



Generating synthesized computed tomography from CBCT using a conditional generative adversarial network for head and neck cancer patients

Technology in Cancer Research & Treatment
Volume 21: 1–8
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/15330338221085358
journals.sagepub.com/home/tct


Yun Zhang, PhD¹ , Sheng-gou Ding, MD¹, Xiao-chang Gong, MD¹, Xing-xing Yuan, MD¹, Jia-fan Lin, MD¹, Qi Chen, MD², and Jin-gao Li, MD, PhD^{1,3,4}

Abstract

Purpose: To overcome the imaging artifacts and Hounsfield unit inaccuracy limitations of cone-beam computed tomography, a conditional generative adversarial network is proposed to synthesize high-quality computed tomography-like images from cone-beam computed tomography images. **Methods:** A total of 120 paired cone-beam computed tomography and computed tomography scans of patients with head and neck cancer who were treated during January 2019 and December 2020 retrospectively collected; the scans of 90 patients were assembled into training and validation datasets, and the scans of 30 patients were used in testing datasets. The proposed method integrates a U-Net backbone architecture with residual blocks into a conditional generative adversarial network framework to learn a mapping from cone-beam computed tomography images to pair planning computed tomography images. The mean absolute error, root-mean-square error, structural similarity index, and peak signal-to-noise ratio were used to assess the performance of this method compared with U-Net and CycleGAN. **Results:** The synthesized computed tomography images produced by the conditional generative adversarial network were visually similar to planning computed tomography images. The mean absolute error, root-mean-square error, structural similarity index, and peak signal-to-noise ratio calculated from test images generated by conditional generative adversarial network were all significantly different than CycleGAN and U-Net. The mean absolute error, root-mean-square error, structural similarity index, and peak signal-to-noise ratio values between the synthesized computed tomography and the reference computed tomography were 16.75 ± 11.07 Hounsfield unit, 58.15 ± 28.64 Hounsfield unit, 0.92 ± 0.04 , and 30.58 ± 3.86 dB in conditional generative adversarial network, 20.66 ± 12.15 Hounsfield unit, 66.53 ± 29.73 Hounsfield unit, 0.90 ± 0.05 , and 29.29 ± 3.49 dB in CycleGAN, and 16.82 ± 10.99 Hounsfield unit, 58.68 ± 28.34 Hounsfield unit, 0.92 ± 0.04 , and 30.48 ± 3.83 dB in U-Net, respectively. **Conclusions:** The synthesized computed tomography generated from the cone-beam computed tomography-based conditional generative adversarial network method has accurate computed tomography numbers while keeping the same anatomical structure as cone-beam computed tomography. It can be used effectively for quantitative applications in radiotherapy.

Keywords

conditional generative adversarial network, CBCT, synthesized CT image, deep learning

¹ Department of Radiation Oncology, Jiangxi Cancer Hospital of Nanchang University, Nanchang, Jiangxi, People's Republic of China

² MedMind Technology Co. Ltd, Beijing, People's Republic of China

³ Key Laboratory of Personalized Diagnosis and Treatment of Nasopharyngeal Carcinoma Nanchang, Jiangxi, People's Republic of China

⁴ Medical College of Nanchang University, Nanchang, Jiangxi, People's Republic of China

Corresponding author:

Jin-gao Li, Department of Radiation Oncology, Jiangxi Cancer Hospital of Nanchang University Nanchang, Jiangxi 330029, People's Republic of China.
Email: ndzhlyy0005@ncu.edu.cn



Abbreviations

HU, Hounsfield unit; CBCT, cone-beam computed tomography; cGAN, conditional generative adversarial network; sCT, synthesized CT; rCT, reference CT; CNN, convolutional neural network; MAE, mean absolute error; RMSE, root-mean-square error; SSIM, structural similarity index; PSNR, peak signal-to-noise ratio; DIR, deformation image registration.

Introduction

Cone-beam computed tomography (CBCT) is currently widely available in external beam image-guided radiotherapy,^{1,2} and it is routinely used for patient position verification and setup displacement correction.¹ However, it could not be directly used for quantitative applications such as dose calculation and adaptive treatment planning because the CT numbers of CBCT are not accurate due to cupping and scattering artifacts.³ Even though the CT numbers in CBCT can be restored by deforming the planning CT through deformable image registration,⁴ this method still requires a rescanned CT of the patient when there are anatomical differences between CBCT and the planning CT.⁵

