



OPEN

Cone-beam CT image quality improvement using Cycle-Deblur consistent adversarial networks (Cycle-Deblur GAN) for chest CT imaging in breast cancer patients

Hui-Ju Tien^{1,2,7}, Hsin-Chih Yang^{1,3,7}, Pei-Wei Shueng^{2,4} & Jyh-Cheng Chen^{1,5,6}✉

Cone-beam computed tomography (CBCT) integrated with a linear accelerator is widely used to increase the accuracy of radiotherapy and plays an important role in image-guided radiotherapy (IGRT). For comparison with fan-beam computed tomography (FBCT), the image quality of CBCT is indistinct due to X-ray scattering, noise, and artefacts. We proposed a deep learning model, "Cycle-Deblur GAN", combined with CycleGAN and Deblur-GAN models to improve the image quality of chest CBCT images. The 8706 CBCT and FBCT image pairs were used for training, and 1150 image pairs were used for testing in deep learning. The generated CBCT images from the Cycle-Deblur GAN model demonstrated closer CT values to FBCT in the lung, breast, mediastinum, and sternum compared to the CycleGAN and RED-CNN models. The quantitative evaluations of MAE, PSNR, and SSIM for CBCT generated from the Cycle-Deblur GAN model demonstrated better results than the CycleGAN and RED-CNN models. The Cycle-Deblur GAN model improved image quality and CT-value accuracy and preserved structural details for chest CBCT images.

The techniques of radiotherapy have developed rapidly from three-dimensional radiotherapy to volumetric modulated radiotherapy in recent decades^{1–3}. The dose distribution conformed to tumours accompanied by rapid dose falloff to critical organs. Image-guided radiotherapy (IGRT) was used to increase treatment accuracy during the radiotherapy course^{4–9}. Cone-beam computed tomography (CBCT) integrated into modern linear accelerators is the most widely used volume imaging system in radiotherapy^{10,11}. For the body scan of the X-ray Volumetric Imager (XVI system, Elekta company, Stockholm, Sweden), the average dose was in the range of 0.1–3.5 cGy^{12,13}. The artefacts of CBCT, including extinction artefacts, beam hardening artefacts, partial volume effects, exponential edge-gradient effects (EEGEs), aliasing artefacts, ring artefacts, and motion artefacts, influenced the image quality. In addition, noise and scatter are well known to produce additional artefacts¹⁴. However, the CT values in CBCT images may fluctuate because of scattering contamination, depending on the shape, positioning, size, and inner tissue structure^{15,16}. The original CT values of CBCT could not be used for dose calculation unless some correction methods were applied^{16–18}. For the multiple purposes of image quality improvement and CT value correction, we adopted the deep learning method to resolve these problems. In recent years, some deep learning methods have been designed to improve image quality in medical imaging. Convolutional neural networks (CNNs) and generative adversarial networks (GANs) are two main kinds of deep learning methods used to improve image quality. Hu et al. proposed a residual encoder-decoder convolutional neural network (RED-CNN)¹⁹, which was designed to remove noise from low-dose CT. The ground truth data, normal-dose CT, came from the National Biomedical Imaging Archive (NBIA), and the input, low-dose CT, was

¹Department of Biomedical Imaging and Radiological Sciences, National Yang-Ming University, B306, Experimental building No. 155, Sec. 2, Linong Street, Taipei 112, Taiwan. ²Division of Radiation Oncology, Department of Radiology, Far Eastern Memorial Hospital, New Taipei City, Taiwan. ³Data Science Degree Program, College of Electrical Engineering and Computer Science, National Taiwan University and Academia Sinica, Taipei, Taiwan. ⁴Faculty of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan. ⁵Department of Medical Imaging and Radiological Technology, Yuanpei University of Medical Technology, Hsinchu, Taiwan. ⁶School of Medical Imaging, Xuzhou Medical University, Jiangsu, China. ⁷These authors contributed equally: Hui-Ju Tien and Hsin-Chih Yang. ✉email: jyhcheng2010@gmail.com

produced by adding noise into the sinograms simulated from the normal-dose images. The output result of the CT images preserved structural details and reduced image noise.

The generative adversarial network (GAN), defined by Goodfellow et al.²⁰, contains two networks: generator and discriminator networks. The generator produced generated images from a convolutional neural network and scored by the discriminator compared to the ground truth images. After the training stage, the generator can generate images closer to the ground truth images. The adversarial loss function of the generator and the discriminator plays an important role in GAN.

GANs are known for their image quality improvement, but the vanilla version has numerous problems, such as non-convergence, model collapse and diminished gradient in the training step, because of Jensen–Shannon divergence between the model distribution and data distribution. To solve these problems, Arjovsky et al.²¹ provided the Wasserstein distance in GANs' loss function.

Because of 1-Lipschitz, WGAN suffers from the weight clipping problem that the weight may constrain on the boundary of the clipping constraint number. Therefore, Gulrajani et al.²² provides an alternative method for enforcing the Lipschitz constraint, called WGAN-GP.

The conditional GAN (cGAN)²³ is similar to an extension of the vanilla GAN. cGAN adds y between the generator and discriminator as additional information, such as word vectors, images or masks, to constrain the generator to generate the desired image. With the additional information y , the generator may be as similar to the traditional supervised network that we set input data as an image, not just noise.

The U-net^{24,25} architecture is developed for fast and precise object segmentation in a 2D image. With the shortcut in the U-net structure, previous layer characteristics can be transferred to the following layers, and backpropagation can avoid weight decay. That is, because of the shortcut, the key in U-net, we can use parameters to create better images. Kida et al.²⁶ took U-net as their model to improve image quality for CBCT. They took low-dose CBCT images as input data and planning computed tomography (pCT) as a ground truth for modelling.

Deblur-GAN²⁷ made a blurred image into a clear image. It also designed a blur algorithm to create blurred images as input data. Zhu et al.²⁸ recently designed CycleGAN as a more amazing image style transfer. It is unprecedented for the CycleGAN structure that two generators generate different domain images that can serve as inputs to another generator, and two generators can compose each other. In our objective, we can imagine that the model can make CBCT with the FBCT style. Kida et al.²⁹ used CycleGAN to synthesize improved CBCT as planning CT to improve the image quality of CBCT for pelvic images with soft tissue and bony structures. The purpose of our study was to improve the image quality of CBCT in truncated chest CT images. Our proposed method combines Deblur-GAN and CycleGAN to achieve more precise image transfer in chest CBCT images.

Methods

Data pre-processing. Fifteen breast cancer patients were enrolled in this study. Before radiotherapy, each patient underwent planning CT, that is, the FBCT acquired by a Big Bore CT scanner (Discovery CT590 RT, GE company, Boston, USA), for treatment planning. The acquisition parameters of the GE CT scanner were detector rows of 16, helical scan pitch of 0.938:1, slice thickness of 2.5 mm, and FOV of 50 cm. The adaptive statistical iterative reconstruction (ASiR) algorithm of 30% (SS30) was selected to reconstruct FBCT images. The SS30 denotes the selected ASiR mode as slice statistical reconstruction mode with 30% of the 100% ASiR, which was reconstructed with the original image³⁰. The reconstructed FBCT images were used for this study. During every treatment fraction, CBCT was performed for image registration. There are 185 CBCT image datasets, in light of 9856 CBCT images acquired by X-ray Volumetric Imager^{10,11} (XVI system, version R5.0, Elekta company, Stockholm, Sweden) using an optimization of Feldkamp backprojection reconstruction algorithm for training and testing. The acquisition parameters of XVI VolumeView were voltage of the X-ray tube of 120 kVp, current of 40 mA, acquisition time of 120 s, frame rate of 5.5 frames per second, and voxel size of 1 mm × 1 mm × 1 mm for chest CBCT images. An M20 protocol was selected for XVI acquisition, “M” was the FOV of 42.5 cm × 42.5 cm at the kV detector panel, and “20” was 27.67 cm in length at the isocenter of the field projections. FBCT, which was performed once during CT simulation, was used as the ground truth for each CBCT set for this study. Among the images from the fifteen patients, those from three patients (1150 images) were kept for testing, and those from the remaining 12 patients (8706 images) were used to train the network. Pre-processing to build an image pair of CBCT and FBCT was as follows. The CBCT images for each patient were three-dimensionally pre-aligned to each of the FBCT images by rigid registration by PMOD software (Version 3.7, PMOD Technologies, Zurich, Switzerland). To avoid any adverse impact from non-anatomical structures on a CBCT to FBCT registration and as a model training procedure, binary masks were created to separate the body region from non-anatomical regions. These masks were created by finding the maximum convex hull with a threshold CT value of − 1000. CycleGAN does not need paired data, and we could use unpaired images, which may provide various characteristics for model training²⁸. To accelerate the model training time, we still used paired data for modelling. We also clipped all image sizes from 512 × 512 down to 264 × 336 to minimize the anatomical region to accelerate the calculation time. We normalized the CT values of CBCT and FBCT images from a range of − 950 to 500 into a range of 0–1. Hence, the pixels with CT values of less than − 950 were assigned to 0, and those with CT values of higher than 500 were assigned to 1. The scale range for modelling is 0–1.

