt mit Genehmigung der Medizinischen Fakultät
der Ludwig–Maximilians–Universität München

Erstgutachter: Prof. Dr. Guillaume Landry
Zweitgutachter: Prof. Dr. Paola Coan
Dekan: Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung: 17. Oktober 2025

Affidavit



I hereby declare that the submitted thesis entitled:

Generative AI for cone-beam CT dose reduction and intensity correction in adaptive radiotherapy

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given. I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree. Finally, I declare the conformity between the printed and the electronic version of the dissertation.

Munich, 04.12.2025	Yan Chi Chan
Place, date	Signature doctoral candidate

Abstract

In image-guided radiotherapy (IGRT), cone-beam computed tomography (CBCT) is used to align patients in the treatment position. CBCT scans administer radiation exposure and subject patients to secondary cancer risk. However, lowering CBCT imaging dose continues to be challenging as the image quality degrades. In current practice, the frequency of scanning can be limited, leading to a potential decrease in localisation precision. Moreover, in cases of adaptation, patients are required to undergo an addition planning CT due to insufficient CBCT image quality, resulting in extra radiation exposure and prolonged treatment time.

This thesis reports on investigations towards the minimum CBCT imaging dose without loss of accuracy in terms of synthetic CT (sCT) image generation using generative artificial intelligence (AI) for adaptive radiotherapy (ART). Many studies translated full dose CBCT images into sCT images using deep learning (DL) algorithms such as U-Net, cycle-consistent generative adversarial network (cycleGAN) or contrastive unpaired translation (CUT). However, only few studies investigated the potential of low imaging dose CBCT. In this thesis, the lowest achievable CBCT imaging dose for online adaptation was investigated. Compared to the previous studies, this work provides a structured investigation on imaging dose levels (100%, 25%, 15%, 10%) and evaluations with IGRT-related metrics, including patient positioning, treatment dose calculations and organ contouring.

Online adaptation in IGRT is currently limited by the CBCT image quality. In the first part of this thesis, the basics of cancer and IGRT in chapter 1, the adaptation workflow in chapter 2 and the physics of CBCT in chapter 3 are discussed, respectively. This will serve as an introduction to explain how CBCT-guided IGRT works and to identify the challenges of using CBCT in ART.

While DL is used to enhance full dose CBCT images, generating sCT images from low dose CBCT requires additional under-sampling streaks removal. The improvements and recent studies for DL-enabled full dose CBCT-to-CT translation are discussed in chapter 4. In chapter 5, the significance of low imaging dose CBCT, and the synthesis of low imaging dose sCT images and the DL algorithms that can be used are discussed. In chapter 6, the metrics for evaluating sCT images are discussed.

To investigate the minimum CBCT imaging dose for IGRT adaptation, we conducted two studies with generative AI models using a retrospective prostate patient dataset. In chapter 7, the patient database and the contributions of the two studies are explained. In chapter 8, the published papers for each of the studies for reference are attached. Especially in the second study, it was found that 25% is the minimum CBCT imaging dose for accurate treatment dose calculation and organ contouring when using the AI methods selected in this project.

Finally, in chapter 9 Discussion, the findings and limitations in this work, the challenges that hinder the development of low imaging dose CBCT, and possible future works that can extend this study and facilitate clinical implementations of the new low imaging dose CBCT technique in the ART workflow are discussed.

Zusammenfassung

Bei der bildgesteuerten Strahlentherapie (IGRT) wird die cone-beam Computertomographie (CBCT) eingesetzt, um die Patienten in der Behandlungsposition auszurichten. CBCT-Scans führen zu einer Strahlenbelastung und setzen die Patienten einem sekundären Krebsrisiko aus. Die Reduzierung der Dosis bei CBCT-Aufnahmen bleibt jedoch eine Herausforderung, da sich die Bildqualität verschlechtert. In der gegenwärtigen Praxis könnte die Häufigkeit des Scannens beschränkt, was zu einer potenziellen Abnahme der Lokalisierungsgenauigkeit führt. Darüber hinaus müssen sich die Patienten im Falle einer Behandlungsadaptation einem zusätzlichen Planungs-CT unterziehen, was zu einer zusätzlichen Strahlenbelastung und einer verlängerten Behandlungszeit führt.

In dieser Arbeit werden Untersuchungen zur Minimierung der CBCT-Bilddosis ohne Qualitätsverlust in Bezug auf die Erzeugung synthetischer CTs (sCT) unter Verwendung generativer KI für die adaptive Strahlentherapie (ART) vorgestellt. Mit Deep Learning (DL) Verbesserungen von U-Net über cycleGAN zu Contrastive Unpaired Translation (CUT), haben viele Studien CBCT-Bilder mit voller Dosis in sCT-Bilder übersetzt. Nur wenige Studien untersuchten jedoch die Möglichkeit, CBCT-Bilder mit niedriger Dosis zu nutzen. Durch Unterabtastung von CBCT-Projektionen haben wir die niedrigste erreichbare CBCT-Bildgebungsdosis für die Online-Anpassung untersucht. Im Vergleich zu früheren Studien bietet diese Arbeit eine strukturierte Untersuchung der Bildgebungsdosis und Bewertungen mit IGRT-bezogenen Metriken, einschließlich Patientenpositionierung, Dosisberechnung und Organkonturierung.

Die Online-Anpassung in der IGRT ist derzeit durch die Qualität der CBCT-Bilder eingeschränkt. Im ersten Teil dieser Arbeit werden die Grundlagen von Krebs und IGRT in Kapitel 1, der Adaptationsworkflow in Kapitel 2 und die Physik der CBCT in Kapitel 3 diskutiert. Dies dient als Einführung, um zu erklären, wie die CBCT-geführte IGRT funktioniert, und um die Herausforderungen bei der Verwendung von CBCT in der ART aufzuzeigen.

Während DL zur Verbesserung von CBCT-Bildern mit voller Dosis verwendet wird, erfordert die Erzeugung von sCT-Bildern aus CBCT-Bildern mit niedriger Dosis eine zusätzliche Entfernung von Streifenartefakten. Im zweiten Teil der Arbeit werden die Verbesserungen und die jüngsten Studien zur DL-gestützten Volldosis-CBCT-zu-CT-Übersetzung in Kapitel 4 diskutiert. In Kapitel 5 werden die Bedeutung von CBCT mit niedriger Bilddosis, die Synthese von CT mit niedriger Bilddosis und die DL-Algorithmen, die für die Übersetzung von CBCT mit niedriger Bilddosis in CT-Bilder verwendet werden können, diskutiert. In Kapitel 6 werden die Metriken für die Bewertung von sCT erörtert.

Um die minimale CBCT-Bildgebungsdosis für die IGRT-Anpassung zu untersuchen, haben wir zwei Studien mit generativen KI-Modellen unter Verwendung von Prostata Patientendatensätzen durchgeführt. In Kapitel 7 werden die Patientendatenbank und die Beiträge der beiden Studien erläutert. In Kapitel 8 fügen wir für jede der Studien die Publikation als Referenz bei. Insbesondere in der zweiten Studie haben wir herausgefunden, dass 25% die Mindestdosis für die CBCT-Bildgebung ist, um eine genaue Berechnung der Behandlungsdosis und Organkonturierung zu ermöglichen.

Abschließend diskutieren wir in der Diskussion (Kapitel 9) die Ergebnisse und Einschränkungen dieser Arbeit, die Herausforderungen, die die Entwicklung von CBCT mit niedriger Bildgebungsdosis behindern, und mögliche zukünftigen Arbeiten, die diese Studie erweitern und die klinische Implementierung der neuen CBCT-Technik mit niedriger Bildgebungsdosis in ART-Workflows erleichtern könnten.

Contents

Chapter 1

Introduction

1.1 Cancer and radiotherapy

Cancer continues to present significant challenges in our modern era. According to the latest estimation by the World Health Organization's (WHO) cancer agency, the International Agency for Research on Cancer (IARC), cancer is the second most common cause of deaths worldwide [1]. There were 20 million new cases and nearly 10 million deaths in 2022. As illustrated in fig. 1.1, lung cancer emerged as the prevailing cancer, with 2.5 million new instances, representing 12.4% of all new cases. Following closely was breast cancer with 2.3 million cases (11.6%), trailed by colorectal cancer at 1.9 million cases (9.6%), prostate cancer at 1.5 million cases (7.3%), and stomach cancer at 970,000 cases (4.9%).

According to the Robert-Koch-Institute, there are around 500,000 new cancer patients every year in Germany. In the latest report "Cancer in Germany 2019/2020" [2], there were more than 262,000 men and over 231,000 women diagnosed with cancer in 2020. The most frequent tumor sites (not including non-melanoma skin cancer) were prostate for men and breast for women, accounting for 25.1% and 30.5% respectively.

1.1.1 Cancer definition and diagnosis

According to the WHO, cancer comprises a collection of diseases distinguished by the uncontrollable proliferation and dissemination of abnormal cells. In [3, 4], cancer is defined as a stepwise malignant transformation that can be characterised by a sequence of biological hallmarks as acquired functional capabilities, as illustrated in fig. 1.2. This heuristic conceptualization condenses the complexity of phenotypes and genotypes of cancer into a provisional set of core principles.

Various diagnostic tools such as laboratory tests, imaging, endoscopic examinations, biopsy and histopathology examinations can be used to detect cancer. Imaging is one of the most frequently used diagnostic tools as it precisely locates and visualizes the tumor cells in an non-invasive way. Modern imaging techniques for cancer diagnosis include x-ray, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET).

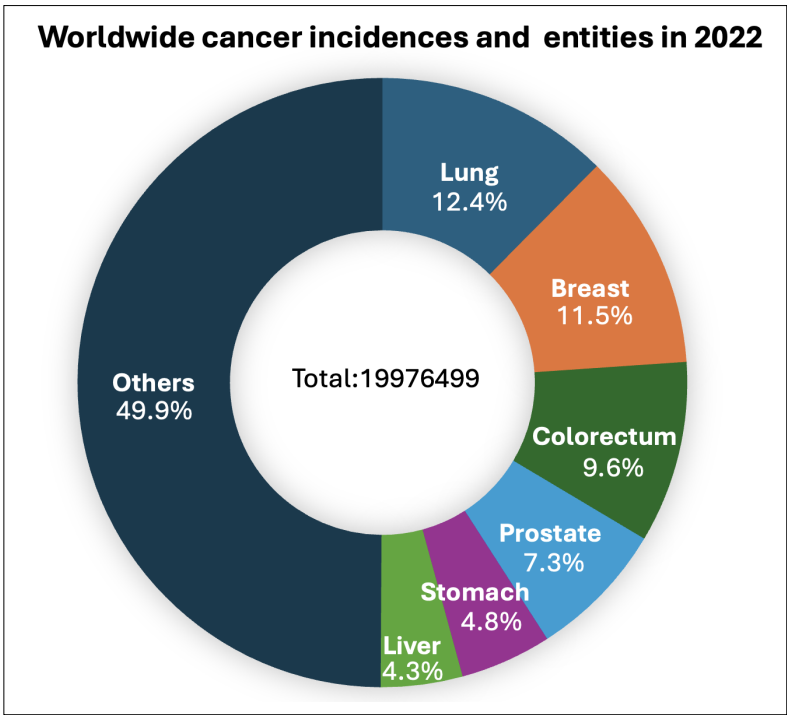


Figure 1.1: The worldwide cancer incidences and the most common entities in 2022. Own figure with the information from the International Agency for Research on Cancer (IARC)[1]. (Own figure)

1.1.2 Treating cancer with radiotherapy

Traditional cancer treatments are radiotherapy, surgery and chemotherapy, or a combination. Since the discovery of x-ray by Wilhelm Conrad Röntgen in 1895, radiotherapy has taken a crucial role for cancer treatments, with nearly 50% of all cancer patients receiving radiotherapy in their course of treatment [5]. The main goal of radiotherapy is to deprive cancer cells of their cell division potential with maximum dose while avoiding altering the cell cycles of surrounding healthy cells with minimum dose. Radio-biologically, radiation damages deoxyribonucleic acid (DNA) and triggers cell repair or cell death mechanisms. Healthy cells are more efficient in repairing themselves at a faster rate and retaining their normal function status than the cancer cells. Conversely, for cancer cells, such radiation-induced damage leads to higher rates of sterilization. Due to this, cancer cells have smaller fractionation sensitivity than normal cells. In other words, cancer cells repair and repopulate more slowly than normal cells after irradiation. Due to these fractionation sensitivity differences between normal and cancer cells, modern radiotherapy methods irradiate patients in a course of fractions with a low radiation dose. Dose prescription and fractionation are discussed in section 2.1.3.

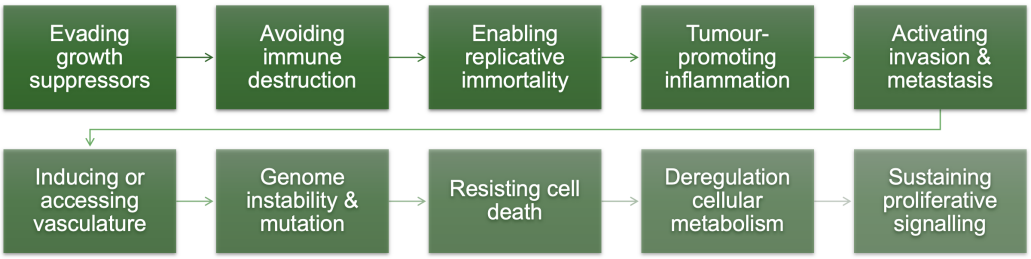


Figure 1.2: The ten hallmarks of cancer underline the fundamental principles of cancer development. Own figure adapted from [4].

1.2 Image-guided radiotherapy

The primary objective of image-guided radiotherapy (IGRT) is to integrate in-room imaging scans with radiotherapy systems for patient positioning and target localization. This allows for verification of the tumor and surrounding organs before and during each treatment, and accurate treatment delivery.

The three major classes of in-room imaging techniques using x-rays integrated into IGRT are radiographic imaging, fluoroscopic imaging and tomographic imaging. In radiographic imaging, 2D projections or planar images are acquired. In fluoroscopic imaging, a continuous stream of planar x-ray images is acquired to monitor intra-fractional patient motion based on anatomical landmarks or fiducial markers. In tomographic imaging, projections are acquired at different gantry angles which allows to reconstruct volumetric cone beam computed tomography (CBCT) images.

The in-room imaging employed in this project is kilovoltage (kV) CBCT, which can for example be mounted as an onboard X-ray Volume Imaging (XVI) system on the gantry of a Synergy medical linear accelerator (version 5.52, Elekta, Sweden). Figure 1.3 shows one of the XVI systems installed in the Department of Radiation Oncology, LMU Munich University Hospital (Großhadern campus).

Before initiating IGRT, the patient undergoes a planning CT (pCT) scan to visualize the tumor and surrounding anatomies. Based on the planning images, radiation oncologists delineate the target volumes and organ-at-risk (OAR) that need to be spared (see details in section 2.1.2). Medical physicists create a treatment plan in the treatment planning system (TPS), which outlines the radiation dose, beam angles, and treatment techniques to be used. The final plan has to be approved by a radiation oncologist.

To deliver conformal doses of radiation to the tumor while sparing surrounding healthy tissue, external beam radiotherapy techniques such as intensity modulated photon radiotherapy (IMRT) and volumetric modulated arc therapy (VMAT) [6] are applied. In IMRT, the intensity of each beam from static positions is modulated to achieve the desired dose distribution through the use of multileaf collimators (MLC). VMAT is a form of rotational IMRT that delivers radiation in a continuous arc around the patient. In the presence of inter-fractional anatomical changes over a course of treatments, the difference between

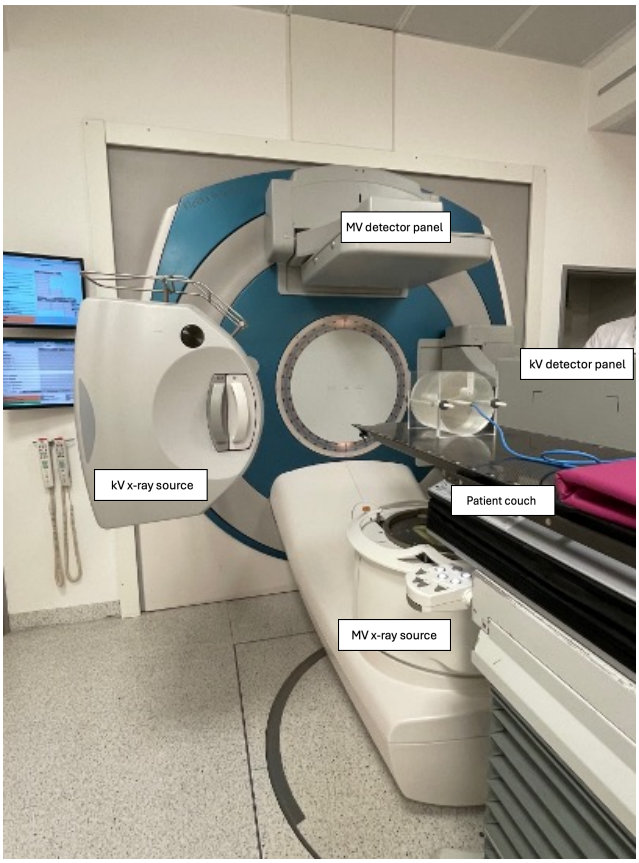


Figure 1.3: One of the Elekta medical accelerators with a gantry-mounted CBCT scanner (XVI) installed in the LMU Munich University Hospital (Großhadern). kV x-ray source and kV flat panel detector are used for CBCT imaging. MV x-ray source and MV flat detector panel are used to deliver and verify the treatment radiation. (Own figure)

acquisition date of the pCT and the actual treatment day reduces the accuracy of dose delivery.

With in-room CBCT imaging, the patient undergoes a scan on the linac treatment table right before the treatment. These updated images are then registered to the pCT to obtain transformation parameters for positioning. Figure 1.4 illustrates the general schematic CBCT-guided radiotherapy workflow with patient positioning.

As a result, IGRT enhances geometric accuracy as it provides an updated measurement of the patient treatment position. Such verified consistency of planned and actual position can reduce the extra margins of contours and the treatment uncertainties in various treatment sites [7, 8, 9, 10].

Alongside position verification, the primary objective of the thesis is to enhance IGRT by unlocking the potential use of in-room CBCT for treatment dose adaptation, while concurrently reducing the x-ray imaging dose in each CBCT scan.

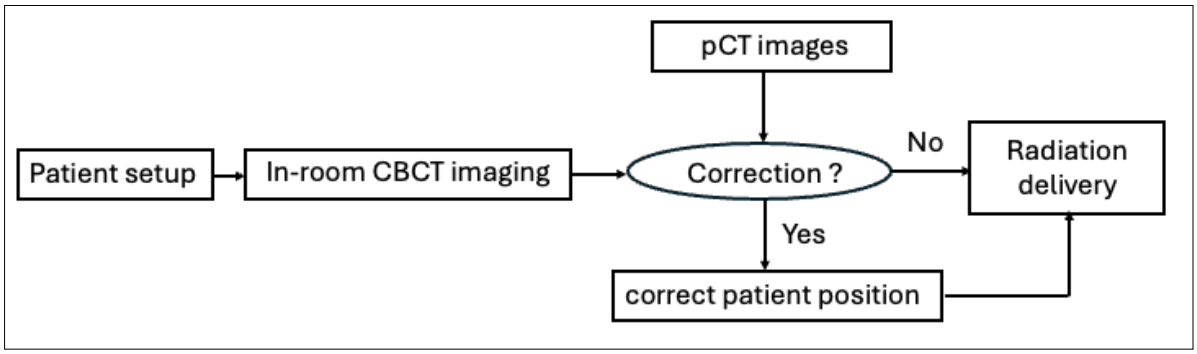


Figure 1.4: A general schematic workflow for in-room IGRT with patient position correction. The in-room CBCT images would be registered to pCT images. When correction is needed, the patient position is corrected by adjusting the couch position. If no correction is needed or when correction is completed, treatment radiation would be delivered. (Own figure)

Chapter 2

Workflow of CBCT-guided adaptive radiation therapy

In IGRT, CBCT imaging is used to assess the daily patient anatomy and to align patients in the treatment position. As the fractionated treatment progresses, anatomical changes may occur. Particularly in prostate cancer patients, inter-fractional changes due to bladder filling and rectum air pockets can be substantial (as shown in fig. 2.1). These changes may shift the surrounding tissues and diverge from the original planning contours, affecting radiation dose received by these OARs [11]. In such cases, plan adaptation becomes essential to ensure accurate delivery of radiation while minimizing exposure of healthy tissue. In CBCT-guided adaptive radiotherapy (ART), anatomical changes during the course of treatment, which can be visualised by in-room CBCT scans, are taken into account by adapting the treatment plan. Therefore it is actively utilized in clinical practice across various organ sites [12, 13, 14, 15].

This section summarizes briefly the main clinical workflow of CBCT-guided ART. These steps include initial treatment planning (section 2.1), in-room CBCT imaging (section 2.2), online adaptation (section 2.3), and radiation delivery (section 2.4).

2.1 Initial treatment planning

2.1.1 Planning image acquisition

As mentioned in section 1.2, IGRT begins with an initial pCT scan to define the tumor locations to be treated and to identify surrounding critical structures (such as bladder and rectum) to be spared. Each patient is scanned typically two to three weeks before the start of the treatment course for initial treatment planning.

In CT scans, each image pixel intensity is represented by a CT number, or a Hounsfield unit (HU) value, which represents the radiodensity of the tissues. Suppose a voxel has a mean linear attenuation coefficient μ, its corresponding HU is defined by the following formula:

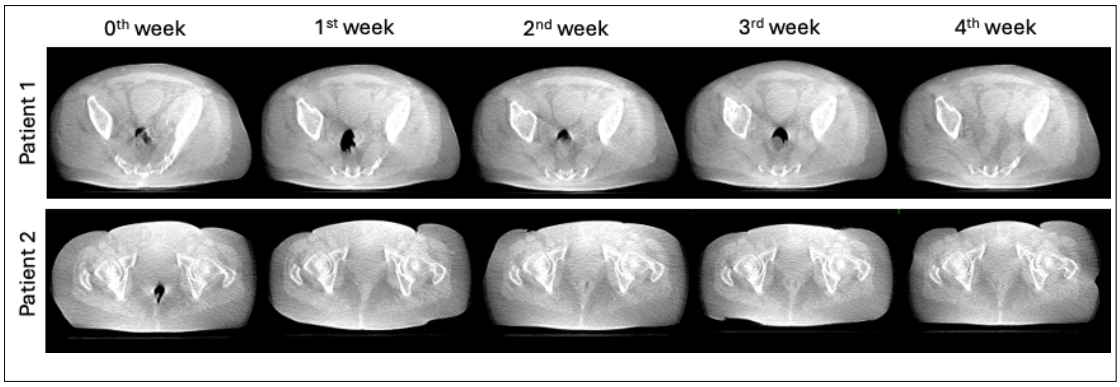


Figure 2.1: CBCT images of two exemplary prostate cancer patients over the course of fractionated treatment. In both patients, the filling of the bladder and the shape of air pockets in the rectum varied from fraction to fraction. (Own figure)

$$\mathrm{HU} = 1000 \times \frac{\mu - \mu_{\mathrm{water}}}{\mu_{\mathrm{water}} - \mu_{\mathrm{air}}}, \tag{2.1}$$

where μ_{water} and μ_{air} denote linear attenuation coefficients of water and air. The HU scale assumes water as 0 HU and air as -1000 HU.

2.1.2 Target and organ delineations

Based on the pCT images and potentially additional images such as MRI or PET images, radiation oncologists identify and manually contour the target volumes and OARs, which are defined as the following scheme according to [16]. Figure 2.2 illustrates the contouring concepts of the target volumes and OARs.

- Gross Tumor Volume (GTV): The GTV includes the visible extent and location of malignancy. The GTV can be further classified as primary tumor (GTV-T), metastatic lymphadenopathy (GTV-N), or other metastases (GTV-M).
- Clinical Target Volume (CTV): The CTV describes the tissue volume that encompasses a GTV and/or microscopic malignant disease. Such tissue volume must be irradiated to achieve the therapeutic objective.
- Internal Target Volume (ITV): The ITV includes the CTV along with an internal margin. Such margin accounts for the changes in the position and shape of the CTV relative to the reference frame of the patient, typically defined by the bony anatomy. These changes can be caused e.g., by respiratory motion.
- Planning Target Volume (PTV): The PTV is generated by incorporating a margin around the CTV or ITV. This margin accounts for various factors such as uncertainties in patient setup, organ motion, and variations in radiation delivery. The purpose

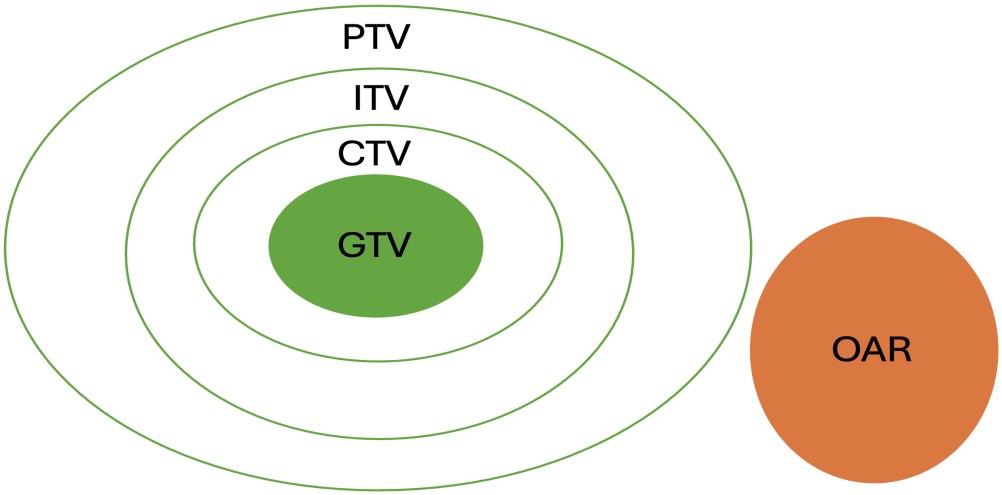


Figure 2.2: Contouring concepts of the target volumes and OARs in radiation oncology. (Own figure)

of the PTV is to ensure that the prescribed radiation dose is delivered to the entire CTV.

- Organ-at-risk (OAR): The OAR refers to the healthy organ that is near the treatment region and could be inadvertently irradiated during the treatment. The OAR exhibits a level of sensitivity to radiation where the dose administered by a treatment plan could be substantial in comparison to its tolerance level. For instance, in prostate cancer radiotherapy, among others, bladder, rectum and urethra are considered as OARs.

2.1.3 Dose prescription and fractionation

Based on the characteristics of the tumor and surrounding tissues, as well as the treatment goals, the radiation oncologist determines the treatment prescription for each patient, which specifies the total therapeutic dose to be delivered to the target volumes over the course of treatment. The dose to the OARs adhere to guidelines in the Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) report [17], which recommends the maximum tolerable dose to individual OARs.

The prescribed dose is administered over multiple radiotherapy sessions rather than in a single session, with partial doses delivered throughout several weeks. This treatment scheme, known as fractionated radiotherapy, is based on the understanding that healthy tissue cells recover more quickly from radiation-induced damage compared to tumor cells. With low doses per fraction, the risk of severe side effects can be reduced in the surroundings of dose-sensitive organs. The biological effect of a radiation dose considering the

fractionation scheme and the characteristic radiosensitivity of the tissue being treated can be formulated as biologically effective dose (BED):

$$\text{BED} = n \times d \times [1 + \frac{d}{\alpha/\beta}]. \tag{2.2}$$

where n denotes the number of fractions, d denotes the dose per fraction and α/β denotes the radiation biological value. Compared to normal cells, cancer cells have higher α/β values and are less sensitive to the sparing effect by fractionation. In clinical practice, a prostate cancer patient, for example, is usually prescribed with 37 fractions of 2 Gy. This approach balances the goal of killing tumor cells while minimizing damage to surrounding healthy cells and allowing time for normal tissue repair between treatments.

2.1.4 Treatment plan optimization and evaluation

With such a prescription and CT images, a patient-specific radiation treatment plan can be created using a TPS. The aims of such a plan is to deliver a homogeneous dose in the target and a low dose outside of the target, ideally with steep dose gradients.

In the TPS, the plan is created on the pCT images, after specifying the isocenter, gantry angles, collimator angles and the gantry rotation direction. Next, in the plan optimization step, objectives and constraints for PTV and OARs are specified. Various dose calculation algorithms, such as Monte Carlo methods [18] and collapsed cone [19] are readily available in the TPS. All algorithms assume a discretized patient model as a grid of voxels to calculate the photon interaction with the tissue. The probabilities of these interactions depend on the electron density and elemental composition, which is calculated from HU values on the pCT images using a calibration curve or conversion table for each individual tissue. The dose calculation also takes into account factors such as beam attenuation, the fluence modulation with the MLC, scatter, and tissue heterogeneity. Based on the entered beam parameters, the TPS performs dose calculations with the selected algorithm, yielding a preliminary dose distribution in the patient's anatomy. With the initial dose calculation, the plan is adjusted to define the optimal beam and MLC settings in an iterative process.

To evaluate the treatment dose planning, a dose-volume histogram (DVH) is computed to visualize graphically the relationship between the dose received by a particular volume of tissue, such as PTV, CTV, rectum, bladder, and the percentage of that tissue volume receiving a specific dose.

2.2 In-room CBCT scan

To ensure that the tumor's position relative to the beam is accurate, a CBCT scan is acquired using the onboard imaging system of the linear accelerator to align the patient prior to treatment delivery.

Based on the CBCT scan, an updated three-dimensional patient's anatomy can be accessed, which allows us to register the patient position with the pCT images in the TPS. The transformation parameters are computed with six degrees of freedom and can be used to align the patient couch with the treatment position. Figure 2.3 illustrates an example of rigid registration of the in-room CBCT with the pCT images, and the resulting transformation parameters.

In addition to patient positioning, CBCT images visualize interfractional variations of anatomical structures. If the CBCT scan reveals that the tumor or critical structures have deviated from their planned position, the treatment plan would have to be adapted.

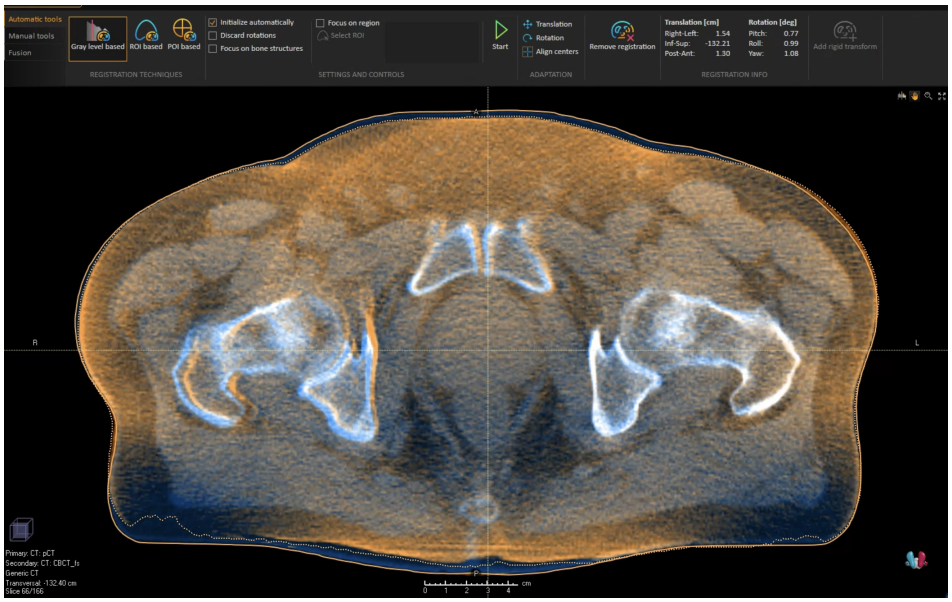


Figure 2.3: An exemplary rigid registration of a pelvic in-room CBCT to pCT in a TPS (RayStation, version 10.01, RaySearch, Sweden). The transformation parameters (top right panel) are computed to align the patient with the treatment position. (Own figure)

2.3 Online adaptation

Over the course of fractionated treatments, the anatomy of prostate cancer patients can have considerable deviations from the initial plan due to the bladder filling and rectum air pockets (as shown in fig. 2.1). The contours of the target volume and OARs on pCT images might not match with the in-room CBCT images. Consequently, the efficacy of the highly conformal treatment plan diminishes when the target volume fails to correspond to the patient's updated anatomy. If we apply only one treatment plan throughout all fractions, errors in dose distribution might lead to radiation-induced toxicities in surrounding normal cells or underdosage of the tumor.

Since late 1990s, adaptive radiotherapy is generally described as modifying the radiation treatment plan offline to accommodate inter-fractional anatomical changes [20]. For CBCT-guided online adaptation, changes in the patient's anatomy are identified with in-room CBCT scans and a new treatment plan is generated accordingly with the patient on the table.

The online adaptation workflow is illustrated in fig. 2.4.

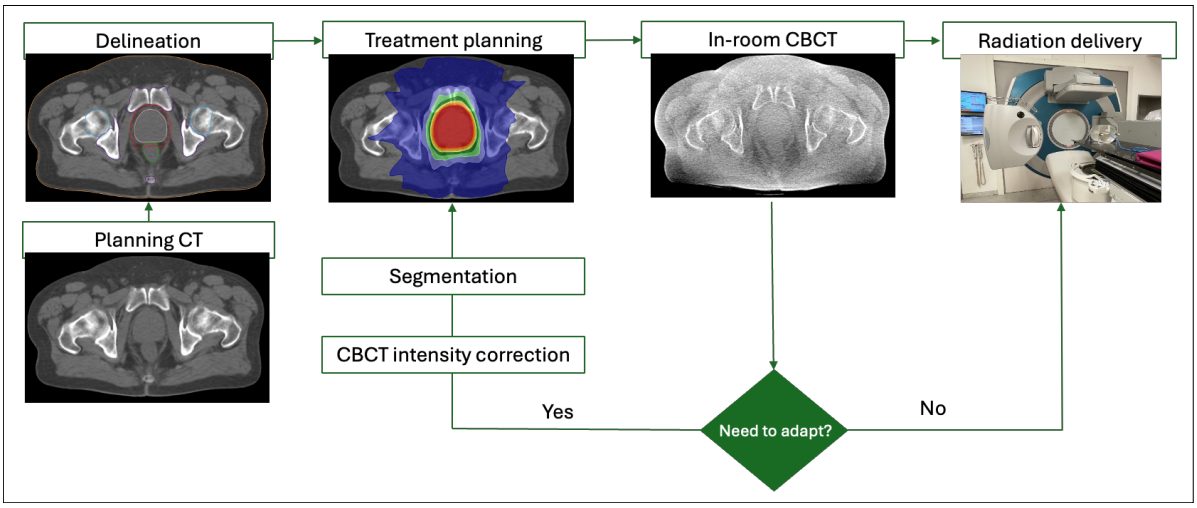


Figure 2.4: Workflow of CBCT-guided online adaptation radiotherapy. (Own figure)

2.3.1 The importance of correcting CBCT

Conventionally, treatment plan adaptation often requires a new offline pCT scan of the patient. This requires a significant amount of time for arranging a new CT scan and a new treatment planning. Ideally, we can use the in-room CBCT images to assess the daily patient anatomy and re-optimize a treatment plan accordingly when the patient is still positioned on the table. However, CBCT images are typically insufficient to infer treatment dose [21]. First, CBCT images contain various artifacts and inaccurate CT numbers (refer to section 3.3). Second, CBCT images often have lower soft tissue contrast compared to pCT images. This presents a greater challenge and extends the duration to accurately delineate target volumes and OAR structures, especially in small structures or complex regions with ambiguous boundaries. One can generate a virtual CT (vCT) with deformable image registration (DIR)(refer to section 3.4), however, geometrical inaccuracies are usually seen in the organ shapes due to registration uncertainties or large discrepancies such as the bowel air pockets, ultimately affecting the treatment accuracy.

A few conventional non-AI ways to correct CBCT images for plan adaptation are discussed in section 3.4. However, these corrections require a considerable amount of time. Therefore, using CBCT images for online adaptation remains challenging.

2.3.2 The current usecases of corrected CBCT in online adaptation

While traditional CBCT correction methods may fall short in achieving online adaptation, a new commercial CBCT linac system, Ethos (Varian Medical Systems, Palo Alto, California, USA), is able to correct CBCT scans by registering pCT images into it and instantaneously visualize the propagated contours on the daily patient anatomy with CT diagnostic quality. Studies [22, 23] have demonstrated that the application of Ethos system is feasible in clinical workflows to account for inter- and intra-fractional variations. In addition, studies [24, 25] have shown that Ethos online adaptive recontouring and replanning processes improved treatment dose accuracy. More recently, an upgraded CBCT imager HyperSight (Varian Medical Systems) was launched with the possibility to be mounted in a Halcyon or Ethos ring-gantry system [26, 27] and to scan patients within six seconds and output volumetric images with accurate HU for treatment dose adaptation. These studies suggested that it is possible to use full dose CBCT images for online adaptation in the clinical practice.

2.4 Radiation delivery

Radiation is delivered by rotational irradiation in which the accelerator head rotates around the patient. VMAT is commonly used in hospitals, simultaneously varying the dose rate and shape of the radiation beam during the rotation. The radiation can be delivered from multiple arcs from different angles. During each arc, the accelerator system modulates the gantry speed, beam intensity, dose rate, and MLC shape according to the treatment plan.

Chapter 3

Physics of cone-beam computed tomography

In radiotherapy, CBCT scanners can be integrated into the linear accelerator system and used as an image guidance technique to align the patient and verify the target position right before the treatment. In general, a CBCT scan varies from a CT scan, which uses a fan-shaped x-ray beam in a helical scanning trajectory to acquire each image slice of the field-of-view (FOV). CBCT uses a pyramid shaped scanning geometry and a flat panel detector which has extended rows perpendicular to the beam, which covers the entire FOV with only one rotation of the gantry. In the following four sections, image acquisition (section 3.1), reconstruction (section 3.2), image noise and artefacts (section 3.3), non-AI image correction (section 3.4) and imaging radiation dose (section 3.5) are discussed.

3.1 Image acquisition

3.1.1 CBCT system

CBCT imaging is performed using a rotating gantry on which an x-ray source and a flat panel detector are mounted. As illustrated in fig. 3.1, an ionizing radiation source emitting a cone shaped x-ray beam is directed towards the center of the region of interest. The x-rays are detected by a large two-dimensional (2D) flat panel detector situated on the opposite side. The x-ray source and detector rotate around a central rotation point of an object fixed within the area of interest. The cone beam is defined by the following parameters: the size of the flat panel detector, the distance of the source to the detector (SD) and the source to iso center distance (SC). The cone angle (θ) which is proportional to the height of the panel size also defines the cone beam.

3.1.2 Flat panel detector

In radiotherapy, the flat panel detector is used to measure the x-rays attenuation in the CBCT system [28, 29, 30]. The detector uses a structured Cesium iodide (CsI) scintillator

to convert x-rays into light. Light is then converted into electrical charge in the photodiode, which is integrated in each pixel of a thin-film transistor (TFT) array. This array is used to collect and process these electrical signals.

In terms of the readout scheme, one TFT row is turned on at a time to collect the charges, typically yielding rapid readout with 30 frames per second (fps) at half resolution or 15 fps at full resolution and the pixel size of 1024 × 768 in a representative flat panel detector [31]. These flat panel images cannot be directly used as dead pixels exist and the pixel sensitivities varies across the panel. Dark and flood signals can be measured and calibrated in the flat panel without an object while the x-ray source is off and on, respectively. For the Elekta XVI CBCT scanners that are used in our hospital, the signal S_{XVI} is saved inverted and can be formulated as

$$S_{\mathrm{XVI}} = 2^{16} - I_{\mathrm{XVI}}, \tag{3.1}$$

where I_{XVI} denotes the dark and flood corrected measurement. The signals are saved as unsigned 16-bit integers, ranging from 0 to $(2^{16} - 1)$.

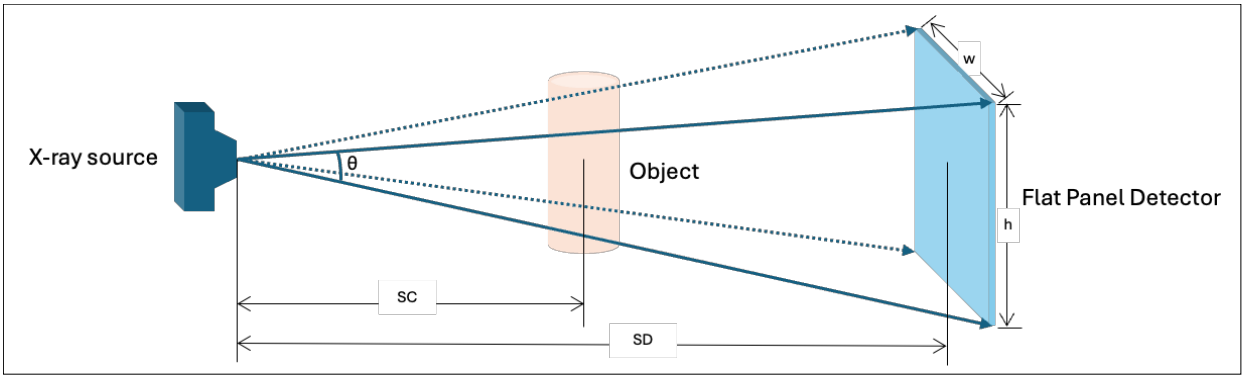


Figure 3.1: A CBCT system consists of an x-ray source and a detector. SD is the distance of the source to the detector, and SC is the source to iso center distance. θ denotes the cone angle proportional to the height of the panel h. (Own figure)

3.2 Image reconstruction

In CBCT, a reconstructed image depicts the spatial distribution of x-ray linear attenuation coefficients. The signal collected on the detector is the x-ray intensity value attenuated after an object according to the Lambert-Beer law. The line integral of the attenuation coefficient distribution over the path of an x-ray beam, or projection, can be computed by applying a negative logarithm to the ratio of the attenuated to the initial x-ray intensity:

$$P_{\mathrm{XVI}} = -\ln\left(\frac{I_{\mathrm{XVI}}}{I_0}\right), \tag{3.2}$$

where P_{XVI} denotes the projection, I_0 and I_{XVI} denote the initial and the measured x-ray intensity, respectively.

For parallel beam projections, we can apply the filtered backprojection (FBP) algorithm based on the Fourier slice theorem to obtain a cross-sectional image of the object using its projections. According to [32], the object can be computed as

$$f(x,y) = \int_0^{\pi} p^F(\theta, s) d\theta, \tag{3.3}$$

where (x, y) is an image point coordinate, and s denotes the distance between the origin of the coordinates and the x-ray line that passes through the point (x, y) with the angle θ.

$$s = x\cos\theta + y\sin\theta. \tag{3.4}$$

$p^F(\theta, s)$ denotes the filtered projection data. Assuming that a projection is a signal with limited bandwidth in the frequency domain, the filtered projection data can be expressed in [32] as

$$p^F(\theta, s) = \int_{-\infty}^{+\infty} p(\theta, s') \cdot h(s - s') ds', \tag{3.5}$$

where $p(\theta, s)$ denotes the measured projection at location s. $h(s)$ denotes the ramp filter.

