IPEM
Institute of Physics and
Engineering in Medicine

PAPER • OPEN ACCESS

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

View the article online for updates and enhancements.

Physics in Medicine & Biology

IPEM
Institute of Physics and
Engineering in Medicine

CrossMark

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

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 × 1.074 mm × 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.
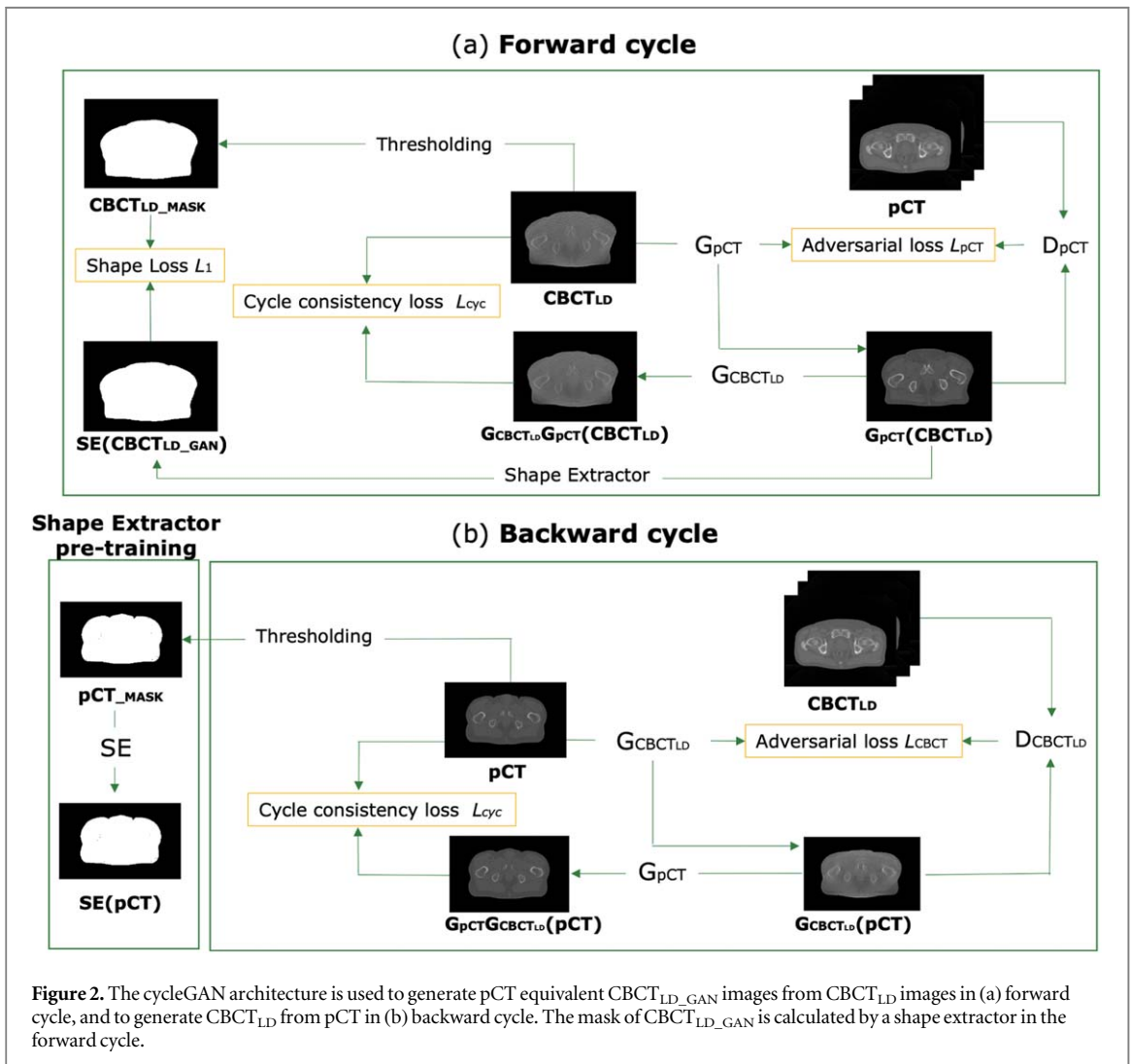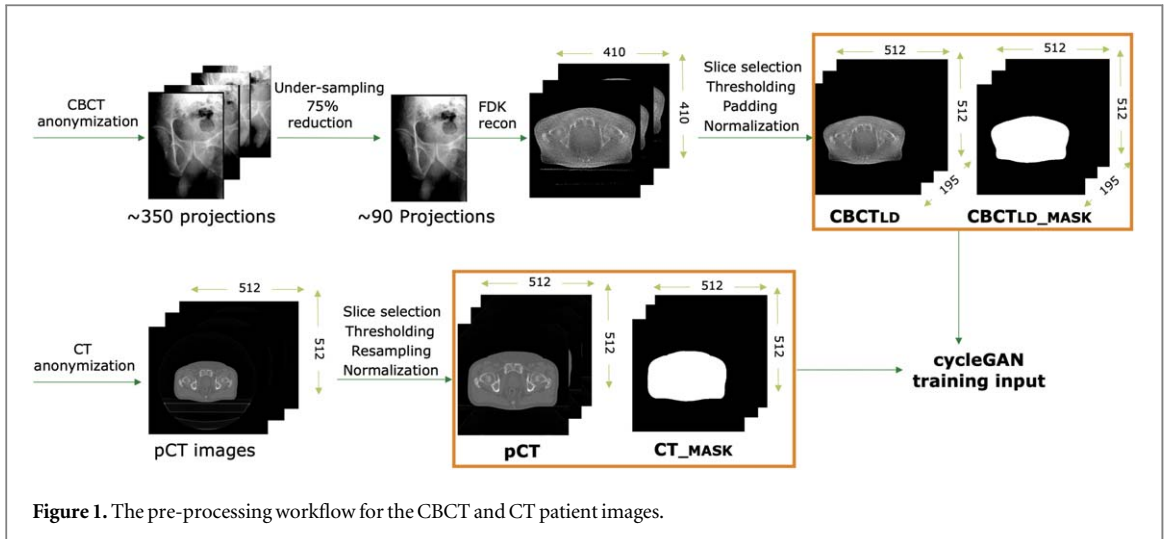
2.1.2. Data preparation