Image modelling. An overview of the model architecture is illustrated in Fig. 1. For generator G, we adopted the architecture for our generative networks from Kupyn et al.²⁷. They proposed Deblur-GAN, which showed impressive results for generating synthetic clear images from blurred images. The generator architecture is shown in Fig. 2. The inception block³¹ adapted by GoogLeNet can extract different kinds of features by different sizes of convolutions. In our study, the features from the large and small ranges can provide specific results and make the boundary clearer. The generative network contained one convolution block with stride two, two

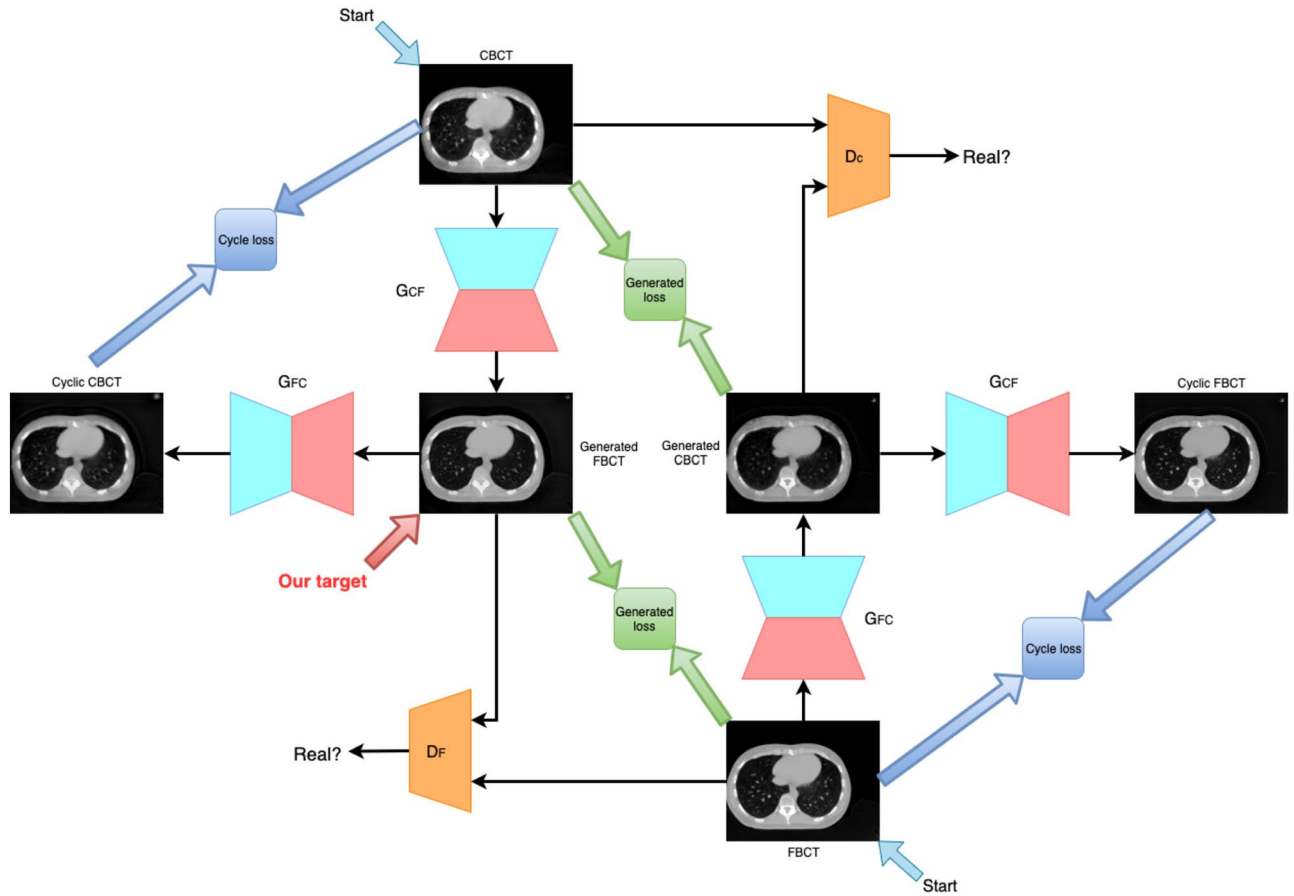


Figure 1. Architecture of Cycle-Deblur GAN.

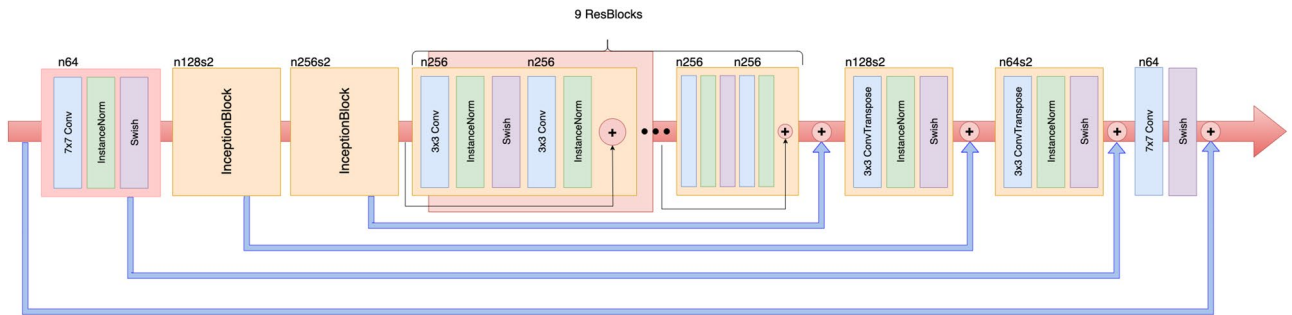


Figure 2. Generator architecture. n denotes the number of convolutional kernels, and s denotes stride. The default stride is 1, i.e., $n64s2$ denotes the convolution layer of 64 channels with stride 2.

inception blocks with stride two, nine residual blocks and two transposed convolution blocks. Each residual block consisted of a convolution layer, an instance normalization layer, and Swish activation^{32,33}. The inception block, a collection of convolution layers of different sizes, such as the 1×1 convolution layer, 5×5 convolution layer, 9×9 convolution layer and 13×13 convolution layer, could capture detailed and brief characters without changing the image size. To concatenate different convolution layers in the inception block with different priorities, as shown in Fig. 3, we multiply different weights by 1, 5, 9 and 13.

For discriminator D , shown in Fig. 4, inspired by Ledig et al.³⁴, we followed architectural guidelines and dropped the last full connection from the last layer to the convolution layer with the input patch image, which aims to classify whether overlapping image patches were real or generated. Such a patch-level discriminator architecture had fewer parameters than a full-image discriminator and could work on arbitrarily sized images in a fully convolutional method.