3.2.1 Fan beam reconstruction

In clinical CBCT scanners, near-point x-ray sources are used to emit x-ray in a fan beam (2D) or cone beam (3D) projection. In fan beam reconstruction, the x-ray data is required to be sorted into parallel coordinates before applying FBP. In the parallel beam geometry, it is sufficient to scan only $\theta \in [0, \pi]$. As illustrated in fig. 3.2, an x-ray source rotates in a circle trajectory $R_F(-\sin\beta, \cos\beta)$, where R_F and β denotes the radius and the rotational angle respectively, an equispatial fan beam projection $g(\beta, t)$ can be produced with respect to t-axis on a virtual detector. According to [32], this fan beam projection can be expressed as a parallel beam projection using this formulation:

$$\begin{aligned} p(\theta, s) &= g(\beta, t) \\ &= g(\theta - \arcsin\frac{s}{R_F}, \frac{sR_F}{\sqrt{R_F^2 - s^2}}). \end{aligned} \tag{3.6}$$

After rebinning the projections, one can reconstruct the image using the parallel FBP algorithm as shown in eq. (3.3) and eq. (3.5). On the other hand, one can also compute a reconstruction algorithm using fan beams. We can pre-weight and filter the fan beam

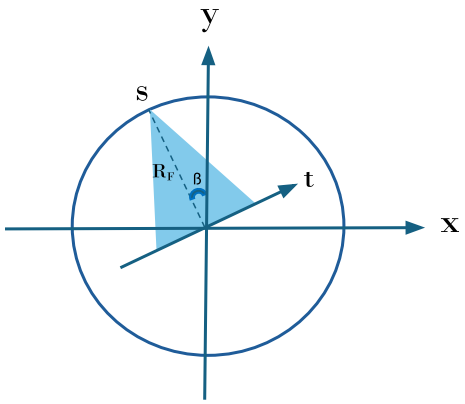


Figure 3.2: An illustration of an equispatial fan beam projection $g(\beta, t)$. In fan beam reconstruction, the virtual detector is illustrated along t-axis line. (Own figure)

projection signal with a weighted ramp filter before applying FBP. According to [32], the FBP algorithm for fan beams in the equispatial plane can be written as

$$f(x,y) = \frac{1}{2}\int_0^{2\pi} \frac{{R_F}^2}{(R_F + x\sin\beta - y\cos\beta)^2} \times \int_{-\infty}^{+\infty} g(\beta,t)\frac{R_F}{\sqrt{{R_F}^2 - s^2}} \cdot h(t'-t)dtd\beta, \quad (3.7)$$

where

$$t' = \frac{(y\sin\beta + x\cos\beta)R_F}{R_F + x\sin\beta - y\cos\beta}. \quad (3.8)$$

Similar to the computational structure for the FBP, the fan beam reconstruction involves two weighting factors during the convolution and backprojection steps as formulated in eq. (3.7).

3.2.2 Cone beam reconstruction using FDK algorithm

In CBCT scans, projection data is acquired in a 3D cone beam geometry. Such a circular trajectory, however, does not sufficiently cover the Radon space. Therefore we need an approximate reconstruction method: the Feldkamp, Davis, Kress (FDK) algorithm [33], in which we can simplify the cone beam reconstruction to a fan beam reconstruction with a flat panel detector. As illustrated in fig. 3.3, this transformation involves re-projecting the cone beam data onto a virtual plane (u, v) parallel to the real detector plane. This virtual plane is located at the SC away from the x-ray source at the axis of rotation (z-axis). This process flattens the cone beam data into a set of fan beam projections.

The cone beam reconstruction algorithm is similar to fan beam filtered backprojection with adjusted weighting schemes. According to [32, 33], for a circular trajectory $R_F(-\sin\beta, \cos\beta)$, the FDK algorithm is formulated as

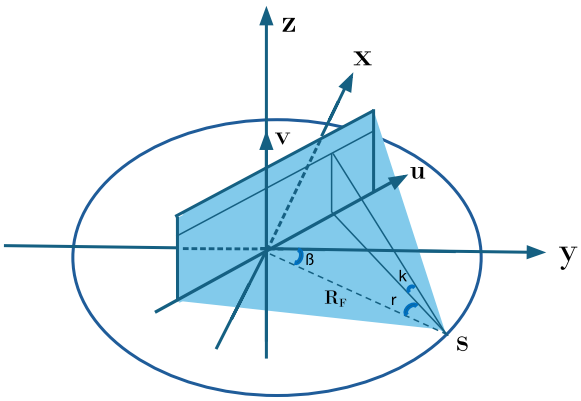


Figure 3.3: An illustration of FDK reconstruction with a virtual detector plane. In FDK reconstruction, the virtual detector is illustrated on (u, v)-axis plane. (Own figure)

$$f(x,y,z) = \frac{1}{2}\int_0^{2\pi} \frac{{R_F}^2}{(R_F + x\sin\beta - y\cos\beta)^2} \times \int_{-\infty}^{+\infty} p(\beta,u,v') \frac{R_F}{\sqrt{{R_F}^2 + u^2 + v'^2}} \cdot h(u'-u)dud\beta, \tag{3.9}$$

where

$$v' = \frac{zR_F}{R_F + x\sin\beta - y\cos\beta}. \tag{3.10}$$

Similar to eq. (3.7), the pre-weighting factor in eq. (3.9) is multiplied by the cosine of the cone angle. According to [32, 33], it can be computed as

$$\frac{R_F}{\sqrt{{R_F}^2 + u^2 + v'^2}} = \frac{\sqrt{{R_F}^2 + u^2}}{\sqrt{{R_F}^2 + u^2 + v'^2}} \cdot \frac{R_F}{\sqrt{{R_F}^2 + u^2}} = \cos k \cdot \cos\gamma. \tag{3.11}$$

where k and γ denote the cone angle and the fan angle for a specific x-ray, respectively.

3.3 Image noise and artefacts

CBCT image quality is generally lower than normal CT due to the noise and artefacts. The image noise can be understood as the stochastic variation in voxel values, i.e. the

fluctuations in voxel values which arise independently of the changes in attenuation coefficient of the images structure. Depending on the position in the image, the noise in CBCT is non-stationary, due to the fact that the mean and variance of radiation intensity at the central region of the detector vary from those at the edge of the detector. The contrast, which represents the signal difference among regions in a CBCT image is also lower than in a CT image. In the followings, several factors that cause noise and artefacts are discussed.

- X-ray scatter: Scattering is one of the main sources of artefacts in CBCT. When x-rays pass through a patient's body, photons can lose energy due to Compton interactions with objects and produce scattered radiation. Most photons have been scattered at least once before reaching the detector. In addition, the flat panel detectors usually do not have anti-scatter grids, capturing more scatter than a narrow CT detector. These scatter artefacts can vary depending on the air gap in the system geometry, object size, or the FOV. As a result, artifacts occur in a form of low-frequency shading and streaks, especially in the regions of structures with high attenuation.

- Beam hardening: The lower energy fraction of the x-ray spectrum is absorbed faster. This results in an increase in mean beam energy and non-linear variations in the x-ray spectrum as it passes through different tissue and thickness. Especially when passing through dense tissues or materials (e.g. bone or metal), the beam becomes harder and these regions are reconstructed with incorrect attenuation values, exhibiting low frequency nonlinear attenuation distortions (cupping artefact) throughout the image.

- Aliasing artefacts: The Nyquist sampling theorem states that the sampling frequency should be greater than twice the highest frequency present in the signal. The aliasing artefacts in the CBCT image are mainly due to the divergence of the cone-shaped beam. The voxels near the source will be sampled more densely than those near the detector, therefore causing aliasing at the periphery.

- Image lag and ghosting: In the detector, residual x-ray signals exist in subsequent frames (lag). Or the sensitivity of the scintillator changes after exposure (ghosting).

- Cone beam artefacts: From a circular source–detector orbit, CBCT scans the volume with incomplete sampling. The edges parallel to the source-detector orbital plane show bright or dark signal smearing, which intensify as the distance from the central axial plane increases.

- Lateral truncation: It is caused by the FOV being smaller than the patient's lateral extent. Some patients may exceed the FOV dimensions in spite of the use of shifted detectors. The artefacts appear as bright rings and dark cups at the periphery of the patient outline in the reconstructed image.

- Motion artefacts: It is caused by respiratory motion of the patients. Especially in the lung region, the breathing causes the location of the anatomy to change, causing inconsistency in the the continuously acquired projections. One correction to this in

4D CBCT applications is to extract the breathing signal and bin the projections into certain breathing frames, so as to reconstruct "static" 3D images along the breathing cycle.

3.4 Non-AI scatter corrections

In order to correct CBCT image intensity, CBCT scatter correction techniques have been developed, including a look-up-table based solution [21], the use of vCT from pCT-to-CBCT registration [34, 35, 36, 37, 38, 39] and Monte-Carlo (MC) based methods [40, 41, 42]. Although several methods have been successful in demonstrating accurate CBCT-based dose calculation in various anatomical sites [43, 44, 45, 46], there are certain constraints. For example, DIR based methods that achieve accurate dose calculation in the head and neck (HN) [21, 36], might have more DIR inaccuracies for a pelvic scan, which has greater inter-fractional changes in anatomy (as discussed in section 2.2). In addition, the time required for correcting images using vCT or MC based methods, which takes up several minutes to hours, restricts the implementation for online adaption.

Here we discuss the CBCT scatter correction technique [45, 47] that has been applied to generate ground-truth images in this thesis, in the following referred to as $\mathrm{CBCT}_{\mathrm{cor}}$ (as shown in fig. 5.1). Several studies have validated this method [48, 49], in which a vCT image is generated using a DIR algorithm mapping pCT onto the original CBCT images, referred to as $\mathrm{CBCT}_{\mathrm{org}}$, and used the vCT as a prior for scatter correction of the acquired projections.

In detail, we first forward project the vCT according to the geometry of the CBCT system to obtain primary beam projections (I_{pri}). The scatter and other low frequency deviations (I_{sca}) are computed as the difference between a scaled original $\mathrm{CBCT}_{\mathrm{org}}$ projection (I_{org}) with ISF and (I_{pri}) followed by a smoothing function f (a median filter with dimension 25×25 pixels, followed by a Gaussian filter of 1.5 pixels standard deviation):

$$I_{\mathrm{sca}} = f(\mathrm{ISF} \times I_{\mathrm{org}} - I_{\mathrm{pri}}). \tag{3.12}$$

The scatter corrected projection (I_{cor}) can be estimated as the difference between the original measured $\mathrm{CBCT}_{\mathrm{org}}$ projection and the scatter contribution:

$$I_{\mathrm{cor}} = \mathrm{ISF} \times I_{\mathrm{org}} - I_{\mathrm{sca}}, \tag{3.13}$$

With I_{cor}, a presumably scatter-free CBCT, in the following referred to as $\mathrm{CBCT}_{\mathrm{cor}}$ (as shown in fig. 5.1) can be computed. $\mathrm{CBCT}_{\mathrm{cor}}$ has HU values equivalent to the pCT, and ideally shares identical anatomical information with $\mathrm{CBCT}_{\mathrm{org}}$.

3.5 Radiation dose

While x-ray beams penetrate the body, ionizing radiation can damage DNA and cause gene mutations during cell repair, potentially leading to cancer. Here we discuss the

standardized metrics according to the International Commission on Radiological Protection (ICRP) [50, 51] that can describe such radiation in section 3.5.1, and the specific metrics [52, 53, 32] for CT and CBCT in section 3.5.2.

3.5.1 Radiation dose units

- Exposure: It refers to the quantity of charge per unit mass of air generated by an x-ray beam. Its SI unit is expressed as coulomb per kilogram (Ckg^{-1}).

- Kerma: It is the non-stochastic amount of the initial kinetic energy transferred to charged particles by uncharged radiation per unit mass $\mathrm{d}m$. Suppose $\mathrm{d}E_{\mathrm{kinetic}}$ represents the expectation value of the sum of the kinetic energies, kerma can be formulated as

$$K = \frac{\mathrm{d}E_{\mathrm{kinetic}}}{\mathrm{d}m.} \tag{3.14}$$

  The SI unit is expressed as joule per kilogram (Jkg^{-1}) or gray (Gy).

- Absorbed dose: It is the quantity of energy released in matter or tissue per unit mass, with an unit of gray (Gy) or joule per kilogram (Jkg^{-1}).

$$D = \frac{\mathrm{d}\bar{\varepsilon}}{\mathrm{d}m_{\mathrm{tissue}}}. \tag{3.15}$$

  where $d\bar{\varepsilon}$ denotes the mean energy imparted to tissue of mass dm by ionising radiation, e.g. x-ray. The absorbed dose of a patient from a CBCT scan cannot be directly measured.

- Equivalent dose: Even in the identical absorbed dose, the biological effects of different radiation types on different tissue can vary. According to the standardization by ICRP 103 [50], for a certain tissue type, the equivalent dose additionally considers a radiation specific weighting factor w_{r}. For instance, proton or alpha particles have a higher weight than photons. Suppose $D_{\mathrm{t,r}}$ is the energy dose applied to the tissue t by the radiation type r, the equivalent dose Q_{t} can be formulated as

$$Q_{\mathrm{t}} = \sum_{\mathrm{r}} w_{\mathrm{r}} D_{\mathrm{t,r}}. \tag{3.16}$$

  The SI unit is typically written as joule per kilogram (Jkg^{-1}) or Sievert (Sv). In CBCT scan, x-rays have w_{r} equal to 1.

- Effective dose: In addition to the radiation specific weighting factor w_{r}, the effective dose also considers the tissue specific weighting factor w_{t}, which represents the proportional impact of a tissue on the overall health detriment. The tissue weighing factor is determined by the radiosensitivity of the organ and also standardised by ICRP. The effective dose (E) estimates the total weighted equivalent doses in all tissues by all radiation types. It can be formulated as

$$E = \sum_{\mathrm{t}} w_{\mathrm{t}} Q_{\mathrm{t}} = \sum_{\mathrm{t}} w_{\mathrm{t}} \sum_{\mathrm{r}} w_{\mathrm{r}} D_{\mathrm{t,r}}. \tag{3.17}$$

  The SI unit is identical as of the equivalent dose: joule per kilogram (Jkg^{-1}) or Sievert (Sv).

3.5.2 CT and CBCT dose indices

Unlike planar x-ray, CT possesses a helical acquisition geometry as the x-ray source rotates around the patient body. Moreover, the dose distribution of CT exhibits higher radial symmetry than that of a planar x-ray scan. Thus, CT and CBCT require additional dose quantities for quantification.

CT dose index (CTDI) is the amount of radiation dose involved during a CT scan [32]. This index is a comparative measure of the x-ray output and the dose recorded in a specific size phantom. Therefore it is an indication of the patient dose but not the dose absorbed by patients. The most basic common form is CTDI_{100}, for which a pencil-shaped ionization chamber with a length of 100 mm is used to measure the total exposure of a phantom in an axial CT scan.

$$\mathrm{CTDI}_{100} = \frac{1}{L} \int_{-50\mathrm{mm}}^{+50\mathrm{mm}} D(z) dz, \tag{3.18}$$

where L denotes the slice width, and $D(z)$ represents the dose profile along the longitudinal axis z. The dose profile is normalized by the width of slice to approximate the average dose from scans across a 100 mm length of adjacent slices. Depending on the pencil chamber location, this CTDI_{100} term can be weighted with the values when pencil chambers are inserted to the center or to the periphery of the phantom. The average dose $\mathrm{CTDI}_{w}^{\mathrm{avg}}$ can be expressed as

$$\mathrm{CTDI}_{\mathrm{w}}^{\mathrm{avg}} = \frac{1}{3}\mathrm{CTDI}_{\mathrm{w}}^{\mathrm{center}} + \frac{2}{3}\mathrm{CTDI}_{\mathrm{w}}^{\mathrm{peripheral}}, \tag{3.19}$$

where $\mathrm{CTDI}_{\mathrm{w}}^{\mathrm{center}}$ and $\mathrm{CTDI}_{\mathrm{w}}^{\mathrm{peripheral}}$ denotes the central and the peripheral dose of the phantom, respectively. Considering a helical 3D CT scan which has a table movement per rotation d and a width of the x-ray beam s, the volumetric CTDI ($\mathrm{CTDI}_{\mathrm{VOL}}$) can be formulated as

$$\mathrm{CTDI_{VOL}} = \frac{\mathrm{CTDI_w^{avg}}}{\mathrm{pitch}}, \tag{3.20}$$

where the pitch is defined as

$$\mathrm{pitch} = \frac{s}{d}. \tag{3.21}$$

$\mathrm{CTDI_{VOL}}$ can thus reflect the relative dose among various protocols and the CT systems from different vendors.

For CBCT, there is no direct translation from CTDI from the American Association of Physicists in Medicine (AAPM) Therapy Physics Committee Task Group 180 [54]. In 2010, Hyer and Hintenlang have measured and evaluated the cone beam dose index (CBDI) value, representing the mean volumetric dose within the CTDI phantom [53]. For the pelvic CBCT scan considered in this thesis, [53] reported a CBDI value (table 2 in [53], chest protocol) for the identical configuration as the protocol used in this thesis (M20 protocol with 120 kV and a bowtie filter at an Elekta XVI scanner) of 1.62 mGy/100 mAs. A body CTDI phantom with a length of 15 cm and a diameter of 32 cm was used for the measurement using a pencil chamber of a length 100 mm.

Chapter 4

Deep learning-enabled full dose CBCT-to-CT translation

In CBCT-guided online adaptive radiotherapy, the patients come to the treatment room and receive a CBCT scan for positioning. When adaptation is needed due to tumor location or anatomical changes, ideally the CBCT image can be used for treatment re-planning. However, as discussed in section 3.3, CBCT image quality is insufficient for treatment dose calculation. The scatter correction method (section 3.4), which can take up more than 10 minutes, is also impractical for the online adaptation workflow.

In the last decade, leveraging a deep convolutional neural network (CNN) to improve CBCT image correction has attracted a lot of attention. With the fast developments in deep learning (DL) algorithms, many studies have applied various types of network architectures to translate volumetric CBCT to CT image quality, and generate synthetic CT (sCT) images. These trained DL models can correct CBCT images within seconds and enable online adaptation.

In this chapter, we discuss the three main DL algorithms (U-Net, cycleGAN and CUT) and their applications in full dose CBCT-to-CT translation tasks in section 4.1, section 4.2 and section 4.3.

4.1 U-Net

Along the developments in DL algorithms, U-Net, developed by Ronnenberger *et al.* in 2015 [55], has been widely adopted beyond the original application in biomedical image segmentation. Due to its flexibility of architecture design, U-Net has been applied in many other medical image applications such as image denoising, image reconstructions, modality transfer etc.

4.1.1 Network architecture and loss function

An exemplary U-Net architecture is illustrated in fig. 4.1. The network has an encoding and a decoding path. In the encoding path, the image features are typically extracted along 4 convolutional layers. For 2D image training, each layer contains a 3×3 convolution with padding, followed by a batch normalization, a rectified linear unit (ReLU) and a max pooling with stride 2 for downsampling, decreasing the image spatial dimensions by half. As the image passes through the encoder, the number of channels doubles after each downsampling. This allows the network to capture complex features at various scales.

In the decoding path, feature maps are up-sampled by the same number of layers containing a 3×3 transposed convolution that halves the number of channels. To retain details lost during downsampling, skip connections are applied to concatenate feature maps from the corresponding layers in the encoding path to those in the decoding path. Each layer is followed by a a batch normalization and ReLU. At the last layer, a 1×1 convolution is applied to map the feature representations from the final up-sampling layer to the reconstructed output image. The above hyper-parameters can be adjusted and optimized depending on the usecases.

U-Net is typically trained with paired datasets, for which each pixel in the predicted image is compared to a corresponding ground truth. For instance in the image translation task, mean absolute error (MAE) or mean squared error (MSE) can be computed along the training process. A typical loss function can be formulated as

$$\mathcal{L} = \lambda_1 \cdot \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| + \lambda_2 \cdot \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, \tag{4.1}$$

where n represents the number of pixel in an image, y and $\hat{y}$ denote the pixel intensities of the predicted image and of the ground truth, respectively. λ_1 and λ_2 are the weights of the MAE and the MSE terms, respectively.

4.1.2 Related studies

There are a number of studies translating CBCT to CT images using U-Net. In the pelvic region, Kida *et al.* [56] trained a U-Net with CBCT images and vCT images as input and ground truth to convert CBCT images into sCT images. Other U-Nets were trained for projection based CBCT image correction with MC simulated scatter distributions [57, 58] or with corrected projections obtained by a vCT-based algorithm [59, 60]. In the thoracic region, Thummerer *et al.* [61] trained a U-Net based deep CNN to convert 4D CBCT images to sCT images, and demonstrated high HU and dose calculation accuracy for adaptive proton therapy.

The sCT images using U-Nets are typically evaluated in HU, photon and proton dose calculation accuracy. For instance, Landry *et al.* [60] investigated three U-Nets trained with original and corrected CBCT projections, original CBCT and vCT images, and original and corrected CBCT images, and computed HU accuracy of sCT (MAE 48/88/56

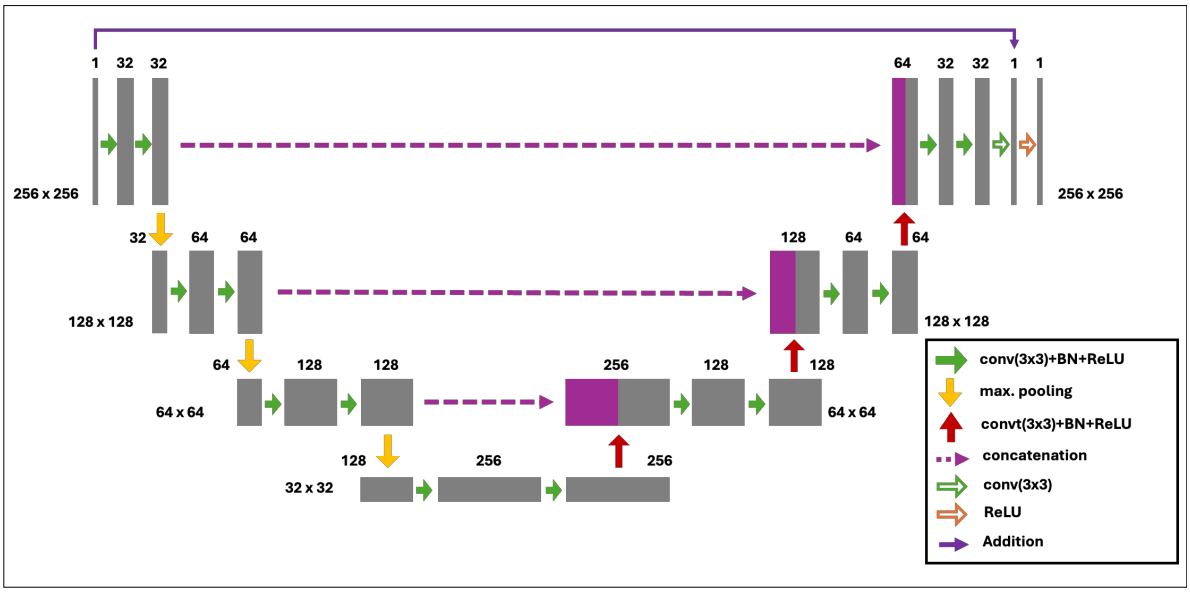


Figure 4.1: An exemplary U-Net architecture. Each box denotes a multi-channel feature map, and the number of channels is shown above the box. The 2D image/feature map sizes are indicated at the lower left corner of the box in the encoding path and lower right corner in the decoding path. Purple boxes represent the concatenated feature maps, and the arrows represents the various operations indicated in the lower right black box. Conv (3×3), BN and ReLU denote convolution with a kernel size of 3×3, batch normalization and rectified linear unit, respectively. Convt (3×3) denotes transpose convolution with a kernel size of 3×3. (Own figure)

HU) from each U-Net compared to the reference corrected CBCT. VMAT and proton pencil beam scanning single field uniform dose plans were optimized on the reference CBCT images and recalculated on sCT images. These sCT images achieved 1% photon dose difference pass rates $\geq 98.4\%$.

4.2 Cycle-consistent generative adversarial network (cycleGAN)

A cycle-consistent generative adversarial network (cycleGAN) can be used for unpaired CBCT-to-CT translation. First developed in 2017 by Zhu and Park *et al.* [62], cycleGAN is able to extract the features from CBCT and CT images and translate CBCT input images into sCT images without corresponding CT image ground truths. This unpaired training scheme eliminates the geometrical inaccuracies that could be caused by pCT to CBCT image registration or vCT generation.

4.2.1 Network architecture and loss function

The network architecture of cycleGAN in this thesis mainly follows the original work by Zhu and Park *et al.* [62] and another implementation by Ge *et al.* [63]. In general, cycleGAN consists of two sets of generators and discriminators. Each set is trained in a forward and a backward cycle, respectively. For the generator, the nine-blocks residual network architecture suggested by Johnson *et al.* [64] is used. The architecture includes a downsampling process, which reduces the input dimensions from 256 × 256 to 32 × 32 using three 2D convolutional layers (each with a stride of 2 and a kernel size of 3). The feature map with reduced spatial dimensions and increased depth is passed into the residual blocks. Each block has convolutional layers to refine and transform the features (e.g. the texture or style changes). The image details are retained by adding the initial feature map to the residual block's output. Such refined feature map is then passed into an upsampling process, which restores the image dimensions from 32 × 32 to 256 × 256 through three 2D transpose convolutional layers (each with a stride of 2 and a kernel size of 3). Instance normalization, suggested by Ulyanov *et al.* [65], is used to allow the networks to learn domain-specific features and to preserve spatial details in each image. For the discriminator, a PatchGAN network, suggested by Isola *et al.* [66], with a receptive field 70 × 70 is employed. Four convolutional layers are used to extract features of the images, each followed by LeakyReLU as the activation function, except for the last layer. The discriminator evaluates overlapping patches across the whole output image predicted by the generator, and determines scalar values between 0 (fake) and 1 (real).

In the context of CBCT-to-CT translation in image domain, two distinct sets of generators and discriminators are trained. In the forward cycle, the generator (G_{CT}) seeks to achieve the most efficient representation of a CBCT image, and generates a corresponding sCT image. Meanwhile, the discriminator (D_{CT}) differentiates between sCT images, labeled as 0, and true CT images, labeled as 1. In the backward cycle, outputs of the generator (G_{CBCT}) and discriminator (D_{CBCT}) are reversed. The two cycles of training are illustrated in fig. 4.2.

The loss functions for both cycles are explained in the following. In the forward cycle, the generator G_{CT} is trained to transform CBCT images into CT images such that the discriminator D_{CT} cannot distinguish sCT images from CT images using an adversarial loss function:

$$L_{CT} = \mathbb{E}_{CBCT}[\log(1 - D_{CT}(G_{CT}(CBCT)))] + \mathbb{E}_{CT}[\log D_{CT}(CT)], \tag{4.2}$$

where G_{CT} focuses on minimizing the first term $\mathbb{E}_{CBCT}[\log(1 - D_{CT}(G_{CT}(CBCT)))]$ by generating sCT images that are highly comparable to CT images. D_{CT} strives to maximize both terms and improve its ability to differentiate between sCT images and true CT images.

In the backward cycle, G_{CBCT} learns to inversely map the sCT images back to the CBCT images, while D_{CBCT} is trained to distinguish the generated CBCT images and the true CBCT images:

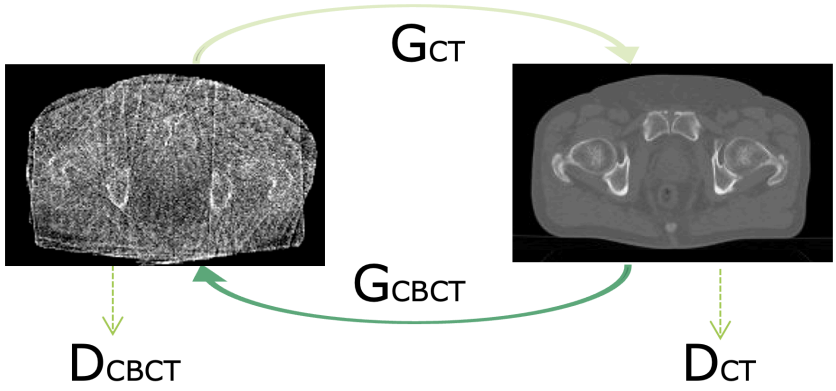


Figure 4.2: An exemplary cycleGAN architecture. G_{CT} and G_{CBCT} denote the generators for mapping CBCT images to CT images and vice versa, respectively. D_{CT} denotes the discriminator for distinguishing sCT images with real CT images. D_{CBCT} denotes the discriminator for distinguishing synthetic CBCT images with real CBCT images respectively. (Own figure)

$$L_{\text{CBCT}} = \mathbb{E}_{\text{CT}}[\log(1 - \text{D}_{\text{CBCT}}(\text{G}_{\text{CBCT}}(\text{CT})))] + \mathbb{E}_{\text{CBCT}}[\log \text{D}_{\text{CBCT}}(\text{CBCT})]. \tag{4.3}$$

Using the adversarial loss formulated in eq. (4.2) and eq. (4.3), the generators G_{CT} and G_{CBCT} learn to generate images of the target domain and deceive the discriminators D_{CT} and D_{CBCT}, respectively.

By chaining the above cycles with a cycle consistency loss, an inverse-consistent mapping between the CBCT and CT domains can be established. During the forward cycle, $L_{\text{cyc}}^{\text{for}}$ calculates the L_1 norm between the output images generated by G_{CBCT} and the input CBCT images:

$$L_{\text{cyc}}^{\text{for}} = \mathbb{E}_{\text{CBCT}}[\|\text{CBCT} - \text{G}_{\text{CBCT}}(\text{G}_{\text{CT}}(\text{CBCT}))\|_1]. \tag{4.4}$$

In the backward cycle, inputs are swapped for the corresponding cycle consistency loss:

$$L_{\text{cyc}}^{\text{back}} = \mathbb{E}_{\text{CT}}[\|\text{CT} - \text{G}_{\text{CT}}(\text{G}_{\text{CBCT}}(\text{CT}))\|_1]. \tag{4.5}$$

Combining the two losses, the loss function in the cycleGAN training is

$$L_{\text{cycleGAN}} = L_{\text{CT}} + L_{\text{CBCT}} + \lambda_1(L_{\text{cyc}}^{\text{for}} + L_{\text{cyc}}^{\text{back}}), \tag{4.6}$$

where λ_1 denotes a weight factor for the cycle consistency loss. Here the objective function can be formulated a an min-max optimization:

$$\mathrm{G_{CT}}, \mathrm{G_{CBCT}} = \arg \min_{\mathrm{G_{CT}}, \mathrm{G_{CBCT}}} \max_{\mathrm{D_{CBCT}}, \mathrm{D_{CT}}} L_{\mathrm{cycleGAN}}(\mathrm{G_{CT}}, \mathrm{G_{CBCT}}, \mathrm{D_{CBCT}}, \mathrm{D_{CT}}), \tag{4.7}$$

The goal of this function is to estimate the model parameters that effectively capture and represent the underlying distribution of the image domains. Instead of a pixel-by-pixel comparison, this approach allows the use of unpaired datasets for training. Once the training is completed, we can apply $\mathrm{G_{CT}}$ to translate CBCT images to CT images.

4.2.2 Related studies

cycleGAN and its variations have been widely used in CBCT-to-CT translation tasks. In the pelvic region, Kurz *et al.* [67] trained a cycleGAN model using unpaired prostate cancer patient datasets of CBCT images and pCT images, and generated sCT images in pCT equivalent image quality with a reduced MAE with respect to corrected CBCT images [49]. Similar training has been done in other pelvic region studies [68, 69]. Kida *et al.* [70] trained a cycleGAN with additional losses, including total variation, gradient loss and idempotent loss (a function to stabilize the mapping by minimizing the difference between $\mathrm{G_{CT}}$(CBCT) and $\mathrm{G_{CT}}(\mathrm{G_{CT}}$(CBCT)), as well as $\mathrm{G_{CBCT}}$(CT) and $\mathrm{G_{CBCT}}(\mathrm{G_{CBCT}}$(CT))). Liu *et al.* [71] suggested a two-step training method with a phantom-based U-Net and a patient-based cycleGAN model to translate CBCT images to CT images. Unpaired training is also applied for other regions, HN [72, 73], lung [73], and breast [73, 74]. Deng *et al.* [75] demonstrated the generalisation of cycleGAN by training a cycleGAN model with HN datasets and improved pelvic CBCT scans.

Apart from unpaired training, cycleGAN has the possibility to be implemented with an additional regularization term to minimize the difference between the sCT images and the corresponding CT images during training. A number of studies trained cycleGAN models using paired datasets (CBCT images as inputs and registered CT images as ground truth) in the pelvic region [76, 77, 78, 79, 80], or for brain or HN [76, 81, 77], pancreatic [82], liver [83], thoracic [77, 84], breast[85], nasopharynx [86, 87] cancers. A few extra losses such as histogram matching loss, gradient loss and perceptual loss have been applied during the training to improve specific anatomy preservation.

Some of the above mentioned studies were compared in a review [88]. Among the studies, Maspero *et al.* [73] reported the highest MAE improvement (more than 100 HU) with one cycleGAN model for unpaired HN, lung and breast cancer datasets. The network was trained with CBCT and CT images which were cropped in a bounding box containing a circular mask. This mask can enforce the network to learn the intensity mapping while occluding the features of specific anatomy.

More recently, a few studies combined a cycleGAN with a transformer in order to constrain the outputs. In Rusanov *et al.* [89], a vision transformer variant ResViT [90] was applied as the generator for a cycleGAN training with paired pelvic datasets. Along with the

modified cycleGAN generator, the study introduced uncertainty estimations (Monte-Carlo Dropout and Test Time Augmentation) and additional losses, including mean structural similarity loss, style loss and conditional L2 loss. Similar addition of a transformer into a cycleGAN was also seen in a recent study [91], which also integrated a vision transformer into a U-Net generator in the cycleGAN training (IViT-CycleGAN). In this study, the transformer incorporated a deep CNN with a self-attention mechanism, and a new gradient penalty [92] was added in the discriminator loss to further improve the stability of the cycleGAN. Moreover, this study trained the two models separately using unpaired HN and chest datasets.

4.3 Contrastive unpaired translation (CUT)

To overcome possible geometric inaccuracies or hallucinations of cycleGAN, a contrastive unpaired translation network (CUT) was developed by the same research team Park *et al.* in 2020 [93] to maximize the mutual information between the inputs and generated images. With the use of a patchwise contrastive loss as discussed below, CUT employs only one set of generator and discriminator. Such one-side translation accelerates the model training and requires fewer computing resources than in cycleGAN, which trains two sets of generator and discriminator.

4.3.1 Network architecture and loss function

To achieve one-side translation training, CUT uses a patchwise contrastive loss L_{PatchNCE} based on noise contrastive estimation (NCE) [94] on image patches instead of L_{cyc}. As illustrated in fig. 4.3, a query patch is sampled from the generator output sCT and compared with the patch at the corresponding location z (denoted as positive z^+) or other patches at different locations (denoted as negatives z^-) of the corresponding CBCT input image. The probability of the positive samples being selected over negatives can be formulated by the following cross-entropy loss:

$$l(\mathbf{v}, \mathbf{v}^+, \mathbf{v}^-) = -\log\left[\frac{e^{\mathbf{v}\cdot\mathbf{v}^+/\tau}}{e^{\mathbf{v}\cdot\mathbf{v}^+/\tau} + \sum_{n=1}^{N} e^{\mathbf{v}\cdot\mathbf{v}_\mathbf{n}^-/\tau}}\right] \tag{4.8}$$

where $\mathbf{v}$, $\mathbf{v}^+$ and $\mathbf{v}^-$ denote the K-dimensional vectors of the query ($\mathbf{v} \in \mathbb{R}^K$), the positive ($\mathbf{v}^+ \in \mathbb{R}^K$), and the N negatives ($\mathbf{v}^- \in \mathbb{R}^{N\times K}$), respectively. These patch vectors ($\mathbf{v}$, $\mathbf{v}^+$ and $\mathbf{v}^-$) are sampled in random locations in the images. The n-th negative is denoted as $\mathbf{v}_\mathbf{n}^- \in \mathbb{R}^K$. For this (N+1) classification problem, τ, which denotes the distances between the query and samples, was set to 0.07. The goal here is to maximize the mutual information between $\mathbf{v}$ and $\mathbf{v}^+$, but to minimize the mutual information between $\mathbf{v}$ and $\mathbf{v}^-$.

The images from the positive and negative samples are passed through the encoder network of the generator (G_{enc}) to obtain embeddings, as illustrated in fig. 4.4. These embeddings are low-dimensional representations of the images that capture their content

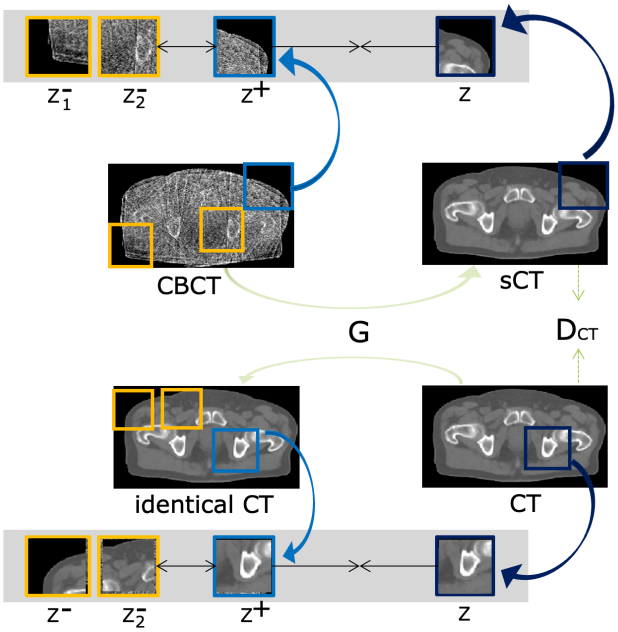


Figure 4.3: An exemplary CUT architecture. A generator (G) is trained to convert a CBCT image into a sCT image. Query patches are sampled from the sCT image and compared with the patches at the corresponding locations z (denoted as positive z^+) or other patches at different locations (denoted as negatives z^-) of the corresponding CBCT input image. Using the same generator (G), an identical CT image is synthesized from a CT image. Similarly, query patches from both images are also sampled and compared. A discriminator ($\mathrm{D_{CT}}$) is used to distinguish sCT images and CT images. (Own figure)

and style information. The layers of interest (L) and the number of spatial locations in each layer (S_l) are selected. The feature maps are processed through a small two-layer multi layer perceptron (MLP) network H_l, producing a stack of features $\{z_l\}_L = \{\mathrm{H}_l(\mathrm{G}^l_{\mathrm{enc}}(\mathrm{CBCT}))\}_L$, where $\mathrm{G}^l_{\mathrm{enc}}$ denotes the output of the l^{th} chosen layer. Likewise, the output image is encoded with the same network into $\{\hat{z}_l\}_L = \{\mathrm{H}_l(\mathrm{G}^l_{\mathrm{enc}}(\mathrm{G}(\mathrm{CBCT})))\}_L$. The other patches within the input can be used as negatives which formulates the following contrastive loss $L_{\mathrm{PatchNCEx}}$.

$$L_{\mathrm{PatchNCEx}}(\mathrm{G}, \mathrm{H}, \mathrm{CBCT}) = \mathbb{E}_{\mathrm{x}\sim\mathrm{CBCT}} \sum_{l=1}^{L} \sum_{s=1}^{S_l} l(\hat{z}_l^s, z_l^s, z_l^{S/s}), \tag{4.9}$$

where $\hat{z}_l^s$, z_l^s and $z_l^{S/s}$ represent the features of the output image, the corresponding feature of the input image ($z_l^s \in \mathbb{R}^{C_l}$) and the negative features ($z_l^{S/s} \in \mathbb{R}^{(S_l-1)\times C_l}$). C_l denotes the number of channels at each layer. Since the generator is trained to identify the similarities across the two domains, the embeddings share the common features. To constrain possible incorrect anatomical changes, CT images are used to generate identical

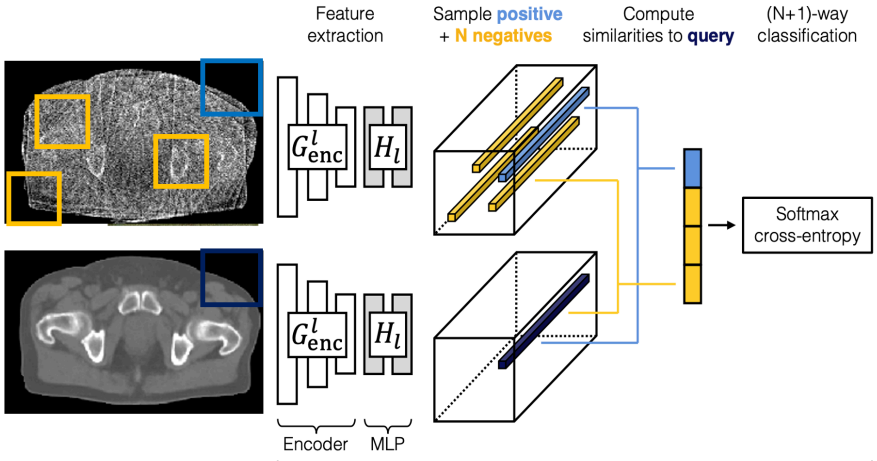


Figure 4.4: The positive and negative samples are passed through the encoder network of the generator (G_{enc}) to obtain embeddings from both the CBCT inputs and sCT images. G^l_{enc} denotes the output of the l^{th} chosen layer and H_l denotes a two-layer multi layer perceptron (MLP) network. (This figure was adapted from the Park *et al.* [93] with a CC BY license.)

CT images using the same generator. The positive and negative samples of these CT images are passed through the same G_{enc} and MLP, which allows us to formulate the following contrastive loss for the CT domain $L_{\text{PatchNCEy}}$:

$$L_{\text{PatchNCEy}}(\text{G}, \text{H}, \text{CT}) = \mathbb{E}_{\text{y}\sim\text{CT}} \sum_{l=1}^{L} \sum_{s=1}^{S_l} l(\hat{z}_l^s, z_l^s, z_l^{S/s}). \tag{4.10}$$

The total loss function is therefore

$$\begin{aligned} L_{\text{CUT}}(\text{G}, \text{D}_{\text{CT}}, \text{CBCT}, \text{CT}) = & L_{\text{adv}}(\text{G}, \text{D}_{\text{CT}}, \text{CBCT}, \text{CT}) + \\ & \lambda_{\text{CBCT}} L_{\text{PatchNCEx}}(\text{G}, \text{H}, \text{CBCT}) + \\ & \lambda_{\text{CT}} L_{\text{PatchNCEy}}(\text{G}, \text{H}, \text{CT})), \end{aligned} \tag{4.11}$$

where λ_{CBCT} and λ_{CT} denote the weights of the patchwise contrastive loss for the CBCT and CT domain, respectively. The main objective for CUT here is to generate realistic CT images, while patches in the input and output images share corresponding information.

4.3.2 Related studies

A few recent studies started to look into the possibility of using CUT in the CBCT-to-CT translation task. In the pelvic region, Wynne *et al.* [95] demonstrated that CUT generates better sCT images than cycleGAN in terms of MAE, structural similarity index measure (SSIM), root mean squared error (RMSE), and Frèchet Inception Distance (FID) (comparing the distribution of the generated and real images in latent space when they reach the deepest layer of the generator). In the HN region, Kang *et al.* [96] demonstrated that a CUT model which was trained with additional losses (including reconstruction, spatially correlative and semantic relation consistency losses) also showed better performance than cycleGAN in terms of FID, peak signal to noise ratio (PSNR), MAE, RMSE. In the lung region, Dong *et al.* [97] trained several CUT models with different hyperparameters using unpaired 4D CBCT and CT datasets and showed better performance than cycleGAN in terms of the image quality metrics.