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 × 410 × 264 voxels on an isotropic 1.0 mm³ 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 × 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:
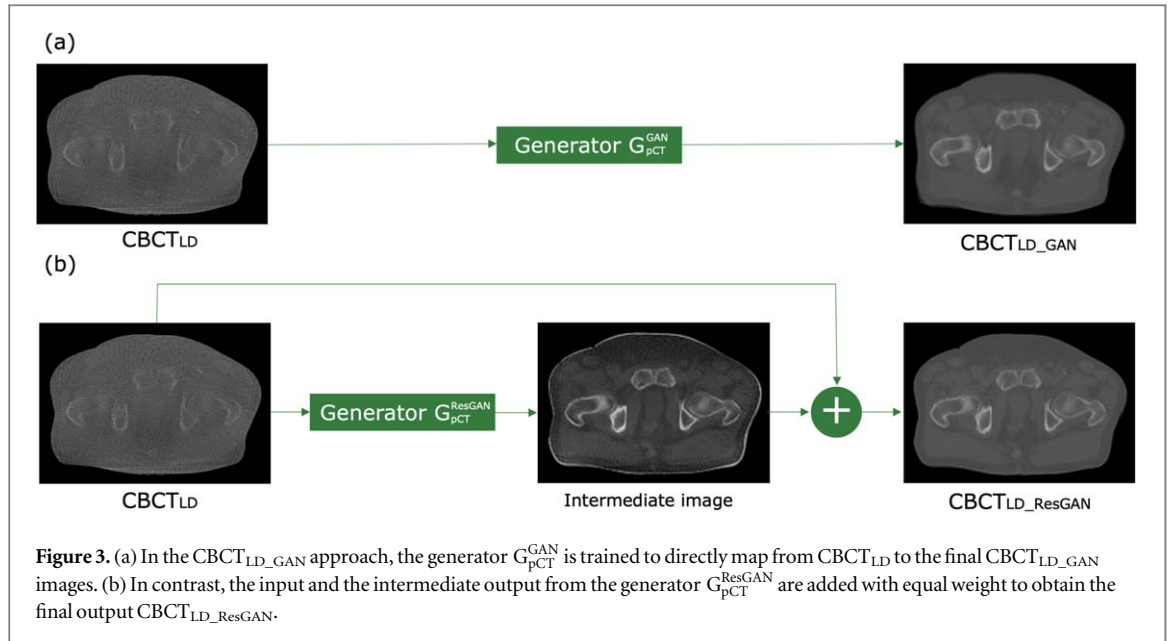
Figure 1. The pre-processing workflow for the CBCT and CT patient images.