Loss functions. Our goal in this study is to define a deep neural network that finds a suitable mapping function that minimizes the loss functions. Let $x \in X$ and X was CBCT that was an input image. Let $y \in Y$ and Y be FBCT, which was the ground truth image. In general, we can define the mapping function as Eq. (1)

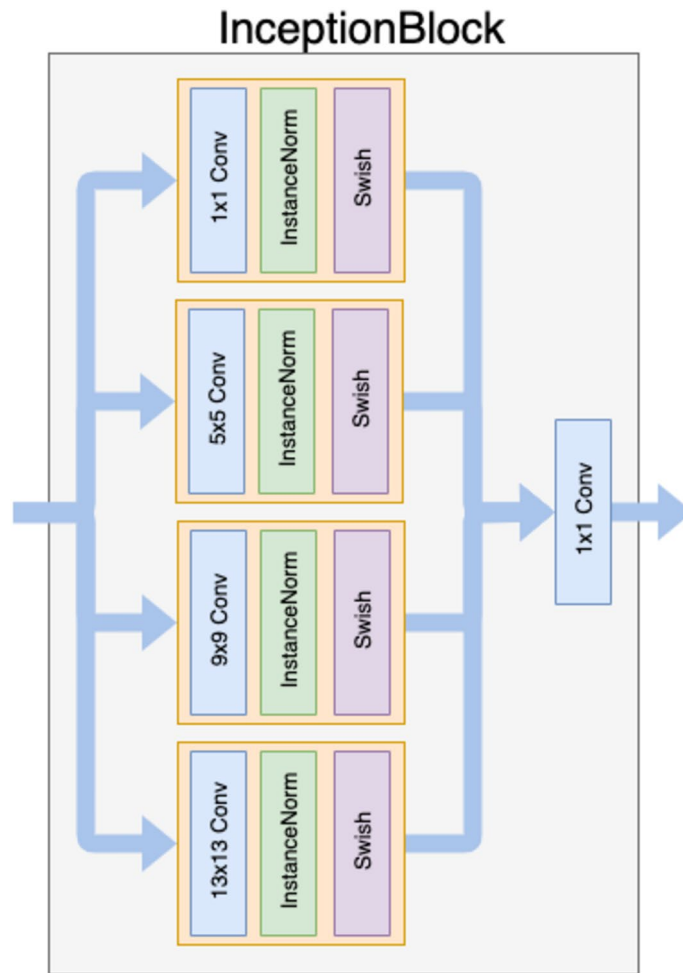


Figure 3. InceptionBlock architecture. It uses different kernel sizes of convolutional layers and deals with detailed and brief features.

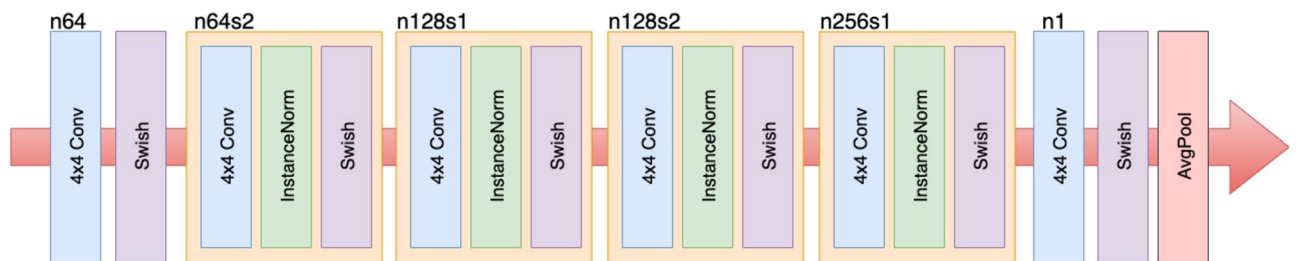


Figure 4. Discriminator architecture.

$$\hat{y} = G_{\theta}(x) \tag{1}$$

where \hat{y} is the synthesis FBCT image and G_{θ} is the generative model that transforms x to y . To obtain decent \hat{y} , the loss function for G_{θ} to generate a synthesis image from the input is shown in Eq. (2)

$$\hat{\theta} = \arg \min_{\theta} \sum_i L(G_{\theta}(x_i), y_i) \tag{2}$$

where (x_i, y_i) are paired CBCT and FBCT images. Inspired by Kida et al.²⁹, our model has two generative models $G_{CF} : X \rightarrow Y$ and $G_{FC} : Y \rightarrow X$ by input image pairs (x_i, y_i) . The two different generative models are trained to synthesize different targets such that G_{CF} generates synthesis FBCT by input CBCT; otherwise, G_{FC} outputs synthesis CBCT by giving FBCT. Moreover, there are two adversarial discriminators D_F and D_C , which aim to distinguish whether the output of the generative model is real or synthesis. For example, given FBCT input y ,

G_{FC} intends to generate synthesis CBCT \hat{x} , which will be as similar as real CBCT x to foolish discriminator D_C . In contrast, D_F will judge the reconstructed \hat{y} from generative model G_{CF} by feeding synthesis CBCT \hat{x} and real x . That is, a cyclic method in which the discriminators are not only discriminator synthesized CBCT (or FBCT) but also reconstructed CBCT. The key idea is that generators and discriminators are trained on each other to enhance their accuracy. Therefore, our objective function is a minimax problem, as shown in Eq. (3):

$$\min_{G,F} \max_{D_F,D_C} L_{gan}(G_{CF}, D_F) + L_{gan}(G_{FC}, D_C) \tag{3}$$

Additionally, our novel networks include five types of loss functions: adversarial loss (*adv*); cycle-consistency loss (*cycle*); generated loss (*generated*); identity loss (*identity*); and Sobel filter loss (*Sobel*)³⁵.

For the discriminators, we adopt WGAN-gp as our objective functions as Eqs. (4) and (5):

$$\min_{G_{CF}} \max_{D_F} L_{WGAN-gp}(G_{CF}, D_F) = \mathbb{E}_x[D_F(G_{CF}(x))] - \mathbb{E}_y[D_F(y)] + \lambda \mathbb{E}_{\curvearrowright} [(\|\nabla D_F(\curvearrowright)\|_2 - 1)^2] \tag{4}$$

$$\min_{G_{FC}} \max_{D_C} L_{WGAN-gp}(G_{FC}, D_C) = \mathbb{E}_y[D_C(G_{FC}(y))] - \mathbb{E}_x[D_C(x)] + \lambda \mathbb{E}_{\curvearrowleft} [(\|\nabla D_C(\curvearrowleft)\|_2 - 1)^2] \tag{5}$$

where $\mathbb{E}[\bullet]$ is the expectation operator, the first two terms are the negative Wasserstein distance, which determines how much better the real term is than the synthesized term, and the last term is the gradient penalty, in which λ is a regularization parameter, $\curvearrowright = \varepsilon y + (1 - \varepsilon)G_{CF}(x)$ and $\curvearrowleft = \varepsilon x + (1 - \varepsilon)G_{FC}(y)$.

Therefore, the overall loss function for the discriminator is shown in Eq. (6):

$$L_D = L_{WGAN-gp}(G_{CF}, D_F) + L_{WGAN-gp}(G_{FC}, D_C) \tag{6}$$

The adversarial loss for generators is as Eq. (7):

$$L_{adv} = -\mathbb{E}_x[D_F(G_{CF}(x))] - \mathbb{E}_y[D_C(G_{FC}(y))] \tag{7}$$

Because of adversarial training, discriminators and generators will handle each best. That is, for the loss of discriminators, they encourage real images to score high and synthesized images as low like $L_{WGAN-gp}(G_{FC}, D_C)$. However, for the generator loss, they intend to let their synthesized images to score much higher, such as $-\mathbb{E}_y[D_C(G_{FC}(y))]$. Thus, they hold different loss function tasks and alternatively train each other.