4.4 CycleGAN and CUT as alternatives to U-Net

As U-Nets are trained on paired datasets, the models learn the direct translation from CBCT to CT images (section 4.1). Compared to the conventional non-AI correction approaches, the computational time using U-Net has been substantially reduced to $\leq$ 5s, which allows online adaptation. However, since the ground truth images are usually generated by DIR, the sCT image might possess a reduced geometrical accuracy due to the uncertainties in DIR. The anatomical shape or location could be adversely affected in the predicted images.

To overcome the potential geometrical inaccuracy in ground truth images, a cycleGAN can be used for unpaired data mapping (section 4.2). With the use of the cycle consistency and the adversarial loss, training cycleGAN does not require ground truth images. As shown from the high HU and treatment dose accuracy in the sCT image generations, cycleGAN can capture complex transformations, however, the output can be geometrically inaccurate. Therefore, some studies implemented cycleGAN with paired datasets. Moreover, since two sets of generators and discriminators have to be trained, the training time is longer with higher computational power demand.

As an alternative, CUT is an one-sided translation with only one set of generator and discriminator (section 4.3), reducing the computational demand and training time by the backward cycle. Moreover, replacing cycle consistency loss with patchwise contrastive loss, CUT constrains the network outputs by learning low-dimensional CBCT and CT features in a shared embedding that could better capture and maximize the mutual information (such as bone and tissue structures) between the CBCT input and the corresponding sCT image. Recent studies demonstrated that CUT achieved better image quality over cycleGAN (section 4.3.2), however, treatment dose calculations using CUT have not yet been investigated (neither photon or proton).

Chapter 5

Deep learning-enabled low dose CBCT-to-CT translation

Full dose CBCT-to-CT translations using DL are well developed, but repeated CBCT scans deliver considerable amounts of radiation dose to patients. What would be the possibilities that DL can bring us in lowering the CBCT imaging dose to the patients?

In this chapter, we briefly discuss the importance of low dose CBCT-to-CT translations in section 5.1, the synthesis of low dose CBCT datasets in section 5.2 and the recent studies in section 5.3.

5.1 The importance of low dose CBCT-to-CT

5.1.1 CBCT imaging dose and secondary cancer risk

According to Ding *et al.* [54], the principle of radiation protection "[as low as reasonably achievable]" (ALARA) for imaging should be pursued in all medical procedures. When the imaging dose is greater than 5% of the treatment dose, treatment planning should include such imaging dose [54]. Among the CBCT imaging dose studies, it is found that daily CBCT scans can lead to additional organ doses [54, 98, 99]. For instance, one CBCT scan can deliver up to 22.7 mSv effective dose in the pelvic region [100]. Depending on the radiosensitivity of each organ and the patient's size as discussed in section 3.5, patients can receive a considerable amount of cumulative dose. Especially for prostate cancer patients, the treatment comprises up to 37 fractions. A recent study also showed that the excess absolute risk (EAR) of secondary cancer incidence grows with higher doses to the organ [101]. Other studies also showed that the imaging radiation can be associated with secondary cancers [102, 103, 104, 105, 106]. Therefore, it is recommended to lower any possible imaging dose in IGRT.

5.1.2 The frequency of accurate patient positioning

To mitigate the risk of secondary cancers by the cumulative CBCT imaging dose, the frequency of CBCT scans is typically limited in clinical practice. This strategy can reduce the cumulative dose, however, it sacrifices the positioning accuracy and decreases OAR sparing. In a multicenter randomized prostate cancer study [107], daily CBCT scans significantly reduced acute rectal bleeding and late rectal toxicity. Moreover, daily CBCT scans improve significantly the biochemical and clinical progression-free intervals. With low dose CBCT, clinicians would have a higher flexibility to prescribe a CBCT scan before the treatment and correct for the position displacements for each fraction which can significantly improve clinical benefits for patients.

5.1.3 Reducing the re-planning CT scan

In current practice, when an in-room CBCT reveals that an adaptation is needed due to anatomical changes, the patient would be prescribed another CT scan for re-planning the treatment. This process requires a substantially greater investment of person-hours, prolongs the treatment duration and delivers additional radiation to the patient. If the treatment planning can be adapted online using low dose CBCT-to-CT translation, the new CT scan can be avoided which further lowers the imaging dose to the patients over the treatment course.

5.2 Low dose CBCT synthesis

5.2.1 Lowering imaging dose

The imaging dose is mainly determined by the dose per projection and the number of projections. For example, the imaging dose to the center of a water cylinder from the exposure measured at the detector can be formulated by Shaw *et al.* [32] as

$$\text{Dose} = N_{\text{projection}}\text{mAs}_{\text{projection}}\left(\frac{X}{\text{mAs}}\right)\left(\frac{\text{SD}}{\text{SC}}\right)^2 e^{\mu_{\text{water}}R} f_{\text{water}}\text{BSF}, \tag{5.1}$$

where $N_{\text{projection}}$ denotes the number of projections, $\text{mAs}_{\text{projection}}$ denotes the product of tube current and time per projection, and $(\frac{X}{\text{mAs}})$ is the exposure per unit $\text{mAs}_{\text{projection}}$, which is measured in air at the detector. μ_{water} and R denotes the attenuation coefficient and the radius of the water cylinder, respectively. $(\frac{\text{SD}}{\text{SC}})^2$ geometrically scales to the center by the inverse square law, and $e^{\mu_{\text{water}}R}$ scales to the center of the water cylinder. f_{water} denotes a factor to convert the exposure to dose, and BSF is a scaling factor for the increased dose due to back scatter inside the water cylinder. It is noted that the exponent of $e^{\mu_{\text{water}}R}$ is positive, as the measurement of the dose here is in the center of the water cylinder but not on the detector. The followings can explain this scale factor of the beam intensity $(e^{+\mu_{\text{water}}R})$ further.

Suppose an x-ray beam has an intensity I_0 at the source, I_{center} and I_{detector} is the beam intensity at the center of the water cylinder and the detector, respectively. According to the Beer-Lambert law, I_{center} can be derived as

$$\begin{aligned} I_{\text{center}} &= I_0 e^{-\mu_{\text{water}} R} \\ I_{\text{detector}} &= I_0 e^{-\mu_{\text{water}} 2R} \\ I_{\text{center}} &= I_{\text{detector}} e^{+\mu_{\text{water}} R}. \end{aligned} \tag{5.2}$$

During a CBCT scan, the gantry rotates around the patients by 360° and projections are collected. If the number of projection frames during the rotation is reduced by a factor of x, the radiation dose is reduced by the same factor x. Moreover, it is also possible to lower the imaging dose by reducing the mAs per projection.

5.2.2 Poisson noise and streaking artefacts

In the context of the effect of imaging dose on image quality, Poisson noise occurs due to the variations of detected x-ray photons. In the Poisson distribution, suppose N_{photons} is the number of photons detected in the panel, or signal, the variance (σ^2) of the N_{photons} is equal to its mean (μ), formulating the following equation:

$$\sigma^2 = \mu = N_{\text{photons}}. \tag{5.3}$$

That gives a characterisation of Poisson noise (σ) as

$$\sigma = \sqrt{N_{\text{photons}}}. \tag{5.4}$$

With this above equation, we can establish a signal and noise relationship by signal-to-noise ratio (SNR):

$$\text{SNR} = \frac{N_{\text{photons}}}{\sigma} = \frac{N_{\text{photons}}}{\sqrt{N_{\text{photons}}}} = \sqrt{N_{\text{photons}}}. \tag{5.5}$$

With lower number of photons, both imaging dose and SNR decrease. The Poisson noise exists in the projection image and can propagate in the CBCT image reconstruction. During FBP which amplifies high frequency components (due to the ramp filter as shown in eq. (3.9)), the Poisson noise in the projection image can be amplified and decreases the image quality of the resulting CBCT image.

In this thesis, the imaging dose was lowered by reducing the number of projections. The $\text{mAs}_{\text{projection}}$ for each projection remained unchanged. To illustrate the relationship between the variance (σ_f^2) at the FBP reconstructed image $f(x, y)$ and the total number of

projections P, we could consider a reconstruction chain using FBP in a 2D discrete form. The FBP reconstructed image can be defined as

$$f(x,y) = \frac{\pi}{P}\sum_{p=1}^{P} h_{\beta_p}(x\cos\beta_p + y\sin\beta_p), \tag{5.6}$$

where β_p and h_{β_p} denote the rotational angle and the filtered projection, respectively. According to [52], the variance of the central pixel of a reconstructed image at $(x,y) =$ (0,0) could be derived as

$$\begin{aligned}\sigma_f^2(0,0) &= \left(\frac{\pi}{P}\right)^2 \sum_{p=1}^{P} \sigma_{h_{\beta_p}}^2(0) \\ &= \frac{\pi^2}{12(\Delta\xi)^2 P N_{\text{photon}}(0)} \\ &\propto \frac{1}{P},\end{aligned} \tag{5.7}$$

where $\sigma_{h_{\beta_p}}^2(0)$ is the variance of the central filtered projection elements, $\Delta\xi$ denotes the sampling parameter (i.e. the width of one detector element) and $N_{\text{photon}}(0)$ is the average number of x-ray beam photons measured in one detector element. With the above formulation, the noise in the reconstructed image σ_f increases by factor of 2 when the total number of projections (P) decreases by a factor of 4.

The under-sampling of projections in this thesis also resulted in additional streak artefacts due to the Nyquist sampling theorem. Figure 5.1 demonstrates the effect on the image quality of a lower dose in CBCT FDK reconstruction images by under-sampling of projections. The streak artefacts increased when a lower number of projections was used for reconstruction.

It is also possible to estimate the dose reduction using the CBDI estimation of the protocol that was used for this thesis (M20 protocol with 120 kV and a bowtie filter at an Elekta XVI scanner) as discussed in (section 3.5). By under-sampling the projections frames, for instance using 90 of the 350 projection frames, the estimated CBDI was reduced from 2.27 mGy to 0.57 mGy (from an overall exposure of 140 mAs to 36 mAs) per CBCT scan.

5.3 Related studies using U-Net, cycleGAN and CUT

Implementing low dose CBCT-to-CT translation in the adaptation workflow would require the DL model to simultaneously remove streaking artefacts and correct HU intensity. Currently, a limited number of studies investigated sCT image generation from low dose CBCT images. In the HN region, Yuan *et al.* [108] acquired low dose CBCT with 182 frames for a 205° rotation (a total exposure of 18.2 mAs), and trained a U-Net with a CT ground

truth to translate a low dose CBCT to sCT images. The sCT achieved an image quality comparable to other full dose supervised CBCT-to-CT studies in pelvis scans [59, 56, 60] or an unsupervised study in the HN region [72]. In the thoracic region, Gao *et al.* [109] under-sampled a clinical chest CBCT scan by 50% (660 projections to 330 projections) and trained an attention-guided cycleGAN with unpaired CT datasets to generate sCT images. Dong *et al.* [97] investigated lung 4D CBCT images, which has a low number of projections in each respiratory phase (50 projections per test patient) and has similar image quality (streaks and noise) as 15% low dose CBCT in fig. 5.1, and generated sCT images using cycleGAN and CUT models.

Although the above studies demonstrated that sCT images from a low dose CBCT can achieve good HU and VMAT dose calculation accuracy, it remains unclear how much we can reduce the CBCT dose for actual clinical use.

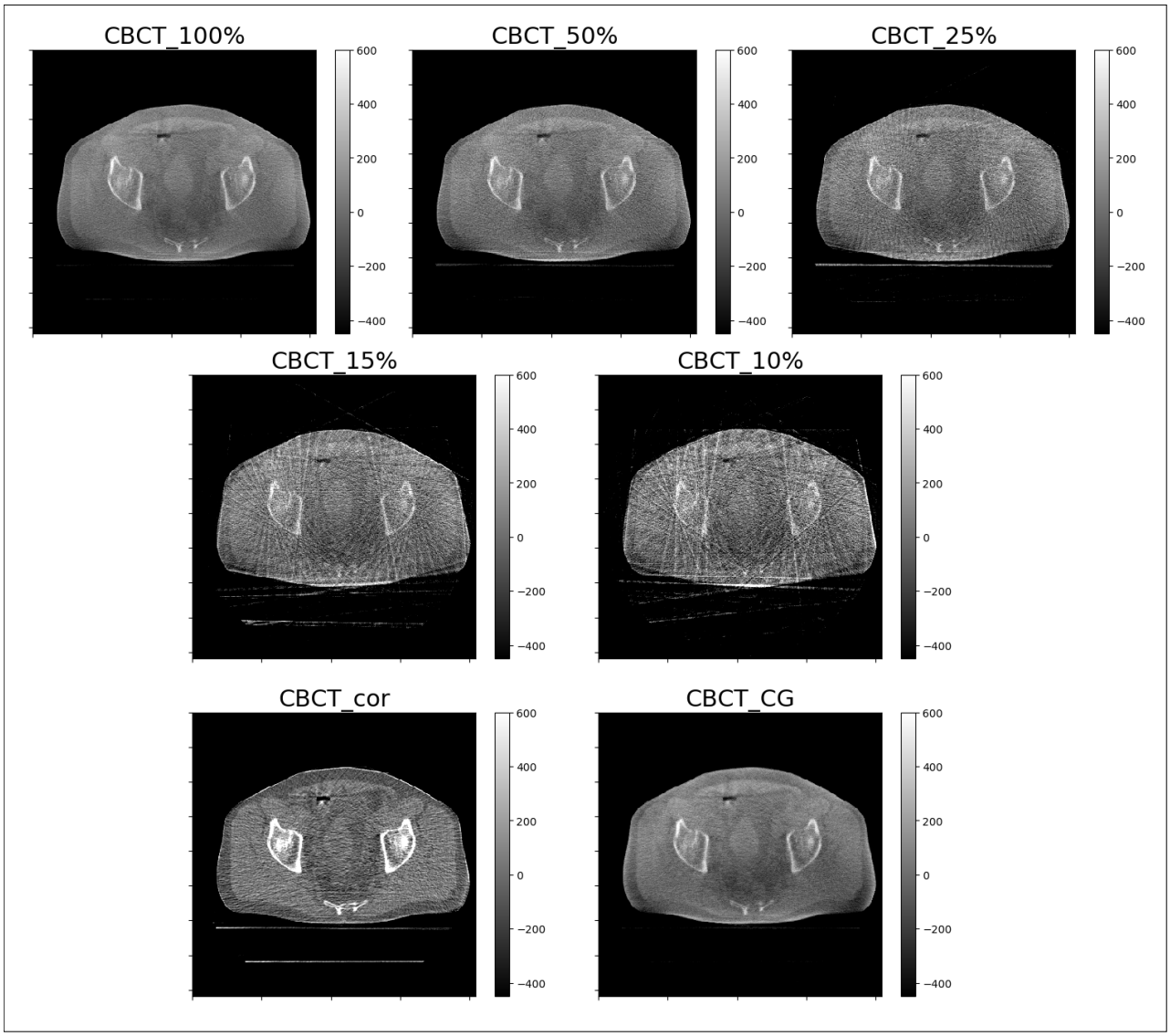


Figure 5.1: Axial view of CBCT images using FDK reconstruction (section 3.2.2) with 100% (352 projections), 50%, 25%, 15%, 10% projections, a scatter-corrected CBCT reconstruction (section 3.4), and an iterative reconstruction using conjugate gradient (CG) with 100% projections of an exemplary prostate cancer patient dataset. All images were reconstructed with 410 × 410 × 264 voxels on an isotropic 1.0 mm^3 grid. All intensities are in HU. (Own figure)

Chapter 6

Metrics for sCT evaluations

One of the potential problems in AI generated images is that the generator can fail to capture and preserve all relevant anatomical details, especially in the pelvic region where organ shapes can vary substantially. As a result, the generated images are prone to losing anatomical details. Therefore, evaluating AI generated sCT images is important to ensure that they are accurate, safe, and effective for use in the clinical practice. A few metrics are typically used to evaluate sCT images and the model performances, including image quality (section 6.1), treatment dose accuracy (section 6.2) and positioning accuracy (section 6.3). However, geometrical accuracy which is crucial for contouring and treatment planning has not been commonly assessed. In section 6.4, the importance and a method to evaluate geometrical accuracy is discussed.

6.1 Image Quality

To evaluate image quality of sCT images, the voxel intensities can be compared to reference images, vCT images or CBCT_{cor} (section 3.4) images. Since these images are supposed to possess identical anatomical structures and corrected HU values, the disparity reflects the HU accuracy of the sCT images. To determine the HU accuracy of anatomical structures, voxels outside the joint body outline of reference and sCT images are excluded. Suppose n_{body} denotes the number of voxel inside the joint body outline, y and $\hat{y}$ denote the voxel intensities of the sCT and of the reference images, respectively. The following metrics are typically used:

- Mean absolute error (MAE): it is a measurement of the mean magnitude of errors between the sCT and the reference images. It can be formulated as

$$\text{MAE} = \frac{1}{n_{\text{body}}} \sum_{i=1}^{n_{\text{body}}} |y_i - \hat{y}_i|. \tag{6.1}$$

- Mean error (ME): Similar to MAE, it is a measurement of the mean values of errors between the sCT and the reference images. It can be formulated as

$$\mathrm{ME} = \frac{1}{n_{\mathrm{body}}} \sum_{i=1}^{n_{\mathrm{body}}} y_i - \hat{y}_i. \tag{6.2}$$

- Structural similarity index measure (SSIM): Compared to MAE or ME, SSIM is to compare the texture and the structural similarity between two images. The SSIM can be formulated as in [110]:

$$\mathrm{SSIM} = \frac{(2\mu_y\mu_{\hat{y}} + C_1)(2\sigma_{y\hat{y}} + C_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + C_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2)}, \tag{6.3}$$

  where μ_y and $\mu_{\hat{y}}$ denotes the mean intensities of the sCT and the reference images, σ_y^2 and $\sigma_{\hat{y}}$ denotes the variances of the sCT and the reference images. C_1 and C_2 are the constants to stabilize the equation when $\mu_y + \mu_{\hat{y}}$ is near zero. The definitions of C_1 and C_2 are $C_1 = (K_1R)^2$ and $C_2 = (K_2R)^2$, where R is the voxel intensity range. K_1 and K_2 are constants that are empirically set as 0.01 and 0.03 according to [110], respectively.

- Peak signal-to-noise ratio (PSNR): It is a measurement of the maximum signal relative to the maximum background noise. PSNR is represented as a logarithmic value on the decibel (dB) scale via mean square error (MSE) of the voxels values between the sCT and the reference images. Suppose n_{body} denotes the number of voxel inside the joint body outline, R denotes the voxel intensity range, y and $\hat{y}$ denote the voxel intensities of the sCT and of the reference images, respectively. PSNR can be formulated as in [111]:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{R^2}{\mathrm{MSE}}\right), \tag{6.4}$$

  where MSE is defined as

$$\mathrm{MSE} = \frac{1}{n_{\mathrm{body}}} \sum_{i=1}^{n_{\mathrm{body}}} (y_i - \hat{y}_i)^2. \tag{6.5}$$

6.2 Treatment dose accuracy

Evaluating treatment dose accuracy is important to verify the electron densities on sCT images. Similar to the downstream tasks in treatment planning as discussed in section 2.1.4, the treatment dose accuracy of sCT images is evaluated by comparing to the reference images ($\mathrm{CBCT}_{\mathrm{cor}}$ or vCT) using the research version of TPS (RayStation, version 10.01, RaySearch, Sweden). In this thesis, we evaluated the accuracy with a photon-based radiation therapy technique VMAT, which is frequently used in the hospitals.

In details, contours of target and OAR structures were propagated from the pCT images to the reference and sCT images via DIR. On reference images, VMAT plans were generated

and optimized on a 3.0 mm isotropic dose grid. A collapsed-cone dose engine was employed. The plans were subsequently recomputed on the sCT images.

For prostate cancer patient data used in this thesis, the prescription was 2 Gy in 37 fractions. A CTV $V_{95\%}$ of 100% was targeted, and PTV should receive at least 95% of the prescription dose for more than 95% of its volume. The DVH constraints for OARs such as bladder and rectum were adhered to the recommendation in the QUANTEC report [17]. The VMAT dose distributions on sCT images were evaluated against the reference images, with consideration given to DVH parameters. For target structures, CTV and PTV $D_{98\%}$ and $D_{2\%}$, as well as PTV $D_{50\%}$ and $V_{95\%}$ were compared. For OARs, bladder (V_{60Gy} and V_{65Gy}) and rectum (V_{50Gy} and V_{60Gy} and V_{65Gy}) were used for the comparison. In addition, the voxels meeting the treatment dose difference (DD) analysis criteria of 1% and 2% (with a 10% threshold) were compared.

For each of the above dose parameters, values from sCT images were statistically compared to the reference images using Wilcoxon signed-rank tests. A significant difference could be observed when a p-value was less than 0.05.

6.3 Positioning accuracy

One of the main objectives of in-room CBCT imaging is daily patient positioning. To assess positioning accuracy when using sCT images, these images can be rigidly registered to the pCT images with the use of the research TPS, similar to the procedure as shown in the fig. 2.3. In this thesis, sCT images were registered to pCT images (sCT-to-pCT), which yielded a set of rigid transformation parameters (in terms of millimeters (mm): inferior–superior (IS), right–left (RL), posterior–anterior (PA); in terms of degree (°): pitch, roll and yaw).

The resulting transformations were evaluated against those obtained by registering the original full dose CBCT images ($CBCT_{org}$) to the pCT images (($CBCT_{org}$-to-pCT)). The mean absolute differences of each transformation parameter between sCT-to-pCT and $CBCT_{org}$-to-pCT were computed to reflect the positioning accuracy of sCT images.

6.4 Geometrical accuracy

As mentioned in the beginning of this section, AI generated images can create hallucinations in anatomy and lower the precision of the treatment planning. Evaluating geometrical accuracy can reveal how well the sCT images replicate the true spatial relationships and structures found in reference images. However, geometrical accuracy is often missing in many studies.

One of the methods to quantify the anatomical fidelity is to segment the organs in both sCT images and reference images, either manually or with automatic algorithms, and to compare the contours of organs. In this thesis, the following metrics were used:

- Dice similarity coefficient (DSC): It measures the coincidence between two sets of

binary contours, with a score that ranges from 0 (no spatial overlap) to 1 (complete overlapping). The formulation can be written as

$$\frac{2|a \cap \hat{a}|}{|a| + |\hat{a}|}, \tag{6.6}$$

where a and $\hat{a}$ denotes the set of voxels in the contour of the organ on sCT images and the reference images, respectively. $|a|$ and $|\hat{a}|$ are the number of voxels in contours on sCT images and reference images respectively. $|a \cap \hat{a}|$ denotes the number of overlapping voxels between contours. Since DSC mainly measures the proportion of overlapping region, it does not fully capture the difference in boundary and shape.

- Hausdorff distance (HD): It measures the spatial discrepancy between two sets of segmentations by calculating the largest distance from one point on sCT contour to the nearest point on the reference contour. In this thesis, we used boundary HD [112] in order to evaluate the contour shapes. Suppose the contour of an organ on sCT images has a set of points $\partial A = a_1, ..., a_{N_a}$ and the contour on the reference images has another set of points $\partial \hat{A} = \hat{a_1}, ..., \hat{a}_{N_{\hat{a}}}$, the average distance between boundaries of the contours $d_{HD_{avg}}(\partial A, \partial \hat{A})$ can be formulated as

$$\mathrm{HD}_{\mathrm{avg}}(\partial A, \partial \hat{A}) = \frac{\vec{d}_{\mathrm{HD}_{\mathrm{avg}}}(\partial A, \partial \hat{A}) + \vec{d}_{\mathrm{HD}_{\mathrm{avg}}}(\partial \hat{A}, \partial A)}{2}, \tag{6.7}$$

where $\vec{d}_{\mathrm{HD}_{\mathrm{avg}}}(\partial A, \partial \hat{A})$ and $\vec{d}_{\mathrm{HD}_{\mathrm{avg}}}(\partial \hat{A}, \partial A)$ denote the directed average Hausdorff measures from ∂A to $\partial \hat{A}$, and from $\partial \hat{A}$ to ∂A, respectively. In addition to $\mathrm{HD}_{\mathrm{avg}}$, the 95^{th} percent ranked distance HD_{95} was also calculated to report the contour differences in this thesis.

6.5 The overview of sCT evaluations

The lack of clinical evaluations of DL generated sCT images has been hindering the actual implementation in the ART workflow. Most of the CBCT-to-CT works have evaluated their sCT images in terms of image quality. Some U-Net-generated and cycleGAN-generated sCT images have been evaluated with treatment dose calculations, however, not yet CUT-generated sCT images. Moreover, only a limited number of studies have evaluated the anatomical fidelity of sCT images which can have high influence on the downstream tasks such as the organs contouring and the treatment planning. Patient positioning accuracy was neither consistently evaluated in many studies.

Chapter 7

Contributions of this thesis

During a course of radiotherapy treatment, daily CBCT scans are important for accurate treatment radiation delivery and sparing OARs (section 5.1.2). However, repeated full dose CBCT scans also expose the patients to a considerable level of extra radiation dose, leading to additional secondary cancer risk. As discussed in chapter 4, DL algorithms in full dose CBCT-to-CT translations have been comprehensively explored (section 4.1, section 4.2, section 4.3) and have recently been adopted in the clinical ART workflow (section 2.3.2). In a recent multi-center sCT challenge "synthRad 2023" [113], the best three models applied variants of U-Net and transformers to generate sCT images from full dose CBCT images and achieved good image quality and treatment dose accuracy. Adhering to the principle of radiation protection "[as low as reasonably achievable]" (ALARA), it is time to explore how we can leverage DL advancements for reducing the imaging dose to patients.

Several separate low dose CBCT-to-CT studies have been conducted within the last four years, as shown in section 5.3. Among these studies, CBCT scans of limited anatomical locations have been investigated, limited to HN or chest regions. As revealed by a number of clinical studies mentioned in section 5.1, prostate cancer patients, who can benefit from daily CBCT scans, also have higher EAR of secondary cancer incidence due to the daily CBCT scans. Therefore, in our first step in DL-enabled low imaging dose CBCT-to-CT, we have selected prostate cancer patient datasets from the LMU University Hospital (section 7.1). In this first study, we investigated the feasibility of substantially lowering the imaging dose of CBCT in the pelvic region with only 25% projections (approximately 90 projections), which could reduce the CBCT imaging dose more than other studies as listed in section 5.3. For example, the number of projections of our low dose CBCT is almost half of the projection frames used in [108] or 70% fewer than in [109]. To enhance such low dose CBCT images into diagnostic quality sCT images, a state-of-the-art DL style transfer model cycleGAN (section 4.2) and a variant model with a residual connection were implemented. The first paper is introduced in the following section 7.2.1.

Among all low dose CBCT-to-CT research works (section 5.3), however, there is a lack of structured investigations of the maximum imaging dose reduction level that is achievable by DL. The majority of the studies have generated sCT images from CBCT scans with only one specific radiation dose. With good full dose sCT image quality, we expect that DL

could further reduce the imaging dose of CBCT scans. To fully unlock DL's potential, in the second study, we have investigated various imaging dose levels in terms of sCT image generation with lower number of CBCT projections (25%, 15% and 10%). The aim was to find the achievable lowest imaging dose that allows accurate treatment adaption using two state-of-the-art DL style transfer models, namely cycleGAN (section 4.2) and CUT (section 4.3). The second paper is introduced in the following section 7.2.2.

To address the lack of clinical evaluations on sCT images (section 6.5), both studies in this thesis evaluated sCT comprehensively with all CBCT-guided adaptive radiotherapy related metrics including image quality (section 6.1), VMAT treatment dose accuracy (section 6.2), positioning accuracy (section 6.3) and organ contouring accuracy (section 6.4). These evaluations aimed to enhance the clinical impact and a possible integration of the low dose CBCT technique in the hospital setting.

7.1 Patient database

CBCT and pCT image datasets of 41 prostate cancer patients who were prescribed with VMAT treatment to an overall dose of 70 Gy–76 Gy in 2 Gy fractions at the Department of Radiation Oncology, LMU Munich University Hospital were retrospectively included. All datasets were previously collected and extracted by Kurz *et al.* [67].

7.1.1 CBCT

For CBCT patient datasets, CBCT images were acquired by the lowest radiation dose protocol for the pelvic region from the hospital. These CBCT images were acquired in the treatment position using the XVI system (version 5.52) of a Synergy medical linear accelerator (Elekta, Sweden), as shown in fig. 1.3. The main selection criteria was: 120 kV tube voltage, 20 ms exposure time, 20 mA X-ray tube current per projection. The detailed scan parameters are in the following table 7.1:

As illustrated in fig. 7.1, panel saturation from high intensities through thin patient sections can cause underestimation of the body outline. This protocol (20 mA and 20 ms) was chosen to have the patient body outline more preserved. During the CBCT acquisition, a laterally-shifted detector panel in M position and a bowtie filter were applied to enlarge the lateral FOV.

The raw data in projections files (.his) and scan files (.ini) were previously collected and extracted in [67]. Reconstructed images with lateral FOV truncation despite the enlarged FOV were not included.

7.1.2 CT

The pCT images were acquired with a Toshiba Acquilion LB CT scanner (Canon Medical Systems, Japan). No contrast agent was used. The scanner reconstructed images in

Table 7.1: A exemplary CBCT scanning protocol that were selected for this study. The number of projections can vary among patients, ranging from 346 to 357 projections.

Scan parameter	**Value**
Tube Voltage	120 kV
Tube mA	20
Tube KV Length	20
KV Filter	F1 (bow tie filter)
Collimator Name	M20
FOV	Medium
FloodImageFilterNorm	29504
FloodImageFilterMA	20
FloodImageFilterMS	20
FloodImageOpenNorm	37732
FloodImageOpenMA	16
FloodImageOpenMS	20
Number of Projections	357
Scan rotation	360°
Patient Position	Head First-Supine (HFS)
Reconstruction Dimension (X, Y, Z)	410 pixels, 410 pixels, 264 pixels
Pixel Spacing (X, Y, Z)	1.0 mm, 1.0 mm, 1.0 mm

DICOM format were previously collected and extracted from the clinical TPS Monaco by Kurz *et al.* [67]. The detailed scan parameters are in the following table 7.2:

7.2 Introduction to the two studies

7.2.1 Study 1: Feasibility of cycleGAN enhanced low dose CBCT imaging for prostate radiotherapy dose calculation

In the first study, we investigated the feasibility of low imaging dose CBCT with only 25% projections by simultaneously removing under-sampling artefacts and correcting image intensities with two cycleGAN models (original cycleGAN implementation and with a generator residual connection (see figure 3 in the first publication)).

In the first step, the full dose CBCT images ($CBCT_{org}$) and low dose CBCT images were reconstructed using the FDK algorithm. We adopted a cycleGAN algorithm and implemented the variants with a patient body shape loss term (see section 2.2.1 and equation 5 in the first publication for details) to train the model using unpaired 4-fold cross-validation (33 patients). For evaluation references, we generated vCT images on 8 test patients. VMAT plans were optimized on vCT images, and recalculated on sCT images. To evaluate the positioning accuracy, we calculated residual shifts by registering sCT images and $CBCT_{org}$ images to pCT images, respectively. For anatomical fidelity, we manually con-

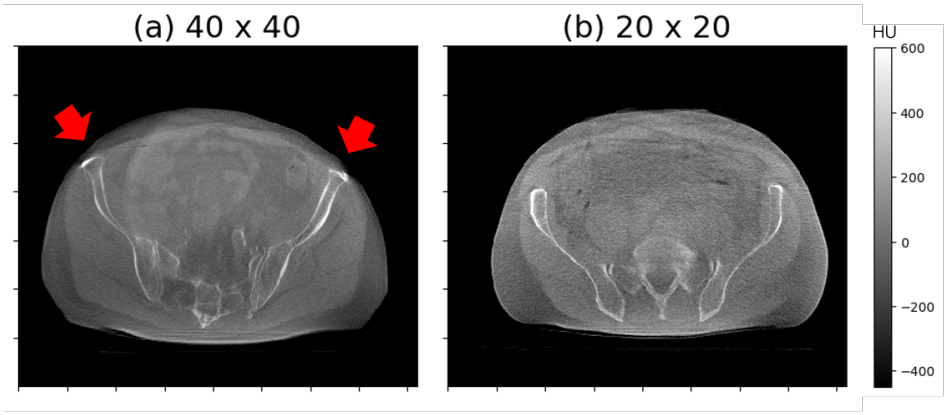


Figure 7.1: A comparison between scanning parameters in two representative prostate cancer patients CBCT scanner images. The images were acquired with (a) 40 ms exposure time, 40 mA per projection, 370 projections, (b) 20 ms exposure time, 20 mA per projection, 351 projections. Red arrows show the body outline artefacts in (a). Both were acquired in the hospital but not used in the patient cohort in this thesis. (Own figure)

toured bladder and rectum in sCT images and original CBCT images and compared in terms of the geometrical metrics (i.e. DSC and HD_{avg}, HD_{95}).

In the test patient datasets, the average MAE (section 6.1) was lowered from 126 HU to 44 HU compared to the vCT images. High treatment dose accuracy (section 6.2) was found, with 2% dose difference pass rates of 99% (10% dose threshold). Positioning accuracy (section 6.3) was high, as most of mean absolute differences of rigid transformation parameters (sCT-to-pCT − $CBCT_{org}$-to-pCT) were less than 0.20 mm/0.20 °. Contouring accuracy (section 6.4) was high with DSC/HD_{avg}/HD_{95} equal to 0.9/0.9 mm/4.1 mm for bladder and for 0.9/1.1 mm/3.9 mm for rectum. The computational time to generate sCT images was approximately 2 seconds for each volumetric CBCT scan.

This work demonstrated the feasibility of adapting two cycleGAN models to simultane-

Table 7.2: An exemplary CT scanning protocol that were selected for this study.

Scan parameter	**Value**
Tube Voltage	120 kV
Exposure Time	750 ms
Tube Current	234 mA
Exposure	175 Ckg^{-1}
Data Collection Diameter	550 mm
Scan Options	Helical CT
Patient Position	Head First-Supine (HFS)
Reconstruction Dimension (X, Y, Z)	512 pixels, 512 pixels, 264 pixels
Pixel Spacing (X, Y, Z)	1.074 mm, 1.074 mm, 3.074 mm

ously clear under-sampling artefacts and correct HU intensities of 25% dose pelvic CBCT images.

7.2.2 Study 2: Investigation of imaging dose reduction levels in pelvic CBCT-based synthetic CT generation using deep learning: dose calculation and segmentation accuracy

In the second study, three imaging dose levels of CBCT scans were generated and image corrections using cycleGAN and CUT models were investigated. The objective of this work is to identify the minimum achievable imaging dose for accurate VMAT dose calculation and organ delineation in online adaptation. A visual summary of the paper can be found in fig. 7.2.

CBCT images of the prostate cancer patient datasets were reconstructed with lower number of projections (100%, 25%, 15% and 10%) using the FDK algorithm. The CBCT-to-CT models were trained with a similar scheme as in the first study. For each low dose levels, we optimized and compared the performance of a cycleGAN model with residual connection and a CUT model to generate sCT images from reduced imaging dose CBCTs. Comprehensive details of the hyper-parameters and the network implementations can be found in the supplementary material of this publication. CT number, treatment dose, positioning accuracy and anatomical fidelity were evaluated in comparison to the reference $CBCT_{cor}$ images (described in section 3.4). Based on contouring accuracy, 25% is the minimum CBCT imaging dose.

In addition, this paper was selected as one of the Physics Highlights in the conference ESTRO 2024: https://www.phiro.science/highlights2024.

7.3 First author's contribution

The 41 CBCT and CT patient datasets were previously selected and extracted by Kurz *et al.* [67]. I implemented an anonymization Matlab script to remove patient identifications from the CBCT datasets (.INI and .XML files) and CT datasets (DICOM files). I under-sampled and reconstructed different imaging dose level CBCT images using FDK reconstruction algorithms [114], implemented image pre-processing Python scripts (including couch removal, re-sampling and intensity normalization) for CBCT and CT images, modified the cycleGAN algorithm with the residual connection, implemented and dockerized the cycleGAN and CUT algorithms on the research server for training with GPUs. I trained and optimised the models to find the best hyper-parameters. In addition, I implemented Python scripts to evaluate the image quality, generated VMAT plans in the TPS, evaluated the treatment dose with DVH parameters and the positioning accuracy. Moreover, I received contouring training from a radiation oncologist and contoured the OARs of the testing patient datasets which were then validated by the radiation oncologist. I implemented the contouring analysis script and wrote the original manuscripts for the two papers. Finally, for the studies I have delivered oral presentations in international

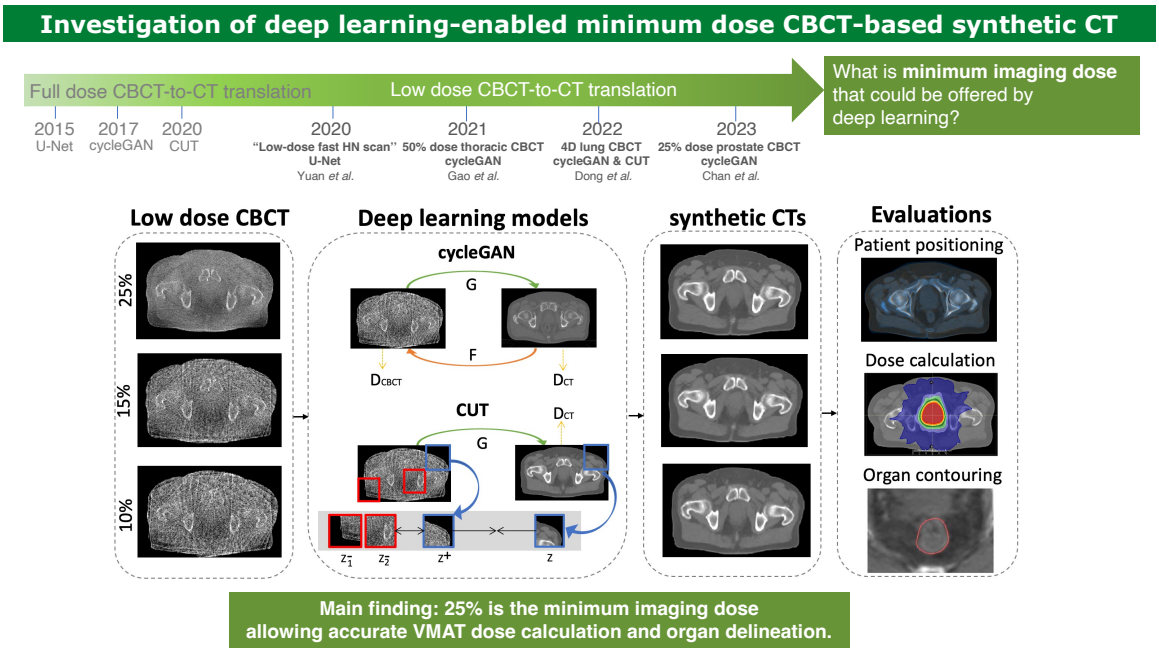


Figure 7.2: A visual summary of the second study to investigate the minimum CBCT imaging dose for accurate treatment dose calculation and OAR contouring. Copyright: © 2024 Published by Elsevier B.V. on behalf of European Society of Radiotherapy & Oncology under CC BY 4.0 license.

conferences, poster presentations as well as data science slams in joint graduate training schools.

Chapter 8

Studies

8.1 Paper 1: Feasibility of CycleGAN enhanced low dose CBCT imaging for prostate radiotherapy dose calculation

Reprinted with permission from "Feasibility of CycleGAN enhanced low dose CBCT imaging for prostate radiotherapy dose calculation." by Yan Chi Ivy Chan, Minglun Li, Katia Parodi, Claus Belka, Guillaume Landry and Christopher Kurz; Physics in Medicine & Biology. 2023 May 11;68(10):105014.

IOP Publishing Phys. Med. Biol. 68 (2023) 105014 https://doi.org/10.1088/1361-6560/acccce

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER

Feasibility of CycleGAN enhanced low dose CBCT imaging for prostate radiotherapy dose calculation

Y Chan[1], M Li[1], K Parodi[2], C Belka[1,3], G Landry[1,4] and C Kurz[1,4]

[1] Department of Radiation Oncology, University Hospital, LMU Munich, D-81377 Munich, Germany
[2] Department of Medical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München (LMU Munich), D-85748 Garching, Germany
[3] German Cancer Consortium (DKTK), D-81377 Munich, Germany
[4] Senior authorship is shared equally.

E-mail: Yanchi.Chan@med.uni-muenchen.de

Keywords: Low dose CBCT, cycleGAN, adaptive therapy

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
7 December 2022

REVISED
27 March 2023

ACCEPTED FOR PUBLICATION
13 April 2023

PUBLISHED
11 May 2023

Abstract

Daily cone beam computed tomography (CBCT) imaging during the course of fractionated radiotherapy treatment can enable online adaptive radiotherapy but also expose patients to a non-negligible amount of radiation dose. This work investigates the feasibility of low dose CBCT imaging capable of enabling accurate prostate radiotherapy dose calculation with only 25% projections by overcoming under-sampling artifacts and correcting CT numbers by employing cycle-consistent generative adversarial networks (cycleGAN). Uncorrected CBCTs of 41 prostate cancer patients, acquired with ~350 projections ($CBCT_{org}$), were retrospectively under-sampled to 25% dose images ($CBCT_{LD}$) with only ~90 projections and reconstructed using Feldkamp–Davis–Kress. We adapted a cycleGAN including shape loss to translate $CBCT_{LD}$ into planning CT (pCT) equivalent images ($CBCT_{LD\_GAN}$). An alternative cycleGAN with a generator residual connection was implemented to improve anatomical fidelity ($CBCT_{LD\_ResGAN}$). Unpaired 4-fold cross-validation (33 patients) was performed to allow using the median of 4 models as output. Deformable image registration was used to generate virtual CTs (vCT) for Hounsfield units (HU) accuracy evaluation on 8 additional test patients. Volumetric modulated arc therapy plans were optimized on vCT, and recalculated on $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ to determine dose calculation accuracy. $CBCT_{LD\_GAN}$, $CBCT_{LD\_ResGAN}$ and $CBCT_{org}$ were registered to pCT and residual shifts were analyzed. Bladder and rectum were manually contoured on $CBCT_{LD\_GAN}$, $CBCT_{LD\_ResGAN}$ and $CBCT_{org}$ and compared in terms of Dice similarity coefficient (DSC), average and 95th percentile Hausdorff distance (HD_{avg}, HD_{95}). The mean absolute error decreased from 126 HU for $CBCT_{LD}$ to 55 HU for $CBCT_{LD\_GAN}$ and 44 HU for $CBCT_{LD\_ResGAN}$. For PTV, the median differences of $D_{98\%}$, $D_{50\%}$ and $D_{2\%}$ comparing both $CBCT_{LD\_GAN}$ to vCT were 0.3%, 0.3%, 0.3%, and comparing $CBCT_{LD\_ResGAN}$ to vCT were 0.4%, 0.3% and 0.4%. Dose accuracy was high with both 2% dose difference pass rates of 99% (10% dose threshold). Compared to the $CBCT_{org}$-to-pCT registration, the majority of mean absolute differences of rigid transformation parameters were less than 0.20 mm/0.20°. For bladder and rectum, the DSC were 0.88 and 0.77 for $CBCT_{LD\_GAN}$ and 0.92 and 0.87 for $CBCT_{LD\_ResGAN}$ compared to $CBCT_{org}$, and HD_{avg} were 1.34 mm and 1.93 mm for $CBCT_{LD\_GAN}$, and 0.90 mm and 1.05 mm for $CBCT_{LD\_ResGAN}$. The computational time was ~2 s per patient. This study investigated the feasibility of adapting two cycleGAN models to simultaneously remove under-sampling artifacts and correct image intensities of 25% dose CBCT images. High accuracy on dose calculation, HU and patient alignment were achieved. $CBCT_{LD\_ResGAN}$ achieved better anatomical fidelity.