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

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

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

In figure 2 (panel (b)), the second generator G$_{\mathrm{CBCT_{LD}}}$ was trained to establish the inverse mapping from pCT to CBCT$_{\mathrm{LD}}$ with the help of the second discriminator D$_{\mathrm{CBCT_{LD}}}$ 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 $CBCT_{LD\_ResGAN}$.

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

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 $CBCT_{LD}$:

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

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

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

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 $CBCT_{LD}$:

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

Therefore the total loss used was:

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

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

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

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.
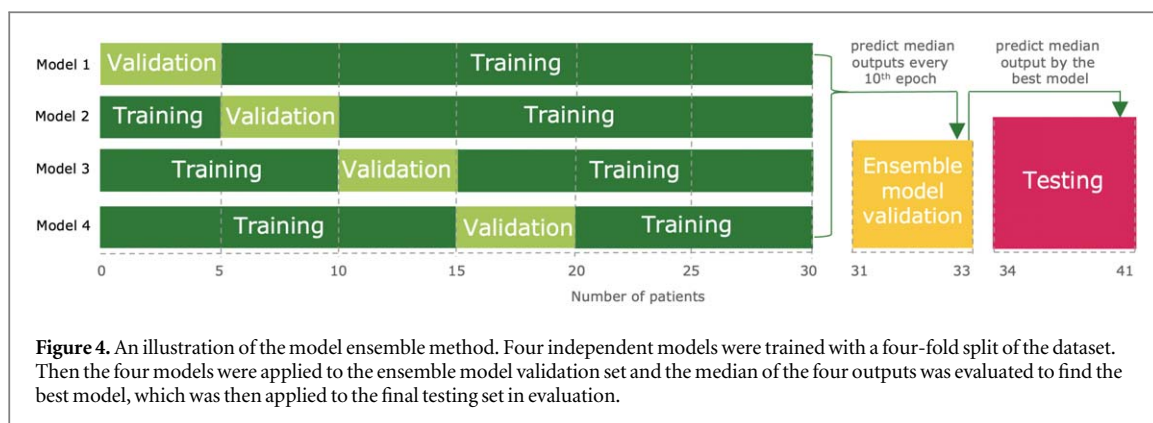
$CBCT_{LD}$ to an intermediate image, which has reversed intensities that suppress the streak artifacts from the $CBCT_{LD}$ input image as shown in panel (b). In the backward cycle, the other generator $G_{CBCT_{LD}}$ in the $CBCT_{LD\_ResGAN}$ approach was also trained to obtain the final output with 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.
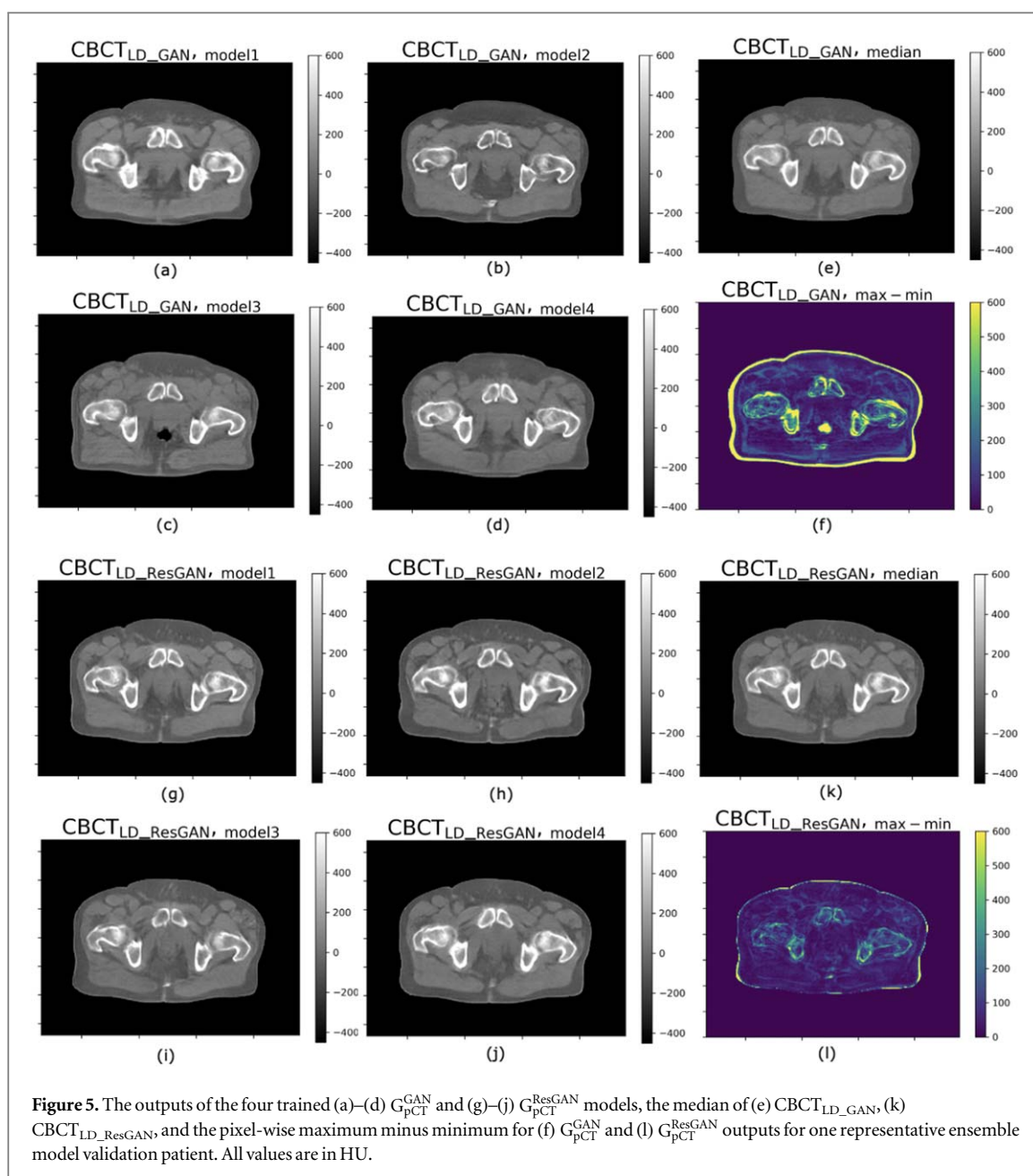
2.3.4. Positioning accuracy

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.