The loss function of cycle consistency is shown in Eq. (8):

$$L_{cycle} = \mathbb{E}_x[\|x - G_{FC}(G_{CF}(x))\|_2] + \mathbb{E}_y[\|y - G_{CF}(G_{FC}(y))\|_2] \tag{8}$$

The loss maps $x \rightarrow G_{CF}(x) \rightarrow G_{FC}(G_{CF}(x)) \approx x$ and $y \rightarrow G_{FC}(y) \rightarrow G_{CF}(G_{FC}(y)) \approx y$, which are referred to forward cycle consistency loss and backward cycle consistency, respectively.

To make generators able to generate synthesis images, we define the generated loss, which can be expressed as Eq. (9):

$$L_{generated} = \mathbb{E}_{x,y}[\|y - G_{CF}(x)\|_1] + \mathbb{E}_{x,y}[\|x - G_{FC}(y)\|_1] \tag{9}$$

The term can maintain the mapping of $\hat{y} \approx y$ and $\hat{x} \approx x$, and it is our final objective. Identity loss is shown in Eq. (10):

$$L_{identity} = \mathbb{E}_x[\|G_{CF}(x) - G_{CF}(G_{CF}(x))\|_2] + \mathbb{E}_y[\|G_{FC}(y) - G_{FC}(G_{FC}(y))\|_2] \tag{10}$$

The idea for identity loss is that the generative model will transform to the input image style regardless of the input, i.e., given synthesis FBCT to generative model G_{CF} , the model should still output the image with FBCT style, even if the input is not a real CBCT. A similar method for G_{FC} is given synthesis CBCT with the same style output as synthesis CBCT. We use a regularization term to keep the training from overfitting.

Sobel filter loss³⁵ is shown in Eq. (11):

$$L_{sobel} = \mathbb{E}_{x,y}[\delta_1 \|y - G_{CF}(x)\|_1] + \mathbb{E}_{x,y}[\delta_2 \|y - G_{CF}(x)\|_1] + \mathbb{E}_{x,y}[\delta_1 \|x - G_{FC}(y)\|_1] + \mathbb{E}_{x,y}[\delta_2 \|x - G_{FC}(y)\|_1], \tag{11}$$

where δ_1 and δ_2 are Sobel gradient operators. The Sobel operator filters the gradient of image colour intensity by δ_1 and δ_2 and keeps the edges blurred.

The total objective function for generators can be defined as Eq. (12):

$$L_G = \lambda_{adv} L_{adv} + \lambda_{cycle} L_{cycle} + \lambda_{generated} L_{generated} + \lambda_{identity} L_{identity} + \lambda_{sobel} L_{sobel} \tag{12}$$

The hyper-parameters λ_{adv} , λ_{cycle} , $\lambda_{generated}$, $\lambda_{identity}$, λ_{sobel} are changed during the training time, and λ_{gan} always has a weight of 1. In the first 10 epochs, the loss function is similar to a vanilla GAN, which means that all of the hyper-parameters are 0 except L_{gan} . In the next 10 epochs, we assigned $\lambda_{cycle} = 5$, $\lambda_{identity} = 5$, $\lambda_{generated} = 1$, $\lambda_{sobel} = 0$ as the cycle-consistent period to be the main target. After 20 epochs, we adopt $\lambda_{cycle} = 10$, $\lambda_{identity} = 10$, $\lambda_{generated} = 10$, $\lambda_{sobel} = 10^{-4}$ as our main loss function parameters and because if the Sobel gradient loss is too high, the total loss may be misleading; thus, we adopt $\lambda_{sobel} = 10^{-4}$ as the optimal weight. Our goal is not only for the model to learn style transfer between CBCT and FBCT by cycle consistency but also to make the model fit another style by obtaining direct loss, such as $\lambda_{generated} = 10$ and $\lambda_{sobel} = 10^{-4}$. The networks are trained with a learning rate of 10^{-4} , with the Adam optimizer³⁶ and with

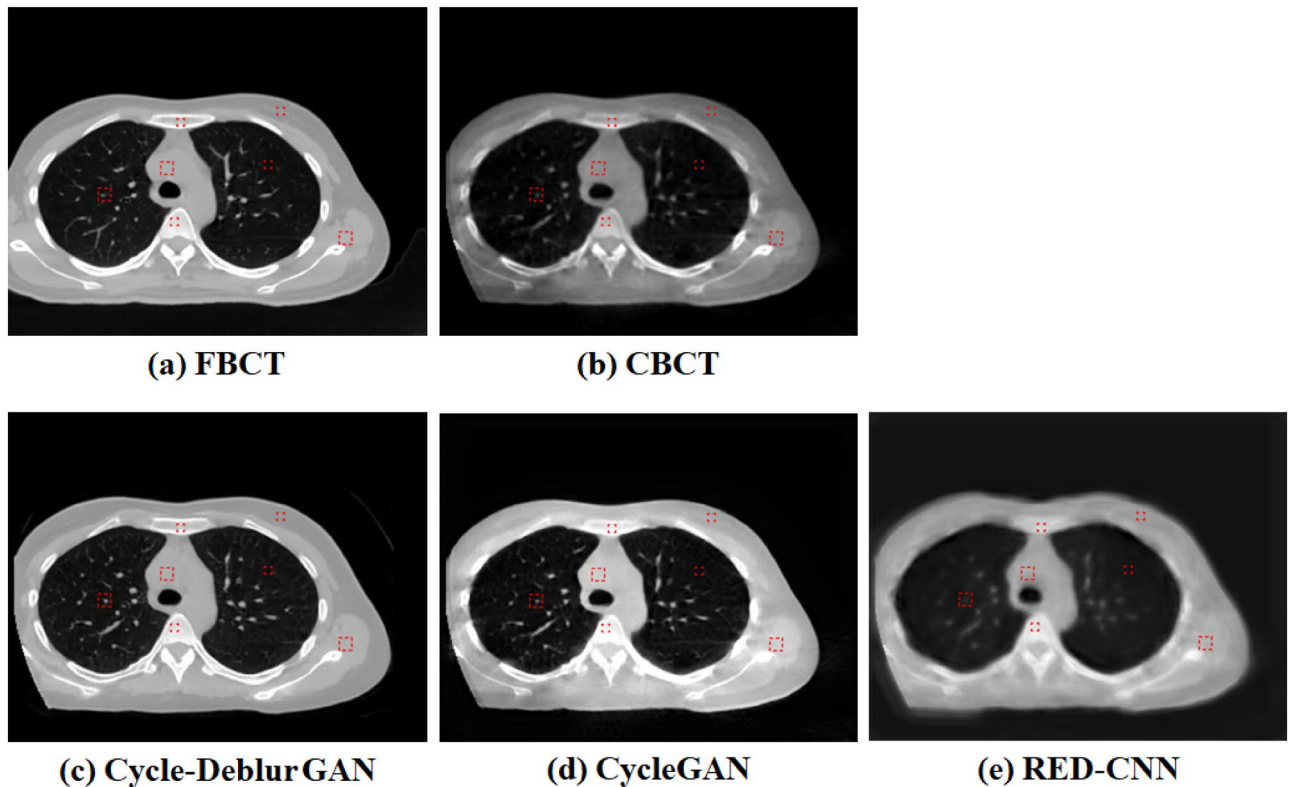


Figure 5. (a) FBCT, (b) CBCT, and the modeling results of (c) Cycle-Deblur GAN, (d) CycleGAN, and (e) RED-CNN were performed by XVI software and PMOD software. ($W = 1450$, $L = 225$ for all CT images).

a batch size of 8. Since GAN has difficulty finding the best minimal loss, we decay the learning rate by a cosine annealing scheduler while keeping the same learning rate in the first 20 epochs.

CT images used for training contain a large range of black backgrounds around the body. The black border will cause the model to be less sensitive to edge pixels. To improve the model stability and prevent overfitting, data augmentation is applied during the training time. Every image pair (CBCT and FBCT), loaded from a dataset, will be synchronously randomly cropped into 128×128 sizes. Second, the image pairs are synchronously randomly rotated angles between -20° and 20° and horizontally and vertically flipped. Then, the image pairs are generated.