1. Introduction

In modern image-guided radiotherapy (IGRT), cone beam computed tomography (CBCT) is used as a routine in-room imaging technique. Most radiotherapy centers have medical linear accelerators equipped with a kilovoltage CBCT (kV-CBCT) scanner, which provides full three-dimensional (3D) information about the patient's anatomy at every treatment fraction. In the presence of inter-fractional anatomical changes between acquisition of the planning CT (pCT) and the treatment day, CBCT imaging data would be suitable for treatment adaptation and enabling accurate dose delivery (de Jong *et al* 2021, Moazzezi *et al* 2021, Sibolt *et al* 2021, Byrne *et al* 2022).

One primary problem which arises in using CBCT for treatment adaptation is that CBCT image quality is typically insufficient to infer and adapt the applied daily dose (Kurz *et al* 2015). Typically, CBCT intensity correction techniques on a standard full dose scan have been investigated in current literature. The wide range of techniques include look-up-table based solutions (Kurz *et al* 2015), the use of pCT-to-CBCT virtual CT (vCT) (Peroni *et al* 2012, Landry *et al* 2014, 2015, Veiga *et al* 2015, 2016, Wang *et al* 2016) yielding a so-called virtual CT (vCT) and the application of Monte-Carlo (MC) based methods (Mainegra-Hing and Kawrakow 2010, Thing *et al* 2016, Zöllner *et al* 2017) for scatter correction. While some of these methods have demonstrated accurate CBCT-based dose calculation in different treatment sites (Ding *et al* 2007, Fotina *et al* 2012, Niu *et al* 2012, Veiga *et al* 2014), there are limitations corresponding to the methods. For instance, DIR based approaches that enabled good dose calculation accuracy in head and neck (HN) (Kurz *et al* 2015, Landry *et al* 2015), might struggle in the pelvic region owing to the more pronounced and complex inter-fractional changes in anatomy. While the DIR inaccuracies could be improved by means of using vCT as prior for projection based intensity correction (Niu *et al* 2010, 2012, Park *et al* 2015, Kurz *et al* 2016), the time for generating corrected images, which takes several minutes, hinders the use of the obtained corrected CBCT images for online treatment adaption. Similarly, MC based methods which take up to several hours are not suitable.

Recently, the use of deep convolutional neural network (CNN) to speed up CBCT correction has received substantial interest. The U-Net architecture (Ronneberger *et al* 2015) has been employed to translate images across domains and correct CBCT intensities. In Kida *et al* (2018), a U-Net was trained using CBCT and vCT as input and target to translate the CBCT into a pCT equivalent image. Other U-Nets were trained for projection based image correction using MC simulated scatter distributions (Maier *et al* 2018, 2019) or corrected projections retrieved with a previously validated algorithm based on a vCT prior (Hansen *et al* 2018, Landry *et al* 2019). Apart from the U-Net, generative adversarial networks (GAN) (Goodfellow *et al* 2014) have been applied to translate CBCT into pCT images. In particular, the cycle-consistent GAN (cycleGAN) (Zhu *et al* 2017) architecture has seen considerable attention for unpaired training. For example, in the brain and the pelvic region (Harms *et al* 2019) (however using an additional paired loss term), in the HN region (Liang *et al* 2019) and the pelvic region (Kida *et al* 2019, Kurz *et al* 2019), dosimetric analysis of the cycle-consistent generative adversarial networks (cycleGAN) based corrected CBCT images were included, highlighting high dose calculation accuracy for photon therapy. The majority of deep learning based correction methods take less than a minute.

Using CBCT in IGRT increases the precision of the treatment, but also adds to the dose delivered to healthy tissues. One additional concern is thus that the imaging dose received from repeated CBCT scans at 20–35 fractions might be considerable and increase the risk of secondary malignancies. Kan *et al* (2008) measured, with thermoluminescent dosimeters, the dose from CBCT in a female anthropomorphic phantom and reported the effective and absorbed doses to 26 organs with standard and low-dose imaging modes. Effective doses to the whole body from standard mode CBCT for imaging of the pelvis were 22.7 mSv per scan. They concluded that CBCT on a daily basis could add an additional 2%–4% to the absolute secondary cancer risk. The radiation-induced cancer risk due to organ doses from kV-CBCT was also estimated by Kim *et al* (2013). Absorbed dose measurements in a cylindrical and in an anthropomorphic phantom yielded 170–187 mGy for the pelvic scan protocol, for which they concluded that 70% of additional secondary cancer risk from radiotherapy treatment of prostate patients can be attributed to CBCT imaging. Therefore, the excess radiation-induced cancer risk from CBCT is not negligible.

According to the Report of the American Association of Physicists in Medicine (AAPM) Therapy Physics Committee Task Group 180 (Ding *et al* 2018), imaging dose should be considered in the treatment planning process if larger than 5% of the therapeutic target dose, and in general the principle of 'as low as reasonably achievable' (ALARA) for imaging should be pursued. In the current clinical practice, radiation oncologists typically use the lowest possible dose of radiation to obtain the CBCT images, or try to to limit the frequency of CBCT imaging during treatment to reduce the risk of secondary cancers from cumulative CBCT dose. Lower dose CBCT at equivalent image quality could thus be favourable as it offers a higher flexibility of in terms of pre-treatment imaging frequency. Reducing dose, however, could be challenging since the CBCT image quality is further degraded, leading among others to potential loss of anatomical information.

Prior research has thoroughly investigated CBCT correction, however it remains to be investigated whether advances in deep learning can be leveraged to substantially reduce CBCT dose while jointly correcting CBCT image intensity and retaining therapeutic dose calculation accuracy. To address the needs of (1) CBCT dose reduction and (2) improving image quality for dose adaptation, our study investigates a cycleGAN-based low dose CBCT approach that translates a CBCT from a reduced number of projections (approximately 90), namely $CBCT_{LD}$, to a pCT equivalent image, referred to as $CBCT_{LD\_GAN}$, by simultaneously removing under-sampling artifacts and correcting image intensities while preserving anatomy fidelity. In parallel to $CBCT_{LD\_GAN}$, we also implemented an alternative cycleGAN with a generator residual connection to improve anatomical fidelity, referred to as $CBCT_{LD\_ResGAN}$.

2. Materials and methods

2.1. Patient data

2.1.1. Data acquisition

In this study, pCT and CBCT imaging datasets of 41 prostate cancer patients who received volumetric modulated arc therapy (VMAT) treatment to a total dose of 70–76 Gy in 2 Gy fractions at the Department of Radiation Oncology of the LMU Munich University Hospital were collected. All patients were advised to follow an in-house bladder and rectum filling protocol. The pCTs were acquired with a Toshiba Acquilion LB CT scanner (Canon Medical Systems, Japan). Tube voltage was set to 120 kV. An image grid of 1.074 mm $\times$ 1.074 mm $\times$ 3.000 mm was used in combination with a 55 cm lateral ield of view (FOV). No contrast agent was used.

To prevent the saturation of the detector panel and body outline artifacts, all retrospectively selected CBCT images were acquired in treatment position with a scan protocol of 120 kV tube voltage, exposure time of 20 ms and x-ray tube current of 20 mA per projection using the XVI system (version 5.52) of a Synergy medical linear accelerator (Elekta, Sweden). This is the lowest dose pelvic protocol at our institution. The lateral FOV was increased by using a laterally-shifted detector panel in M position and a bowtie filter. Images with body outline truncation in spite of the increased fov were excluded from the study. Around 350 projections [346, 357] were acquired in each 360° scan.

2.1.2. Data preparation

To generate a low dose $CBCT_{LD}$ from the full dose $CBCT_{org}$, CBCT projection data were uniformly under-sampled by a factor of 4 (keeping 25% of the projections) from about 350 to 90 projections, followed by a reconstruction using the Feldkamp–Davis–Kress (FDK) implementation of Reconstruction ToolKit (RTK) (Rit *et al* 2014) with 410 $\times$ 410 $\times$ 264 voxels on an isotropic 1.0 mm^3 grid. By thresholding and morphological masking, the patient couch was removed from the CBCT image, which was then converted to an image size of 512 $\times$ 512 by zero padding with the pixel intensity in the attenuation coefficient value (μ) range [0, 0.04] (values above 0.04 were set to 0.04). The first and last 35 image slices in superior-inferior direction with partial FOV cone truncation were excluded. pCTs were re-sampled to the same grid and image size using a linear interpolator from the SimpleITK library. The table was also removed from the images. The pixel intensity of the CT images was empirically converted to the range of the CBCT images ((HU + 1024)/65536) (Park *et al* 2015). The resulting intensities were mapped to the range [0, 0.05] (values above 0.05 were set to 0.05). Patients were instructed to lay with arms down and forearms folded up during acquisition. Since pCT slices showing limbs were excluded, the data used for training covered the pelvis and lower abdomen. To incorporate patient outline information in the training, a binary mask of each pCT and CBCT image was created by thresholding. All images were stored in 16 bit format before training. The data pre-processing workflow is illustrated in figure 1.

2.2. CycleGAN architecture and training

2.2.1. Forward and backward cycles and loss function

To correct the intensity of low dose $CBCT_{LD}$, we adapted a cycleGAN architecture (Zhu *et al* 2017, Ge *et al* 2019) to learn the image translation between low dose $CBCT_{LD}$ (input) and pCT equivalent images (output) with unpaired patient data (planning and fraction images). The framework chains two sets of a generator and discriminator networks. The generator aims to obtain the most efficient representation of $CBCT_{LD}$ from which a synthetic pCT can be generated slice by slice in the forward cycle. The discriminator is used to distinguish synthetic pCT with output label 0 and true pCT with label 1 in the forward cycle. In the backward cycle, outputs of the generator and discriminator are reversed. The loss function for both generators and discriminators consists of the terms described below.

In figure 2 (panel (a)), a generator G_{pCT} learns a mapping from $CBCT_{LD}$ to pCT such that the distribution of images from $G_{pCT}(CBCT_{LD})$ is indistinguishable from the distribution of pCT by a discriminator D_{pCT} using an adversarial loss in the forward cycle:

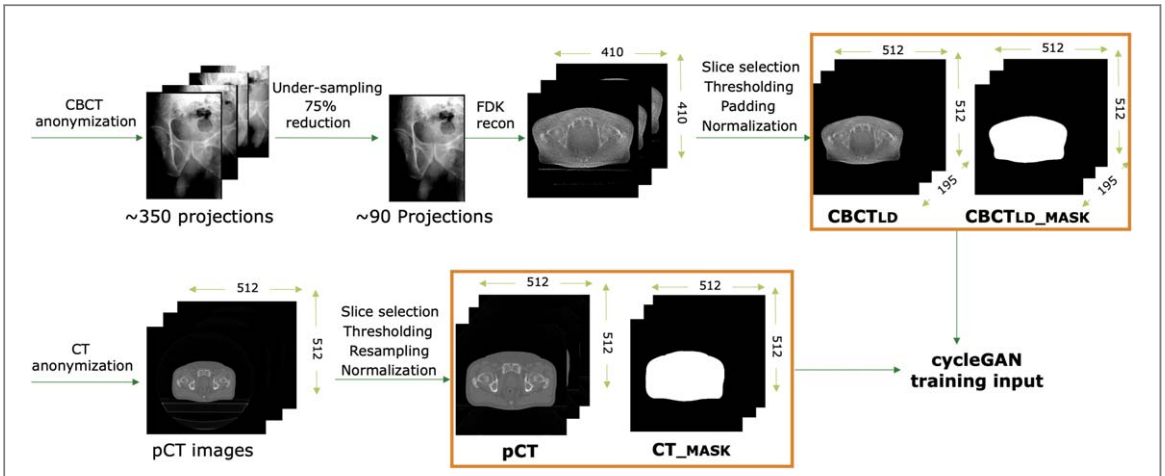


Figure 1. The pre-processing workflow for the CBCT and CT patient images.

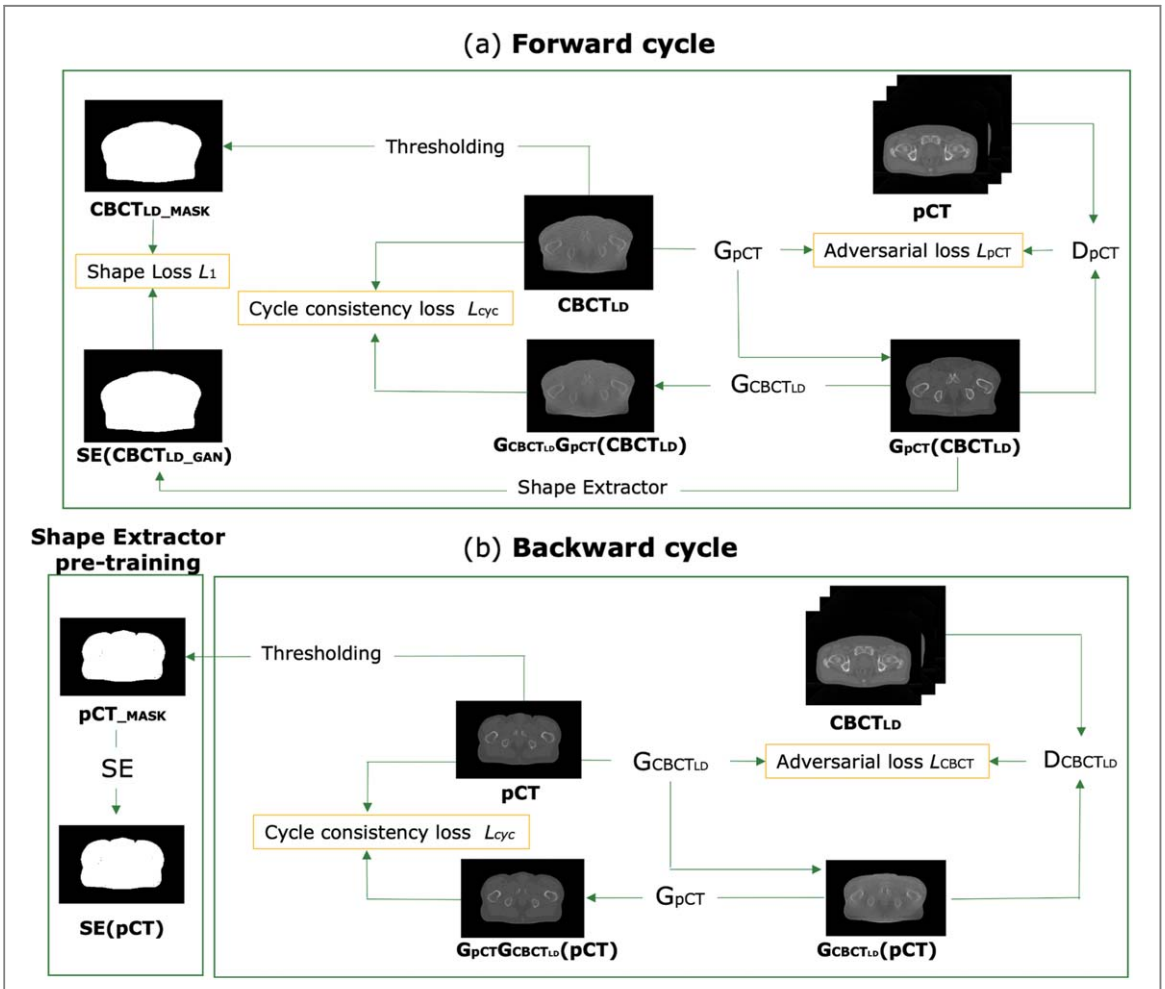


Figure 2. The cycleGAN architecture is used to generate pCT equivalent $CBCT_{LD\_GAN}$ images from $CBCT_{LD}$ images in (a) forward cycle, and to generate $CBCT_{LD}$ from pCT in (b) backward cycle. The mask of $CBCT_{LD\_GAN}$ is calculated by a shape extractor in the forward cycle.

$$L_{pCT} = \mathbb{E}_{CBCT_{LD}}[\log(1 - D_{pCT}(G_{pCT}(CBCT_{LD})))] + \mathbb{E}_{pCT}[\log D_{pCT}(pCT)], \quad (1)$$

where G_{pCT} aims to minimize the first term $\mathbb{E}_{CBCT_{LD}}[\log(1 - D_{pCT}(G_{pCT}(CBCT_{LD})))]$ by generating synthetic images $G_{pCT}(CBCT_{LD})$ that closely resemble pCT, while D_{pCT} aims to maximize both terms and become as good as possible in distinguishing between synthetic images $G_{pCT}(CBCT_{LD})$ and real pCTs.

In figure 2 (panel (b)), the second generator $G_{CBCT_{LD}}$ was trained to establish the inverse mapping from pCT to $CBCT_{LD}$ with the help of the second discriminator $D_{CBCT_{LD}}$ in the backward cycle:

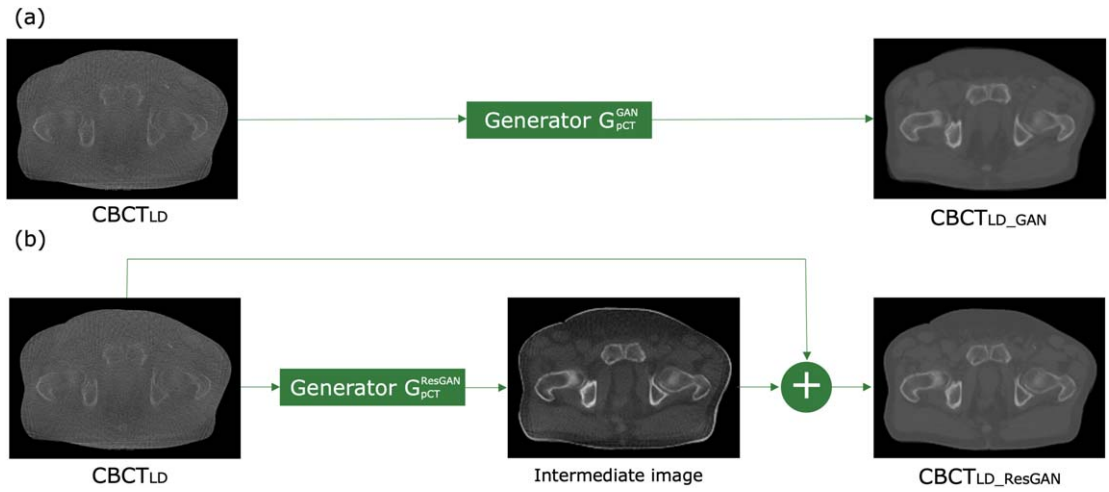


Figure 3. (a) In the $CBCT_{LD\_GAN}$ approach, the generator G_{pCT}^{GAN} is trained to directly map from $CBCT_{LD}$ to the final $CBCT_{LD\_GAN}$ images. (b) In contrast, the input and the intermediate output from the generator G_{pCT}^{ResGAN} are added with equal weight to obtain the final output $CBCT_{LD\_ResGAN}$.

$$L_{CBCT_{LD}} = \mathbb{E}_{pCT}[\log(1 - D_{CBCT_{LD}}(G_{CBCT_{LD}}(pCT)))] + \mathbb{E}_{CBCT_{LD}}[\log D_{CBCT_{LD}}(CBCT_{LD})]. \quad (2)$$

With the above adversarial loss, the generators G_{pCT} and $G_{CBCT_{LD}}$ are encouraged to generate realistic images of the target domain in order to fool the discriminators D_{pCT} and $D_{CBCT_{LD}}$.

To stabilize the training and ensure the inverse-consistent mappings with respect to the two image domains, a cycle consistency loss L_{cyc} is introduced to enforce $G_{CBCT_{LD}}(G_{pCT}(CBCT_{LD})) \approx CBCT_{LD}$ and $G_{pCT}(G_{CBCT_{LD}}(pCT)) \approx pCT$. In the forward cycle, L_{cyc} computes the L_1 norm of the output from $G_{CBCT_{LD}}$ with the generated synthetic pCT as input and the input low dose $CBCT_{LD}$:

$$L_{cyc}^{for} = \mathbb{E}_{CBCT_{LD}}[\|CBCT_{LD} - G_{CBCT_{LD}}(G_{pCT}(CBCT_{LD}))\|_1]. \quad (3)$$

In the backward cycle, the roles of $CBCT_{LD}$ and pCT are again swapped and the corresponding cycle consistency loss function is:

$$L_{cyc}^{back} = \mathbb{E}_{pCT}[\|pCT - G_{pCT}(G_{CBCT_{LD}}(pCT))\|_1]. \quad (4)$$

The cycle consistency loss, however, does not directly enforce the structural similarity between the input $CBCT_{LD}$ and the generated CT images. A previous CBCT-to-CT study has shown that there are measurable deviations in the patient body outline (Kurz *et al* 2019). To incorporate patient outline information and geometrically constrain the generator, we have adapted a shape loss as suggested in Ge *et al* (2019). A U-Net shape extractor (SE) was first trained for 5 epochs with paired pCT as input and the corresponding binary masks as the ground truth output. During the cycleGAN training, the shape extractor segments the patient outline of the generated $CBCT_{LD\_GAN}$ image from G_{pCT} and computes the L_1 loss between this new mask and its corresponding ground truth mask from the input low dose $CBCT_{LD}$:

$$L_{shape} = L_1(CBCT_{LD\_mask}, SE(G_{pCT}(CBCT_{LD}))). \quad (5)$$

Therefore the total loss used was:

$$L_{LD\_GAN} = L_{pCT} + L_{CBCT_{LD}} + \lambda_1(L_{cyc}^{for} + L_{cyc}^{back}) + \lambda_2 L_{shape}. \quad (6)$$

where λ_1 and λ_2 are hyperparameters that were empirically set to 25 and 1 in this study. The objective function to be solved was

$$G_{pCT}, G_{CBCT_{LD}} = \arg \min_{G_{pCT}, G_{CBCT_{LD}}} \max_{D_{CBCT_{LD}}, D_{pCT}} L_{LD\_GAN}(G_{pCT}, G_{CBCT_{LD}}, D_{CBCT_{LD}}, D_{pCT}). \quad (7)$$

Since this min-max optimization aims to find the model parameters that could describe the distribution of the image domain instead of using pixel-wise comparison, unpaired datasets could be used for this study.

We additionally trained a cycleGAN variant where a residual skip connection was used for the generator (see figure 3). This approach has been reported to improve geometric fidelity to the input image in the field of histopathology (de Bel *et al* 2021) and used in a previous CBCT-to-CT study (Deng *et al* 2022). Since anatomical fidelity is critical in our application, we have adopted this approach. As shown in figure 3, G_{pCT}^{GAN} was trained to convert $CBCT_{LD}$ directly to $CBCT_{LD\_GAN}$ in panel (a). For $CBCT_{LD\_ResGAN}$, G_{pCT}^{ResGAN} was trained to convert

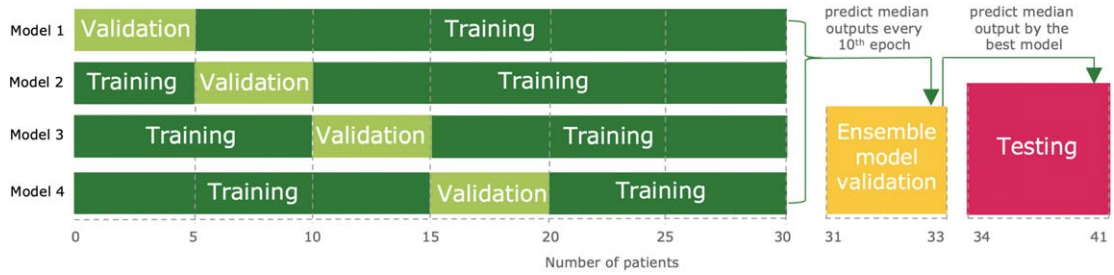


Figure 4. An illustration of the model ensemble method. Four independent models were trained with a four-fold split of the dataset. Then the four models were applied to the ensemble model validation set and the median of the four outputs was evaluated to find the best model, which was then applied to the final testing set in evaluation.

$CBCT_{LD}$ to an intermediate image, which has reversed intensities that suppress the streak artifacts from the $CBCT_{LD}$ input image as shown in panel (b). In the backward cycle, the other generator $G_{CBCT_{LD}}$ in the $CBCT_{LD\_ResGAN}$ approach was also trained to obtain the final output with the addition of the pCT input. Hyperparameters λ_1 and λ_2 were empirically set to 25 and 0 for $CBCT_{LD\_ResGAN}$. It was observed that the shape loss did not improve the performance of $CBCT_{LD\_ResGAN}$, as opposed to $CBCT_{LD\_GAN}$. Supplementary figure S1 and supplementary figure S2 illustrate the λ_2 experiments for one exemplary ensemble model validation patient (section 2.2.2) for $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$, respectively.

2.2.2. Network training

In a geometric augmentation pipeline, we employed two-dimensional (2D) horizontal flipping and affine transformations including rotation of $[-5°, 5°]$ and scaling by [0.9, 1.1] with a bicubic interpolation over 4×4 neighboring pixels to the CBCT and pCT inputs and their masks to enhance the generalisability of the model.

For the generator, the encoder contains two convolutional layers with stride 2 and the decoder contains two deconvolutional layers with stride 2. Nine residual blocks between encoding and decoding operations were used (Johnson *et al* 2016). For the discriminator, 70×70 PatchGAN (Isola *et al* 2017) was employed with a downsampling scheme from 256×256 to 32×32 by applying four series of 2D convolutional layers followed by instance normalization (Ulyanov *et al* 2016), except for the first and last layer, and LeakyReLU with a slope of 0.2 as nonlinearity, except for the last layer. The receptive field of the network was 70×70 and each pixel in the output was evaluated as a scalar in the range [0, 1]. The networks were implemented in PyTorch (v1.12.0).

Training was performed starting from the pre-trained model provided by Ge *et al* (2019). Results from training without the pre-trained model did not show convergence at the same number of epochs as for the pre-trained model. The adam optimizer was used for both generator and discriminator. The learning rate was set to 0.0002 during the first 100 epochs, and gradually reduced to zero over the next 100 epochs. For input to the network, the image patch was resampled to 256×256 pixels for the data augmentation. The batch size was set to one. A RTX A6000 graphics processing unit (GPU) (NVIDIA, California USA) was used.

Among a total of 41 patient datasets, a subset of 30 patients using four single folds, each containing 25 patients were used to perform the training with unpaired datasets. Three patient datasets were used as an ensemble model validation set and eight were used as a final test set. After the training, the generators G_{pCT}^{GAN} and G_{pCT}^{ResGAN} were used to correct $CBCT_{LD}$ intensity by translating $CBCT_{LD}$ slice-by-slice into pCT equivalent images, labelled $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$. As illustrated in figure 4, since four different folds were used for training the cycleGAN, four G_{pCT}^{GAN} and G_{pCT}^{ResGAN} with identical training hyper-parameters were obtained and applied to the ensemble model validation set. The median of the four models was used as the final output. For every 10th epoch, we computed the mean absolute error (MAE) and mean error (ME) for the three ensemble model validation cases in comparison to the reference vCT (section 2.3.1) and compared the appearance of soft tissues, bones, air cavities and body outline visually to find the optimal stopping epoch.

2.3. Data evaluation

2.3.1. Reference vCT and scatter corrected CBCT

Since there could be substantial anatomical differences between pCT and $CBCT_{LD}$ due to changes in bladder and rectum filling, as well as in patient positioning, the obtained images were not directly compared to the pCT for determining the accuracy of $CBCT_{LD\_GAN}$ or $CBCT_{LD\_ResGAN}$. Instead, we generated a vCT by mapping the pCT to the daily CBCT via a dedicated DIR approach. As described in Hofmaier *et al* (2017), we aim for (1) image similarity which is computed by normalized gradient fields, and (2) deformation regularity which is computed

by curvature regularization. The optimization problem is solved in a discretize-then-optimize scheme using a quasi-Newton L-BFGS optimizer.

A CBCT correction technique that had been validated in Park *et al* (2015) and Kurz *et al* (2016) was employed as an alternative reference for evaluating the network results and their comparison to vCT for the eight test cases. This reference correction approach was fully described in the original publications of Niu *et al* (2010) and Niu *et al* (2012) and in follow-up studies from Hansen *et al* (2018) and Landry *et al* (2019). We first forward project the VCT according to the geometry of the CBCT scanner to retrieve primary beam projections (I_{pri}). The scatter and other low frequency deviations (I_{sca}) are calculated as the difference between a scaled original $CBCT_{org}$ projection (I_{org}) with ntensity scaling factor (ISF) and (I_{pri}) followed by a generous smoothing function *f*. The scatter corrected projection (I_{cor}) was estimated by subtracting the scatter contribution from the original measured $CBCT_{org}$ projections. With I_{cor}, we could reconstruct a scatter-corrected CBCT, in the following referred to as $CBCT_{cor}$ with HU values equivalent to the pCT, and with ideally the same anatomy as $CBCT_{org}$. In line with $CBCT_{LD}$, $CBCT_{cor}$ was reconstructed using the FDK algorithm with the same reconstruction settings.

2.3.2. CT number accuracy

For the eight test cases, $CBCT_{LD}$, $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ were compared to vCT in terms of the MAE and ME in HU. All pixel intensities were scaled from model output in μ to HU using the inverse empirical scaling used for the pCT. Pixels outside the joint body outline of vCT and $CBCT_{LD\_GAN}/CBCT_{LD}$ or $CBCT_{LD\_ResGAN}/CBCT_{LD}$ were excluded.

2.3.3. Dosimetric analysis

To determine dosimetric accuracy, we generated and recalculated VMAT plans on vCT, $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ for the eight test patients in a research version of a commercial treatment planning system (TPS) (RayStation, version 10.01, RaySearch, Sweden). Contours of target structures and organs-at-risks (OARs) were transferred via DIR from pCT to vCT, on which VMAT plans using one arc were optimized on an isotropic dose grid of 3.0 mm using a collapsed-cone dose engine. These plans were then recalculated on $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$. The generic Elekta Synergy beam model with Agility multi-leaf-collimator in the TPS was employed. The prescription was 74 Gy in 37 fractions and we aimed at clinical target volume (CTV) $V_{95\%}$ of 100%, and planning target volume (PTV) $V_{95\%}$ better than 95% of the prescription dose. We aimed at fulfilling the dose-volume histogram (DVH) constraints that are given in the QUANTEC report (Marks *et al* 2010) for the rectum and the bladder. Identical generic CT number to electron density conversion tables were employed for vCT, $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ in all cases. The dose distributions on vCT, $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ were then compared in terms of a 1%, 2% and 3% dose difference criterion. Voxels with less than 10% of the prescribed dose were excluded. In addition, the VMAT dose distributions for vCT, $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ were compared with regard to DVH parameters of clinically relevant target structures and OARs. CTV and PTV $D_{98\%}$ and $D_{2\%}$, together with PTV $D_{50\%}$ and $V_{95\%}$ were analyzed. For the rectum $V_{50/60/65\ Gy}$ and for the bladder $V_{60/65\ Gy}$ were analyzed.

To evaluate the robustness of the dosimetric results to the reference image, the VMAT plans were additionally recalculated on $CBCT_{cor}$ and the dose distribution compared to the one from vCT with a 1% dose difference criterion.

2.3.4. Positioning accuracy

Daily patient positioning is one of the primary purposes of in-room CBCT. To evaluate registration accuracy when using $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$, we rigidly registered these images to the pCT using the research TPS. The transformations were compared to the one obtained from registering $CBCT_{org}$ to the pCT. Automated gray level rigid registration was used with six degrees of freedom.

2.3.5. Anatomical fidelity

To evaluate the networks' capability for preserving the anatomy correctly, we evaluated the shapes of organs geometrically. Two OARs, bladder and rectum, were segmented manually using the research TPS on $CBCT_{org}$, $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ for this purpose. All contours were thoroughly validated by a radiation oncologist with expertise in prostate cancer radiotherapy. Dice similarity coefficient (DSC), average and 95th percentile Hausdorff distance (HD_{avg}, HD_{95}) of the contours on $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ were computed to determine the fidelity of the organ shape in the network output, using $CBCT_{org}$ as ground truth.

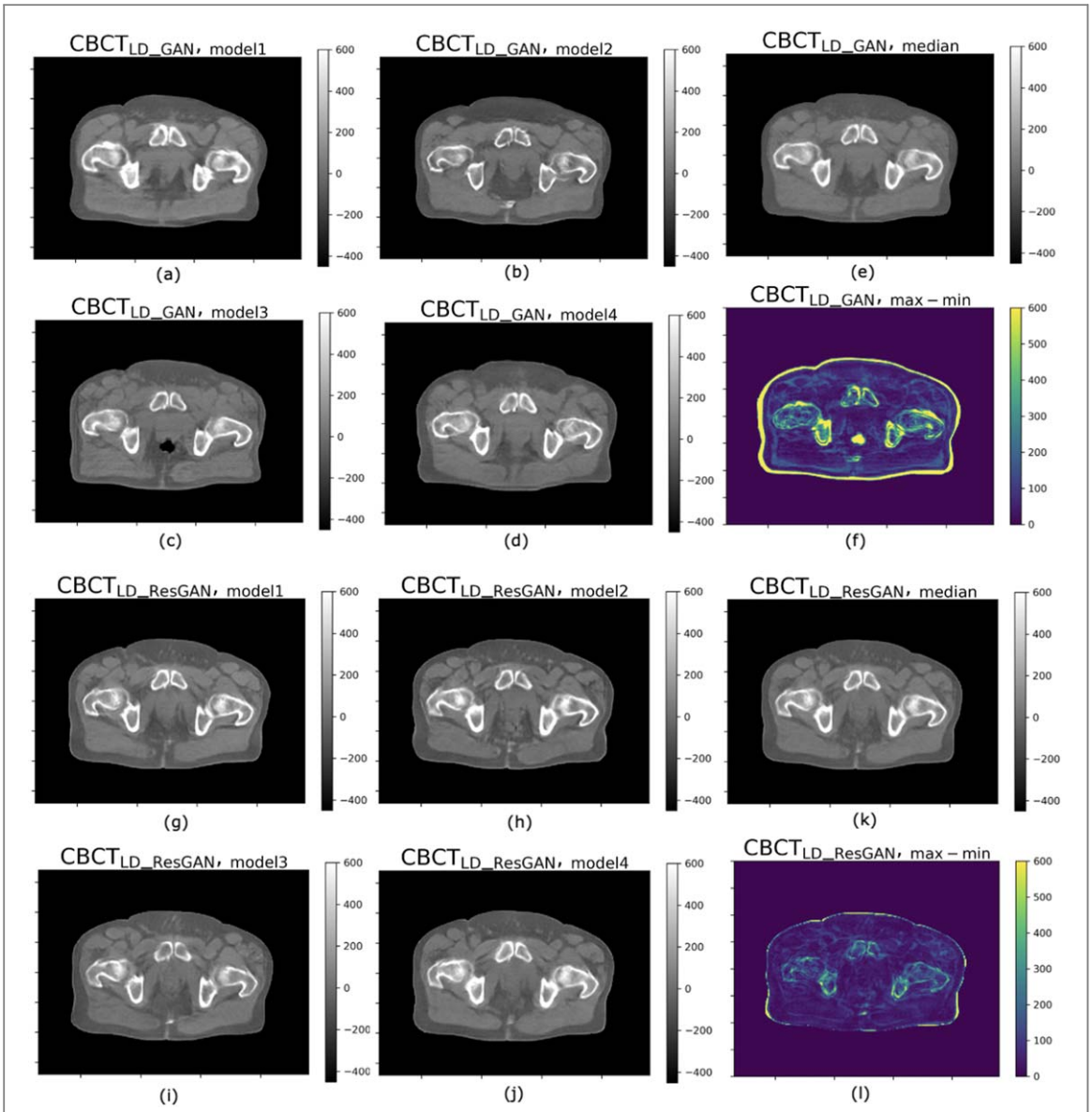


Figure 5. The outputs of the four trained (a)–(d) G_{pCT}^{GAN} and (g)–(j) G_{pCT}^{ResGAN} models, the median of (e) $CBCT_{LD\_GAN}$, (k) $CBCT_{LD\_ResGAN}$, and the pixel-wise maximum minus minimum for (f) G_{pCT}^{GAN} and (l) G_{pCT}^{ResGAN} outputs for one representative ensemble model validation patient. All values are in HU.

3. Results

3.1. Model selection based on ensemble validation

The model of epoch 50 for $CBCT_{LD\_GAN}$ and the model of epoch 60 for $CBCT_{LD\_ResGAN}$ which had the lowest MAE and ME and high soft-tissue geometric fidelity upon visual inspection of the validation cases were selected. In figure 5, the output images from the four trained G_{pCT}^{GAN} and G_{pCT}^{ResGAN} are shown for an exemplary ensemble model validation patient (panel (a)–(d) and panel (g)–(j)), together with the calculated median images (panel (e) and (k)) and the pixel-wise difference between maximum and minimum HU values (panel (f) and (l)). For $CBCT_{LD\_GAN}$, deviations between the four different models were most pronounced at the edges of the bony anatomy, as well as at the patient body outline. We also observed variations in the bowels with occasional generation of air pockets (panel (c)). For $CBCT_{LD\_ResGAN}$, deviations were generally less pronounced as in $CBCT_{LD\_GAN}$, and no random large air pocket was generated. In the following analysis, only the median images were considered.

3.2. Computational details

The training to the best model at epoch 50 of a single fold took about 9 h for $CBCT_{LD\_GAN}$, and at epoch 60 took about 10.5 h for $CBCT_{LD\_ResGAN}$. The average time to convert a complete 3D $CBCT_{LD}$ of one patient with 195 slices into $CBCT_{LD\_GAN}$ or $CBCT_{LD\_ResGAN}$ was about 2 s (about 10 ms per slice) on a GPU.

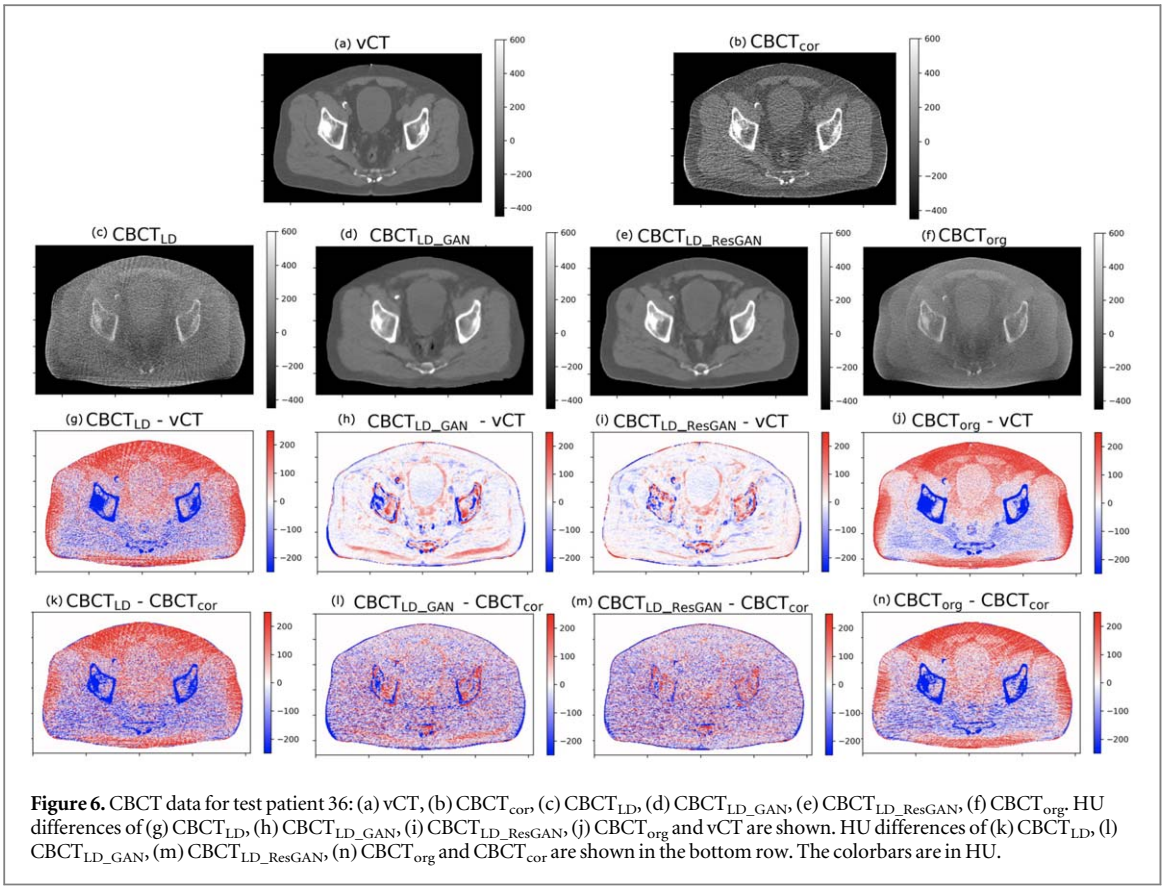


Figure 6. CBCT data for test patient 36: (a) vCT, (b) $CBCT_{cor}$, (c) $CBCT_{LD}$, (d) $CBCT_{LD\_GAN}$, (e) $CBCT_{LD\_ResGAN}$, (f) $CBCT_{org}$. HU differences of (g) $CBCT_{LD}$, (h) $CBCT_{LD\_GAN}$, (i) $CBCT_{LD\_ResGAN}$, (j) $CBCT_{org}$ and vCT are shown. HU differences of (k) $CBCT_{LD}$, (l) $CBCT_{LD\_GAN}$, (m) $CBCT_{LD\_ResGAN}$, (n) $CBCT_{org}$ and $CBCT_{cor}$ are shown in the bottom row. The colorbars are in HU.

3.3. Image analysis

We evaluated $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ on eight test patients. CBCT images of test patient 36 and their HU differences are shown in figure 6. In $CBCT_{LD}$ (panel (c)), streaks and undersampling artifacts are clearly observed when compared to $CBCT_{org}$ (panel (f)). In panel (d) and (e), $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ have successfully removed these artifacts. Figure 6 also shows the HU differences of all CBCT results with respect to vCT. $CBCT_{LD}$ (panel (g)) and $CBCT_{org}$ (panel (j)) show larger underestimated regions and larger overestimated regions, as well as pronounced deviations in the bony structures. As seen from the reduced differences to vCT, $CBCT_{LD\_GAN}$ (panel (h)) and $CBCT_{LD\_ResGAN}$ (panel (i)) improved image intensities compared to $CBCT_{org}$. The remaining differences between $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ with respect to vCT are observed at the patient body outline and bone interfaces. In addition, figure 6 also shows the HU differences of all CBCT results with respect to $CBCT_{cor}$. All HU differences to $CBCT_{cor}$ are similar to the differences to vCT but with remaining increased noise.

To quantify the image quality, we computed the average ME and MAE in HU of $CBCT_{LD\_GAN}$, $CBCT_{LD\_ResGAN}$ and $CBCT_{LD}$ compared to vCT for training, validation and test patients as shown in figure 7. In panels (a) to (c), the ME of $CBCT_{LD}$ had positive values in almost all patients while $CBCT_{LD\_GAN}$ had negative values in the majority of datasets. $CBCT_{LD\_ResGAN}$ had slightly more negative values than positive ones. The MEs of all datasets were comparable within the correction method. In panels (d) to (f), $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ showed a substantially reduced MAE for all datasets compared to $CBCT_{LD}$.