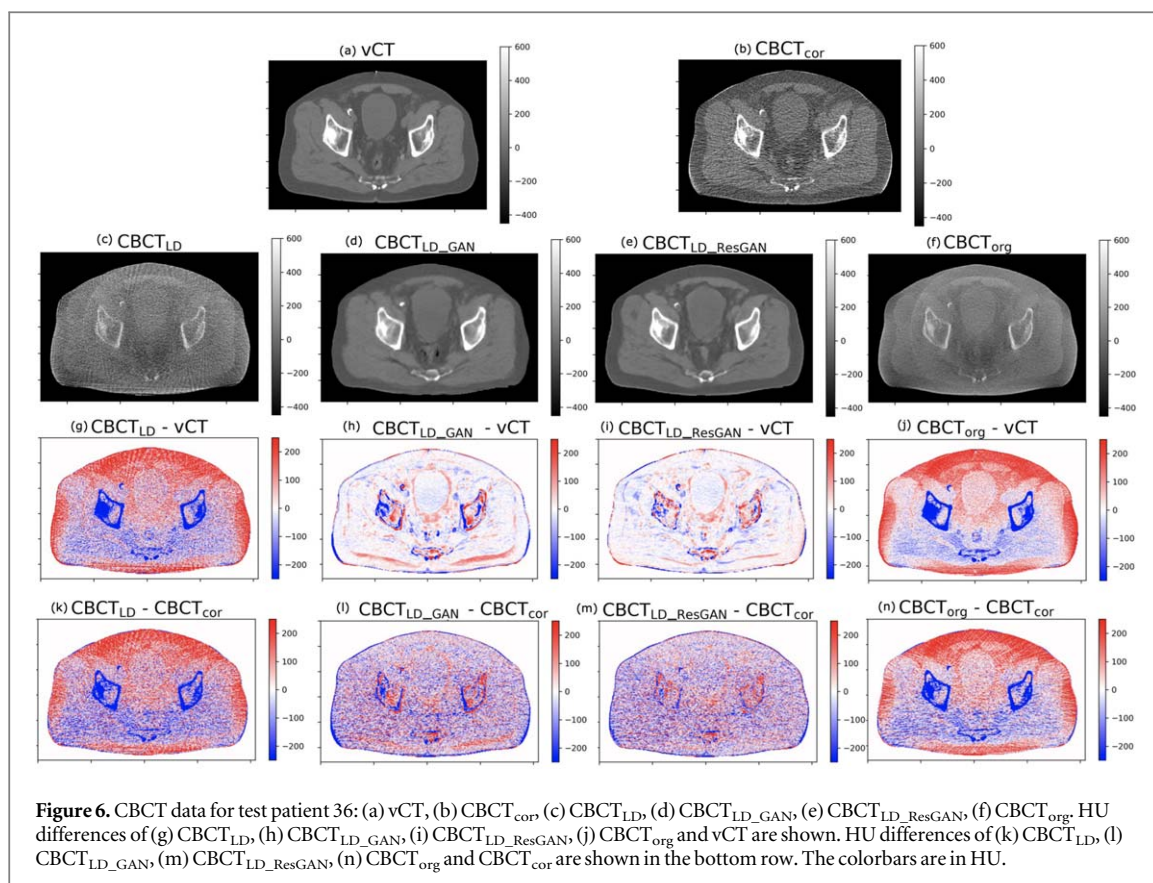## 3. Results

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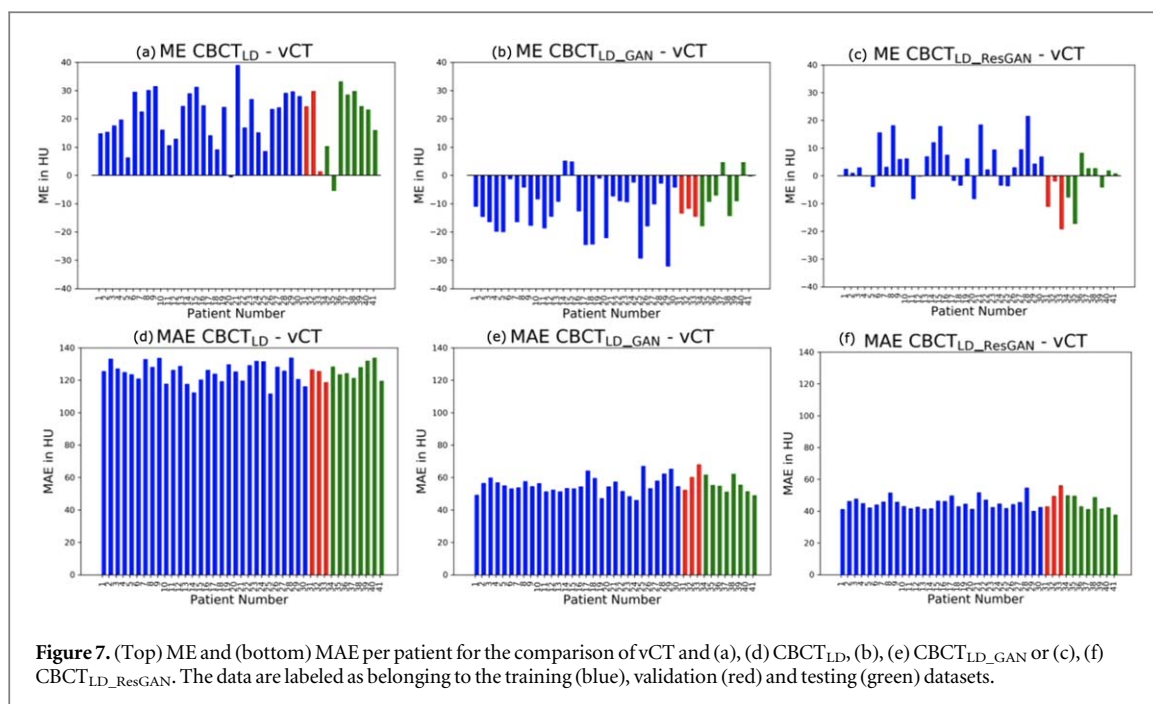
Figure 7. (Top) ME and (bottom) MAE per patient for the comparison of vCT and (a), (d) CBCT$_{LD}$, (b), (e) CBCT$_{LD\_GAN}$ or (c), (f) CBCT$_{LD\_ResGAN}$. The data are labeled as belonging to the training (blue), validation (red) and testing (green) datasets.