We used a personal computer with a single GPU (Nvidia Titan XP) and a CPU (Intel Xeon E5-2620 v4 @ 2.10 GHz) with 64 GB memory, running Ubuntu 18.04 LTS. We implemented our method with Python 3.6.7 and PyTorch 1.0.0. The training time for 200 epochs needs approximately 3 days.

Quantitative evaluation. The CT value is a linear transformation of the original linear attenuation coefficient measurement into one in which the radiodensity of distilled water at standard pressure and temperature (STP) is defined as the zero CT value, while the radiodensity of air at STP is defined as -1024 HU. The CT values of the different regions for CBCT, RED-CNN model images, CycleGAN model images, and CycleDeblur GAN images were compared to FBCT. We chose three kinds of soft tissue, breast, muscle, and mediastinum; two kinds of bony structures, sternum and spine; and lung tissue to compare CT values. To evaluate the performance of the proposed method, Cycle-Deblur GAN, we chose existing metrics such as the peak-signal-to-noise-ratio (PSNR), which is measured to capture the reduction in noise, and the structural similarity index measure (SSIM)³⁷, which is one of the human visual system-based metrics and to evaluate different attributes such as luminance, contrast, and structure comprehensively. The mean absolute error (MAE) is one of the quantitative evaluations and is also used in our objective (loss) function. The PSNR is calculated from the mean square error (MSE), which is commonly used to measure distortion. There were seven regions of interest (ROIs) shown in Fig. 5, which were used to compare the CT value, MSE, PSNR, and SSIM. We define MSE, PSNR, and SSIM as Eqs. (13)–(21):

$$MAE = \frac{1}{N} \sum |y - G_{CF}(x)| \quad (13)$$

$$MSE = \frac{1}{N} \sum (y - G_{CF}(x))^2 \quad (14)$$

$$PSNR = 10 \times \log_{10} \left(\frac{MAX_I^2}{MSE} \right), \text{ where } MAX = 65,535 \text{ for } 16\text{-bit image} \quad (15)$$

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (16)$$

$$\mu_x = \frac{1}{MN} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} x_{i,j} \quad (17)$$

$$\mu_y = \frac{1}{MN} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} y_{i,j} \quad (18)$$

$$\sigma_x = \sqrt{\frac{1}{MN-1} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (x_{i,j} - \mu_x)^2} \quad (19)$$

$$\sigma_y = \sqrt{\frac{1}{MN-1} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (y_{i,j} - \mu_y)^2} \quad (20)$$

$$\sigma_{xy} = \frac{1}{MN-1} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (x_{i,j} - \mu_x)(y_{i,j} - \mu_y) \quad (21)$$

Blind image observer study. The FBCT, CBCT, generated CBCT images from the Cycle-Deblur GAN model, CycleGAN model and RED-CNN model were scored by thirteen medical imaging professionals, including seven radiation oncologists and six medical physicists, using a five-grade scoring method. FBCT was defined as five out of five as the ground truth. The CBCT and generated CBCT images from the Cycle-Deblur GAN model, Cycle-GAN model and RED-CNN model were scored by comparison to FBCT.

Ethical statement. We confirmed that all methods were carried out in accordance with relevant guidelines and regulations, and informed consent for patients was waived by the Research Ethics Review Committee of Far Eastern Memorial Hospital (FEMH). Images were provided by FEMH and approved by the Research Ethics Review Committee of FEMH (107144-E).

Results

The generated CBCT images from RED-CNN, CycleGAN, and our proposed Cycle-Deblur GAN with seven ROIs are shown in Fig. 5. The generated CBCT images from Cycle-Deblur GAN performed better visualization than those from CycleGAN and RED-CNN. The CT images of the lung, soft tissue, and bone in seven specific ROIs are shown in Figs. 5, 6, 7 and 8. For the lung ROIs, Cycle-Deblur GAN demonstrated more lung detail preservation than other methods. The ROIs were analysed by the CT value, MAE, PSNR, and SSIM. In Table 1, the CT values of seven ROIs in Cycle-Deblur GAN, CycleGAN, and RED-CNN are shown as the mean values with standard deviations.

The CT values of the ROI in the right lung for FBCT, CBCT, generated CBCT from Cycle-Deblur GAN, CycleGAN, and RED-CNN were -775 ± 123 , -862 ± 77 , -778 ± 166 , -807 ± 121 , and -780 ± 33 , respectively. The CT values of the ROI in breast tissue for FBCT, CBCT, generated CBCT from Cycle-Deblur GAN, CycleGAN, and RED-CNN were -80 ± 20 , -336 ± 36 , -81 ± 6 , 26 ± 30 , and -42 ± 10 , respectively. The CT values of the ROI in the sternum for FBCT, CBCT, generated CBCT from RED-CNN, CycleGAN, and Cycle-Deblur GAN were 245 ± 151 , 66 ± 101 , 230 ± 103 , 411 ± 52 , and 237 ± 30 , respectively. In Table 2, it can be seen that the MAE of our proposed Cycle-Deblur GAN had the smallest value compared to CycleGAN and RED-CNN for different ROIs.

In Table 3, the PSNRs for the CBCT, Cycle-Deblur GAN, CycleGAN, and RED-CNN models in the right lung were 21.86, 25.05, 25.38, and 22.88, respectively. The PSNRs of the Cycle-Deblur GAN and CycleGAN models in the lung showed comparable results. The PSNRs for the CBCT, Cycle-Deblur GAN, CycleGAN, and RED-CNN models for breast tissue were 15.01, 36.49, 22.49, and 29.75, respectively. The PSNRs for the CBCT, Cycle-Deblur GAN, CycleGAN, and RED-CNN models for the sternum were 17.74, 25.39, 17.45, and 20.71, respectively. The Cycle-Deblur GAN was shown to have a better PSNR in the breast, mediastinum, muscle, sternum, and spine than the other models.

In Table 4, the SSIM results of CBCT, Cycle-Deblur GAN, CycleGAN, and RED-CNN for the left lung were 0.9922, 0.9993, 0.9992, and 0.9986, respectively. The SSIM results of CBCT, Cycle-Deblur GAN, CycleGAN, and RED-CNN for the sternum were 0.8759, 0.9118, 0.5681, and 0.6849, respectively. The SSIMs of the Cycle-Deblur