Table 1 reports the quantitative results in terms of the average ME and MAE of all patient images in training, validation and testing datasets. For the testing datasets, the average ME changed from 20 HU for $CBCT_{LD}$ to −6 HU for $CBCT_{LD\_GAN}$ and −2 HU for $CBCT_{LD\_ResGAN}$. The average MAE reduced from 126 HU for $CBCT_{LD}$ to 55 HU for $CBCT_{LD\_GAN}$ and 44 HU for $CBCT_{LD\_ResGAN}$.

3.4. Dosmetric analysis

The quantitative results of the dose difference analysis of the VMAT plans comparing $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ to vCT are given in table 2 for all test datasets and the investigated dose difference (DD) levels. For $CBCT_{LD\_GAN}$, the average 1% DD pass-rate was 95.9%, with a value range from 87.3% to 98.7%. For $CBCT_{LD\_ResGAN}$, the average 1% DD pass-rate was 97.0%, with a value range from 92.0% to 98.6%. This shows that a high agreement of $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ to the reference vCT was found. In addition, the average 1% DD pass-rate comparing vCT to $CBCT_{cor}$ for all test datasets was 98.4%, indicating excellent dosimetric agreement between the two benchmark images.

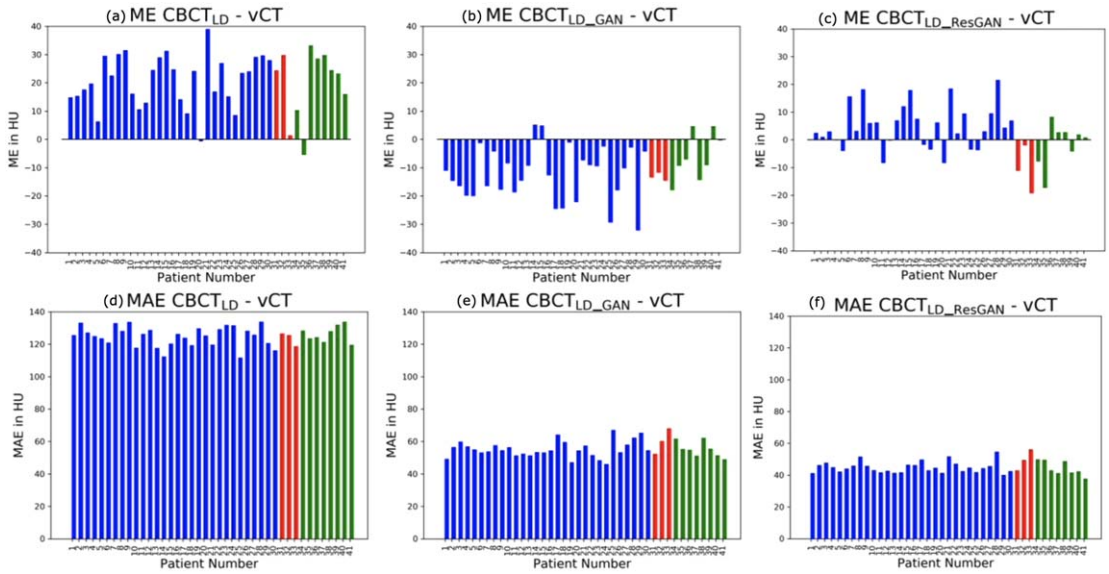


Figure 7. (Top) ME and (bottom) MAE per patient for the comparison of vCT and (a), (d) $CBCT_{LD}$, (b), (e) $CBCT_{LD\_GAN}$ or (c), (f) $CBCT_{LD\_ResGAN}$. The data are labeled as belonging to the training (blue), validation (red) and testing (green) datasets.

Table 1. Average HU ME and MAE of all patient images in training, validation and testing datasets for the comparison of $CBCT_{LD}$, $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ with vCT, respectively. The number in square brackets represent [min, max] values among all patients in the corresponding groups.

Dataset	ME $CBCT_{LD}$	ME $CBCT_{LD\_GAN}$	ME $CBCT_{LD\_ResGAN}$
Training	21 [−1, 39]	−12 [−32, 5]	5 [−8, 21]
Validation	19 [1, 30]	−13 [−15, −12]	−10 [−18, −1]
Test	20 [−5, 33]	−6 [−18, 5]	−2 [−17, 8]
Dataset	**MAE $CBCT_{LD}$**	**MAE $CBCT_{LD\_GAN}$**	**MAE $CBCT_{LD\_ResGAN}$**
Training	125 [112, 134]	55 [46, 67]	45 [40, 55]
Validation	123 [118, 126]	60 [52, 68]	49 [42, 55]
Test	126 [119, 134]	55 [49, 62]	44 [38, 50]

Table 2. Dose Differences (DD) of the eight test patients for the VMAT plans recalculated on $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ with respect to vCT . All values are in percent.

	$CBCT_{LD\_GAN}$			$CBCT_{LD\_ResGAN}$		
Test patient	1%DD	2%DD	3%DD	1%DD	2%DD	3%DD
34	92.7	98.1	98.8	92.0	98.3	98.9
35	97.1	98.9	99.5	97.2	99.0	99.6
36	97.4	99.1	99.8	98.0	99.4	99.8
37	98.3	99.5	99.8	98.4	99.6	99.9
38	87.3	97.1	98.4	95.7	97.8	98.8
39	97.4	99.1	99.6	97.8	99.2	99.7
40	98.7	99.7	99.9	98.6	99.7	99.9
41	97.9	99.2	99.7	98.0	99.2	99.7
Average	95.9	98.8	99.4	97.0	99.0	99.5

The dose distribution and difference of test patient 38 are depicted in figure 8. Only minor dose differences in the planning target volume (PTV) region between $CBCT_{LD\_GAN}$, $CBCT_{LD\_ResGAN}$ and vCT were found. The dose difference for $CBCT_{LD\_ResGAN}$ has smaller magnitude than for $CBCT_{LD\_GAN}$.

Figure 9 shows target and OAR DVH parameter differences with respect to vCT as boxplots over all patients. For most of the considered parameters in both $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$, differences were within 1.5 Gy for dose DVH parameters (D_x) and below 1.5% for volume DVH parameters (V_x). All deviations were

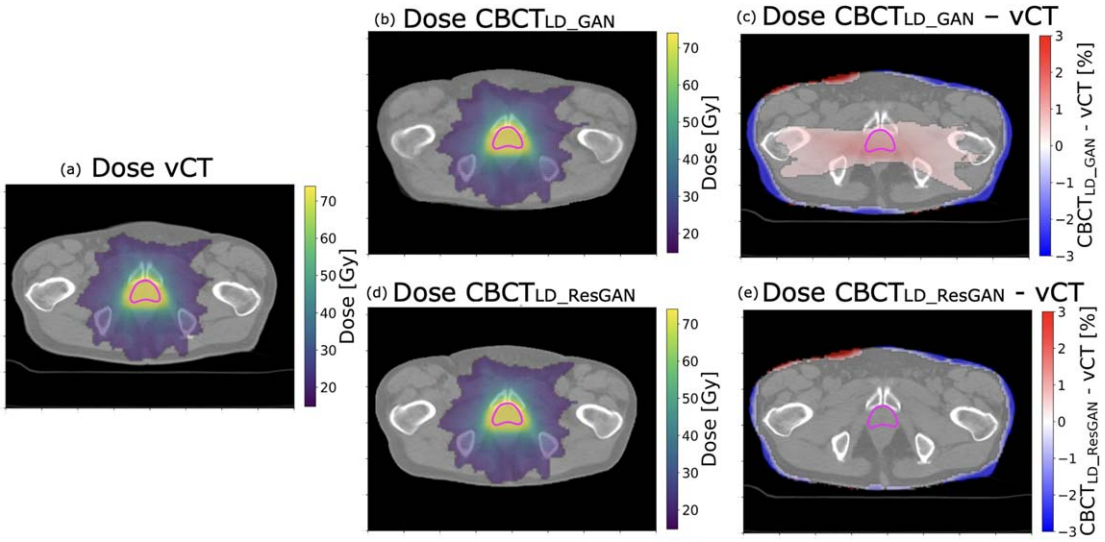


Figure 8. VMAT dose distributions of the test patient 38. Dose distributions optimized on (a) vCT and recalculated on (b) $CBCT_{LD\_GAN}$ and (d) $CBCT_{LD\_ResGAN}$ are shown together with (c), (e) their corresponding differences. The PTV is shown in magenta. Dose differences below 0.4% are not shown for better visualization.

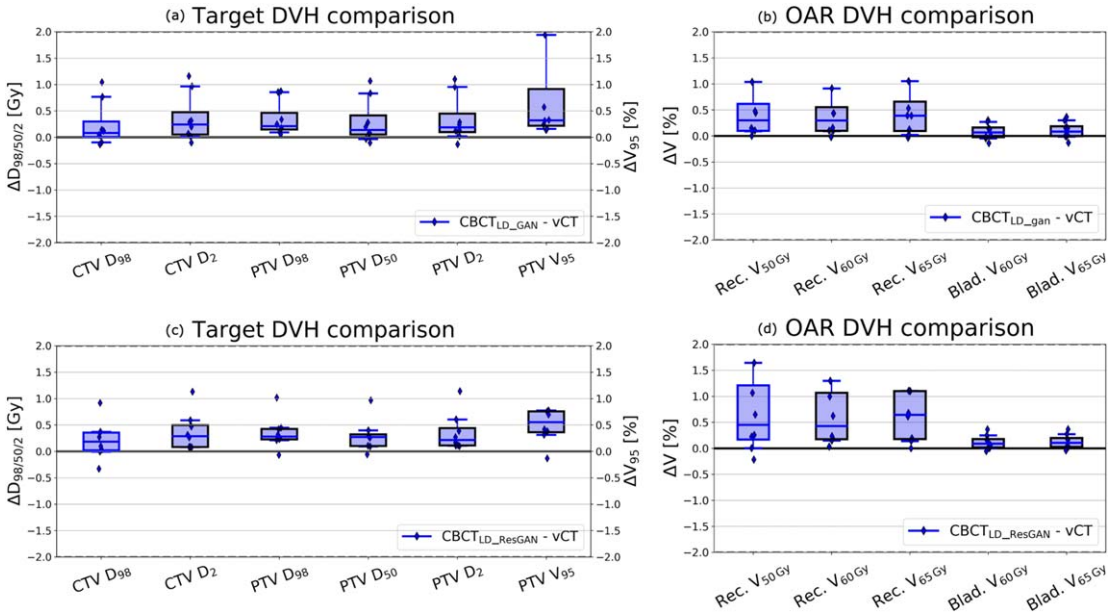


Figure 9. Clinically relevant DVH parameter differences of $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ with respect to vCT for (a), (c) target and (b), (d) OAR structures. Each data point represents a test patient. Whiskers correspond to the 5th–95th percentile. All dose values correspond to the total dose of the fractionated treatment.

below 2 Gy/2%. Particularly in the target DVH comparison, the median differences of $D_{98\%}$, $D_{50\%}$ and $D_{2\%}$ comparing $CBCT_{LD\_GAN}$ with respect to vCT were 0.3%, 0.3% and 0.3% for the PTV. In $CBCT_{LD\_ResGAN}$, the median differences of $D_{98\%}$, $D_{50\%}$ and $D_{2\%}$ with respect to vCT were 0.4%, 0.3% and 0.4% for the PTV.

3.5. Positioning accuracy

With respect to $CBCT_{org}$-to-pCT, the mean absolute difference of rigid transformation parameters were 0.07 mm (right–left) (RL), 0.05 mm (inferior–superior) (IS), 0.01 mm (posterior–anterior) (PA), 0.17° (pitch), 0.15° (roll) and 0.24° (yaw) for $CBCT_{LD\_GAN}$-to-pCT, and similarly, the mean absolute differences were 0.03 mm (RL), 0.05 mm (IS), 0.04 mm (PA), 0.16° (pitch), 0.19° (roll) and 0.26° (yaw) for $CBCT_{LD\_ResGAN}$-to-pCT. The majority of differences were thus less than 0.20 mm or 0.20°, except the pitch of patient 34 was 0.32° for $CBCT_{LD\_GAN}$, the yaw of patient 38 was 0.82° for $CBCT_{LD\_GAN}$ and 0.77° for $CBCT_{LD\_ResGAN}$, the roll of patient 39 was −0.60° for $CBCT_{LD\_GAN}$ and −0.79° for $CBCT_{LD\_ResGAN}$, the yaw of

Table 3. The anatomical fidelity results of bladder in terms of Dice similarity coefficient (DSC), average and 95th percentile Hausdorff distance (HD_{avg}, HD_{95}) in the test patients from $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$.

	$CBCT_{LD\_GAN}$			$CBCT_{LD\_ResGAN}$		
Test patient	DSC	HD_{avg} (mm)	HD_{95} (mm)	DSC	HD_{avg} (mm)	HD_{95} (mm)
34	0.83	1.68	7.39	0.93	0.65	3.86
35	0.90	1.27	5.79	0.93	0.83	3.29
36	0.91	1.12	6.11	0.93	0.81	4.42
37	0.84	1.46	4.42	0.85	1.43	6.10
38	0.94	0.78	4.42	0.94	0.69	3.67
39	0.83	1.82	6.40	0.90	1.08	3.79
40	0.91	1.52	4.45	0.94	1.04	3.37
41	0.89	1.06	6.82	0.94	0.65	3.90
Average	0.88	1.34	6.03	0.92	0.90	4.05

Table 4. The anatomical fidelity results of rectum in terms of Dice similarity coefficient (DSC), average and 95th percentile Hausdorff distance (HD_{avg}, HD_{95}) in the test patients from $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$.

	$CBCT_{LD\_GAN}$			$CBCT_{LD\_ResGAN}$		
Test patient	DSC	HD_{avg} (mm)	HD_{95} (mm)	DSC	HD_{avg} (mm)	HD_{95} (mm)
34	0.75	2.55	8.67	0.83	1.32	6.17
35	0.72	2.17	7.52	0.85	0.98	3.56
36	0.80	2.01	6.83	0.82	1.91	7.11
37	0.85	1.11	4.03	0.90	0.68	2.12
38	0.89	1.10	4.06	0.92	0.84	3.15
39	0.74	1.98	6.68	0.90	0.63	2.12
40	0.62	2.86	8.12	0.87	0.89	3.15
41	0.79	1.62	5.56	0.85	1.15	3.70
Average	0.77	1.93	6.43	0.87	1.05	3.89

patient 39 was $-0.42°$ for $CBCT_{LD\_GAN}$ and $-0.69°$ for $CBCT_{LD\_ResGAN}$ and the pitch of patient 41 was $-0.65°$ for $CBCT_{LD\_ResGAN}$.

3.6. Anatomical fidelity

As shown in table 3, the average DSC of bladder was 0.88 for $CBCT_{LD\_GAN}$ and 0.92 for $CBCT_{LD\_ResGAN}$ with respect to $CBCT_{org}$. HD_{avg} and HD_{95} of bladder were 1.34 mm and 6.03 mm for $CBCT_{LD\_GAN}$, and 0.90 mm and 4.05 mm for $CBCT_{LD\_ResGAN}$. As shown in table 4, the average DSC of rectum was 0.77 for $CBCT_{LD\_GAN}$ and 0.87 for $CBCT_{LD\_ResGAN}$ with respect to $CBCT_{org}$. HD_{avg} and HD_{95} of rectum were 1.93 mm and 6.43 mm for $CBCT_{LD\_GAN}$, and 1.05 mm and 3.89 mm for $CBCT_{LD\_ResGAN}$. In both bladder and rectum, $CBCT_{LD\_ResGAN}$ had a higher DSC and lower HD_{avg} and HD_{95} than $CBCT_{LD\_GAN}$. In addition, bladder had generally higher DSC and lower HD than rectum in both $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$. Figure 10 illustrates that the contour of the rectum in $CBCT_{LD\_GAN}$ (panel (b) and (e)) had a larger shape deviation than in $CBCT_{LD\_ResGAN}$ (panel (c) and (f)) with respect to $CBCT_{org}$ (panel (a) and (d)) due to a small incorrect air pocket generated, which would also be contoured as part of the rectum in clinical practice.

4. Discussion

The daily use of CBCT imaging during a fractionated radiotherapy course could deliver a considerable amount of radiation dose to patients. Due to the insufficient image quality, CBCT also cannot be used for daily dose calculation and adaptation. To address these problems, our study aimed at addressing dose reduction and intensity correction simultaneously. We generated synthetic low dose $CBCT_{LD}$ to train two cycleGAN architectures to tackle the tasks of (1) removing the under-sampling artifacts and (2) correcting the intensity of $CBCT_{LD}$, and evaluated both approaches on a cohort of prostate cancer patients. The key finding of this study is

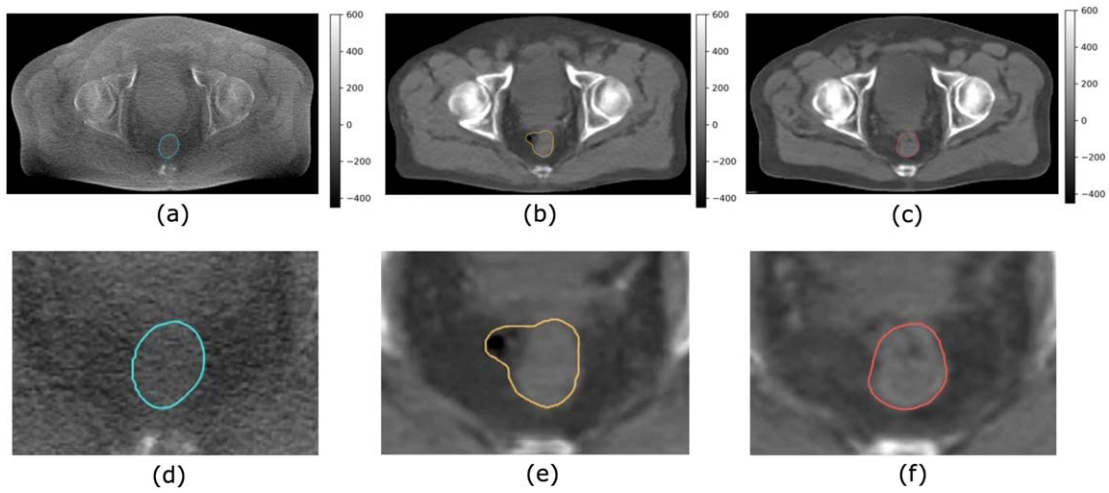


Figure 10. The contours of rectum in (a) $CBCT_{LD}$, (b) $CBCT_{LD\_GAN}$ and (c) $CBCT_{LD\_ResGAN}$, and ((d)–(f)) the corresponding zoom-in contours for the test patient 41.

that it was possible to reduce the CBCT imaging dose by 75% and enable VMAT dose calculation accurately with the use of cycleGAN.

To obtain $CBCT_{LD}$, the number of projections was subsampled by a factor of four, which led to severe streaking in the reconstructed images. The proposed $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ techniques successfully removed all streak artifacts, by training the generators G_{pCT} to map the $CBCT_{LD}$ input to the pCT domain which has no under-sampling noise. In addition, the cycle consistency loss regularized the body structures between $CBCT_{LD}$ and $CBCT_{LD\_GAN}$, and between $CBCT_{LD}$ and $CBCT_{LD\_ResGAN}$. The hyperparameter λ_1 was increased from a default value of 10 to 25, as the relative importance of preserving the anatomical content in the loss function was previously demonstrated in Kurz *et al* (2019) and confirmed in our study. Furthermore, the shape loss was added to incorporate patient body outline information as suggested in Ge *et al* (2019). The hyperparameter λ_2 was adjusted from a default value of 10 to 1 for $CBCT_{LD\_GAN}$. Compared to the default value 10, the smaller λ_2 tends to output soft tissue and organs with more correct shapes in our experiments. For $CBCT_{LD\_GAN}$, λ_2 of 1 was empirically found beneficial in comparison to using no shape loss as shown in the supplementary figure S1. For $CBCT_{LD\_ResGAN}$, λ_2 of 0 gives the least variation in the min-max plots and thus a higher stability of the model outputs, as shown in supplementary figure 2.

Compared to previous unpaired CBCT-to-CT correction works using cycleGAN in pelvic scans, our model has achieved a slightly higher MAE reduction. This could be explained by the fact that the input $CBCT_{LD}$ has more noise than the usual standard full dose CBCT input in other studies. The MAE in comparison to vCT was substantially reduced from 126 HU for $CBCT_{LD}$ to 55 HU for $CBCT_{LD\_GAN}$ and to 44 HU for $CBCT_{LD\_ResGAN}$. Liu *et al* (2022) proposed a two-step method with phantom-based and patient-based models, and reduced MAE of well-matched slices from 67 to 32 HU with respect to a deformably registered reference CT. In Deng *et al* (2022), the model that had a similar generator residual connection reduced MAE from 29 to 18 HU. Harms *et al* (2019) trained a cycleGAN model with paired CBCT and pCT datasets and reduced MAE from 56 to 18 HU. In another study with a similar patient cohort, Kurz *et al* (2019) reduced MAE from 103 to 87 HU with respect to $CBCT_{cor}$ (Kurz *et al* 2016) as reference, which has higher anatomical fidelity to $CBCT_{org}$ but more noise than vCT.

In terms of dose calculation accuracy, good results were achieved for VMAT when comparing $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ to vCT. For a 2% dose difference criterion, a mean pass-rate of 99% was determined for the test patients for both proposed approaches. Despite the additional under-sampling artifacts in the low dose CBCT input, the $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ dosimetric results are still comparable to the previous work by Kurz *et al* (2019) which used a fully sampled prostate dataset with a similar cycleGAN architecture (without shape loss or a generator residual connection). In line with this, for most cases a very good agreement of $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ with respect to vCT in terms of clinically relevant DVH parameters was achieved. For VMAT, a trend of marginally overestimated doses on $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ was found in the target structures and OARs, with deviations below 1 Gy for dose DVH parameters (D_x) and below 1.5% for volume DVH parameters (V_x) for 7 out of 8 test cases.

In order to investigate the anatomical fidelity, two OARs in the network-generated images were contoured and compared to a ground truth contour on $CBCT_{org}$. The DSC in rectum was lower than in bladder, possibly due to the higher variability of the rectum shape and the random natural occurrence of air pockets in the rectum.

In addition, it is more difficult to segment the rectum, thus increasing the uncertainties for rectum contours. It is notable that $CBCT_{LD\_ResGAN}$ still yielded generally higher DSC and lower HD_{avg} and HD_{95} than $CBCT_{LD\_GAN}$ in the two OARs. This demonstrated that $CBCT_{LD\_ResGAN}$ can achieve improved geometrical accuracy, and indicated a positive effect from a generator residual connection.

While having high treatment dose calculation accuracy and enhanced anatomical fidelity, the proposed low dose CBCT techniques could deliver at least 75% lower dose in a pelvic scan. To estimate the reduced patient dose, we have chosen the cone beam dose index (CBDI) value which provides a single number that represents the mean volumetric dose in the CT dose index (CTDI) phantom as reported in (Hyer and Hintenlang 2010). They reported a CBDI value (table 2 in Hyer and Hintenlang (2010), chest protocol) for the same configuration as our protocol (M20 protocol with 120 kV and a bowtie filter at an Elekta XVI scanner) of 1.62 mGy/100 mAs. By selecting only 90 out of 350 projection frames, our $CBCT_{LD}$ has thus reduced the patient dose from 2.27 to 0.57 mGy (from a total exposure of 140 mAs to 36 mAs) per scan. For reference, another Elekta XVI CBCT-to-CT work using cycleGAN with a regular full dose scan in prostate cancer reported a total exposure of 288 mAs without providing complete acquisition details such as kV collimator type or the use of a bowtie filter (Kida *et al* 2019). In a recent deep learning CBCT low-dose study using a U-Net, Yuan *et al* (2020) used a clinical HN protocol with 182 projections over 205°, which would correspond to 319 projections over 360°, and thus to a considerably higher sampling rate than our approach by a factor of 3.5.

The computational time of the investigated low dose CBCT techniques for correcting a 3D pelvic scan per patient was shorter when compared to vCT or the projection-based scatter correction approach $CBCT_{cor}$ in Kurz *et al* (2016), which have correction times in the order of 6–10 min per patient. The correction time per slice of 10 ms in $CBCT_{LD\_GAN}$ or $CBCT_{LD\_ResGAN}$ is identical to the other prostate CBCT-to-CT works by Landry *et al* (2019) using a U-Net, and by Kurz *et al* (2019) using a similar cycleGAN. It should be noted that there are also iterative reconstruction works using compressed sensing, e.g. in Choi *et al* (2010), Lee *et al* (2012) and Park *et al* (2012) or total variation in Song *et al* (2014) to remove under-sampling artefacts in CBCT images. However, one more prior scatter correction step would be required to convert the CBCT image intensities to CT diagnostic intensities. Since the proposed $CBCT_{LD\_GAN}$ or $CBCT_{LD\_ResGAN}$ techniques allow fast image correction within 2 s per patient (195 slices), they have the potential to be applied for CBCT-based online treatment plan adaptation.

There are some limitations in this study. First, the evaluation of the HU and dose calculation accuracy rely on vCT. The advantage of using vCT as a reference is that it has correct intensity and ideally identical anatomy to $CBCT_{LD}$. However, vCT might not be a perfect ground truth due to uncertainties in DIR. This might be one of the potential causes for the small deviation found in the patient body outline in figure 6 panel (h) and (i), and in the dose difference maps in figure 8 panel (c) and (e). This is also the reason why we compared the network results with an alternative ground truth $CBCT_{cor}$ for inspecting the deviations that might have been caused by the DIR uncertainties. As shown in (figure 6 panel (l) and (m)), similar deviations in the patient body outline were also found in the comparison to $CBCT_{cor}$, which implies that the uncertainties in DIR did not affect HU accuracy analysis. In addition, the average 1% DD pass-rate comparing vCT to $CBCT_{cor}$ was 98.4% as reported in section 3.4, which also implies that employing either vCT or $CBCT_{cor}$ as ground truth has only minimal impact on the dosimetric comparison for the network results.

Second, it is observed that the prediction from some single models before ensembling can be geometrically unstable, especially for $CBCT_{LD\_GAN}$. Our approach is to stabilize the output by taking the median of the 4 model outputs. Yet this does not control variability of each individual model. In $CBCT_{LD\_ResGAN}$, the variability has been reduced due to the generator residual connection.

In future work, we would like to investigate the feasibility of further reducing CBCT dose and explore under-sampling schemes that might provide the opportunity to selectively avoid irradiating critical organs. Besides, we would extend the proposed low dose CBCT imaging technique to other anatomical locations.

5. Conclusion

This study showed that it is possible to reduce the CBCT imaging dose by 75% in pelvic scans while enabling accurate VMAT dose calculation with the use of a cycle-consistent generative adversarial network. The network was successfully trained to simultaneously remove streaking artifacts and translate low dose $CBCT_{LD}$ to CT equivalent images using unpaired training data. The resulting low dose $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$ images resemble planning CTs in HU accuracy and the daily in-room $CBCT_{org}$ in anatomy. Clinically relevant DVH parameters were accurately predicted. $CBCT_{LD\_ResGAN}$ has improved the anatomical fidelity in comparison to $CBCT_{LD\_GAN}$. Compared to the reference technique (vCT), $CBCT_{LD\_GAN}$ and $CBCT_{LD\_ResGAN}$, which allow substantially faster correction and are not affected by DIR uncertainties in the presence of pronounced inter-fractional changes, have thus the potential to be applied for online treatment adaptation.

Acknowledgments

This work was supported by German Research Foundation (DFG) project number 399148265 and Research Training Group GRK 2274.

Data availability statement

The data cannot be made publicly available upon publication because they contain sensitive personal information.

Ethical statement

This retrospective study was exempt from requiring ethics approval. Bavarian state law (Bayrisches Krankenhausgesetz/Bavarian Hospital Law §27 Absatz 4 Datenschutz (Data protection)) allows the use of patient data for research, provided that any personâs related data are kept anonymous. German radiation protection laws request a regular analysis of outcomes in the sense of quality control and assurance, thus in the case of purely retrospective studies no additional ethical approval is needed under German law.

ORCID iDs

Y Chan https://orcid.org/0000-0002-2364-535X
G Landry https://orcid.org/0000-0003-1707-4068

References

Byrne M, Archibald-Heeren B, Hu Y, Teh A, Beserminji R, Cai E, Liu G, Rijken J, Collett N and Aland T 2022 Varian ethos online adaptive radiotherapy for prostate cancer: early results of contouring accuracy, treatment plan quality, and treatment time *J. Appl. Clin. Med. Phys.* **23** e13479

Choi K, Wang J, Zhu L, Suh T-S, Boyd S and Xing L 2010 Compressed sensing based cone-beam computed tomography reconstruction with a first-order method a *Med. Phys.* **37** 5113–25

de Bel T, Bokhorst J-M, van der Laak J and Litjens G 2021 Residual cyclegan for robust domain transformation of histopathological tissue slides *Med. Image Anal.* **70** 102004

de Jong R, Visser J, van Wieringen N, Wiersma J, Geijsen D and Bel A 2021 Feasibility of conebeam ct-based online adaptive radiotherapy for neoadjuvant treatment of rectal cancer *Radiat. Oncol.* **16** 1–11

Deng L, Zhang M, Wang J, Huang S and Yang X 2022 Improving cone-beam ct quality using a cycle-residual connection with a dilated convolution-consistent generative adversarial network *Phys. Med. Biol.* **67** 145010

Ding G X, Alaei P, Curran B, Flynn R, Gossman M, Mackie T R, Miften M, Morin R, Xu X G and Zhu T C 2018 Image guidance doses delivered during radiotherapy: quantification, management, and reduction: report of the aapm therapy physics committee task group 180 *Med. Phys.* **45** e84–e99

Ding G X, Duggan D M, Coffey C W, Deeley M, Hallahan D E, Cmelak A and Malcolm A 2007 A study on adaptive imrt treatment planning using kv cone-beam CT *Radiother. Oncol.* **85** 116–25

Fotina I, Hopfgartner J, Stock M, Steininger T, Lütgendorf-Caucig C and Georg D 2012 Feasibility of CBCT-based dose calculation: comparative analysis of hu adjustment techniques *Radiother. Oncol.* **104** 249–56

Ge Y, Wei D, Xue Z, Wang Q, Zhou X, Zhan Y and Liao S 2019 Unpaired MR to CT synthesis with explicit structural constrained adversarial learning *2019 IEEE 16th Int. Symp. on Biomedical Imaging (ISBI 2019)* (IEEE) pp 1096–99

Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial nets *Advances in Neural Information Processing Systems* vol 27

Hansen D C, Landry G, Kamp F, Li M, Belka C, Parodi K and Kurz C 2018 Scatternet: a convolutional neural network for cone-beam ct intensity correction *Med. Phys.* **45** 4916–26

Harms J, Lei Y, Wang T, Zhang R, Zhou J, Tang X, Curran W J, Liu T and Yang X 2019 Paired cycle-gan-based image correction for quantitative cone-beam computed tomography *Med. Phys.* **46** 3998–4009

Hofmaier J *et al* 2017 Multi-criterial patient positioning based on dose recalculation on scatter-corrected cbct images *Radiother. Oncol.* **125** 464–9

Hyer D E and Hintenlang D E 2010 Estimation of organ doses from kilovoltage cone-beam CT imaging used during radiotherapy patient position verification *Med. Phys.* **37** 4620–6

Isola P, Zhu J-Y, Zhou T and Efros A A 2017 Image-to-image translation with conditional adversarial networks *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition* pp 1125–34

Johnson J, Alahi A and Fei-Fei L 2016 Perceptual losses for real-time style transfer and super-resolution *European Conf. on Computer Vision* (Springer) pp 694–711

Kan M W, Leung L H, Wong W and Lam N 2008 Radiation dose from cone beam computed tomography for image-guided radiation therapy *Int. J. Radiat. Oncol.* Biol.* Phys.* **70** 272–9

Kida S, Kaji S, Nawa K, Imae T, Nakamoto T, Ozaki S, Ohta T, Nozawa Y and Nakagawa K 2019 Visual enhancement of Cone-beam CT by use of CycleGAN *Med. Phys.* **47** 998–1010

Kida S, Nakamoto T, Nakano M, Nawa K, Haga A, Kotoku J, Yamashita H and Nakagawa K 2018 Cone beam computed tomography image quality improvement using a deep convolutional neural network *Cureus* **10** e2548

Kim D W, Chung W K and Yoon M 2013 Imaging doses and secondary cancer risk from kilovoltage cone-beam ct in radiation therapy *Health Phys.* **104** 499–503
Kurz C, Dedes G, Resch A, Reiner M, Ganswindt U, Nijhuis R, Thieke C, Belka C, Parodi K and Landry G 2015 Comparing cone-beam ct intensity correction methods for dose recalculation in adaptive intensity-modulated photon and proton therapy for head and neck cancer *Acta Oncol.* **54** 1651–7
Kurz C *et al* 2016 Investigating deformable image registration and scatter correction for cbct-based dose calculation in adaptive impt *Med. Phys.* **43** 5635–46
Kurz C, Maspero M, Savenije M H, Landry G, Kamp F, Pinto M, Li M, Parodi K, Belka C and Van den Berg C A 2019 Cbct correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation *Phys. Med. Biol.* **64** 225004
Landry G *et al* 2014 Phantom based evaluation of CT to CBCT image registration for proton therapy dose recalculation *Phys. Med. Biol.* **60** 595
Landry G, Hansen D, Kamp F, Li M, Hoyle B, Weller J, Parodi K, Belka C and Kurz C 2019 Comparing unet training with three different datasets to correct cbct images for prostate radiotherapy dose calculations *Phys. Med. Biol.* **64** 035011
Landry G *et al* 2015 Investigating ct to cbct image registration for head and neck proton therapy as a tool for daily dose recalculation *Med. Phys.* **42** 1354–66
Lee H, Xing L, Davidi R, Li R, Qian J and Lee R 2012 Improved compressed sensing-based cone-beam ct reconstruction using adaptive prior image constraints *Phys. Med. Biol.* **57** 2287
Liang X, Chen L, Nguyen D, Zhou Z, Gu X, Yang M, Wang J and Jiang S 2019 Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using cyclegan for adaptive radiation therapy *Phys. Med. Biol.* **64** 125002
Liu Y, Chen X, Zhu J, Yang B, Wei R, Xiong R, Quan H, Liu Y, Dai J and Men K 2022 A two-step method to improve image quality of cbct with phantom-based supervised and patient-based unsupervised learning strategies *Phys. Med. Biol.* **67** 084001
Maier J, Berker Y, Sawall S and Kachelrieß M 2018 Deep scatter estimation (DSE): feasibility of using a deep convolutional neural network for real-time x-ray scatter prediction in cone-beam CT *Medical Imaging 2018: Physics of Medical Imaging*105731L
Maier J, Eulig E, Vöth T, Knaup M, Kuntz J, Sawall S and Kachelrieß M 2019 Real-time scatter estimation for medical ct using the deep scatter estimation: method and robustness analysis with respect to different anatomies, dose levels, tube voltages, and data truncation *Med. Phys.* **46** 238–49
Mainegra-Hing E and Kawrakow I 2010 Variance reduction techniques for fast monte carlo cbct scatter correction calculations *Phys. Med. Biol.* **55** 4495
Marks L B, Yorke E D, Jackson A, Ten Haken R K, Constine L S, Eisbruch A, Bentzen S M, Nam J and Deasy J O 2010 Use of normal tissue complication probability models in the clinic *Int. J. Radiat. Oncol.* Biol.* Phys.* **76** S10–9
Moazzezi M, Rose B, Kisling K, Moore K L and Ray X 2021 Prospects for daily online adaptive radiotherapy via ethos for prostate cancer patients without nodal involvement using unedited cbct auto-segmentation *J. Appl. Clin. Med. Phys.* **22** 82–93
Niu T, Al-Basheer A and Zhu L 2012 Quantitative cone-beam CT imaging in radiation therapy using planning ct as a prior: first patient studies *Med. Phys.* **39** 1991–2000
Niu T, Sun M, Star-Lack J, Gao H, Fan Q and Zhu L 2010 Shading correction for on-board cone-beam ct in radiation therapy using planning MDCT images *Med. Phys.* **37** 5395–406
Park J C, Song B, Kim J S, Park S H, Kim H K, Liu Z, Suh T S and Song W Y 2012 Fast compressed sensing-based cbct reconstruction using barzilai-borwein formulation for application to on-line igrt *Med. Phys.* **39** 1207–17
Park Y-K, Sharp G C, Phillips J and Winey B A 2015 Proton dose calculation on scatter-corrected cbct image: feasibility study for adaptive proton therapy *Med. Phys.* **42** 4449–59
Peroni M, Ciardo D, Spadea M F, Riboldi M, Comi S, Alterio D, Baroni G and Orecchia R 2012 Automatic segmentation and online virtualct in head-and-neck adaptive radiation therapy *Int. J. Radiat. Oncol.* Biol.* Phys.* **84** e427–33
Rit S, Oliva M V, Brousmiche S, Labarbe R, Sarrut D and Sharp G C 2014 The reconstruction toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the insight toolkit (ITK) *J. Phys. Conf. Ser.* **489** 012079
Ronneberger O, Fischer P and Brox T 2015 U-net: convolutional networks for biomedical image segmentation *Int. Conf. on Medical Image Computing and Computer-assisted Intervention* (Springer) pp 234–241
Sibolt P, Andersson L M, Calmels L, Sjöström D, Bjelkengren U, Geertsen P and Behrens C F 2021 Clinical implementation of artificial intelligence-driven cone-beam computed tomography-guided online adaptive radiotherapy in the pelvic region *Phys. Imaging Radiat. Oncol.* **17** 1–7
Song B, Park J C and Song W Y 2014 A low-complexity 2-point step size gradient projection method with selective function evaluations for smoothed total variation based cbct reconstructions *Phys. Med. Biol.* **59** 6565
Thing R S, Bernchou U, Mainegra-Hing E, Hansen O and Brink C 2016 Hounsfield unit recovery in clinical cone beam ct images of the thorax acquired for image guided radiation therapy *Phys. Med. Biol.* **61** 5781
Ulyanov D, Vedaldi A and Lempitsky V 2016 Instance normalization: the missing ingredient for fast stylization arXiv:1607.08022
Veiga C, Alshaikhi J, Amos R, Lourenćo A M, Modat M, Ourselin S, Royle G and McClelland J R 2015 Cone-beam computed tomography and deformable registration-based 'dose of the day' calculations for adaptive proton therapy *Int. J. Part. Ther.* **2** 404–14
Veiga C *et al* 2016 First clinical investigation of cone beam computed tomography and deformable registration for adaptive proton therapy for lung cancer *Int. J. Radiat. Oncol.* Biol.* Phys.* **95** 549–59
Veiga C, McClelland J, Moinuddin S, Lourenço A, Ricketts K, Annkah J, Modat M, Ourselin S, D'Souza D and Royle G 2014 Toward adaptive radiotherapy for head and neck patients: feasibility study on using CT-to-CBCT deformable registration for 'dose of the day' calculations *Med. Phys.* **41** 031703
Wang P *et al* 2016 Quantitative assessment of anatomical change using a virtual proton depth radiograph for adaptive head and neck proton therapy *J. Appl. Clin. Med. Phys.* **17** 427–40
Yuan N, Dyer B, Rao S, Chen Q, Benedict S, Shang L, Kang Y, Qi J and Rong Y 2020 Convolutional neural network enhancement of fast-scan low-dose cone-beam ct images for head and neck radiotherapy *Phys. Med. Biol.* **65** 035003
Zhu J-Y, Park T, Isola P and Efros A A 2017 Unpaired image-to-image translation using cycle-consistent adversarial networks *Proc. of the IEEE Int. Conf. on Computer Vision* pp 2223–32
Zöllner C, Rit S, Kurz C, Vilches-Freixas G, Kamp F, Dedes G, Belka C, Parodi K and Landry G 2017 Decomposing a prior-ct-based cone-beam ct projection correction algorithm into scatter and beam hardening components *Phys. and Imaging Radiat. Oncol.* **3** 49–52

8.2 Paper 2: Investigation of imaging dose reduction levels in pelvic CBCT-based synthetic CT generation using deep learning: dose calculation and segmentation accuracy

Reprinted with permission from "Minimum imaging dose for deep learning-based pelvic synthetic computed tomography generation from cone beam images." by Yan Chi Ivy Chan, Minglun Li, Adrian Thummerer, Katia Parodi, Claus Belka, Christopher Kurz, and Guillaume Landry; Physics and Imaging in Radiation Oncology (2024), 30, p.100569.

Physics and Imaging in Radiation Oncology 30 (2024) 100569

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/physics-and-imaging-in-radiation-oncology

Original Research Article

Minimum imaging dose for deep learning-based pelvic synthetic computed tomography generation from cone beam images

Yan Chi Ivy Chan [a], Minglun Li [a,b], Adrian Thummerer [a], Katia Parodi [c], Claus Belka [a,d,e], Christopher Kurz [a,1], Guillaume Landry [a,*,1]

[a] Department of Radiation Oncology, LMU University Hospital, LMU Munich, Munich 81377, Germany
[b] Department of Radiation Oncology, Lueneburg Hospital, Lueneburg 21339, Germany
[c] Department of Medical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München, Garching b. München 85748, Germany
[d] German Cancer Consortium (DKTK), partner site Munich, a partnership between DKFZ and LMU University Hospital Munich, Munich 81377, Germany
[e] Bavarian Cancer Research Center (BZKF), Munich 81377, Germany

ARTICLE INFO

Keywords:
Deep learning
Generative networks
Synthetic CT
Low dose CBCT
Online adaptation
cycleGAN
Contrastive unpaired translation

ABSTRACT

Background and purpose: Daily cone-beam computed tomography (CBCT) in image-guided radiotherapy administers radiation exposure and subjects patients to secondary cancer risk. Reducing imaging dose remains challenging as image quality deteriorates. We investigated three imaging dose levels by reducing projections and correcting images using two deep learning algorithms, aiming at identifying the lowest achievable imaging dose.
Materials and methods: CBCTs were reconstructed with 100%, 25%, 15% and 10% projections. Models were trained (30), validated (3) and tested (8) with prostate cancer patient datasets. We optimized and compared the performance of 1) a cycle generative adversarial network (cycleGAN) with residual connection and 2) a contrastive unpaired translation network (CUT) to generate synthetic computed tomography (sCT) from reduced imaging dose CBCTs. Volumetric modulated arc therapy plans were optimized on a reference intensity-corrected full dose $CBCT_{cor}$ and recalculated on sCTs. Hounsfield unit (HU) and positioning accuracy were evaluated. Bladder and rectum were manually delineated to determine anatomical fidelity.
Results: All sCTs achieved average mean absolute mean absolute error/structural similarity index measure/peak signal-to-noise ratio of $\leqslant$59HU/$\geqslant$0.94/$\geqslant$33 dB. All dose-volume histogram parameter differences were within 2 Gy or 2%. Positioning differences were $\leqslant$0.30 mm or 0.30°. cycleGAN with Dice similarity coefficients (DSC) for bladder/rectum of $\geqslant$0.85/$\geqslant$0.81 performed better than CUT ($\geqslant$0.83/$\geqslant$0.76). A significantly lower DSC accuracy was observed for 15% and 10% sCTs. cycleGAN performed better than CUT for contouring, however both yielded comparable outcomes in other evaluations.
Conclusion: sCTs based on different CBCT doses using cycleGAN and CUT were investigated. Based on segmentation accuracy, 25% is the minimum imaging dose.

1. Introduction

In image-guided radiotherapy (IGRT), deep learning (DL) algorithms have been widely employed to enhance radiotherapy treatments. Particularly for the pelvic region, where the anatomy exhibits inter- and intra-fractional variations, the adaptive workflow relies on accurate cone beam computed tomography (CBCT)-to-CT translation [1,2], and organ segmentation on synthetic CTs (sCT) [3,4]. CBCT imaging dose has often been disregarded, viewed as negligible compared to the therapeutic dose. However, studies suggested that daily CBCT potentially results in considerable additional organ doses in the pelvic region [5–7]. Each pelvic scan can deliver up to 22.7 mSv effective dose [8].