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 $CBCT_{LD\_ResGAN}$.

Test patient	$CBCT_{LD\_GAN}$			$CBCT_{LD\_ResGAN}$		
	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 $CBCT_{LD\_ResGAN}$.

Test patient	$CBCT_{LD\_GAN}$			$CBCT_{LD\_ResGAN}$		
	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 $CBCT_{LD\_ResGAN}$.
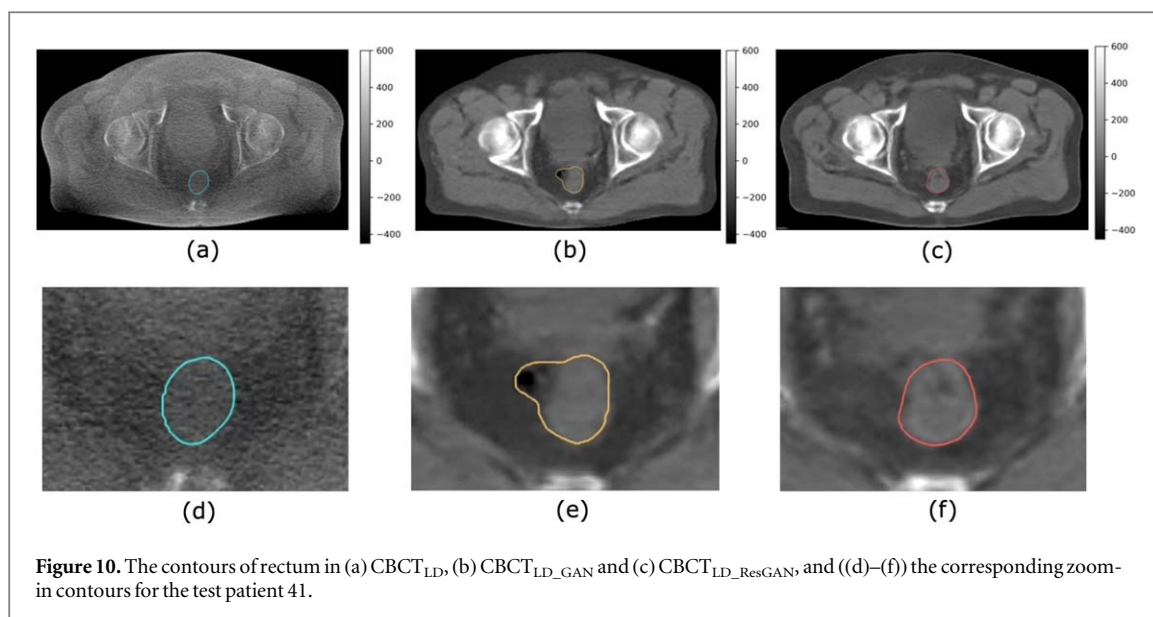
3.6. Anatomical fidelity

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

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