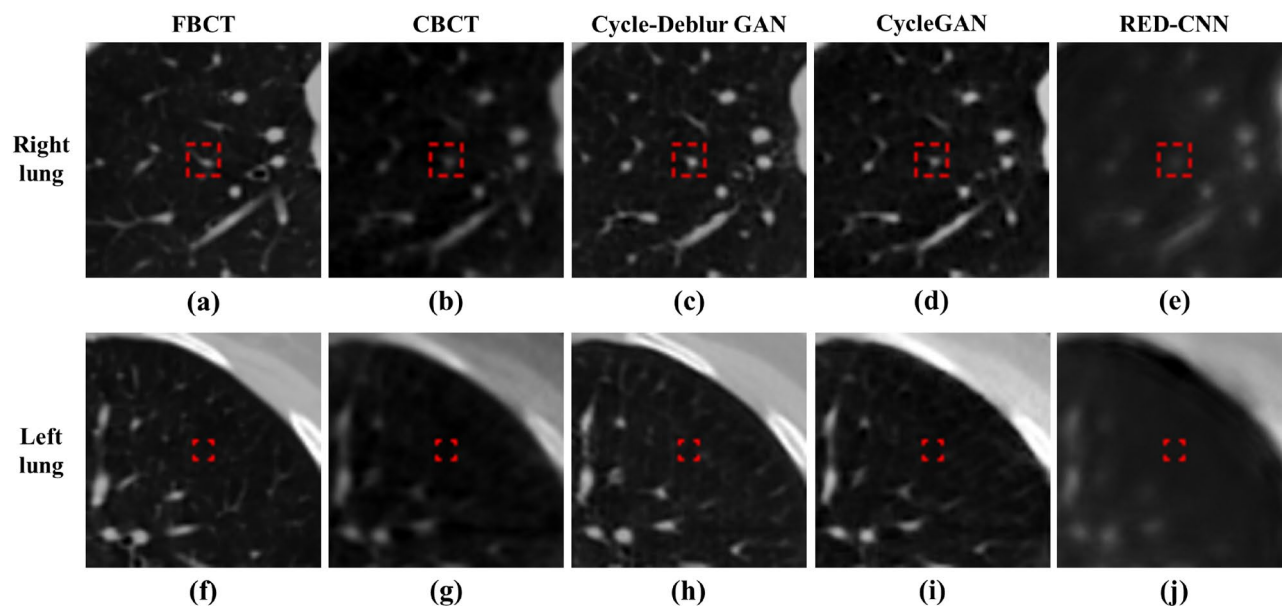


Figure 6. The right and left lung ROIs for quantitative evaluation are shown in (a,f): FBCT, (b,g): CBCT, (c,h): Cycle-Deblur GAN, (d,i): CycleGAN and (e,j): RED-CNN were performed by XVI software and PMOD software. ($W = 1450$, $L = -225$ for all CT images).

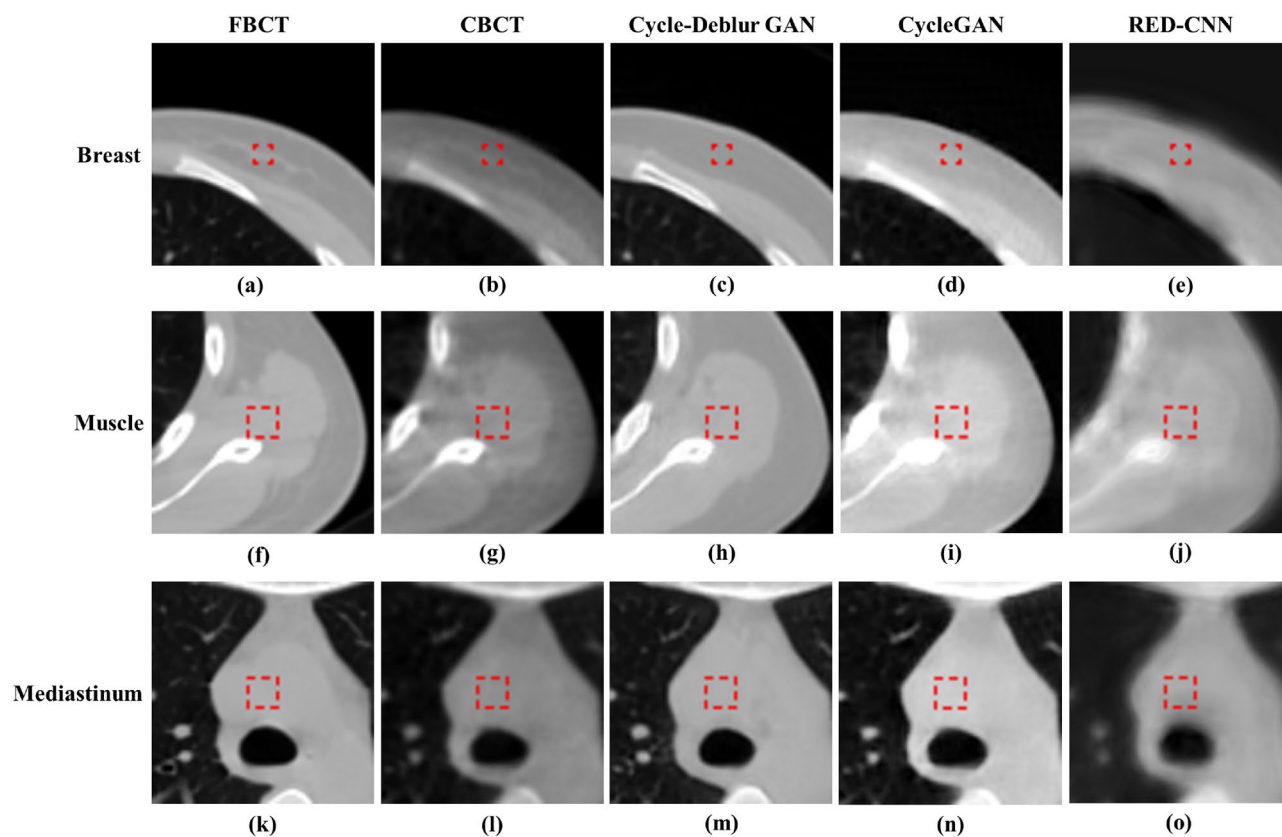


Figure 7. The soft tissue ROIs of the breast, muscle and mediastinum for quantitative evaluation are shown in (a,f,k): FBCT, (b,g,l): CBCT, (c,h,m): Cycle-Deblur GAN, (d,i,n): CycleGAN and (e,j,o): RED-CNN were performed by XVI software and PMOD software. ($W = 1450$, $L = -225$ for all CT images).

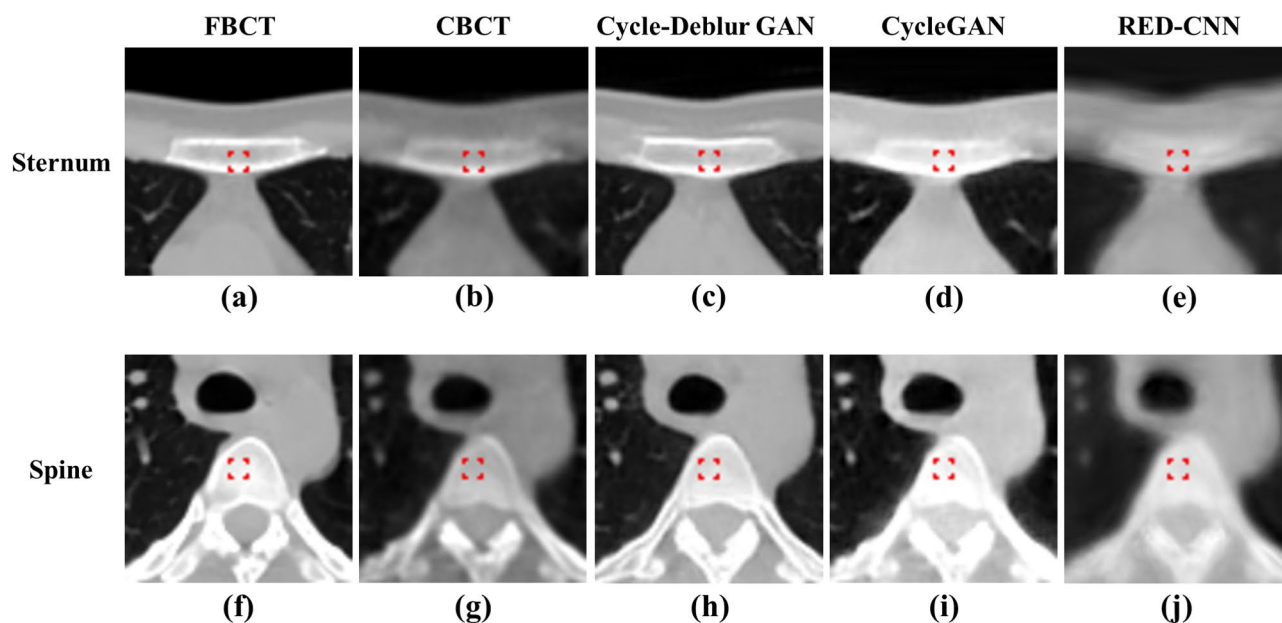


Figure 8. The bony structure ROIs of the sternum and spinal process for quantitative evaluation are shown in (a,f): FBCT, (b,g): CBCT, (c,h): Cycle-Deblur GAN, (d,i): CycleGAN and (e,j): RED-CNN were performed by XVI software and PMOD software. ($W = 1450$, $L = -225$ for all CT images).