Adhering to "as low as reasonably achievable" (ALARA), radiation oncologists use the lowest possible imaging dose or restrict the frequency of CBCT scans to reduce secondary cancer risk. Further reducing imaging dose, however, remains impractical since the image quality would degrade to unusable levels with potential loss of anatomical information. Lower imaging dose CBCTs with enhanced image quality

* Corresponding author at: Marchioninistr. 15, 81377 Munich, Germany.
E-mail address: Guillaume.Landry@med.uni-muenchen.de (G. Landry).
[1] Senior authorship is equally shared.

https://doi.org/10.1016/j.phro.2024.100569
Received 26 January 2024; Received in revised form 8 March 2024; Accepted 11 March 2024
Available online 22 March 2024

could not only mitigate the secondary cancer risk concern, but also offer a higher flexibility in terms of in-room imaging frequency and enable online treatment dose adaptation. With sufficient sCT quality, one could also avoid acquiring new planning CTs for plan adaptation, thus further reducing imaging dose.

DL-enabled CBCT-to-CT translation has mostly been developed for standard full dose CBCT. Three DL architectures have been applied to pelvic scans: U-Net [9], cycleGAN [10] and contrastive unpaired translation (CUT) [11]. U-Nets were trained with paired data in image [12–14] or projection domain [14–17]. To overcome potential misalignments, cycleGAN has been used for unpaired training [18–23]. In recent studies [24–26], CUT demonstrated better performance over cycleGAN. Treatment dose calculation on CUT, however, remained unexplored.

Limited studies explored the possibility of using low imaging dose CBCT. Our previous study [27] investigated the feasibility of removing under-sampling artifacts and correcting intensities of 25% imaging dose CBCT using cycleGAN. sCT from 25% imaging dose CBCT (≈0.6 mGy) showed high accuracy in therapeutic photon dose calculation, anatomical fidelity (in terms of Dice similarity coefficient (DSC) and Hausdorff distance (HD) of contours) and positioning. In [24], cycleGAN and CUT removed streaks from 4D CBCT which is comparable to low dose CBCT. There are a few low dose CBCT-to-CT studies in other anatomies [28,29]. Among low dose CBCT-to-CT studies, there is a scarcity of systematic investigation of the maximum imaging dose reduction level that DL could offer. In most CBCT-to-CT studies, organ segmentation is rarely evaluated except [18,21,22] in pelvic and [30] head and neck region.

In this study, we aim at finding the achievable lowest imaging dose using cycleGAN and CUT in terms of all metrics relevant to CBCT-guided adaptive radiotherapy: image quality, positioning, organs-at-risk (OAR) contouring accuracy and therapeutic photon dose calculation. We investigated imaging dose levels in terms of sCT generation from a CBCT with reduced number of projections (25%, 15% and 10%) by removing under-sampling artefacts and correcting image intensities. Dose reduction is achieved via the reduction of the number of projections.

2. Materials and methods

The workflow of CBCT restoration at different dose levels is illustrated in Fig. 1. In general, imaging dose was reduced by retroactively reducing the number of acquired projections.

2.1. Patient database for model training

The database contained 41 prostate cancer patients who received volumetric modulated arc therapy (VMAT) at Department of Radiation Oncology of LMU University Hospital. One planning CT (pCT) acquired before treatment course and CBCT images of one arbitrary fraction of each patient were collected.

Bavarian state law (Bayrisches Krankenhausgesetz/Bavarian Hospital Law §27 Absatz 4 Datenschutz) allows the use of patient data for research, provided that any person's related data are kept anonymous. All patient data were fully anonymised. Identification from pelvic CT data is not possible. German radiation protection laws request a regular analysis of outcomes in the sense of quality control and assurance, thus in the case of purely retrospective studies no additional ethical approval is needed under German law.

Only CBCT datasets acquired with the lowest dose pelvic protocol (120 kV tube voltage, 20 ms exposure time, 20 mA X-ray tube current per projection) in treatment position using the XVI system (version 5.52) of a Synergy medical linear accelerator (Elekta, Sweden) were selected. For each fully sampled (FS) scan, approximately 350 projections [346, 357] were acquired over 360° with a shifted panel and reconstructed using Feldkamp–Davis–Kress (FDK) implementation of Reconstruction ToolKit (RTK) [31], referred to as $CBCT_{FS}$. To investigate achievable dose reduction levels, CBCTs were under-sampled to 25% (~90 projections), 15% (~52 projections) and 10% (~35 projections) and reconstructed using the same settings. Since each projection was acquired with a fixed dose, reducing the number of projections results in a dose reduction.

pCTs were acquired without contrast agent on a Toshiba Acquilion LB CT scanner (Canon Medical Systems, Japan). A virtual CT (vCT) was generated using a dedicated deformable image registration (DIR) algorithm mapping the pCT onto the daily $CBCT_{FS}$ [32]. For reference, a intensity-corrected $CBCT_{cor}$ was generated using a projection-based

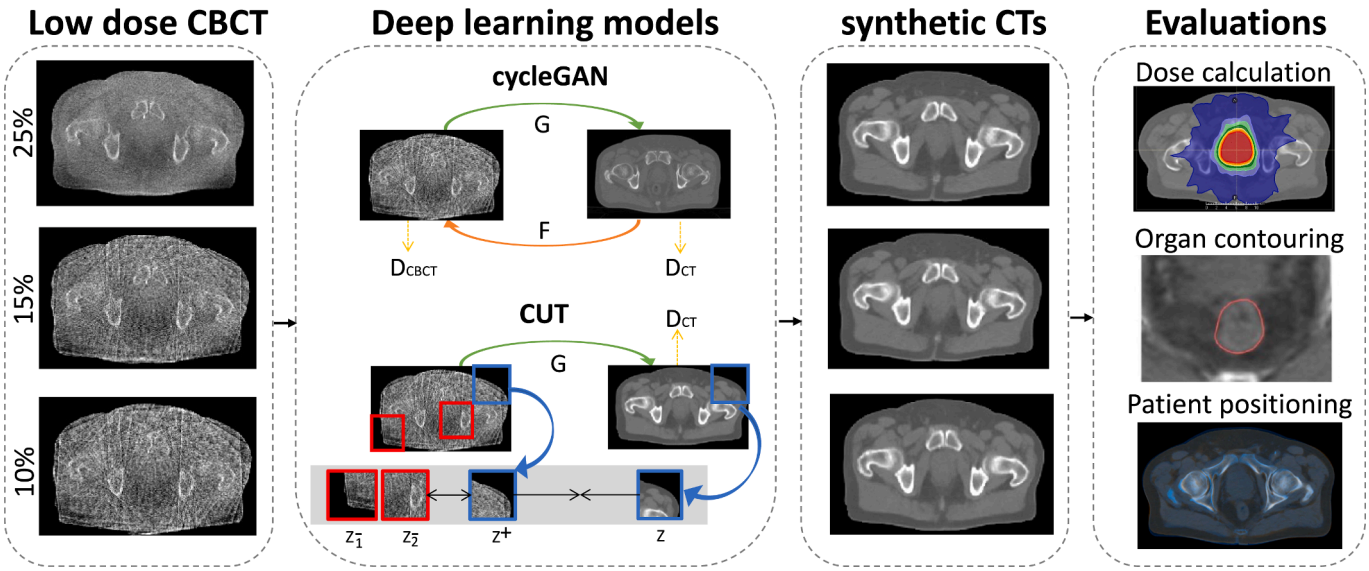


Fig. 1. Workflow of CBCT restoration at different dose reduction levels investigated in this study. Low dose CBCTs were set as inputs in the cycleGAN and CUT algorithms to generate sCTs, which were then evaluated by means of patient positioning, dose calculation and contouring accuracy. The black arrows denote the sequence of the investigation steps in this study (low dose CBCT generation, deep learning model training, sCT generation, and finally the evaluation of therapeutic photon dose calculation, contouring and positioning). Models were trained separately for each dose reduction level. G and F denote generators, D_{CBCT} and D_{CT} denote discriminators, z denotes the image patches used in CUT.

scatter correction technique [14,17,33–36] based on DIR of pCT to CBCT and forward projection followed by conjugate gradient iterative reconstruction [31].

CBCTs, vCTs and $CBCT_{cor}$ were padded to an axial size of 512 × 512 pixels (1 mm × 1 mm) with a slice thickness of 1 mm. Details of the data acquisitions and pre-processing can be found in the Supplementary material.

2.2. Deep learning algorithms

Low dose CBCT-to-CT translation can be formulated as:

$$sCT = G(CBCT_{LD})$$

where G is an encoder-decoder based model that simultaneously converts $CBCT_{LD}$ to sCT while preserving the anatomical content. In this study, cycleGAN and CUT algorithms were employed to train G.

2.2.1. cycleGAN

We applied the cycleGAN algorithm that was implemented in a previous study [27]. This training process involved two sets of generator and discriminator networks. A cycle consistency loss (L_{cyc}) is computed to stabilise anatomical mappings between CBCT and CT. A residual skip connection was used for both generators to attain higher anatomical fidelity. vCT was used in the training instead of pCT to evaluate the efficacy of an additional paired loss term.

2.2.2. CUT

We adapted the CUT algorithm proposed by Park et al.[11]. Only one set of generator and discriminator is required, since L_{cyc} is replaced by a patchwise contrastive loss ($L_{PatchNCEx}$, $L_{PatchNCEy}$). As shown in Fig. 1, a sCT patch should match more with its corresponding input CBCT patch (denoted as positives), in comparison with other random CBCT patches (denoted as negatives). The encoder part of the generator (G_{enc}) followed by a two-layer multilayer perceptron (MLP) network is employed, which allows the model to learn and project both patches to a shared feature embedding space.

Training of each model used identical data, pre-processing and data augmentation. Details are provided in Supplementary material.

2.3. Training details

For each of the CBCT dose reduction levels, cycleGAN and CUT models were trained with 4-fold cross-validation with 25 out of 30 patients per fold, from which the median of the four predicted images was used. We determined hyper-parameters for each model through ensemble validation on three patient datasets. Subsequently, we preserved the model weights associated with the highest validation performance and applied them for testing.

The test set consisted of 8 patient datasets. The generators for every imaging dose level were applied to convert CBCTs into sCTs. Details are provided in Supplementary material.

2.4. Evaluation

2.4.1. Image quality

sCTs of different imaging dose levels for the test set were compared to $CBCT_{cor}$ in terms of the mean absolute error (MAE), mean error (ME), structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR). Only voxels within the joint body outline of $CBCT_{cor}$ and sCTs were included.

2.4.2. Treatment dose calculation

VMAT plans on $CBCT_{cor}$ for the test patients were generated in a research version of a commercial treatment planning system (TPS) (RayStation, version 10.01, RaySearch, Sweden). Contours of target structures and OAR were transferred from the pCT to sCTs and $CBCT_{cor}$ using DIR, VMAT plans were optimized on an isotropic dose grid of 3.0 mm using a collapsed-cone dose engine. These plans were then recalculated on all sCTs. The prescription was 74 Gy in 37 fractions and we aimed at a clinical target volume (CTV) $V_{95\%}$ of 100%, and planning target volume (PTV) $V_{95\%}$ better than 95% of the prescription dose. The dose-volume histogram (DVH) constraints for the bladder and the rectum were pursued as suggested in the QUANTEC report [37]. The VMAT dose distributions were compared with the $CBCT_{cor}$ reference considering DVH parameters of clinically relevant target structures and OAR. CTV and PTV $D_{98\%}$ and $D_{2\%}$, together with PTV $D_{50\%}$ and $V_{95\%}$ were analyzed. For the bladder $V_{60/65\,Gy}$ and for the rectum $V_{50/60/65\,Gy}$ were analyzed. Moreover, the voxels passing a therapeutic dose difference (DD) analysis with a 1% and 2% criterion (10% threshold) were compared. For each dose parameter, results from sCTs were compared to $CBCT_{cor}$ using Wilcoxon signed-rank tests. Similarly, low imaging dose sCTs were compared to FS sCT for both models. P-values less than 0.05 were considered significant.

2.4.3. Segmentation accuracy

To determine the anatomical fidelity of all sCTs, bladder and rectum were contoured manually under the supervision of a radiation oncologist using the research TPS on $CBCT_{FS}$ and sCTs. Please keep in mind that these contours were unrelated to the contours used to generate the treatment plans used for the treatment dose evaluation of Section 2.4.2. DSC, average and 95th percentile HD (HD_{avg}, HD_{95}) of the contours of all sCTs were compared with $CBCT_{FS}$ as reference. sCTs from all imaging dose reduction levels, as well as from both models at the same dose reduction, were statistically analysed using Wilcoxon signed-rank tests.

2.4.4. Positioning accuracy

To evaluate positioning accuracy at different CBCT dose reduction levels, all sCT images were rigidly registered to the pCT using TPS (automated, gray level, six degrees of freedom). The transformations were compared to the one obtained from registering $CBCT_{FS}$ to pCT.

3. Results

The average time to generate a sCT slice from CBCT was 6 ms for both models. Detailed epoch selection and the corresponding training time are shown in Supplementary Table 1.

3.1. Image comparison

Fig. 2 illustrates sCTs, $CBCT_{cor}$ of a representative test patient and their corresponding HU differences. Both cycleGAN and CUT removed streak artifacts from all CBCTs, and simultaneously converted them into diagnostic quality. Compared to inputs, all sCTs show reduced differences to $CBCT_{cor}$. The remaining differences are observed at body outline and bone interfaces. The coronal view is illustrated in Supplementary material Fig. 1.

All metrics are substantially enhanced by both models (Supplementary Table 2). The average MAE of all sCTs with respect to $CBCT_{cor}$ were improved from ⩾102 HU to ⩽59 HU. The average ME of the majority of the sCTs has decreased by ⩾7 HU. SSIM/PSNR were enhanced from ⩽0.91/⩽33 dB on CBCTs to ⩾0.94/⩾33 dB on sCTs.

3.2. Treatment dose calculation

Fig. 3 shows the treatment dose distribution and difference of an exemplary test patient. Compared sCTs to $CBCT_{cor}$, only minor dose differences were found in the PTV region (<3%). The remaining treatment dose differences were mainly in patient outline.

In Fig. 4, target and OAR DVH parameter differences with respect to $CBCT_{cor}$ over all test patients are depicted. Deviations were within 2 Gy

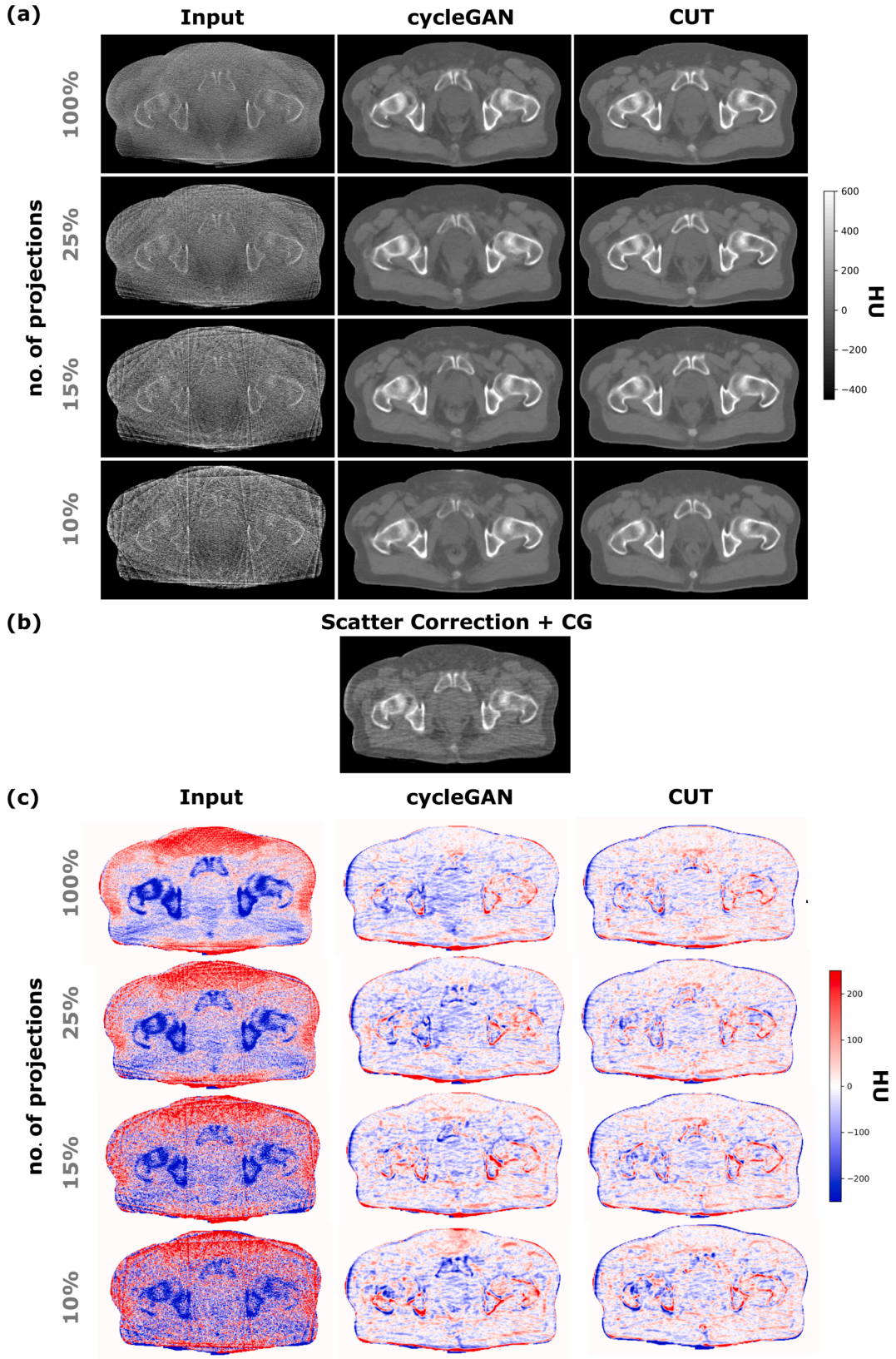


Fig. 2. Axial view of (a) the CBCT inputs and sCTs generated by cycleGAN and CUT with 100%, 25%, 15%, 10% projections; (b) the scatter corrected $CBCT_{cor}$ reconstructed with conjugate gradient (CG); (c) HU difference between inputs and corresponding sCTs with $CBCT_{cor}$ of a test patient.

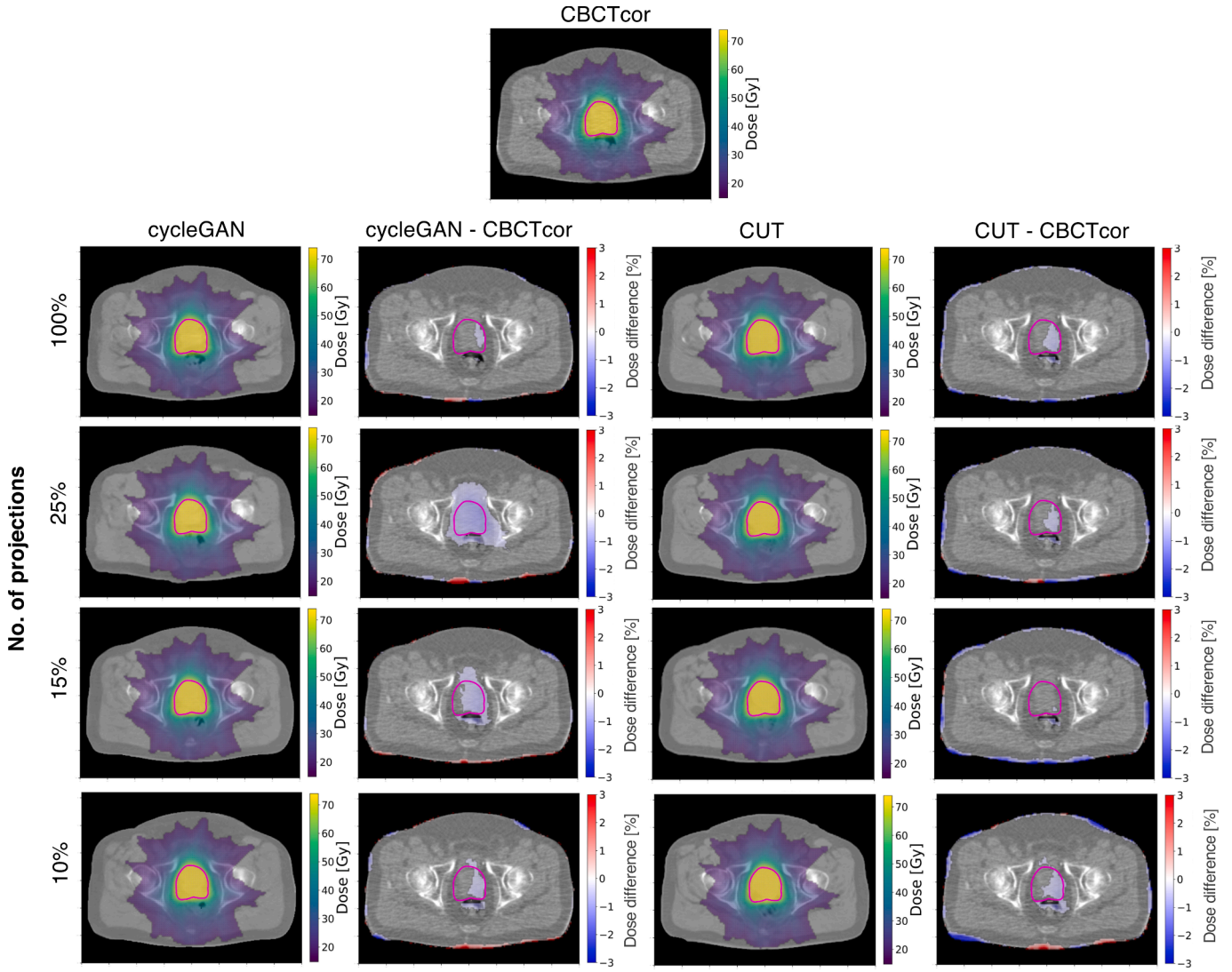


Fig. 3. VMAT dose distributions of an exemplary patient. Dose distributions optimized on $CBCT_{cor}$ and recalculated on sCTs at difference dose reduction levels, and their corresponding differences for cycleGAN and CUT. The PTV is shown in magenta. Dose differences below 0.4% are not shown for better visualization.

for dose DVH parameters ($D_{2/50/98\%}$) and below 2% for volume DVH parameters ($V_{50/60/65\,Gy}$). Particularly in the target DVH comparison, the mean differences of $D_{2/50/98\%}$ comparing all sCTs with respect to $CBCT_{cor}$ were ⩽0.5% for the PTV. For cycleGAN, no significant differences were found in the majority of the low dose sCT, except CTV $D_{2\%}$ and PTV $D_{2\%}$ of the 10% sCT. Statistically significant differences were observed for all FS sCTs, but most magnitudes were constrained by 1 Gy. For CUT, significant differences were observed in 15% and 10% sCTs for most of target and OAR DVH parameters, except rectum $V_{50\%}$ and bladder $V_{65\%}$ for 10% sCTs.

Fig. 5 illustrates the quantitative results of the treatment dose difference analysis of the VMAT plans comparing sCTs to $CBCT_{cor}$ with a 1% criterion. The average 1% DD pass-rates of all sCTs were above 95% for cycleGAN and 97% for CUT. Statistically significant differences were observed comparing the dose reduced sCTs to the FS sCT for CUT, the 10% sCT to the FS sCT for cycleGAN. CycleGAN performed significantly better than CUT for 10% sCT. The average 2% DD pass-rates were higher than 98% for both models, indicating an excellent agreement of all sCTs to the reference $CBCT_{cor}$.

3.3. Anatomical accuracy

Fig. 6 shows the (a-c) bladder and (d-f) rectum contouring results. For bladder, the average DSC was above 0.80 in all sCTs with respect to $CBCT_{FS}$. HD_{avg}/ HD_{95} of bladder were ⩽1.5 mm/⩽8.0 mm for cycleGAN, and ⩽2.0 mm/⩽8.3 mm for CUT in all sCTs. For cycleGAN, significant differences were observed comparing 15% and 10% to FS sCTs in all metrics. For CUT, significant differences only in DSC were observed comparing 15% and 10% to FS sCTs. CycleGAN performed significantly better than CUT in FS and 25% sCT for all metrics.

For rectum, the average DSC was ⩾0.80/⩾0.75 for cycleGAN/CUT in all sCTs. HD_{avg}/ HD_{95} of rectum were ⩽1.7 mm/⩽6.0 mm for cycleGAN, and ⩽1.9 mm/⩽6.6 mm for CUT. For cycleGAN, significant differences were observed comparing 15% and 10% to FS sCTs in all metrics. For CUT, significant differences were observed comparing 25%, 15% and 10% to FS sCTs. CycleGAN performed significantly better than CUT in FS, 25% and 10% sCT for all metrics except HD_{95}.

In both organs, FS sCT has the highest DSC, lowest HD_{avg} and HD_{95} among all sCTs for cycleGAN and CUT. In addition, bladder had higher DSC and lower HD_{avg} and HD_{95} than rectum in both models.

3.4. Positioning accuracy

Compared to $CBCT_{FS}$-to-pCT rigid registration, the mean absolute differences of rigid transformation parameters in all sCTs-to-pCT registrations were less than 0.30 mm or 0.30° for both models, demonstrating

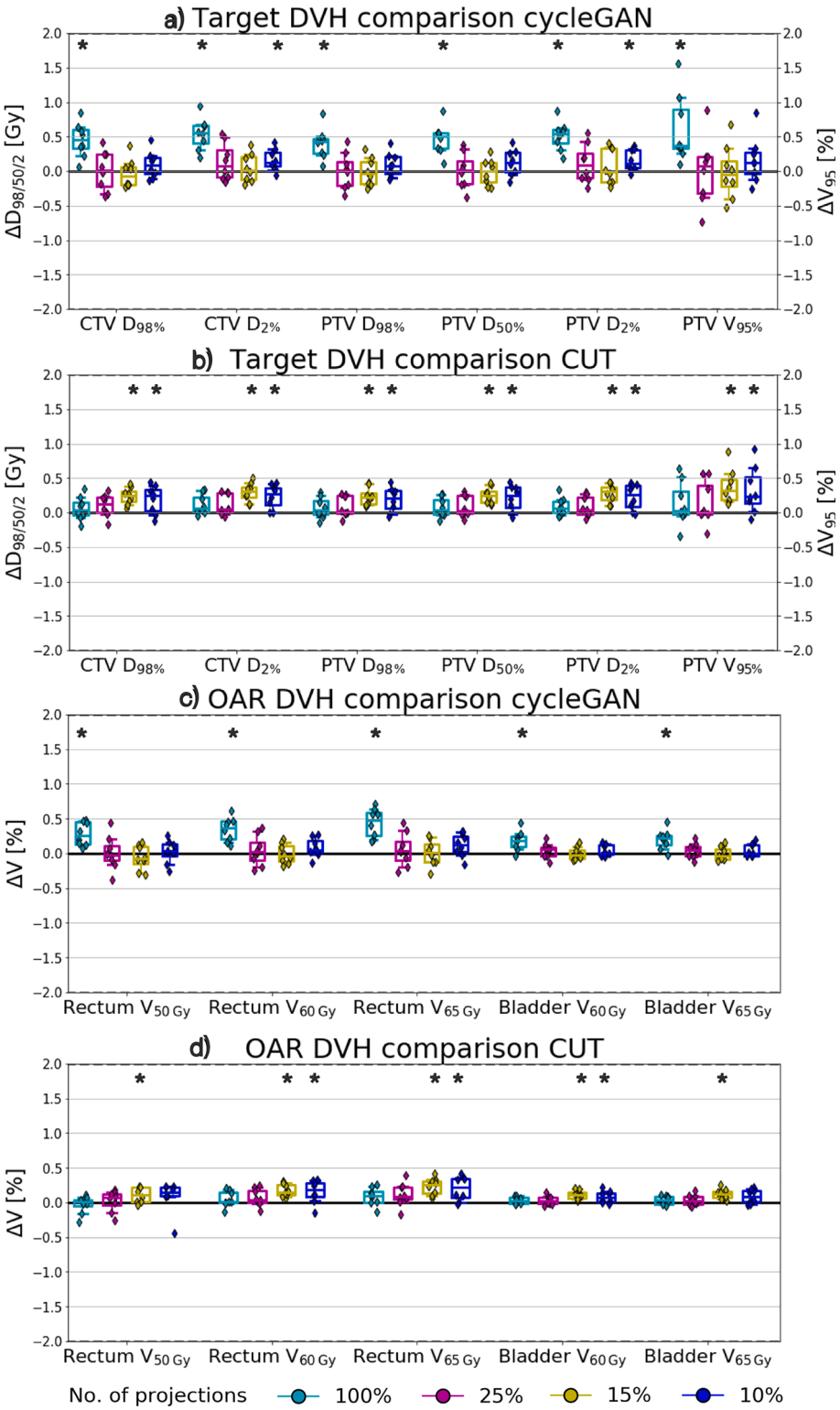


Fig. 4. Clinically relevant DVH parameter differences of sCTs in different dose reduction levels with respect to $CBCT_{cor}$ for target and OAR structures using (a, c) cycleGAN and (b, d) CUT. Each data point represents a test patient. Whiskers correspond to the 5th–95th percentile. The box denotes the interquartile range, and a horizontal line inside the box is used to represent the median. All values correspond to the total treatment dose of the fractionated treatment. Significant difference between sCT and $CBCT_{cor}$ is indicated by a star (p-value <0.05).

sCTs from the investigated low dose CBCT have the potential to align patients accurately. Detailed results are provided in Supplementary material Table 3.

4. Discussion

We investigated sCT generation based on different CBCT imaging dose reduction levels achieved by reducing the number of projections using cycleGAN and CUT, and evaluated image quality, dose calculation

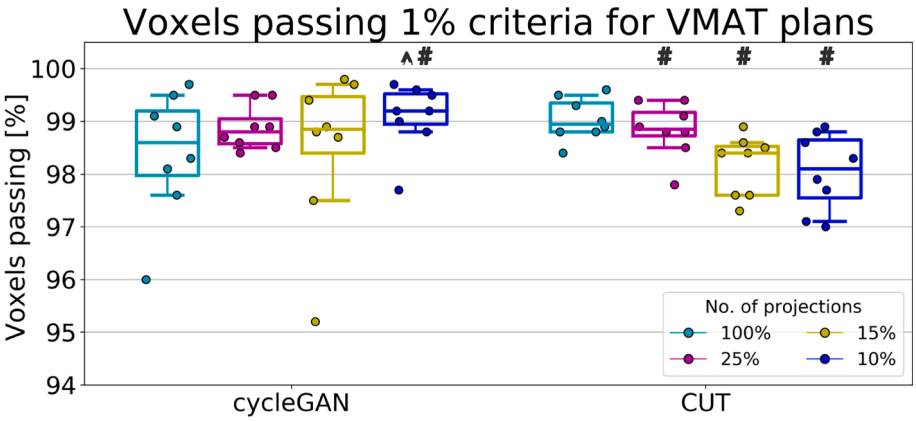


Fig. 5. Voxels passing a 1% dose difference criterion in the eight test patients for the VMAT plans recalculated on sCTs from different dose reduction levels using cycleGAN and CUT with respect to $CBCT_{cor}$. Each data point represents a test patient. Whiskers correspond to the 5th–95th percentile. The box denotes the interquartile range, and a horizontal line inside the box is used to represent the median. Significant difference comparing dose reduced sCT to FS sCT is indicated by a hash sign (p-value <0.05). Significant difference comparing cycleGAN and CUT for identical dose sCT is indicated by a circumflex (p-value <0.05).

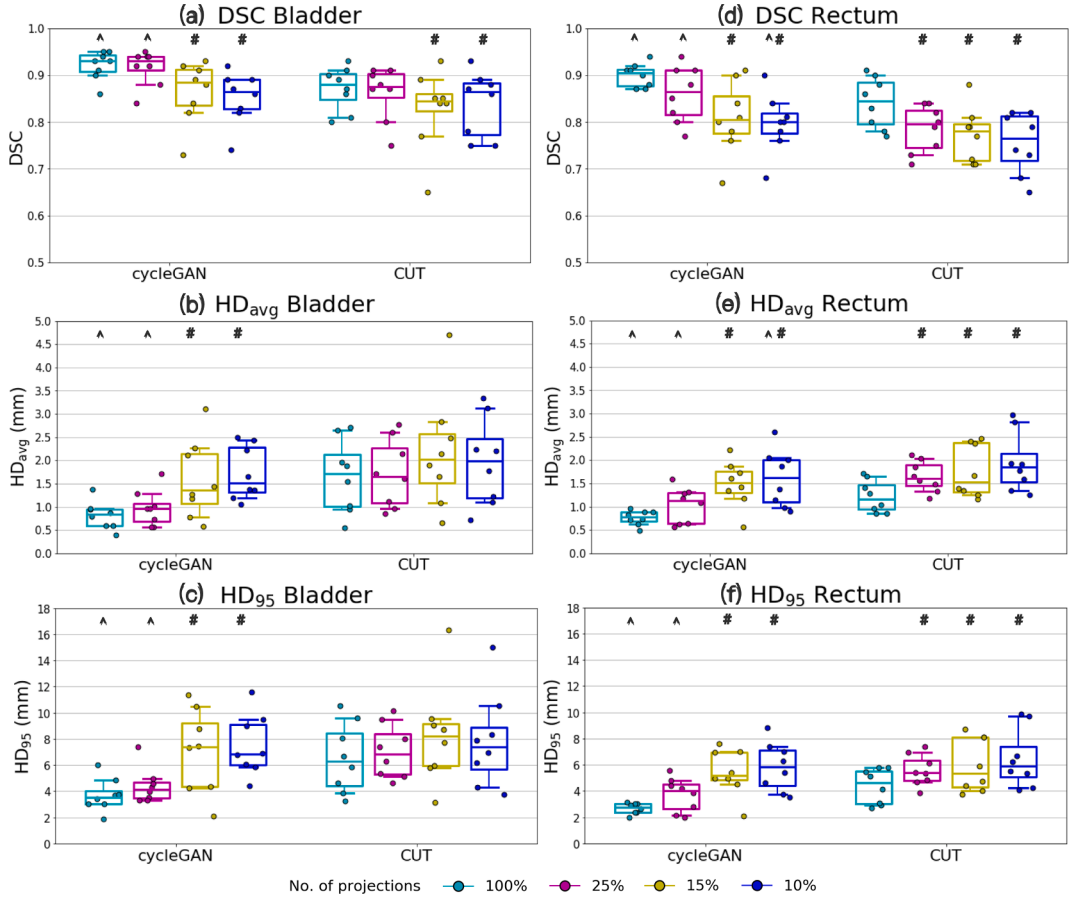


Fig. 6. The anatomical fidelity results of (a–c) bladder and (d–f) rectum in terms of Dice similarity coefficient (DSC), average and 95th percentile Hausdorff distance (HD_{avg}, HD_{95}) comparing sCTs with $CBCT_{FS}$ in the test patients. Each data point represents a test patient. Whiskers correspond to the 5th–95th percentile. The box denotes the interquartile range, and a horizontal line inside the box is used to represent the median. Significant difference comparing dose reduced sCT to FS sCT is indicated by a hash sign (p-value <0.05). Significant difference comparing cycleGAN and CUT for identical dose sCT is indicated by a circumflex (p-value <0.05).

and organ segmentation accuracy. The CBCT inputs were initially reconstructed with 100%, 25%, 15% and 10% projections. Our primary objective was to determine the largest imaging dose reduction without loss of accuracy.

Over the evaluation metrics for image quality, treatment dose calculation and positioning accuracy, limited deviations were noted among all imaging dose reduction levels. However, organ segmentation showed differentiation among the sCTs for both models. From the DSC, HD_{avg} and HD_{95} results (Fig. 6), the performance drops at 15% dose, where the cycleGAN model began to exhibit degraded accuracy in generating accurate bladder and rectum shapes, as observed from the significant differences appearing at 15% and 10% dose sCT compared to $CBCT_{FS}$. For CUT, the performance declined at 15% dose for the bladder, and already at 25% for the rectum. In addition, as revealed by the

significant differences in all metrics for organ contours, cycleGAN performed slightly better than CUT in FS and 25% dose, while in further reduced dose levels both models demonstrated similar inferior performance.

Compared to previous pelvic sCT segmentation studies [18,21,22], this is the first time sCTs based on CBCTs at different imaging dose levels were compared. A direct comparison to other studies is not easily possible because the imaging dose of CBCT testing data is different. However, mean DSC of bladder and rectum for FS sCT, from cycleGAN (0.92, 0.90) and CUT (0.88, 0.84) agree with the other studies (0.89–0.92, 0.81–0.87) [18,21,22]. No rectum DSC was reported in [18].

Regarding therapeutic dose accuracy, DVH parameters difference were within 2 Gy or 2%, which aligned with previous studies using cycleGAN [20,27]. Photon dose calculation using CUT is however not yet found in other studies. For CUT, the drop of performance at 15% sCTs was manifested for OAR DVH parameters. Using 1% DD criteria, we observed a significant decrease of voxels passing at 10% for cycleGAN and 25% for CUT. However all values were still above 97%, indicating high agreement which allowed accurate dose calculation.

Regarding image quality, both models substantially enhanced all CBCTs and CUT performed slightly better than cycleGAN. Coronal views showed slight jittering in the internal organs along slices, since the training was conducted in 2D. Compared to [24], our sCTs yielded higher PSNR and SSIM but higher MAE. These differences are mainly due to the use of a deformed CT as reference in [24], which might have more uncertainties from DIR but less scatter noise than our reference $CBCT_{cor}$.

While this study illustrated minimum imaging dose at 25%, it is limited by the number of patients in the test datasets. Besides, the models are not anatomical-site-agnostic as only pelvic datasets were used. Moreover, DL-generated images may suffer from anatomical inaccuracies. Despite the use of $L_{PatchNCEx}$ and $L_{PatchNCEy}$ in CUT, accurately predicting organs, particularly in those with variable shapes like rectum, remains challenging. Low dose CBCTs can yield high positioning accuracy [38] or a small dosimetric deviation using a water-density override. However, it is still meaningful to generate sCTs which enable organ contouring for adaptation.

Unlike prior DL-enabled CBCT-to-CT works [12–16,18–21,23–26], this study investigated anatomical fidelity in sCT by manual OAR contouring. This aspect revealed a performance threshold for imaging dose reduction. Our results suggest that a CBCT imaging dose as low as 25% is clinically feasible, enabled by the optimized cycleGAN or CUT model. Further reduction to 15% or 10% requires additional DL advancements.

sCTs based on different CBCT imaging dose reduction levels (100%, 25%, 15% and 10%) using cycleGAN and CUT were investigated. While all sCTs demonstrated very good dosimetric, HU and positioning accuracy for both models, considerable differences were found in terms of contouring accuracy. In line with all evaluations, 25% is the minimum imaging dose without loss of anatomical accuracy.

CRediT authorship contribution statement

Yan Chi Ivy Chan: Conceptualization, Data curation, Methodology, Investigation, Software, Formal analysis, Writing – original draft, Visualization. **Minglun Li:** Data curation, Writing – review & editing. **Adrian Thummerer:** Methodology, Software, Writing – review & editing. **Katia Parodi:** Writing – review & editing. **Claus Belka:** Writing – review & editing. **Christopher Kurz:** Conceptualization, Data curation, Methodology, Software, Writing – review & editing, Supervision. **Guillaume Landry:** Conceptualization, Data curation, Methodology, Software, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The Department of Radiation Oncology of LMU University Hospital of LMU Munich has a research agreement with Elekta.

Acknowledgments

This work was supported by German Research Foundation (project 399148265) and Research Training Group GRK2274.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.phro.2024.100569.

References

[1] de Jong R, Visser J, van Wieringen N, Wiersma J, Geijsen D, Bel A. Feasibility of cone beam CT-based online adaptive radiotherapy for neoadjuvant treatment of rectal cancer. Radiat Oncol J 2021;16:1–11. https://doi.org/10.1186/s13014-021-01866-7.
[2] Sibolt P, Andersson LM, Calmels L, Sjöström D, Bjelkengren U, Geertsen P, et al. Clinical implementation of artificial intelligence-driven cone-beam computed tomography-guided online adaptive radiotherapy in the pelvic region. Phys Imaging Radiat Oncol 2021;17:1–7. https://doi.org/10.1016/j.phro.2020.12.004.
[3] Moazzezi M, Rose B, Kisling K, Moore KL, Ray X. Prospects for daily online adaptive radiotherapy via ethos for prostate cancer patients without nodal involvement using unedited CBCT auto-segmentation. J Appl Clin Med Phys 2021; 22:82–93. https://doi.org/10.1002/acm2.13399.
[4] Byrne M, Archibald-Heeren B, Hu Y, Teh A, Beserminji R, Cai E, et al. Varian ethos online adaptive radiotherapy for prostate cancer: early results of contouring accuracy, treatment plan quality, and treatment time. J Appl Clin Med Phys 2022; 23:e13479. https://doi.org/10.1002/acm2.13479.
[5] Alaei P, Spezi E. Imaging dose from cone beam computed tomography in radiation therapy. Phys Med 2015;31:647–58. https://doi.org/10.1016/j.ejmp.2015.06.003.
[6] Ding GX, Alaei P, Curran B, Flynn R, Gossman M, Mackie TR, et al. Image guidance doses delivered during radiotherapy: quantification, management, and reduction: report of the AAPM Therapy Physics Committee Task Group 180. Med Phys 2018; 45:e84–99. https://doi.org/10.1002/mp.12824.
[7] Ordonez-Sanz C, Cowen M, Shiravand N, MacDougall ND. CBCT imaging: a simple approach for optimising and evaluating concomitant imaging doses, based on patient-specific attenuation, during radiotherapy pelvis treatment. Brit J Radiol 2021;94:20210068. https://doi.org/10.1259/bjr.20210068.
[8] Kan MW, Leung LH, Wong W, Lam N. Radiation dose from cone beam computed tomography for image-guided radiation therapy. Int J Radiat Oncol Biol Phys 2008; 70:272–9. https://doi.org/10.1016/j.ijrobp.2007.08.062.
[9] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Med Image Comput Comput Assist Interv 2015;9351:234–41. https://doi.org/10.48550/arXiv.1505.04597.
[10] Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proc IEEE Int Conf Comput Vis 2017::2223–32. https://doi.org/10.48550/arXiv.1703.10593.
[11] Park T, Efros AA, Zhang R, Zhu JY. Contrastive learning for unpaired image-to-image translation. Comput Vis ECCV 2020;12354:319–45. https://doi.org/10.48550/arXiv.2007.15651.
[12] Kida S, Nakamoto T, Nakano M, Nawa K, Haga A, Kotoku J, et al. Cone beam computed tomography image quality improvement using a deep convolutional neural network. Cureus 2018;10:e2548. https://doi.org/10.7759/cureus.2548.
[13] Chen L, Liang X, Shen C, Jiang S, Wang J. Synthetic CT generation from CBCT images via deep learning. Med Phys 2020;47:1115–25. https://doi.org/10.1002/mp.13978.
[14] Landry G, Hansen D, Kamp F, Li M, Hoyle B, Weller J, et al. Comparing Unet training with three different datasets to correct CBCT images for prostate radiotherapy dose calculations. Phys Med Biol 2019;64:035011. https://doi.org/10.1088/1361-6560/aaf496.
[15] Maier J, Berker Y, Sawall S, Kachelrieß M. Deep scatter estimation (DSE): feasibility of using a deep convolutional neural network for real-time x-ray scatter prediction in cone-beam CT. Proc SPIE Int Soc Opt Eng 2018;10573:393–8. https://doi.org/10.1117/12.2292919.
[16] Maier J, Eulig E, Vöth T, Knaup M, Kuntz J, Sawall S, et al. Real-time scatter estimation for medical CT using the deep scatter estimation: method and robustness analysis with respect to different anatomies, dose levels, tube voltages, and data truncation. Med Phys 2019;46:238–49. https://doi.org/10.1002/mp.13274.
[17] Hansen DC, Landry G, Kamp F, Li M, Belka C, Parodi K, et al. ScatterNet: a convolutional neural network for cone-beam CT intensity correction. Med Phys 2018;45:4916–26. https://doi.org/10.1002/mp.13175.
[18] Zhao J, Chen Z, Wang J, Xia F, Peng J, Hu Y, et al. MV CBCT-based synthetic CT generation using a deep learning method for rectal cancer adaptive radiotherapy. Front Oncol 2021;11:655325. https://doi.org/10.3389/fonc.2021.655325.