	FBCT (mean \pm SD)	CBCT (mean \pm SD)	Cycle-Deblur GAN (mean \pm SD)	CycleGAN (mean \pm SD)	RED-CNN (mean \pm SD)
Right lung	-775 ± 123	-862 ± 77	-778 ± 166	-807 ± 121	-780 ± 33
Left lung	-806 ± 34	-893 ± 33	-811 ± 28	-852 ± 39	-795 ± 16
Breast	-80 ± 20	-336 ± 36	-81 ± 6	26 ± 30	-42 ± 10
Mediastinum	48 ± 6	-137 ± 12	40 ± 7	215 ± 12	-10 ± 45
Muscle	66 ± 9	-112 ± 23	56 ± 12	238 ± 25	75 ± 25
Sternum	245 ± 151	66 ± 101	230 ± 103	411 ± 52	237 ± 30
Spine	304 ± 19	83 ± 23	308 ± 33	443 ± 27	334 ± 22

Table 1. The CT value comparison of different ROIs.

MAE	Right lung	Left lung	Breast	Mediastinum	Muscle	Sternum	Spine
CBCT	0.06333	0.04346	0.17647	0.12745	0.12325	0.12342	0.15261
Cycle-Deblur GAN	0.03647	0.01242	0.01166	0.00659	0.00894	0.04532	0.01634
CycleGAN	0.03933	0.01427	0.07277	0.11502	0.11804	0.11449	0.09553
RED-CNN	0.04757	0.01841	0.02865	0.04075	0.01510	0.07865	0.02244

Table 2. MAE comparison of different models at different sites.

PSNR	Right lung	Left lung	Breast	Mediastinum	Muscle	Sternum	Spine
CBCT	21.86	27.01	15.01	17.87	18.11	17.74	16.28
Cycle-Deblur GAN	25.05	36.07	36.49	40.92	39.18	25.39	32.88
CycleGAN	25.38	35.98	22.49	18.76	18.47	17.45	20.24
RED-CNN	22.88	33.85	29.75	25.64	34.29	20.71	31.25

Table 3. PSNR comparison of different models at different sites.

GAN and CycleGAN models in the lung showed comparable results. The Cycle-Deblur GAN was shown to have better SSIM in the breast, mediastinum, muscle, sternum, and spine than the other models.

SSIM	Right lung	Left lung	Breast	Mediastinum	Muscle	Sternum	Spine
CBCT	0.9811	0.9922	0.9134	0.9345	0.9323	0.8759	0.7566
Cycle-Deblur GAN	0.9923	0.9993	0.9972	0.9988	0.9981	0.9118	0.9469
CycleGAN	0.9926	0.9992	0.9726	0.8953	0.8728	0.5681	0.5172
RED-CNN	0.9870	0.9986	0.9924	0.9722	0.9922	0.6849	0.9476

Table 4. SSIM comparison of different models at different sites.

	CBCT	Cycle-Deblur GAN	CycleGAN	RED-CNN
RO 1	3	5	4	1
RO 2	2	4	3	1
RO 3	2	4	3	1
RO 4	4	5	4	1
RO 5	2	4	3	1
RO 6	2	4	3	1
RO 7	4	5	3	2
MP 1	4	5	4	3
MP 2	2	4	3	1
MP 3	2	4	3	1
MP 4	2	4	3	1
MP 5	4	5	3	2
MP 6	3	5	4	1
Mean	2.8	4.5	3.3	1.3

Table 5. The results of five-grade scoring method. *RO* radiation oncologist, *MP* medical physicist.

In the blind image observer study, the median years of experience of radiation oncologists and medical physicists were 11 years, with a range from 6 to 33 years, and 8 years, with a range from 6 to 22 years. The results are shown in Table 5. The mean scores of the CBCT and generated CBCT images from the Cycle-Deblur GAN model, CycleGAN model and RED-CNN model were 2.8, 4.5, 3.3, and 1.3, respectively. The CBCT generated from the Cycle-Deblur GAN model scored higher than the other models.

Discussion

Our proposed Cycle-Deblur GAN consists of CycleGAN and Deblur-GAN with increasing shortcut numbers and inception blocks to preserve the detailed structure. For the activation layers, since the performance of Swish was better than ReLU in test set accuracy when changing the number of layers³³, we adopted it as our activation function for Cycle-Deblur GAN. Satoshi Kida et al.²⁹ proposed CycleGAN for visual enhancement in pelvic CT images. However, CycleGAN could not perform better to improve image quality in PSNR, SSIM, and MAE for chest CBCT images, as shown in bone and soft tissues. In the RED-CNN¹⁹ model, 14 input images and ground truth images were created from the same projections for comparison. However, in our study, CBCT and FBCT were acquired from one patient on different days. The image registration was needed before modelling. When using the RED-CNN model to train CBCT and FBCT in our study, the misalignment influenced the results of the RED-CNN model, which showed blurred results. For the Cycle-Deblur GAN model, the CBCT and FBCT images were both treated as the input images to derive a more stable model. Hence, the registration error due to the different acquisition dates of FBCT and CBCT for Cycle-Deblur GAN represented less influence, and the Cycle-Deblur GAN model could generate higher image quality images.

CT values of the original CBCT images may fluctuate for the same material in the different relative positions being scanned in the image volume¹⁵. In Table 1, the CT values of the generated CBCT from Cycle-Deblur GAN showed better results than those from RED-CNN and CycleGAN in the breast, lung, muscle, mediastinum, and sternum. For bone tissue, including the spine and sternum, the CBCT generated from the RED-CNN model showed a better result in the quantitative analysis of PSNR and SSIM. However, the visual enhancement of the generated CBCT from the RED-CNN model, as shown in Fig. 8, was blurred. The ROI size of the spine, breast, and sternum was smaller than others due to contouring the same structure in one ROI. The PSNR and SSIM of our proposed method demonstrated better results than other methods and showed more detail preservation, especially in lung tissues, as shown in Fig. 6.

Once the Cycle-Deblur GAN was well trained, the generator of the Cycle-Deblur GAN was used for testing. In the testing process, we input the CBCT image passing through the generator model and receive the generated CBCT. The generated CBCT with high image quality benefits the image verification by the oncologist. The average time of the generator to produce an improved CBCT image was approximately 0.17 s and depended on the hardware used.

Limitations of the study

We proposed the Cycle-Deblur GAN method to model chest CBCT images and obtain better results than the CycleGAN and RED-CNN methods in this study. The input data were all chest CT images, and the modelling generator used to produce the generated CBCT may be limited to the chest region only. Smoothed images of better PSNR with lower noise may be accompanied by lower contrast. Hence, the visualized evaluation of SSIM was also evaluated in our study.