[19] Kida S, Kaji S, Nawa K, Imae T, Nakamoto T, Ozaki S, et al. Visual enhancement of cone-beam CT by use of CycleGAN. Med Phys 2020;47:998–1010. https://doi.org/10.1002/mp.13963.
[20] Kurz C, Maspero M, Savenije MH, Landry G, Kamp F, Pinto M, et al. CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation. Phys Med Biol 2019;64:225004. https://doi.org/10.1088/1361-6560/ab4d8c.
[21] Sun H, Fan R, Li C, Lu Z, Xie K, Ni X, et al. Imaging study of pseudo-ct synthesized from cone-beam CT based on 3D CycleGAN in radiotherapy. Front Oncol 2021;11:603844. https://doi.org/10.3389/fonc.2021.603844.
[22] Eckl M, Hoppen L, Sarria GR, Boda-Heggemann J, Simeonova-Chergou A, Steil V, et al. Evaluation of a cycle-generative adversarial network-based cone-beam CT to synthetic CT conversion algorithm for adaptive radiation therapy. Phys Med 2020;80:308–16. https://doi.org/10.1016/j.ejmp.2020.11.007.
[23] Uh J, Wang C, Acharya S, Krasin MJ, Hua CH. Training a deep neural network coping with diversities in abdominal and pelvic images of children and young adults for CBCT-based adaptive proton therapy. Radiother Oncol 2021;160:250–8. https://doi.org/10.1016/j.radonc.2021.05.006.
[24] Dong G, Zhang C, Deng L, Zhu Y, Dai J, Song L, et al. A deep unsupervised learning framework for the 4D CBCT artifact correction. Phys Med Biol 2022;67:055012. https://doi.org/10.1088/1361-6560/ac55a5.
[25] Wynne JF, Lei Y, Pan S, Wang T, Pasha M, Luca K, et al. Rapid unpaired CBCT-based synthetic CT for CBCT-guided adaptive radiotherapy. J Appl Clin Med Phys 2023;24:e14064. https://doi.org/10.1002/acm2.14064.
[26] Kang SR, Shin W, Yang S, Kim JE, Huh KH, Lee SS, et al. Structure-preserving quality improvement of cone beam CT images using contrastive learning. Comput Biol Med 2023;158:106803. https://doi.org/10.1016/j.compbiomed.2023.106803.
[27] Chan Y, Li M, Parodi K, Belka C, Landry G, Kurz C. Feasibility of CycleGAN enhanced low dose CBCT imaging for prostate radiotherapy dose calculation. Phys Med Biol 2023;68:105014. https://doi.org/10.1088/1361-6560/acccce.
[28] Gao L, Xie K, Wu X, Lu Z, Li C, Sun J, et al. Generating synthetic CT from low-dose cone-beam CT by using generative adversarial networks for adaptive radiotherapy. Radiat Oncol J 2021;16:1–16. https://doi.org/10.1186/s13014-021-01928-w.
[29] Yuan N, Dyer B, Rao S, Chen Q, Benedict S, Shang L, et al. Convolutional neural network enhancement of fast-scan low-dose cone-beam CT images for head and neck radiotherapy. Phys Med Biol 2020;65:035003. https://doi.org/10.1088/1361-6560/ab6240.
[30] Chen W, Li Y, Yuan N, Qi J, Dyer BA, Sensoy L, et al. Clinical enhancement in AI-based post-processed fast-scan low-dose CBCT for head and neck adaptive radiotherapy. Front Artif Intell 2021;3:614384. https://doi.org/10.3389/frai.2020.614384.
[31] Rit S, Oliva MV, Brousmiche S, Labarbe R, Sarrut D, Sharp GC. The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK). J Phys Conf Ser 2014;489:012079. https://doi.org/10.1088/1742-6596/489/1/012079.
[32] Hofmaier J, Haehnle J, Kurz C, Landry G, Maihoefer C, Schüttrumpf L, et al. Multi-criterial patient positioning based on dose recalculation on scatter-corrected CBCT images. Radiother Oncol 2017;125:464–9. https://doi.org/10.1016/j.radonc.2017.09.020.
[33] Park YK, Sharp GC, Phillips J, Winey BA. Proton dose calculation on scatter-corrected CBCT image: feasibility study for adaptive proton therapy. Med Phys 2015;42:4449–59. https://doi.org/10.1118/1.4923179.
[34] Niu T, Sun M, Star-Lack J, Gao H, Fan Q, Zhu L. Shading correction for on-board cone-beam CT in radiation therapy using planning MDCT images. Med Phys 2010;37:5395–406. https://doi.org/10.1118/1.3483260.
[35] Niu T, Al-Basheer A, Zhu L. Quantitative cone-beam CT imaging in radiation therapy using planning CT as a prior: first patient studies. Med Phys 2012;39:1991–2000. https://doi.org/10.1118/1.3693050.
[36] Kurz C, Kamp F, Park YK, Zöllner C, Rit S, Hansen D, et al. Investigating deformable image registration and scatter correction for CBCT-based dose calculation in adaptive impt. Med Phys 2016;43:5635–46. https://doi.org/10.1118/1.4962933.
[37] Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys 2010;76:S10–9. https://doi.org/10.1016/j.ijrobp.2009.07.1754.
[38] Olch AJ, Alaei P. How low can you go? A CBCT dose reduction study. J Appl Clin Med Phys 2021;22:85–9. https://doi.org/10.1002/acm2.13164.

Supplementary material

Data acquisition and pre-processings

planning CT (pCT)s of the selected prostate cancer patients were reconstructed with an image grid of 1.074 mm × 1.074 mm × 3.000 mm in combination with a 55 cm lateral field of view (FOV). In each fraction, these patients were advised to follow an in-house bladder and rectum filling protocol before treatment and cone beam computed tomography (CBCT) scanning. Their corresponding CBCT images were acquired with an increased lateral FOV by using a laterally-shifted detector panel in M position and a bow-tie filter. Images with body outline truncation in spite of the increased FOV were excluded from the study. The mean, min, max time gap between pCT and CBCT are 26, 7, 61 days respectively. CBCT inputs ($CBCT_{FS}$, $CBCT_{LD25}$, $CBCT_{LD15}$ and $CBCT_{LD10}$) were reconstructed identically using Feldkamp–Davis–Kress (FDK) (as described previously in [1]) with 410 × 410 × 264 voxels on an isotropic 1.0 mm^3 grid. The patient couch was removed from the CBCT input images by thresholding and morphological maskings, followed by zero padding to an image size of 512 × 512 pixels. The pixel intensity of all CBCTs was normalised in the attenuation coefficient value (μ) range [0, 0.04] (values above 0.04 were set to 0.04).

vCTs were generated by registering the pCT to the daily CBCT via DIR. As described in [2], we aim for 1) image similarity which is computed by normalized gradient fields, and 2) deformation regularity which is computed by curvature regularization. The optimization problem is solved in a discretize-then-optimize scheme using a quasi-Newton L-BFGS optimizer. Following the generation, vCTs were re-sampled to an isotropic 1.0 mm^3 grid and an image size of 512 × 512 pixels. The table was also removed. The pixel intensity was empirically converted to the range of the CBCT images ((Hounsfield units (HU) + 1024) / 65536) [3]. The resulting intensities were clipped to the range [0, 0.05] (values above 0.05 were set to 0.05).

The reference $CBCT_{cor}$ were generated (as described in [4]) and reconstructed using iterative conjugate gradient (CG) with 410 × 410 × 264 voxels on an isotropic 1.0 mm^3 grid. In the CG algorithm [5], the objective function consists of a data consistency term, Laplacian and Tikhonov regularization as shown in the following formulation:

$$||\mathrm{sqrt}(D)(Rf - p)||_2^2 + \gamma||\nabla \cdot f||_2^2 + \mathrm{T}||f||_2^2 \tag{S1}$$

with R the forward projection operator, f the image to be reconstructed, p the measured projections, D the displaced detector weighting operator, γ the weighting of the Laplacian regularization, ∇ being the spatial derivative of the image f, T being the strength of the Tikhonov regularization. The goal of this iterative CG algorithm is to find the image f that minimizes the above equation. By empirical experiments, the hyper-parameters for reconstructing the CBCT images were set to 100 iterations, $\gamma = 1000$ and T $= 100$ for the optimal image quality.

The beginning and last 35 image slices of all CBCT inputs, CBCT_{cor} and vCT in superior–inferior direction were excluded due to partial FOV cone truncation.

Deep learning algorithms

Two deep learning algorithms were used to investigate each of the dose reduction levels in CBCT-to-CT translation tasks. Both algorithms are based on generative adversarial networks (GAN) and trained with a back-and-forth interaction between a generator and a discriminator. An adversarial loss term $L_{\text{adv}}(\text{G}, \text{D}, \text{CBCT}, \text{CT})$ is computed in both algorithms for which the generator G tries to convert low dose CBCT to synthetic CT (sCT) such that is indistinguishable from real CT according to the discriminator D.

- cycleGAN: the first algorithm was the cycleGAN that we implemented in a previous study [1]. In addition to the 25% dose CBCT, 15% and 10% dose CBCTs were specified as the inputs to train corresponding sets of generators and discriminators in this study. To attain higher anatomical fidelity, a residual skip connection was added for both generators to keep the high resolution features in the input image and reduce the vanishing gradients problem in the encoding process. This approach has been reported to improve geometric fidelity in the field of histopathology [6] and was used in a previous CBCT-to-CT study [1]. A cycle consistency loss L_{cyc} is introduced to stabilise the anatomical mappings between CBCT and CT using L1 norm regularisation. This process involves training an extra set of generator and discriminator, for which CT and CBCT are swapped. vCT was used in the training, as we added an L2 norm between CBCT and corresponding vCT to investigate the efficacy of such a paired loss. The total objective function

can be formulated as

$$\begin{aligned} L_{\text{cycleGAN}}(\text{G}, \text{F}, \text{D}_{\text{CBCT}}, \text{D}_{\text{CT}}) = & L_{\text{adv}}(\text{G}, \text{D}_{\text{CT}}, \text{CBCT}, \text{CT}) + \\ & L_{\text{adv}}(\text{F}, \text{D}_{\text{CBCT}}, \text{CT}, \text{CBCT}) + \\ & \lambda_1 L_{\text{cyc}}(\text{G}, \text{F}) + \\ & \lambda_2 L_{\text{L2}}(\text{CBCT}, \text{CT}). \end{aligned} \tag{S2}$$

where G and F denote the generators in forward and backward cycle respectively, D_{CBCT} and D_{CT} denote the discriminators, λ_1 and λ_2 are hyperparameters that were empirically set to 25 and 0 for the optimal image quality in this study.

- CUT: the second algorithm is an alternative one-side translation to reduce the computing resources of the auxiliary network. This can be achieved by replacing L_{cyc} with a loss on image patches, referred to as patchwise contrastive loss L_{PatchNCE}. In contrastive learning, a query patch is sampled from the sCT output and compared with the patch at the corresponding location (denoted as positive) or other patches at different locations (denoted as negatives) of the corresponding CBCT input image. The probability of the positive samples being selected over negatives can be formulated by the following cross-entropy loss:

$$l(v, v^+, v^-) = -\log \left[\frac{e^{v \cdot v^+ / \tau}}{e^{v \cdot v^+ / \tau} + \sum_{n=1}^{N} e^{v \cdot v_n^- / \tau}} \right] \tag{S3}$$

where v, v^+ and v^- denote the K-dimensional vectors of the query ($v \in \mathbb{R}^K$), the positive ($v^+ \in \mathbb{R}^K$), and N number of negatives ($v^- \in \mathbb{R}^{N \times K}$), respectively. The n-th negative is denoted as $v_n^- \in \mathbb{R}^{N \times K}$. For such an (N+1) classification problem, τ, which denotes the distances between the query and samples was set as 0.07. The goal here is to maximize the mutual information between v and v^+, but minimize between v and v^-.

The images from the positive and negative samples are passed through the encoder network of the generator (G_{enc}) to obtain embeddings. These embeddings are low-dimensional representations of the images that capture their content and style information. The layers of interest (L) and the number of spatial locations in each layer (S) are

selected. The feature maps are passed through a small two-layer multi layer perceptron (MLP) network $\mathrm{H_l}$, yielding a stack of features $\{z_l\}_L = \{H_l(G^l_{enc}(x))\}_L$. Likewise, the output image is encoded with the same network into $\hat{z}^s_l = \{H_l(G^l_{enc}(G(x)))\}_L$. The other patches within the input can be used as negatives and that formulates the following contrastive loss $L_{\mathrm{PatchNCEx}}$.

$$L_{\mathrm{PatchNCEx}}(\mathrm{G}, \mathrm{H}, \mathrm{CBCT}) = \mathbb{E}_{\mathrm{x}\sim\mathrm{CBCT}} \sum_{l=1}^{L} \sum_{s=1}^{S_l} l(\hat{z}^s_l, z^s_l, z^{S/s}_l) \tag{S4}$$

where $\hat{z}^s_l$, z^s_l and $z^{S/s}_l$ represents the feature of the output image, the corresponding feature ($z^s_l \in \mathbb{R}^{C_l}$) and the negative feature ($z^{S/s}_l \in \mathbb{R}^{(S_L-1)\times C_l}$). C_l denotes the number of channels at each layer. Since the generator learns to pay attention to the similarities between the two domains, the embeddings share the common features. To avoid incorrect anatomical changes, CT is used to generate identical CT using the same generator. The positive and negative samples of these CT images are passed through the same $\mathrm{G_{enc}}$ and MLP, which allows us to formulate the following contrastive loss for the CT domain $L_{\mathrm{PatchNCEy}}$:

$$L_{\mathrm{PatchNCEy}}(\mathrm{G}, \mathrm{H}, \mathrm{CT}) = \mathbb{E}_{\mathrm{y}\sim\mathrm{CT}} \sum_{l=1}^{L} \sum_{s=1}^{S_l} l(\hat{z}^s_l, z^s_l, z^{S/s}_l) \tag{S5}$$

The total loss function is therefore

$$\begin{aligned} L_{\mathrm{CUT}}(\mathrm{G}, \mathrm{D_{CT}}, \mathrm{CBCT}, \mathrm{CT}) = & L_{\mathrm{adv}}(\mathrm{G}, \mathrm{D_{CT}}, \mathrm{CBCT}, \mathrm{CT}) + \\ & \lambda_{\mathrm{CBCT}} L_{\mathrm{PatchNCEx}}(\mathrm{G}, \mathrm{H}, \mathrm{CBCT}) + \\ & \lambda_{\mathrm{CT}} L_{\mathrm{PatchNCEy}}(\mathrm{G}, \mathrm{H}, \mathrm{CT})). \end{aligned} \tag{S6}$$

where λ_{cbct} and λ_{ct} were both set to 10 for the optimal image quality as suggested in [7]. The main objective for CUT here is to generate realistic CT images, while patches in the input and output images share corresponding information.

Hyper-parameters and network settings

In the following, we describe the hyper-parameters and the network settings for each model.

- cycleGAN: a majority of the hyper-parameters which gave the best results in our previous study [1] were used, i.e. for the generators, two convolutional layers with stride 2 in the encoder and two deconvolutional layers with stride 2 in the decoder, nine residual blocks [8] between encoding and decoding operations. For the discriminators, we used 70 $\times$ 70 PatchGAN [9] with a downsampling scheme from 256 $\times$ 256 to 32 $\times$ 32 by applying four series of 2D convolutional layers, followed by instance normalization (not for the first and last layer) and LeakyReLU with a slope of 0.2 as nonlinearity (not for the last layer). The receptive field of the network was 70 $\times$ 70 and each pixel in the output was evaluated as a scalar in the range [0, 1]. Both generators and discrimninators were optimized with the Adam algorithm. The learning rate was set to 0.0002 during the first 100 epochs, and gradually reduced to zero over the next 100 epochs. The batch size was set to one.

- CUT: the hyper-parameters which were provided in the original CUT implementation [10] were mainly employed. Similar to the network architecture and hyperparameters of cycleGAN, we used the identical nine residual blocks and the PatchGAN discriminator, batch size of one, Adam optimizer with initial learning rate 0.0002 for first 200 epochs and then reduced linearly to zero over the next 200 epochs. Same as in cycleGAN, the hyper-parameters β_1 and β_2 that were used to calculate the momentum term of Adam were set as 0.5 and 0.999 respectively. For the L_{PatchNCE}, 5 layers of features were extracted (L=5), corresponding to the receptive fields of sizes 1 $\times$ 1, 9 $\times$ 9, 15 $\times$ 15, 35 $\times$ 35, and 99 $\times$ 99. For every layer's features, 256 random locations (S = 256) were sampled, and a 2-layer MLP was used to acquire final features.

An identical geometric augmentation pipeline was employed in both algorithms as described in [1]. Each CBCT and vCT input image was resampled to 256 $\times$ 256 pixels, followed by two dimensional (2D) horizontal flipping and affine transformations including rotation of [-5°, 5°] and scaling by [0.9, 1.1] with a bicubic interpolation over 4 $\times$ 4 neighboring pixels.

Implementation details

All models were trained and evaluated using the PyTorch based framework MONAI 1.1.0. [11] on an NVIDIA RTX A6000 GPU with 48 GB of memory.

Supplementary Figure 1

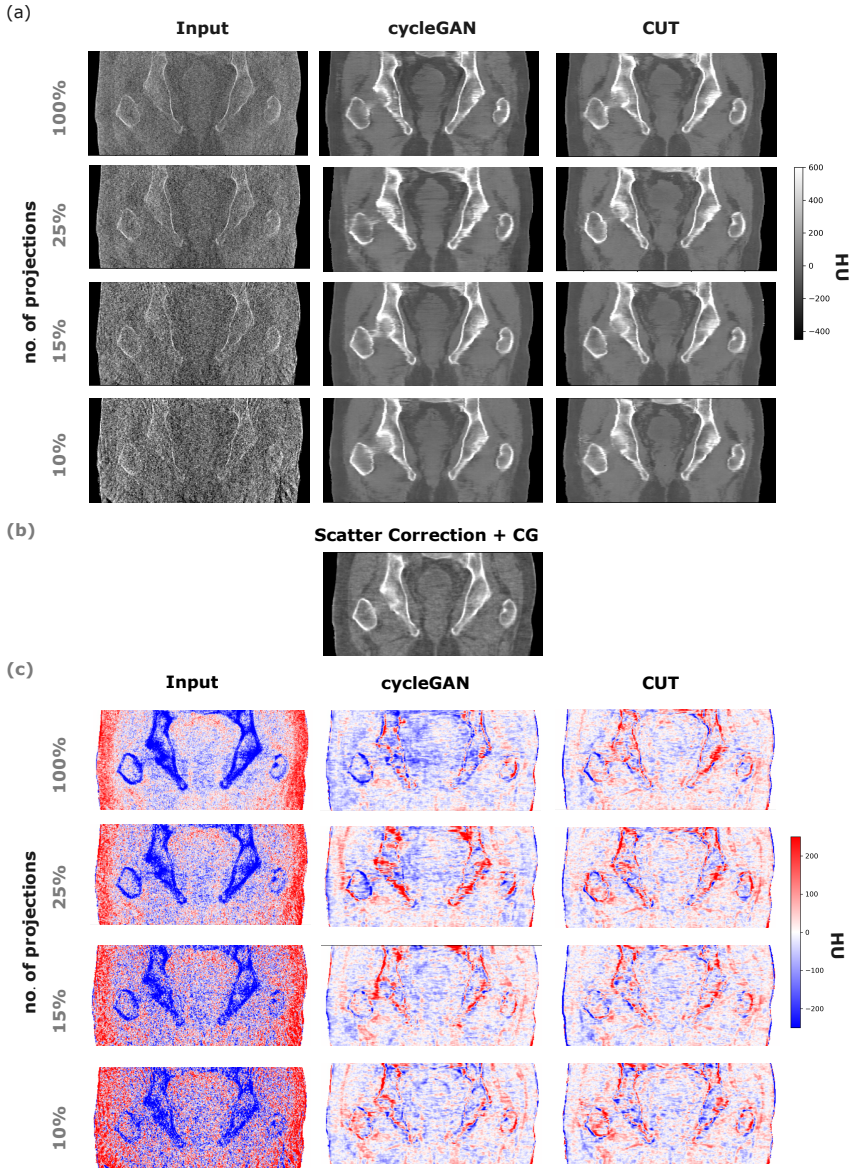


Figure S1: Coronal view of (a) the CBCT inputs and sCTs generated by cycleGAN and CUT with 100%, 25%, 15%, 10% projections; (b) the scatter corrected CBCT_{cor} reconstructed with conjugate gradient; (c) HU difference between corresponding sCTs and CBCT_{cor} of a test patient.

Supplementary Table 1

Table S1: Best epoch selection and the corresponding training time

	cycleGAN				**CUT**			
No. of Projections	100%	25%	15%	10%	100%	25%	15%	10%
Best epoch	40	60	150	160	250	360	340	350
Training time (hours)	7	10	25	27	42	60	57	58

Supplementary Table 2

Table S2: Average HU mean absolute error and mean error of test patient set for the comparison of sCTs and CBCT inputs with $CBCT_{cor}$, respectively. The number in square brackets represent [min, max] values among all patients in the corresponding groups. CBCT inputs were reconstructed using FDK without correction.

Mean Absolute Error (HU)

Dataset	**100**%	**25**%	**15**%	**10**%
cycleGAN	54 [48, 66]	56 [49, 65]	58 [52, 68]	59 [54, 68]
CUT	49 [44, 60]	51 [46, 61]	52 [46, 63]	54 [48, 65]
CBCT inputs	102 [97, 108]	120 [115, 126]	144 [137, 152]	164 [158, 173]

Mean Error (HU)

Dataset	**100**%	**25**%	**15**%	**10**%
cycleGAN	-18 [-35, -6]	-2 [-23, 13]	-21 [-39, -8]	-5 [-8, 16]
CUT	-5 [-12, 6]	-2 [-10, 5]	-9 [-17, -2]	-5 [-15, 4]
CBCT inputs	15 [-10, 33]	15 [-9, 32]	16 [10, 34]	18 [-8, 36]

Structural Similarity Index Measure

Dataset	**100**%	**25**%	**15**%	**10**%
cycleGAN	0.96 [0.94, 0.97]	0.95 [0.92, 0.97]	0.94 [0.92, 0.96]	0.94 [0.91, 0.95]
CUT	0.96 [0.93, 0.97]	0.95 [0.92, 0.97]	0.95 [0.92, 0.96]	0.94 [0.91, 0.96]
CBCT inputs	0.91 [0.87, 0.93]	0.87 [0.81, 0.89]	0.83 [0.77, 0.86]	0.81 [0.75, 0.84]

Peak Signal-to-Noise Ratio (dB)

Dataset	**100**%	**25**%	**15**%	**10**%
cycleGAN	36 [34, 39]	35 [33, 38]	34 [32, 36]	33 [31, 36]
CUT	36 [34, 39]	36 [33, 38]	35 [33, 37]	34 [32, 37]
CBCT inputs	33 [30, 35]	31 [29, 34]	29 [27, 32]	28 [26, 30]

Supplementary Table 3

Table S3: Mean absolute transformation parameter differences comparing sCTs-to-pCT from each dose levels with $CBCT_{FS}$-to-pCT for all test patients.

	cycleGAN				**CUT**			
No. of Projections	100%	25%	15%	10%	100%	25%	15%	10%
Right-left (mm)	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.02
Inferior-superior (mm)	0.04	0.05	0.04	0.04	0.04	0.06	0.05	0.05
Posterior-anterior (mm)	0.02	0.03	0.02	0.03	0.03	0.03	0.02	0.02
Pitch (°)	0.10	0.18	0.15	0.17	0.17	0.17	0.21	0.24
Roll (°)	0.14	0.15	0.17	0.19	0.16	0.18	0.15	0.18
Yaw (°)	0.23	0.25	0.20	0.20	0.26	0.26	0.24	0.19

References

[1] Chan Y, Li M, Parodi K, Belka C, Landry G, Kurz C. Feasibility of CycleGAN enhanced low dose CBCT imaging for prostate radiotherapy dose calculation. Phys Med Biol 2023;68:105014. https://doi.org/10.1088/1361-6560/acccce.

[2] Hofmaier J, Haehnle J, Kurz C, Landry G, Maihoefer C, Schüttrumpf L et al. Multi-criterial patient positioning based on dose recalculation on scatter-corrected CBCT images. Radiother Oncol 2017;125:464–469. https://doi.org/10.1016/j.radonc.2017.09.020.

[3] Park YK, Sharp GC, Phillips J, Winey BA. Proton dose calculation on scatter-corrected CBCT image: feasibility study for adaptive proton therapy. Med Phys 2015;42:4449–4459. https://doi.org/10.1118/1.4923179.

[4] Kurz C, Kamp F, Park YK, Zöllner C, Rit S, Hansen D et al. Investigating deformable image registration and scatter correction for CBCT-based dose calculation in adaptive impt. Med Phys 2016;43:5635–5646. https://doi.org/10.1118/1.4962933.

[5] Rit S, Oliva MV, Brousmiche S, Labarbe R, Sarrut D, Sharp GC. The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK). J Phys Conf Ser 2014;489:012079. https://doi.org/10.1088/1742-6596/489/1/012079.

[6] de Bel T, Bokhorst JM, van der Laak J, Litjens G. Residual cycleGAN for robust domain transformation of histopathological tissue slides. Medical Image Analysis 2021;70:102004.

[7] Dong G, Zhang C, Deng L, Zhu Y, Dai J, Song L et al. A deep unsupervised learning framework for the 4D CBCT artifact correction. Phys Med Biol 2022;67:055012. https://doi.org/10.1088/1361-6560/ac55a5.

[8] Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. European conference on computer vision 2016;:694–711.

[9] Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125–1134.

[10] Park T, Efros AA, Zhang R, Zhu JY. Contrastive learning for unpaired image-to-image translation. Comput. Vis. ECCV 2020;12354:319–345. https://doi.org/10.48550/arXiv.2007.15651.

[11] Cardoso MJ, Li W, Brown R, Ma N, Kerfoot E, Wang Y et al. MONAI: An open-source framework for deep learning in healthcare 2022;:. https://doi.org/10.48550/arXiv.2211.02701.

Chapter 9

Discussion

9.1 Current work

9.1.1 Key findings and limitations

In this thesis, we have achieved the goal of investigating low imaging dose CBCT of the pelvis in terms of sCT image generation. In the first study, the objective was to explore the feasibility of enhancing low imaging dose CBCT images by a cycleGAN network. By undersampling CBCT projections and optimizing two cycleGAN models trained on prostate cancer patient datasets, the study has demonstrated that it is possible to use 25% imaging dose CBCT scans for online plan adaptation. In the second study, we aimed at investigating the minimum imaging dose CBCT (25%, 15%, 10% dose) using a cycleGAN model with residual connection and a CUT model. In both studies, we have carefully evaluated sCT images with image quality measurements and all CBCT-guided ART metrics, including patient positioning, VMAT dose calculations and organs contouring accuracy. As a result, we have successfully found the imaging dose threshold for each of the networks in generating sCT images, and concluded that 25% is the minimum CBCT imaging dose for accurate online plan adaptation.

However, our studies have five minor limitations. First, we have investigated the low dose CBCT technique only on the pelvic datasets. One of the main reasons is the scarcity of low dose CBCT-to-CT studies in the pelvic region, as explained in chapter 7. Moreover, the total number of prostate cancer patients datasets was not large. Compared to the number of patients used in the majority of the CBCT-to-CT studies (ranging from 12 to 200), we have an adequate number of patients datasets for training, validation and testing. With more patient datasets, we might be able to increase the number of testing datasets or possibly increase the DL performance. Additionally, we have only investigated three imaging dose reduction levels in the minimum imaging dose study. Given the fact that we found the contouring performance of cycleGAN and CUT drops significantly when reducing the dose to 15%, the imaging dose threshold can also lie between 25% and 15%. The number of dose reduction levels could be increased for a more accurate determination. However, the manual volumetric contouring took up to 1.5 hours per test patient. For each test patient,

we contoured eight sets of sCT images (four imaging dose levels by two neural networks) and one set of $CBCT_{org}$ images (original full dose FDK reconstructions). With such labour-intensive evaluation procedures, using 8 test patients balanced between statistical power and a realistic workload for this study. Furthermore, we did not investigate the EAR of secondary cancer incidence in prostate cancer patients. Such treatment response or long-term effects would be useful to demonstrate the efficacy of low dose CBCT-to-CT translation, in addition to the general ALARA principle. Finally, we have studied CBCT-to-CT translation using the patient datasets only from one hospital. Generative models are potentially susceptible to dependencies on training data due to the reliance on the statistical patterns present in the dataset. Nevertheless, our comprehensive evaluations demonstrated the feasibility of using sCT images from minimum imaging dose CBCT scan for online adaptation.

9.2 Challenges

Despite the maturity of CBCT-to-CT translation using either U-Net or cycleGAN, it is still outside the clinical workflow. Here we briefly discuss what are the major challenges that hinder such implementation of sCT image generations, from both full dose and low dose, into radiotherapy workflows.

9.2.1 Identifying the best DL network

With the improvements from U-Net over cycleGAN to CUT (as discussed in section 4.4), it remains challenging to select the best DL network architecture or the best training strategies. Despite the development of new DL networks (such as in the synthRad challenge 2023 [61]), contouring analysis was often not included. The main problem in image correction is that generative models can introduce artifacts that may not be present in original CBCT images, or misinterpret artifacts as parts of an organ. For example, the under-sampling artifacts in low dose CBCT introduce bright streaks in various directions which affect the feature of bone intensity and distort the tissue shape.

Several new loss functions were added in order to preserve the anatomies in the full dose CBCT-to-CT cycleGAN training (section 4.2.2), such as shape loss (see the first publication section 2.2.1), gradient loss (see section 2.B (3) in [70]), air pocket loss (see section 2.B (4) in [70]) etc. But not all loss terms can be directly applied in low dose studies. In one of our ablation studies, sCT images had worse MAE with the use of shape loss. It is also rare to see loss functions tackling under-sampling streaks. Total variation loss, which was used in only one full dose study [70], might be useful to remove streaks as seen in various MR image reconstructions with radial k-space trajectory ([115, 116, 117]). In terms of anatomy preservation, contrastive patchwise loss in CUT maximizes the mutual information between CBCT inputs and corresponding sCT images, so that theoretically sCT images have more consistent organ shapes, as demonstrated in other full dose studies. However, in our organ contouring analysis, CUT performed worse than the cycleGAN

model, which has a residual connection in the generator. Similar to the data-consistency concept in iterative reconstruction, this residual connection, which concatenates the CBCT input into its corresponding intermediate sCT image, adds the original anatomies in the translation. The recent full dose CBCT studies of combining a vision transformer into cycleGAN training (as mentioned in section 4.2.2), or the other study [118] which trained a diffusion model to generate sCT images within 2 minutes opened avenues for investigation. Therefore, the low dose CBCT-to-CT translation may yet be improved by novel network architectures.

9.2.2 Geometrical accuracy evaluations

One of the major concerns from clinicians or researchers is: how realistic is the DL prediction? In a CBCT-to-CT translation, the non-linear transformation of a trained generative model maps the intensity values of the same organ in the CBCT to their equivalent HU in the CT. Such intensities in CBCT images can vary due to scattering noise and undersampling streaks. It is possible that the organ shapes are distorted in the resulting sCT images.

In the majority of studies, sCT images are not evaluated with geometrical accuracy, but limited to image quality metrics and treatment dose calculations, which scarcely capture such morphological changes. One of the studies [70] has evaluated sCT images with MAE of multiple small regions-of-interests (ROIs) on muscles or fat area, which was still insufficient to account for the organ shape variations. For organ segmentation, which can compare directly the organ shapes, only limited studies were able to conduct segmentation analysis by manual contouring [77] or by another DL segmentation network [69, 78] comparing to deformed contours from pCT. However, depending on the changes of the organs between the pCT images and CBCT images, the contours propagated from pCT images may not represent precise ground truths due to inherent DIR uncertainties.

To accurately assess the anatomical fidelity, we compared the organs shapes in sCT images directly with those in the original CBCT image in our studies. However, the lack of clinical CBCT contouring labels caused a substantial challenge. Due to the poor contrast of soft tissue and image noise, there is no current practice of contouring organs on CBCT images. Such a new task requires experienced radiation oncologists to create and validate contours for a longer time and a greater commitment. Additionally, in the absence of CBCT contours as ground truths, training an auto-segmentation model on CBCT images remains challenging. A recent work has shown the feasibility to segment organs in CBCT without ground-truth in the HN region, using propagated contours from pCT and a refined model from a CT segmentation training [119]. Furthermore, in the Ethos online adaptation workflow, clinicians are required to rectify the propagated contours [25, 120], potentially serving as a valuable reservoir of ground truths for CBCT segmentation training. Perhaps in the future, the anatomical fidelity evaluation could be facilitated by the expansion of ground truth datasets or the advancement of unsupervised DL models.

9.2.3 Lack of quality assurance tools for sCT in online ART

In online ART, quality assurance (QA) is applied to ensure the adapted treatment plan. This includes checks for anything that is changed from the initial treatment plan and the patient model, such as the new anatomies and contours on sCT images, or any other manual changes. What hinders the applications of sCT images, from both full dose and low dose CBCT, is the potential introduction of various risks. First, sCT images can possess geometric distortions from the CBCT inputs. As there is no ground truth, visual inspections or comparing with deformed CT scans can hardly identify small differences of the organ shapes. Currently, there is no QA assessment scheme and no DL model that could rectify the contours on the fly. Moreover, each sCT generation model from different studies (using different networks, datasets, image pre-processing steps etc.) can yield different geometrical accuracy. The method of establishing a tolerance level for sCT images is unknown. In addition, sCT can still possess HU differences compared to pCT. No QA model is available to check HU accuracy, which can lead to inaccurate treatment dose calculations.

In MR-guided adaptive radiotherapy, treatment dose QA methods are applied to evaluate the sCT images generated from MRI images. A recent MRI-to-CT translation study [121] has evaluated dose calculation accuracy of four QA methods for sCT images generated by cycleGAN: water override of patient body, bulk densities contouring, another set of sCT images generated by another neural network and deformed CT. The study suggested that using the sCT images generated by the additional network has the best QA performance: The verification procedures were completed within 10 minutes and treatment dose deviated within 2%. However, using AI to verify AI would require different sets of data for training in order to avoid bias and error propagation. Furthermore, in MR-linac clinical practice with the ViewRay linac system, an independent secondary dose calculation is only used for treatment plan QA. sCT images depends on manual review, and when inaccuracies are identified, sCT images are corrected using density overrides. There is still no consensus certifying which of the treatment dose QA tools is sufficient. The QA guidelines for sCT images, especially for low dose CBCT-to-CBCT translation, remain unclear.

9.3 Future works

During this doctoral research in radiation oncology and AI, the clinical need of a low dose CBCT-to-CT translation and the proposed generative AI solutions were identified. The feasibility study and the minimum imaging dose investigation study demonstrated that the low dose CBCT-to-CT translation is technically achievable. The following are the suggestions to take this work further into the clinical implementations.

9.3.1 Extension of this work: 4D CBCT and other anatomical sites

This study has found that it is possible to lower the CBCT imaging dose to generate sCT images for accurate dose calculation and organ contouring in the pelvic region. Theoretically, this DL technique which jointly removes streak artefacts and corrects HU intensities can be applied on other anatomical regions. However, the effect of reducing the dose by 75% dose could vary. For example, as the original HN protocol uses a lower imaging dose, reducing the imaging dose by 75% might result in more dominant streak artefacts and noise than in the pelvic protocol. One can reconstruct the CBCT images with under-sampling projections in each region and systematically investigate the minimum imaging dose reduction rate. Alternatively, as shown in a few full-dose studies (section 4.2.2), one could also train a single DL model with the low dose CBCT datasets of multiple regions. This would require another advanced network training to learn the general, but not specific human anatomy features in the presence of more noise and streak artefacts. CycleGAN and CUT would be good candidates for such trials, but more anatomical-related loss functions and modification of the network architectures should be investigated, in order to maximize the anatomical information from the CBCT inputs.

In addition, 4D CBCT images have similar streak artefacts due to insufficient numbers of projections after respiratory binning. As an initial study, we have tried our technique on 48 lung cancer patient 4D CBCT datasets. First we have extracted the respiratory breathing signal using the RTK Amsterdam Shroud image filter [114, 122] and binned the projections into 10 breathing phases. Each phase has around 68 projections (depending on the breathing pattern), and was reconstructed to volumetric images using FDK. This is a more challenging task, as the number of projections in each phase varies, leading to different amount of streaking artefacts in the reconstructed images, or the input images for training a cycleGAN network. As shown in fig. 9.1, our initial over-fitting result has demonstrated that it is possible to extend this low dose CBCT-to-CT technique to enhance 4D CBCT images using a cycleGAN model. Network optimization, hyper-parameters searching and comprehensive evaluations would be the next steps in this extension.

9.3.2 Data-driven DL improvements: patient-specific, data consistency, structure-guided

To address the possible changes in the organ shapes, among the model architectures and loss functions (such as body shape loss, paired loss, patchwise contrastive loss etc.) that were tested during our studies, we found that a cycleGAN model with a residual connection in the generator and an adversarial and a cycle-consistency loss performed the best in the contouring analysis. In the context of low dose CBCT-to-CT translation in radiotherapy workflow, this could be achieved in different ways:

- Patient-specific training: Instead of applying one model for all patients, it is also possible to train a sCT image generation model for individual patients. The concept

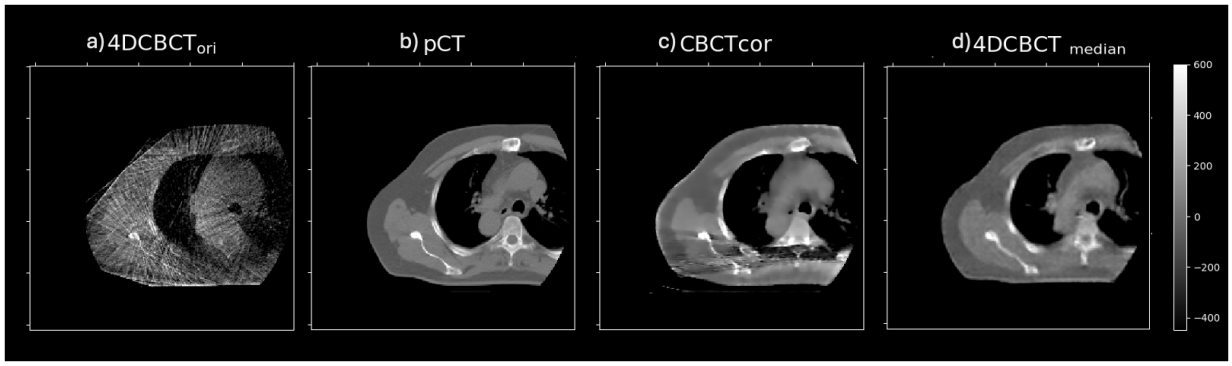


Figure 9.1: Our initial over-fitting result of 4D CBCT images (a single motion phase) of one exemplary lung cancer patient: a) the original 4D CBCT image using FDK reconstruction; b) pCT image; c) the corrected CBCT image using MA-ROOSTER reconstruction [123]; d) the cycleGAN output image of the 4D CBCT. All intensities are in HU. (Own figure)

here is to first train a population model with all patient datasets. When a patient undergoes an in-room CBCT scan in the first fraction, the population model can be fine-tuned and re-trained as a patient-specific model using the pCT images and the daily CBCT images of the patient. The advantage is that the generator has learnt to map the intensity and remove noise in the population model step, and can learn geometrical features of the patient anatomy in the patient-specific step. This training strategy has been investigated and shown to achieve a better performance in a few auto-segmentation studies, for example CBCT in the pelvic and HN regions [124] and MRI in the pelvic region [125]. With the use of prior sequential datasets (pCT and fraction CBCT images), a patient-specific CBCT-to-CT model can account for the unique anatomy and has potential to produce geometrically consistent sCT images.

- Data consistency term: Similar to the data consistency concept in iterative reconstructions, a representation of the input image can be reapplied in the objective function and this drives the training to retain desired features. For example, in conjugate gradient (CG) image reconstruction, the data consistency term is formulated as

$$||Rf - p||_2^2, \tag{9.1}$$

where R the forward projection operator, f is the image to be reconstructed by the network, p is the measured projection. If we can integrate the CBCT reconstruction process and forward projection into the model training, we can add such a term into the total objective function and drive the prediction in consistency with original projection data in each iteration in the training.

- Structure-guided training: Instead of adding the projection data, one can alternatively add patient-specific contours as a training constraint to increase geometrical consistency. Similar to gradient loss, one can extract and minimize the edge difference of the organs in the input and predicted image during the training process,

as shown in an example low dose CT study [126]. However, extracting organ edges correctly from a low dose CBCT might be challenging due to more severe streak artefacts. In recent studies on low dose CBCT correction [127, 128], a patient organ shape map which combines the edge information from the deformed pCT and network intermediate output was used to regulate the training.

Additionally, shape and geometry processing methods also begin to gain importance in the computational medical imaging field. One of the focus workshops of the Medical Image Computing and Computer Assisted Interventions (MICCAI) conference called "[Shape in Medical Imaging]" has attracted new ideas on generating images with high geometrical accuracy, including new network architectures, loss functions and analysis [129, 130, 131]. In the future, we can exchange ideas and collaboratively develop a highly robust technique with this community.

9.3.3 Bridging DL-clinical gap: adaptive QA tools for sCT images

To bridge the gap between the low dose CBCT-to-CT translation technique and the clinical workflow, an adaptive QA tool for sCT images should be in place to ensure that the processes involved in sCT image generation are effective and consistently produce high-quality images for each patient.

- Standardized protocol and documentation: Since generative AI models can have different performance depending on different datasets, it is important to document the development process of such AI models, including datasets, hyper-parameters, model architectures, evaluation performances etc. In parallel to this idea, a recent guideline published jointly by the European Society for Radiotherapy and Oncology (ESTRO) and American Association of Physicists in Medicine (AAPM) for AI model developments in radiotherapy [132] also recommends QA at least at the use case level, such as formally reporting the DL application and the requirements of inputs and outputs. With full details of the model, users can select the best suitable model, trace back the image generation process and adjust the sCT when necessary.

- Anatomical fidelity QA: Organ shapes can vary in sCT images as discussed above. In addition to the treatment dose QA, a QA tool is necessary to ensure the geometrical accuracy of sCT images. As an initial idea, we can develop a QA tool to calculate an uncertainty map, which can visualize and reflect the overall confidence of the sCT images in the organ shapes and HU intensity. A recent study [133] has explored the feasibility of generating uncertainty maps on sCT images for HN patients using Monte Carlo dropout in online MRI-guided proton therapy.