Conclusions

The CBCT generated by our proposed Cycle-Deblur GAN model demonstrated higher PSNR and SSIM results in soft tissue, lung, and bony structures with improved image quality. The generated CBCT images with accurate CT values can be used for adaptive dose calculation in radiotherapy. The overall artefact of CBCT was well removed by using this model. This model enhanced the structural details in the lung, soft tissue, and bony structure and showed better visualization than the original CBCT. The Cycle-Deblur GAN model improved the image quality of CBCT, preserved structural details and provided accurate CT values for dose calculation. The high image quality and accurate CT values of CBCT would assist the development of radiomics in our future work.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 1 June 2020; Accepted: 23 December 2020

Published online: 13 January 2021

References

1. Yoo, S., Blitzblau, R., Yin, F. F. & Horton, J. K. Dosimetric comparison of preoperative single-fraction partial breast radiotherapy techniques: 3D CRT, noncoplanar IMRT, coplanar IMRT, and VMAT. *J. Appl. Clin. Med. Phys.* **16**, 5126. <https://doi.org/10.1120/jacmp.v16i1.5126> (2015).
2. Xu, D., Li, G., Li, H. & Jia, F. Comparison of IMRT versus 3D-CRT in the treatment of esophagus cancer: A systematic review and meta-analysis. *Medicine (Baltimore)* **96**, e7685. <https://doi.org/10.1097/md.0000000000007685> (2017).
3. Gupta, T., Kannan, S., Ghosh-Laskar, S. & Agarwal, J. P. Systematic review and meta-analyses of intensity-modulated radiation therapy versus conventional two-dimensional and/or three-dimensional radiotherapy in curative-intent management of head and neck squamous cell carcinoma. *PLoS One* **13**, e0200137. <https://doi.org/10.1371/journal.pone.0200137> (2018).
4. Lee, J. *et al.* Image-guided study of inter-fraction and intra-fraction set-up variability and margins in reverse semi-decubitus breast radiotherapy. *Radiat. Oncol.* **13**, 254. <https://doi.org/10.1186/s13014-018-1200-1> (2018).
5. Lefkopoulos, D., Ferreira, I., Isambert, A., Le Péchoux, C. & Mornex, F. Present and future of the image guided radiotherapy (IGRT) and its applications in lung cancer treatment. *Cancer Radiother.* **11**, 23–31. <https://doi.org/10.1016/j.canrad.2006.10.001> (2007).
6. Beddok, A. & Blanchard, P. Image-guided radiotherapy for head and neck carcinoma. *Cancer Radiother.* **22**, 617–621. <https://doi.org/10.1016/j.canrad.2018.06.015> (2018).
7. Ippolito, E. *et al.* IGRT in rectal cancer. *Acta Oncol.* **47**, 1317–1324. <https://doi.org/10.1080/02841860802256459> (2008).
8. Rossi, M. *et al.* Dosimetric effects of anatomical deformations and positioning errors in VMAT breast radiotherapy. *J. Appl. Clin. Med. Phys.* **19**, 506–516. <https://doi.org/10.1002/acm2.12409> (2018).
9. Lozano Ruiz, F. J. *et al.* The importance of image guided radiotherapy in small cell lung cancer: Case report and review of literature. *Rep. Pract. Oncol. Radiother.* **25**, 146–149. <https://doi.org/10.1016/j.rpor.2019.12.013> (2020).
10. McBain, C. A. *et al.* X-ray volumetric imaging in image-guided radiotherapy: The new standard in on-treatment imaging. *Int. J. Radiat. Oncol. Biol. Phys.* **64**, 625–634. <https://doi.org/10.1016/j.ijrobp.2005.09.018> (2006).
11. Thorson, T. & Prosser, T. X-ray volume imaging in image-guided radiotherapy. *Med. Dosim.* **31**, 126–133 (2006).
12. Giaddui, T., Cui, Y., Galvin, J., Yu, Y. & Xiao, Y. Comparative dose evaluations between XVI and OBI cone beam CT systems using Gafchromic XRQA2 film and nanoDot optical stimulated luminescence dosimeters. *Med. Phys.* **40**, 062102 (2013).
13. Song, W. Y. *et al.* A dose comparison study between XVI* and OBI* CBCT systems. *Med. Phys.* **35**, 480–486 (2008).
14. Schulze, R. *et al.* Artefacts in CBCT: A review. *Dentomaxillofac. Radiol.* **40**, 265–273 (2011).
15. Abe, T. *et al.* Method for converting cone-beam CT values into Hounsfield units for radiation treatment planning. *Int. J. Med. Phys. Clin. Eng. Radiat. Oncol.* **6**, 361–375 (2017).
16. Rong, Y., Smilowitz, J., Tewatia, D., Tomé, W. A. & Paliwal, B. Dose calculation on kV cone beam CT images: An investigation of the Hu-density conversion stability and dose accuracy using the site-specific calibration. *Med. Dosim.* **35**, 195–207 (2010).
17. Niu, T. *et al.* Shading correction for on-board cone-beam CT in radiation therapy using planning MDCT images. *Med. Phys.* **37**, 5395–5406 (2010).
18. Giacometti, V., Hounsell, A. H. & McGarry, C. K. A review of dose calculation approaches with cone beam CT in photon and proton therapy. *Phys. Med.* **76**, 243–276 (2020).
19. Chen, H. *et al.* Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans. Med. Imaging* **36**, 2524–2535. <https://doi.org/10.1109/tmi.2017.2715284> (2017).
20. Goodfellow, I. J. *et al.* Generative Adversarial Networks. <http://arxiv.org/abs/1406.2661> (arXiv e-prints) (2014). <https://ui.adsabs.harvard.edu/abs/2014arXiv1406.2661G>.
21. Arjovsky, M., Chintala, S. & Bottou, L. Wasserstein GAN. <http://arxiv.org/abs/1701.07875> (arXiv e-prints) (2017). <https://ui.adsabs.harvard.edu/abs/2017arXiv170107875A>.
22. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. & Courville, A. Improved Training of Wasserstein GANs. <http://arxiv.org/abs/1704.00028> (arXiv e-prints) (2017). <https://ui.adsabs.harvard.edu/abs/2017arXiv170400028G>.
23. Mirza, M. & Osindero, S. Conditional Generative Adversarial Nets. <http://arxiv.org/abs/1411.1784> (arXiv e-prints) (2014). <https://ui.adsabs.harvard.edu/abs/2014arXiv1411.1784M>.
24. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. <http://arxiv.org/abs/1505.04597> (arXiv e-prints) (2015). <https://ui.adsabs.harvard.edu/abs/2015arXiv150504597R>.
25. Chen, L., Liang, X., Shen, C., Jiang, S. & Wang, J. Synthetic CT generation from CBCT images via deep learning. *Med. Phys.* **47**, 1115–1125. <https://doi.org/10.1002/mp.13978> (2020).
26. Kida, S. *et al.* *Cureus* **10**, e2548 (2018).
27. 27Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D. & Matas, J. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 8183–8192 (Salt Lake City, UT, 2018).

28. Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. <http://arxiv.org/abs/1703.10593> (arXiv e-prints) (2017). <https://ui.adsabs.harvard.edu/abs/2017arXiv170310593Z>.
29. Kida, S. *et al.* Visual enhancement of Cone-beam CT by use of CycleGAN. *Med. Phys.* **47**, 998–1010 (2020).
30. Silva, A. C., Lawder, H. J., Hara, A., Kujak, J. & Pavlicek, W. Innovations in CT dose reduction strategy: Application of the adaptive statistical iterative reconstruction algorithm. *Am. J. Roentgenol.* **194**, 191–199 (2010).
31. Szegedy, C. *et al.* Going Deeper with Convolutions. <http://arxiv.org/abs/1409.4842> (arXiv e-prints) (2014). <https://ui.adsabs.harvard.edu/abs/2014arXiv1409.4842S>.
32. Ramachandran, P., Zoph, B. & Le, Q. V. Searching for Activation Functions. <http://arxiv.org/abs/1710.05941> (arXiv e-prints) (2017). <https://ui.adsabs.harvard.edu/abs/2017arXiv171005941R>.
33. Ramachandran, P., Zoph, B. & Le, Q. Swish: A self-gated activation function. 2017. <http://arxiv.org/abs/1710.05941> (arXiv preprint).
34. Ledig, C. *et al.* in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4681–4690.
35. Sobel, I. & Feldman, G. A 3x3 isotropic gradient operator for image processing. *A talk at the Stanford Artificial Project in*, 271–272 (1968).
36. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. <http://arxiv.org/abs/1412.6980> (arXiv preprint) (2014).
37. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).

Acknowledgements

This project was supported by the NYMU-FEMH Joint Research Program (107DN13 and 109DN20). We thank FEMH for providing patient data and the Biomedical Imaging Physics and Instrument Lab at NYMU for providing GPU machine services.

Author contributions

H.J.T., H.C.Y. and J.C.C. designed the experiments. H.J.T. and P.W.S. were responsible for data acquisition and processing. H.J.T. and H.C.Y. performed data analyses and drafted the manuscript. All authors critically revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.-C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021