9.3.4 Quantification of CBDI

As discussed in section 3.5.2, CBDI cannot be directly represented by CTDI, since CBCT has a volumetric acquisition with a FOV larger than the dimensions of the standard pencil chamber (100 mm), which is typically used in CTDI measurements. However, in the user manual of "[Elekta instructions for USE Volume 3 - XVI]", the dose information for each protocol is provided only in terms of CTDI and its variants. Amer *et al.* [134] has developed $CBDI_{100}$ by integrating the central 100 mm FOV of the dose profile. By developing a new pencil chamber with a size that matches the FOV, one could measure and standardize the actual CBDI values among CBCT systems and protocols, so as to document and compare the CBCT imaging dose more precisely.

9.3.5 Research translation: model integration into clinical workflow

To ensure successful implementation of a low dose CBCT-to-CT, the workflow can be structured as illustrated in fig. 9.2. In the treatment room, the patient is instructed to lay on the treatment couch as before. The gantry-mounted CBCT system scans only 25% of the original number of projections. In the TPS, the projections are used to reconstruct low dose CBCT images using FDK (approximately 7 seconds on a GPU). These images will be used as an input to the generative AI model which can be implemented in the TPS as a module and generate volumetric sCT images in around 2 seconds on a GPU. On these sCT images, the treatment dose can be recalculated with updated anatomical contours which could be propagated from the contours in the last fraction and corrected by a radiation oncologist. The sCT images can be registered in the TPS and the patient can be positioned using the transformation parameters. With the adaptive QA processes, the treatment dose and geometrical accuracy of such plan can be ensured and adjusted when necessary. Finally, an adapted treatment radiation can be delivered.

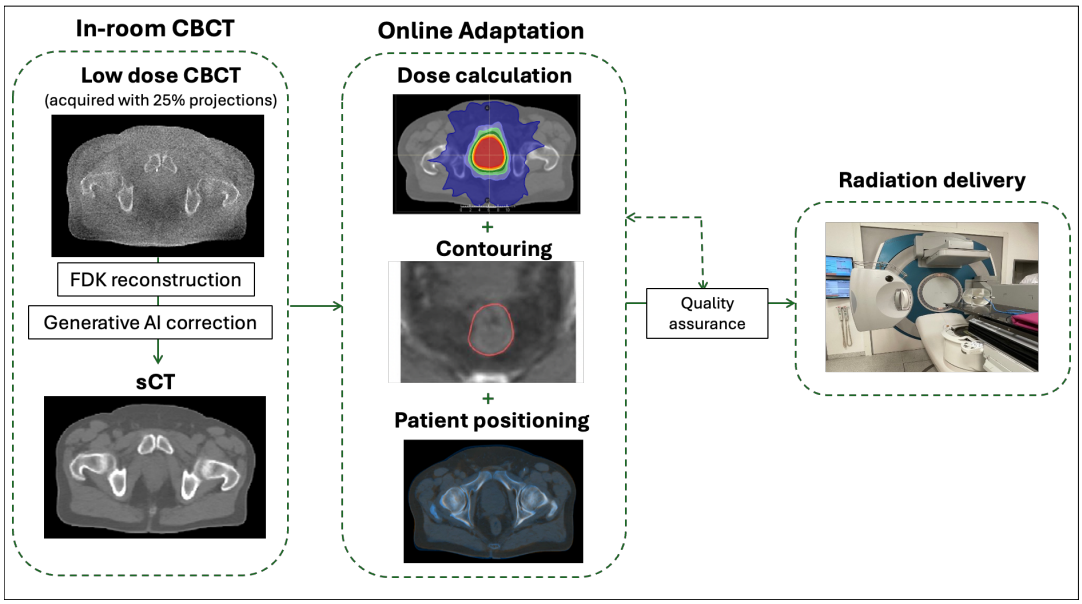


Figure 9.2: The integration of a low dose CBCT pipeline in the CBCT-guided online adaptive radiotherapy. (Own figure)

Chapter 10

Conclusion

In this thesis, two DL algorithms, cycleGAN and CUT, were implemented and compared to investigate the lowest achievable CBCT imaging dose without loss of accuracy in terms of sCT image generations. In both studies, 41 prostate cancer patient datasets (pCT and CBCT) were included. CBCT projections were under-sampled and low dose CBCT images (25%, 15% and 10%) were reconstructed. In the first study, the feasibility of 25% dose CBCT-to-CT by training two cycleGAN models using unpaired datasets was investigated. In the second study, cycleGAN and CUT models were trained and applied to generate sCT images for each of the CBCT imaging dose levels. All sCT images of testing patient datasets were carefully evaluated with respect to image quality, positioning accuracy, VMAT treatment dose calculation and organ contouring (bladder and rectum) accuracy. Over these IGRT-metrics, only slight deviations were observed among sCT images for all CBCT imaging dose reduction levels. In addition, the second study suggested that the cycleGAN model with a residual connection performed better than CUT for organ contouring. However, both achieved similar results in all other metrics. Based on segmentation accuracy, 25% is the minimum pelvic CBCT imaging dose that enables accurate VMAT treatment dose calculation and organ delineation for online adaptation.

Appendix A

List of publications

A.1 Journal paper contributions as first author

- Minimum imaging dose for deep learning-based pelvic synthetic computed tomography generation from cone beam images
  Y.C.I. Chan, M. Li, A. Thummerer, K. Parodi, C. Belka, C. Kurz and G. Landry
  Physics and Imaging in Radiation Oncology. 2024 March: 10.1016/j.phro.2024.100569.

- Feasibility of cycleGAN Enhanced Low Dose CBCT Imaging for Prostate Radiotherapy Dose Calculation
  Y. Chan, M. Li, K. Parodi, C. Belka, G. Landry and C. Kurz
  Physics in Medicine & Biology. 2023 May 11;68(10):105014.

A.2 Conference contributions as first author

- Towards AI-enabled minimum dose CBCT-based synthetic CT: dose calculation and contouring accuracy
  Y.C.I Chan, M. Li, A. Thummerer, K. Parodi, C. Belka, C. Kurz and G. Landry
  European Society for Radiotherapy and Oncology (ESTRO) annual meeting, Glasgow UK [May 3-7, 2024], oral presentation

- Evaluierung einer cycleGAN-basierten low-dose Cone-Beam CT Bildkorrektur zur Dosisberechnung in der adaptiven Prostata-Strahlentherapie
  Y. Chan, M. Li, K. Parodi, C. Belka, G. Landry and C. Kurz
  53. Jahrestagung der Deutschen Gesellscaft für Medizinische Physik (DGMP), Aachen, Germany [Sep 21-24 , 2022], oral presentation

- Feasibility of cycleGAN enhanced low dose CBCT Imaging for prostate radiotherapy dose calculation
  Y. Chan, M. Li, K. Parodi, C. Belka, G. Landry and C. Kurz

64th American Association of Physicists in Medicine (AAPM) annual meeting & exhibition, Washington D.C., USA [Jul 10-14 , 2022], oral presentation

Appendix B

List of abbreviations

Abbreviation	**Meaning**
AI	artificial intelligence
ALARA	as low as reasonably achievable
ART	adaptive radiotherapy
BED	biologically effective dose
BN	batch normalization
CBCT	cone-beam computed tomography
CBDI	CBCT dose index
CNN	convolutional neural network
CT	computed tomography
CTDI	CT dose index
CTV	clinical target volume
CUT	contrastive unpaired translation
cycleGAN	cycle-consistent generative adversarial network
dB	decibel
D_{CBCT}	the discriminator that distinguishes between synthetic CBCT images and true CBCT images in cycleGAN
D_{CT}	the discriminator that distinguishes between sCT images and true CT images in cycleGAN
DIR	deformable image registration
DL	deep learning
DNA	deoxyribonucleic acid
DSC	Dice similarity coefficient
DVH	dose-volume histogram
EAR	excess absolute risk
FBP	filtered back projection
FDK	Feldkamp, Davis, Kress
FID	Fréchet inception distance

fps	frames per second
G_{CT}	the generator that translates CBCT images to CT images in cycleGAN
G_{CBCT}	the generator that translates CT images to CBCT images in cycleGAN
G_{enc}	the encoder network of the generator in CUT
GTV	gross tumor volume
HD	Hausdorff distance
HN	head and neck
HU	Hounsfield unit
IARC	International Agency for Research on Cancer
IGRT	image-guided radiotherapy
IMRT	intensity modulated photon radiotherapy
IS	inferior-superior
ITV	internal target volume
kV	kilovoltage
MAE	mean absolute error
MC	Monte-Carlo
ME	mean error
MLC	multileaf collimators
MLP	multi layer perceptron
MRI	magnetic resonance imaging
MSE	mean squared error
NCE	noise contrastive estimation
OAR	organ-at-risk
PA	posterior-anterior
pCT	planning CT
PET	positron emission tomography
PSNR	peak signal-to-noise ratio
PTV	planning target volume
QUANTEC	Quantitative Analysis of Normal Tissue Effects in Clinic
ReLU	rectified linear unit
RL	right-left
RMSE	root mean squared error
SC	the distance of the x-ray source to the iso center
sCT	synthetic CT
SD	the distance of the x-ray source to the detector
SNR	signal-to-noise ratio
SSIM	structural similarity index measure

TFT	thin-film transistor
TPS	treatment planning system
vCT	virtual CT
VMAT	volumetric modulated arc therapy
WHO	World Health Organization
XVI	x-ray volume imaging

Appendix C

Congruency confirmation

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dekanat Medizinische Fakultät
Promotionsbüro

Confirmation of congruency between printed and electronic version of the doctoral thesis

Chan, Yan Chi

name, first name

I hereby declare that the electronic version of the submitted thesis, entitled:

Generative AI for cone-beam CT dose reduction and intensity correction in adaptive radiotherapy

is congruent with the printed version both in content and format.

Munich, 04.12.2025

Place, Date

Yan Chi Chan

Signature doctoral candidate

Bibliography

[1] Bray F, Laversanne M, Sung H, Ferlay J, Siegel R, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. A Cancer Journal for Clinicans. 2024.

[2] Ronckers C, Spix C, Katalinic A, Christ M, Cicero A, Folkerts J, et al. Cancer in Germany 2019/2020. 14th edition Robert Koch Institute and the Association of Population-based Cancer Registries in Germany. 2023.

[3] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.

[4] Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discovery. 2022;12(1):31-46.

[5] World Health Organization. WHO list of priority medical devices for cancer management. World Health Organization; 2017.

[6] Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. Medical Physics. 2008;35(1):310-7.

[7] Gill SK, Reddy K, Campbell N, Chen C, Pearson D. Determination of optimal PTV margin for patients receiving CBCT-guided prostate IMRT: comparative analysis based on CBCT dose calculation with four different margins. Journal of Applied Clinical Medical Physics. 2015;16(6):252-62.

[8] Kilburn JM, Soike MH, Lucas JT, Ayala-Peacock D, Blackstock W, Isom S, et al. Image guided radiation therapy may result in improved local control in locally advanced lung cancer patients. Practical Radiation Oncology. 2016;6(3):e73-80.

[9] Den RB, Doemer A, Kubicek G, Bednarz G, Galvin JM, Keane WM, et al. Daily image guidance with cone-beam computed tomography for head-and-neck cancer intensity-modulated radiotherapy: a prospective study. International Journal of Radiation Oncology, Biology, Physics. 2010;76(5):1353-9.

[10] Yeung AR, Li JG, Shi W, Newlin HE, Chvetsov A, Liu C, et al. Tumor localization using cone-beam CT reduces setup margins in conventionally fractionated radiotherapy for lung tumors. International Journal of Radiation Oncology, Biology, Physics. 2009;74(4):1100-7.

[11] Pearson D, Gill SK, Campbell N, Reddy K. Dosimetric and volumetric changes in the rectum and bladder in patients receiving CBCT-guided prostate IMRT: analysis based on daily CBCT dose calculation. Journal of Applied Clinical Medical Physics. 2016;17(6):107-17.

[12] Nijkamp J, Pos FJ, Nuver TT, De Jong R, Remeijer P, Sonke JJ, et al. Adaptive radiotherapy for prostate cancer using kilovoltage cone-beam computed tomography: first clinical results. International Journal of Radiation Oncology, Biology, Physics. 2008;70(1):75-82.

[13] Schwartz DL, Garden AS, Thomas J, Chen Y, Zhang Y, Lewin J, et al. Adaptive radiotherapy for head-and-neck cancer: initial clinical outcomes from a prospective trial. International Journal of Radiation Oncology, Biology, Physics. 2012;83(3):986-93.

[14] Kwint M, Conijn S, Schaake E, Knegjens J, Rossi M, Remeijer P, et al. Intra thoracic anatomical changes in lung cancer patients during the course of radiotherapy. Radiotherapy and Oncology. 2014;113(3):392-7.

[15] Lim-Reinders S, Keller BM, Al-Ward S, Sahgal A, Kim A. Online adaptive radiation therapy. International Journal of Radiation Oncology, Biology, Physics. 2017;99(4):994-1003.

[16] Jones D. ICRU report 50 prescribing, recording and reporting photon beam therapy. Medical Physics. 1994;21(6):833-4.

[17] Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. International Journal of Radiation Oncology, Biology, Physics. 2010;76(3):S10-9.

[18] Verhaegen F, Seuntjens J. Monte Carlo modelling of external radiotherapy photon beams. Physics in Medicine & Biology. 2003;48(21):R107.

[19] Mackie TR, Ahnesjø A, Dickof P, Snider A. Development of a convolution/superposition method for photon beams. In: IX International Conference on the use of Computers in Radiation Therapy; 1987. p. p107.

[20] Yan D, Vicini F, Wong J, Martinez A. Adaptive radiation therapy. Physics in Medicine & Biology. 1997;42(1):123.

[21] Kurz C, Dedes G, Resch A, Reiner M, Ganswindt U, Nijhuis R, 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. 2015;54(9):1651-7.

[22] de Jong R, Visser J, van Wieringen N, Wiersma J, Geijsen D, Bel A. Feasibility of Conebeam CT-based online adaptive radiotherapy for neoadjuvant treatment of rectal cancer. Radiation Oncology. 2021;16:1-11.

[23] Sibolt P, Andersson LM, Calmels L, Sjöström D, Bjelkengren U, Geertsen P, et al. Clinical implementation of artificial intelligence-driven cone-beam computed tomography-guided online adaptive radiotherapy in the pelvic region. Physics and Imaging in Radiation Oncology. 2021;17:1-7.

[24] Moazzezi M, Rose B, Kisling K, Moore KL, Ray X. Prospects for daily online adaptive radiotherapy via ethos for prostate cancer patients without nodal involvement using unedited CBCT auto-segmentation. Journal of Applied Clinical Medical Physics. 2021;22(10):82-93.

[25] Byrne M, Archibald-Heeren B, Hu Y, Teh A, Beserminji R, Cai E, 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. 2022;23(1):e13479.

[26] Robar JL, Cherpak A, MacDonald RL, Yashayaeva A, McAloney D, McMaster N, et al. Novel technology allowing cone beam computed tomography in 6 seconds: a patient study of comparative image quality. Practical Radiation Oncology. 2024;14(3):277-86.

[27] Lustermans D, Fonseca GP, Taasti VT, van de Schoot A, Petit S, van Elmpt W, et al. Image quality evaluation of a new high-performance ring-gantry cone-beam computed tomography imager. Physics in Medicine & Biology. 2024;69(10):105018.

[28] Jaffray D, Siewerdsen J. Cone-beam computed tomography with a flat-panel imager: initial performance characterization. Medical physics. 2000;27(6):1311-23.

[29] Siewerdsen JH, Jaffray DA. Cone-beam computed tomography with a flat-panel imager: magnitude and effects of x-ray scatter. Medical physics. 2001;28(2):220-31.

[30] Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. International Journal of Radiation Oncology* Biology* Physics. 2002;53(5):1337-49.

[31] Orth RC, Wallace MJ, Kuo MD, of the Society of Interventional Radiology TAC, et al. C-arm cone-beam CT: general principles and technical considerations for use in interventional radiology. Journal of Vascular and Interventional Radiology. 2008;19(6):814-20.

[32] Shaw CC. Cone beam computed tomography (Imaging in medical diagnosis and therapy). CRC Press; 2014.

[33] Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. Journal of the Optical Society of America A. 1984;1(6):612-9.

[34] Peroni M, Ciardo D, Spadea MF, Riboldi M, Comi S, Alterio D, et al. Automatic segmentation and online virtualCT in head-and-neck adaptive radiation therapy. International Journal of Radiation Oncology, Biology, Physics. 2012;84(3):e427-33.

[35] Landry G, Dedes G, Zöllner C, Handrack J, Janssens G, de Xivry JO, et al. Phantom based evaluation of CT to CBCT image registration for proton therapy dose recalculation. Physics in Medicine & Biology. 2014;60(2):595.

[36] Landry G, Nijhuis R, Dedes G, Handrack J, Thieke C, Janssens G, et al. Investigating CT to CBCT image registration for head and neck proton therapy as a tool for daily dose recalculation. Medical Physics. 2015;42(3):1354-66.

[37] Veiga C, Alshaikhi J, Amos R, Lourenço AM, Modat M, Ourselin S, et al. Cone-beam computed tomography and deformable registration-based "dose of the day" calculations for adaptive proton therapy. International Journal of Particle Therapy. 2015;2(2):404-14.

[38] Veiga C, Janssens G, Teng CL, Baudier T, Hotoiu L, McClelland JR, et al. First clinical investigation of cone beam computed tomography and deformable registration for adaptive proton therapy for lung cancer. International Journal of Radiation Oncology, Biology, Physics. 2016;95(1):549-59.

[39] Wang P, Yin L, Zhang Y, Kirk M, Song G, Ahn PH, et al. Quantitative assessment of anatomical change using a virtual proton depth radiograph for adaptive head and neck proton therapy. Journal of Applied Clinical Medical Physics. 2016;17(2):427-40.

[40] Mainegra-Hing E, Kawrakow I. Variance reduction techniques for fast Monte Carlo CBCT scatter correction calculations. Physics in Medicine & Biology. 2010;55(16):4495.

[41] Thing RS, Bernchou U, Mainegra-Hing E, Hansen O, Brink C. Hounsfield unit recovery in clinical cone beam CT images of the thorax acquired for image guided radiation therapy. Physics in Medicine & Biology. 2016;61(15):5781.

[42] Zöllner C, Rit S, Kurz C, Vilches-Freixas G, Kamp F, Dedes G, et al. Decomposing a prior-CT-based cone-beam CT projection correction algorithm into scatter and beam hardening components. Physics and Imaging in Radiation Oncology. 2017;3:49-52.

[43] Ding GX, Duggan DM, Coffey CW, Deeley M, Hallahan DE, Cmelak A, et al. A study on adaptive IMRT treatment planning using kV cone-beam CT. Radiotherapy and Oncology. 2007;85(1):116-25.

[44] Fotina I, Hopfgartner J, Stock M, Steininger T, Lütgendorf-Caucig C, Georg D. Feasibility of CBCT-based dose calculation: comparative analysis of HU adjustment techniques. Radiotherapy and Oncology. 2012;104(2):249-56.

[45] Niu T, Al-Basheer A, Zhu L. Quantitative cone-beam CT imaging in radiation therapy using planning CT as a prior: first patient studies. Medical Physics. 2012;39(4):1991-2000.

[46] Veiga C, McClelland J, Moinuddin S, Lourenço A, Ricketts K, Annkah J, et al. Toward adaptive radiotherapy for head and neck patients: feasibility study on using CT-to-CBCT deformable registration for "dose of the day" calculations. Medical Physics. 2014;41(3):031703.

[47] Niu T, Sun M, Star-Lack J, Gao H, Fan Q, Zhu L. Shading correction for on-board cone-beam CT in radiation therapy using planning MDCT images. Medical Physics. 2010;37(10):5395-406.

[48] Park YK, Sharp GC, Phillips J, Winey BA. Proton dose calculation on scatter-corrected CBCT image: feasibility study for adaptive proton therapy. Medical Physics. 2015;42(8):4449-59.

[49] Kurz C, Kamp F, Park YK, Zöllner C, Rit S, Hansen D, et al. Investigating deformable image registration and scatter correction for CBCT-based dose calculation in adaptive IMPT. Medical Physics. 2016;43(10):5635-46.

[50] Valentin J, et al. The 2007 recommendations of the international commission on radiological protection. vol. 37. Elsevier Oxford; 2007.

[51] Protection R. ICRP publication 103. Ann ICRP. 2007;37(2.4):2.

[52] Buzug T. Dose. Computed Tomography: From Photon Statistics to Modern Cone-Beam CT. 2008:485-502.

[53] Hyer DE, Hintenlang DE. Estimation of organ doses from kilovoltage cone-beam CT imaging used during radiotherapy patient position verification. Medical Physics. 2010;37(9):4620-6.

[54] Ding GX, Alaei P, Curran B, Flynn R, Gossman M, Mackie TR, et al. Image guidance doses delivered during radiotherapy: Quantification, management, and reduction: Report of the AAPM Therapy Physics Committee Task Group 180. Medical Physics. 2018;45(5):e84-99.

[55] Ronneberger O, Fischer P, Brox T; Springer. U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. 2015:234-41.

[56] Kida S, Nakamoto T, Nakano M, Nawa K, Haga A, Kotoku J, et al. Cone beam computed tomography image quality improvement using a deep convolutional neural network. Cureus. 2018;10(4):e2548.

[57] Maier J, Berker Y, Sawall S, Kachelrieß M; SPIE. Deep scatter estimation (DSE): feasibility of using a deep convolutional neural network for real-time x-ray scatter prediction in cone-beam CT. Medical Imaging 2018: Physics of Medical Imaging. 2018;10573:393-8.

[58] Maier J, Eulig E, Vöth T, Knaup M, Kuntz J, Sawall S, et al. Real-time scatter estimation for medical CT using the deep scatter estimation: Method and robustness analysis with respect to different anatomies, dose levels, tube voltages, and data truncation. Medical Physics. 2019;46(1):238-49.

[59] Hansen DC, Landry G, Kamp F, Li M, Belka C, Parodi K, et al. ScatterNet: a convolutional neural network for cone-beam CT intensity correction. Medical Physics. 2018;45(11):4916-26.

[60] Landry G, Hansen D, Kamp F, Li M, Hoyle B, Weller J, et al. Comparing Unet training with three different datasets to correct CBCT images for prostate radiotherapy dose calculations. Physics in Medicine & Biology. 2019;64(3):035011.

[61] Thummerer A, Seller Oria C, Zaffino P, Visser S, Meijers A, Guterres Marmitt G, et al. Deep learning–based 4D-synthetic CTs from sparse-view CBCTs for dose calculations in adaptive proton therapy. Medical Physics. 2022;49(11):6824-39.

[62] Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE International Conference on Computer Vision. 2017:2223-32.

[63] Ge Y, Wei D, Xue Z, Wang Q, Zhou X, Zhan Y, et al. Unpaired MR to CT synthesis with explicit structural constrained adversarial learning. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE; 2019. p. 1096-9.

[64] Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. In: Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. Springer; 2016. p. 694-711.

[65] Ulyanov D, Vedaldi A, Lempitsky V. Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:160708022. 2016.

[66] Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition; 2017. p. 1125-34.

[67] Kurz C, Maspero M, Savenije MH, Landry G, Kamp F, Pinto M, et al. CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation. Physics in Medicine & Biology. 2019;64(22):225004.

[68] Dong G, Zhang C, Liang X, Deng L, Zhu Y, Zhu X, et al. A deep unsupervised learning model for artifact correction of pelvis cone-beam CT. Frontiers in Oncology. 2021;11:686875.

[69] Zhao J, Chen Z, Wang J, Xia F, Peng J, Hu Y, et al. MV CBCT-based synthetic CT generation using a deep learning method for rectal cancer adaptive radiotherapy. Frontiers in Oncology. 2021;11:655325.

[70] Kida S, Kaji S, Nawa K, Imae T, Nakamoto T, Ozaki S, et al. Visual enhancement of cone-beam CT by use of CycleGAN. Medical Physics. 2020;47(3):998-1010.

[71] Liu Y, Chen X, Zhu J, Yang B, Wei R, Xiong R, et al. A two-step method to improve image quality of CBCT with phantom-based supervised and patient-based unsupervised learning strategies. Physics in Medicine & Biology. 2022;67(8):084001.

[72] Liang X, Chen L, Nguyen D, Zhou Z, Gu X, Yang M, et al. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. Physics in Medicine & Biology. 2019;64(12):125002.

[73] Maspero M, Houweling AC, Savenije MH, van Heijst TC, Verhoeff JJ, Kotte AN, 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. 2020;14:24-31.

[74] Dai Z, Zhang Y, Zhu L, Tan J, Yang G, Zhang B, et al. Geometric and dosimetric evaluation of deep learning-based automatic delineation on CBCT-synthesized CT and planning CT for breast cancer adaptive radiotherapy: a multi-institutional study. Frontiers in Oncology. 2021;11:725507.

[75] Deng L, Hu J, Wang J, Huang S, Yang X. Synthetic CT generation based on CBCT using respath-cycleGAN. Medical Physics. 2022;49(8):5317-29.

[76] Harms J, Lei Y, Wang T, Zhang R, Zhou J, Tang X, et al. Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. Medical Physics. 2019;46(9):3998-4009.

[77] Eckl M, Hoppen L, Sarria GR, Boda-Heggemann J, Simeonova-Chergou A, Steil V, et al. Evaluation of a cycle-generative adversarial network-based cone-beam CT to synthetic CT conversion algorithm for adaptive radiation therapy. Physica Medica. 2020;80:308-16.

[78] Sun H, Fan R, Li C, Lu Z, Xie K, Ni X, et al. Imaging study of pseudo-CT synthesized from cone-beam CT based on 3D CycleGAN in radiotherapy. Frontiers in Oncology. 2021;11:603844.

[79] Rossi M, Cerveri P. Comparison of supervised and unsupervised approaches for the generation of synthetic CT from cone-beam CT. Diagnostics. 2021;11(8):1435.

[80] Uh J, Wang C, Acharya S, Krasin MJ, Hua CH. Training a deep neural network coping with diversities in abdominal and pelvic images of children and young adults for CBCT-based adaptive proton therapy. Radiotherapy and Oncology. 2021;160:250-8.

[81] Lei Y, Wang T, Harms J, Shafai-Erfani G, Dong X, Zhou J, et al. Image quality improvement in cone-beam CT using deep learning. In: Medical Imaging 2019: Physics of Medical Imaging. vol. 10948. SPIE; 2019. p. 556-61.

[82] Liu Y, Lei Y, Wang T, Fu Y, Tang X, Curran WJ, et al. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. Medical Physics. 2020;47(6):2472-83.

[83] Lemus OMD, Wang YF, Li F, Jambawalikar S, Horowitz DP, Xu Y, et al. Dosimetric assessment of patient dose calculation on a deep learning-based synthesized computed tomography image for adaptive radiotherapy. Journal of Applied Clinical Medical Physics. 2022;23(7):e13595.

[84] Qiu RL, Lei Y, Shelton J, Higgins K, Bradley JD, Curran WJ, et al. Deep learning-based thoracic CBCT correction with histogram matching. Biomedical Physics & Engineering Express. 2021;7(6):065040.

[85] Tien HJ, Yang HC, Shueng PW, Chen JC. Cone-beam CT image quality improvement using Cycle-Deblur consistent adversarial networks (Cycle-Deblur GAN) for chest CT imaging in breast cancer patients. Scientific Reports. 2021;11(1):1133.

[86] Xue X, Ding Y, Shi J, Hao X, Li X, Li D, et al. Cone beam CT (CBCT) based synthetic CT generation using deep learning methods for dose calculation of nasopharyngeal carcinoma radiotherapy. Technology in Cancer Research & Treatment. 2021;20:15330338211062415.

[87] Jihong C, Kerun Q, Kaiqiang C, Xiuchun Z, Yimin Z, Penggang B. CBCT-based synthetic CT generated using CycleGAN with HU correction for adaptive radiotherapy of nasopharyngeal carcinoma. Scientific Reports. 2023;13(1):6624.

[88] Rusanov B, Hassan GM, Reynolds M, Sabet M, Kendrick J, Rowshanfarzad P, et al. Deep learning methods for enhancing cone-beam CT image quality toward adaptive radiation therapy: A systematic review. Medical Physics. 2022;49(9):6019-54.

[89] Rusanov B, Hassan GM, Reynolds M, Sabet M, Rowshanfarzad P, Bucknell N, et al. Transformer CycleGAN with uncertainty estimation for CBCT based synthetic CT in adaptive radiotherapy. Physics in Medicine & Biology. 2024;69(3):035014.

[90] Dalmaz O, Yurt M, Çukur T. ResViT: residual vision transformers for multimodal medical image synthesis. IEEE Transactions on Medical Imaging. 2022;41(10):2598-614.

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

[92] Torbunov D, Huang Y, Yu H, Huang J, Yoo S, Lin M, et al. Uvcgan: Unet vision transformer cycle-consistent gan for unpaired image-to-image translation. In: Proceedings of the IEEE/CVF Cinter conference on Applications of Computer Vision; 2023. p. 702-12.

[93] Park T, Efros AA, Zhang R, Zhu JY. Contrastive learning for unpaired image-to-image translation. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16. Springer; 2020. p. 319-45.

[94] Oord Avd, Li Y, Vinyals O. Representation learning with contrastive predictive coding. arXiv preprint arXiv:180703748. 2018.

[95] Wynne JF, Lei Y, Pan S, Wang T, Pasha M, Luca K, et al. Rapid unpaired CBCT-based synthetic CT for CBCT-guided adaptive radiotherapy. Journal of Applied Clinical Medical Physics. 2023;24:e14064.

[96] Kang SR, Shin W, Yang S, Kim JE, Huh KH, Lee SS, et al. Structure-preserving quality improvement of cone beam CT images using contrastive learning. Computers in Biology and Medicine. 2023;158:106803.

[97] Dong G, Zhang C, Deng L, Zhu Y, Dai J, Song L, et al. A deep unsupervised learning framework for the 4D CBCT artifact correction. Physics in Medicine & Biology. 2022;67(5):055012.

[98] Alaei P, Spezi E. Imaging dose from cone beam computed tomography in radiation therapy. Physica Medica. 2015;31(7):647-58.

[99] Ordonez-Sanz C, Cowen M, Shiravand N, MacDougall ND. CBCT imaging: a simple approach for optimising and evaluating concomitant imaging doses, based on patient-specific attenuation, during radiotherapy pelvis treatment. British Journal of Radiology. 2021;94(1124):20210068.

[100] Kan MW, Leung LH, Wong W, Lam N. Radiation dose from cone beam computed tomography for image-guided radiation therapy. International Journal of Radiation Oncology, Biology, Physics. 2008;70(1):272-9.

[101] Yuasa Y, Shiinoki T, Onizuka R, Fujimoto K. Estimation of effective imaging dose and excess absolute risk of secondary cancer incidence for four-dimensional cone-beam computed tomography acquisition. Journal of Applied Clinical Medical Physics. 2019;20(11):57-68.

[102] Kim DW, Chung WK, Yoon M. Imaging doses and secondary cancer risk from kilovoltage cone-beam CT in radiation therapy. Health Physics. 2013;104(5):499-503.

[103] Quinn A, Holloway L, Koh ES, Delaney G, Arumugam S, Goozee G, et al. Radiation dose and contralateral breast cancer risk associated with megavoltage cone-beam computed tomographic image verification in breast radiation therapy. Practical Radiation Oncology. 2013;3(2):93-100.

[104] Zhou L, Bai S, Zhang Y, Ming X, Zhang Y, Deng J. Imaging dose, cancer risk and cost analysis in image-guided radiotherapy of cancers. Scientific Reports. 2018;8(1):10076.

[105] Dzierma Y, Minko P, Ziegenhain F, Bell K, Buecker A, Rübe C, et al. Abdominal imaging dose in radiology and radiotherapy–phantom point dose measurements, effective dose and secondary cancer risk. Physica Medica. 2017;43:49-56.

[106] Dzierma Y, Mikulla K, Richter P, Bell K, Melchior P, Nuesken F, et al. Imaging dose and secondary cancer risk in image-guided radiotherapy of pediatric patients. Radiation Oncology. 2018;13:1-14.

[107] de Crevoisier R, Bayar MA, Pommier P, Muracciole X, Pêne F, Dudouet P, et al. Daily versus weekly prostate cancer image guided radiation therapy: phase 3 multicenter randomized trial. International Journal of Radiation Oncology, Biology, Physics. 2018;102(5):1420-9.

[108] Yuan N, Dyer B, Rao S, Chen Q, Benedict S, Shang L, et al. Convolutional neural network enhancement of fast-scan low-dose cone-beam CT images for head and neck radiotherapy. Physics in Medicine & Biology. 2020;65(3):035003.

[109] Gao L, Xie K, Wu X, Lu Z, Li C, Sun J, et al. Generating synthetic CT from low-dose cone-beam CT by using generative adversarial networks for adaptive radiotherapy. Radiation Oncology. 2021;16:1-16.

[110] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing. 2004;13(4):600-12.

[111] Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM. In: 2010 20th International Conference on Pattern Recognition. IEEE; 2010. p. 2366-9.

[112] Dubuisson MP, Jain AK. A modified Hausdorff distance for object matching. In: Proceedings of 12th International Conference on Pattern Recognition. vol. 1. IEEE; 1994. p. 566-8.

[113] Huijben EM, Terpstra ML, Pai S, Thummerer A, Koopmans P, Afonso M, et al. Generating synthetic computed tomography for radiotherapy: SynthRAD2023 challenge report. Medical Image Analysis. 2024;97:103276.

[114] Rit S, Oliva MV, Brousmiche S, Labarbe R, Sarrut D, Sharp GC; IOP Publishing. The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK). Journal of Physics: Conference Series. 2014;489(1):012079.

[115] Feng L, Grimm R, Block KT, Chandarana H, Kim S, Xu J, et al. Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI. Magnetic Resonance in Medicine. 2014;72(3):707-17.

[116] Feng L, Benkert T, Block KT, Sodickson DK, Otazo R, Chandarana H. Compressed sensing for body MRI. Journal of Magnetic Resonance Imaging. 2017;45(4):966-87.

[117] Haji-Valizadeh H, Rahsepar AA, Collins JD, Bassett E, Isakova T, Block T, et al. Validation of highly accelerated real-time cardiac cine MRI with radial k-space sampling and compressed sensing in patients at 1.5 T and 3T. Magnetic Resonance in Medicine. 2018;79(5):2745-51.

[118] Peng J, Qiu RL, Wynne JF, Chang CW, Pan S, Wang T, et al. CBCT-Based synthetic CT image generation using conditional denoising diffusion probabilistic model. Medical Physics. 2024;51(3):1847-59.

[119] Liang X, Morgan H, Bai T, Dohopolski M, Nguyen D, Jiang S. Deep learning based direct segmentation assisted by deformable image registration for cone-beam CT based auto-segmentation for adaptive radiotherapy. Physics in Medicine & Biology. 2023;68(4):045012.

[120] Kim JY, Tawk B, Knoll M, Hoegen-Saßmannshausen P, Liermann J, Huber PE, 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. 2024;16(6):1210.

[121] Dal Bello R, Lapaeva M, Saint-Esteven ALG, Wallimann P, Günther M, Konukoglu E, et al. Patient-specific quality assurance strategies for synthetic computed tomography in magnetic resonance-only radiotherapy of the abdomen. Physics and Imaging in Radiation Oncology. 2023;27:100464.

[122] Rit S, van Herk M, Zijp L, Sonke JJ. Quantification of the variability of diaphragm motion and implications for treatment margin construction. International Journal of Radiation Oncology, Biology, Physics. 2012;82(3):e399-407.

[123] Mory C, Janssens G, Rit S. Motion-aware temporal regularization for improved 4D cone-beam computed tomography. Physics in Medicine & Biology. 2016;61(18):6856.

[124] Chen Y, Gensheimer MF, Bagshaw HP, Butler S, Yu L, Zhou Y, et al. Patient-Specific Auto-segmentation on Daily kVCT Images for Adaptive Radiation Therapy. International Journal of Radiation Oncology, Biology, Physics. 2023;117(2):505-14.

[125] Kawula M, Vagni M, Cusumano D, Boldrini L, Placidi L, Corradini S, et al. Prior knowledge based deep learning auto-segmentation in magnetic resonance imaging-guided radiotherapy of prostate cancer. Physics and Imaging in Radiation Oncology. 2023;28:100498.

[126] Tian Z, Jia X, Yuan K, Pan T, Jiang SB. Low-dose CT reconstruction via edge-preserving total variation regularization. Physics in Medicine & Biology. 2011;56(18):5949.

[127] Chen Y, Yin FF, Zhang Y, Zhang Y, Ren L. Low dose CBCT reconstruction via prior contour based total variation (PCTV) regularization: a feasibility study. Physics in Medicine & Biology. 2018;63(8):085014.

[128] Chen Y, Yin FF, Zhang Y, Zhang Y, Ren L. Low dose cone-beam computed tomography reconstruction via hybrid prior contour based total variation regularization (hybrid-PCTV). Quantitative Imaging in Medicine and Surgery. 2019;9(7):1214.

[129] Wachinger C, Paniagua B, Elhabian S, Li J, Egger J. Shape in Medical Imaging: International Workshop, ShapeMI 2023, Held in Conjunction with MICCAI 2023, Vancouver, BC, Canada, October 8, 2023, Proceedings. vol. 14350. Springer Nature; 2023.

[130] Reuter M, Wachinger C, Lombaert H, Paniagua B, Goksel O, Rekik I. Shape in Medical Imaging: International Workshop, ShapeMI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Proceedings. vol. 12474. Springer Nature; 2020.

[131] Reuter M, Wachinger C, Lombaert H, Paniagua B, Lüthi M, Egger B. Shape in Medical Imaging: International Workshop, ShapeMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings. vol. 11167. Springer; 2018.

[132] Hurkmans C, Bibault JE, Brock KK, van Elmpt W, Feng M, Fuller CD, et al. A joint ESTRO and AAPM guideline for development, clinical validation and reporting of artificial intelligence models in radiation therapy. Radiotherapy and Oncology. 2024:110345.

[133] Galapon Jr AV, Thummerer A, Langendijk JA, Wagenaar D, Both S. Feasibility of Monte Carlo dropout-based uncertainty maps to evaluate deep learning-based synthetic CTs for adaptive proton therapy. Medical Physics. 2024;51(4):2499-509.

[134] Amer A, Marchant T, Sykes J, Czajka J, Moore C. Imaging doses from the Elekta Synergy X-ray cone beam CT system. The British Journal of Radiology. 2007;80(954):476-82.

Acknowledgements

Science is a team effort, and I am very grateful to be in an amazing team which made a significant difference in my PhD study journey. It is my great pleasure to acknowledge them in this thesis.

First of all, I would like to express my most sincere appreciation to my supervisors Guillaume Landry and Christopher Kurz for giving me this opportunity to work on this project. Guillaume has created an ideal research environment, which effectively fosters our innovations and collaborations. He is always approachable and deeply compassionate, supports my research, professional growth and well-being. He shares with us a clear vision of this interdisciplinary research field and is open to learn further with students. I will always remember his willingness to spend extra time with me in our student office, engaging in discussions, tackling programming challenges together. Christopher has been always generous to share his all-rounded expertise and providing interactive guidance, creating a harmonious blend of a relaxed, diligent and reciprocal trust research atmosphere. His unwavering attention to details, dedication to accurate scientific knowledge and constructive advice to my learning have been invaluable in this journey. I greatly appreciate such high-quality of academic supervision and inspiring interactions with both Guillaume and Christopher. The camaraderie that they cultivate within our group motivated me to overcome the boundaries and contributed significantly to my work. This mentorship has enabled me to achieve more than I expected.

A special thank goes to our radiation oncologist Minglun Li and post-doc researcher Adrian Thummerer. Minglun generously provided me with contouring training and investigated the manual contouring techniques in the CBCT and sCT images. His verification of my contours has become one of the major contributions on evaluating the geometrical accuracy of sCT images. Adrian's advice in CBCT image pre-processings and constructive feedback in my second project on minimum imaging dose CBCT and the ongoing project on 4D CBCT have been immensely helpful.

This PhD study journey would not have been so enjoyable without my fellow teammates in this research group. We have started at a Home-Office period at the first year, yet our bonding and mutual supports have been building up stronger over time. I am profoundly grateful to Maria Rädler and Lili Huang for their invaluable contributions to my projects. They generously devoted additional time to share with me their expertise in physics and programming which were instrumental in shaping my work. A lovely thank to Maria who brought me into a regular sport training in the university sport center. Another heartfelt

thank to Lili who completed a climbing certification course with me. I will always cherish the moments we spent sharing important life events and memories. Moreover, I would like to extend my thanks to Elia Lombardo, Jackie Xiong, Nikolaos Ntelopoulos, Catrin Rodenberg for their unwavering supports and collaborations throughout my research. Elia and Nikos (together with Guillaume, Christopher, Adrian and Jan) set up our computing nodes for GPU allocations and docker usage in the server, and generously conducted user tutorial sessions among us. Elia, Jackie and Catrin have been sharing essential insights from the unique intersection of physics research and clinical experiences. Nikos has continued my master student's project on MRI-to-CT translation and has been taking it into a more realistic clinical application with Jackie. Another special thank goes to our former post-doc researcher Yiling Wang, who encouraged me in important life and academic events. One more special thank goes to our post-doc researcher Moritz Rabe who has shared with me his PhD experience and encouraged my work. I also would like to thank my fellow teammates for their sharing and feedback to my work: Rabea Klaar, Chengtao Wei, Domagoj Radonic, Henning Schmidt and Johanna Albrecht. I will forever treasure the moments all of us shared discussing scientific challenges, presenting projects in conferences and graduate schools and exploring new methods together.

Moreover, I would like to thank my thesis advisory committee members Katia Parodi and Michael Ingrish, who gave critical feedback and guidance on my research studies and professional developments. Many thanks go to the essential supports from Claus Belka, Christian Thieke, Michael Reiner, Jan Hoffmann and Vanessa Filipa da Silva Mendes. Another thank goes to Simon Rit at University of Lyon, who generously supported us in CBCT FDK and iterative reconstructions with the RTK toolkit. I appreciate all the patients who contributed into the CBCT and CT imaging datasets.

Another special thank goes to the graduate school (research training group GRK 2274), which encourages doctoral students at LMU and TUM to exchange ideas and research experiences among experimental and clinical medical physics through seminars and retreats. I appreciate the annual Bioengineering autumn school training with the collaboration of TUM Munich Institute of Biomedical Engineering (MIBE), where we had chances to present our work and obtain new engineering perspectives. In addition, I would like to thank the organizing team at German Cancer Research Center (DKFZ) in Heidelberg for the summer school in Medical Physics (Data Science and Machine Learning in Radiotherapy), where I had opportunities to visit the ion-beam therapy center, to discuss latest developments with students from research teams worldwide and to explore my new skills in data science slam.

What is more, I would like to thank my master thesis supervisor Mehdi H. Moghari for his guidance at Harvard Medical School and Boston Children's Hospital. He inspired my academic journey to enhance medical imaging with deep learning.

Finally, I would like to thank my parents Ka Kuen and Kam Fun, and my sisters for their continuous support. I will always remember how my father dedicated hours each day to engaging in conversations with me during my essential developmental phase. He fully supports my academic journey towards the first PhD degree in our family and coaches my life. In addition, I would like to express my greatest thank to Ryana Fok (and her family)

for her wholehearted and unconditional support during our bachelor, master, doctoral studies and all overseas exchange and internship opportunities that we have participated together. I could not finish my acknowledgement without thanking Fabian Schmid (and his family), who is willing to explore new science with me, and to support me in science and life.