Quickly restoring the correct CT number in CBCT directly for quantitative applications is possible now using deep learning methods.^{6–19} A method based on U-Net has been proved to be an effective way to restore the correct CT number and reduce noise in CBCT.⁷ The skip connection and U-shaped structure of U-Net realize the fusion of low- and high-resolution information and suppress global scattering and local noise in CBCT, so U-Net is very suitable for the process of synthesized CT (sCT) generation using paired CBCT-CT. However, U-Net is shallow and focuses only on CT numbers rather than inclusive image information, which leads to image distortion in small areas and unclearness in soft-tissue boundaries. For instance, Minnema et al. found that U-Net incorrectly labeled background voxels as the bone around the vertebrae in CBCT scans containing vertebrae.¹¹ Another recently studied method

is CycleGAN; it can resolve image distortion and improve soft tissue resolution.^{14–20} Additionally, it can be trained with unpaired CBCT-CT data, which is a great advantage, especially when paired data are hard to prepare. However, the cycle-consistent requirement of CycleGAN means that it keeps some strong characteristics of CBCT, such as metal or ring artifacts.¹⁵ Even the CycleGAN model may not perform well on 512×512 -resolution images if the generator and discriminator are not in balance.¹⁵ In addition, it is also very time-consuming and computationally intensive to train a CycleGAN model.¹⁵

Recently, Maspero et al. generated sCT images from paired MRI-CT data with conditional generative adversarial networks (cGANs).²¹ The results of the study confirmed that the cGAN can explore both global image information and local information because it trains the network in a supervised way by adding conditions based on the GAN,^{22,23} which gives effective supervision feedback to the model and guides the model in generating sCT images. The cGAN incorporates additional information between the generator and discriminator to constrain the generator to generate the desired image. Similarly, CycleGAN incorporates the cycle network into a GAN as additional information, leading to an advanced cGAN. Until now, no research on CT generation from CBCT based on the cGAN method has been reported, and there is little literature comparing the performance of cGAN and CycleGAN for the task of CBCT-to-synthetic-CT transformation. In this study, we applied a cGAN to synthesize CT from the CBCT scans of head and neck cancer patients and quantitatively compared it with other models (U-Net and CycleGAN) by analyzing the synthetic CT images.

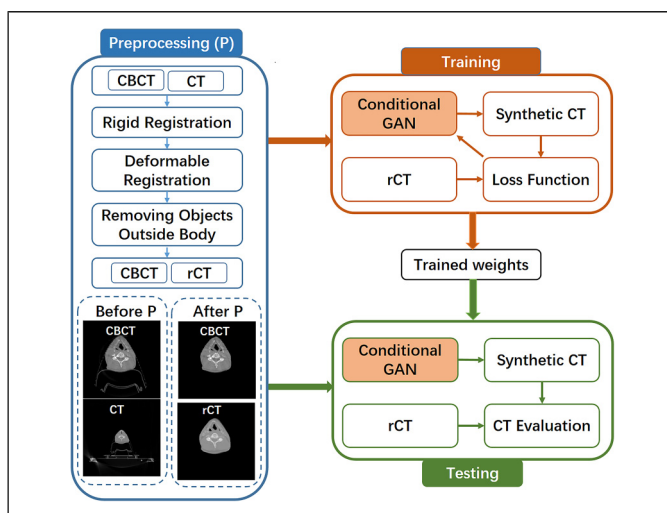


Figure 1. Schematic of the image workflow for applying the trained generator on a new 2D transverse slice of the cone-beam computed tomography (CBCT) of a head and neck cancer to create an sCT.

Materials and Methods

Because head and neck cancer patients can be considered as a quasi-rigid structure with better positional repeatability compared to other cancer patients. Therefore, to reduce the effect of registration, one hundred and twenty patients with head and neck cancer who had received radiotherapy between January 2019 and December 2020 at our Hospital were retrospectively enrolled in this study. This study was approved by the medical ethics committee of our hospital. The requirement for informed consent was waived by the ethics committee because of the study's retrospective nature, and we have de-identified all patient details. Planning CT and CBCT images were acquired from a CT (SOMATOM Definition AS, Siemens Medical Systems) and a TrueBeam linear accelerator (Varian Medical Systems, OBI), respectively. Patients' data were assembled into training (80), validation (10), and testing (30) datasets. The dimensions of both the CT and

CBCT images were 512×512 on the axial plane, while the spatial resolution was $1.27 \text{ mm} \times 1.27 \text{ mm} \times 3.00 \text{ mm}$ for CT scans, and $0.51 \text{ mm} \times 0.51 \text{ mm} \times 1.99 \text{ mm}$ for CBCT scans. Preprocessing was performed with software that was developed in-house for all the scans to generate reference CTs (rCTs) from planning CTs. First, rigid registration was carried out to align the CT images to the CBCT images. The CT images were upsampled to match the resolution of the CBCT images during this process. Then, deformable registration (the combination of free-form registration and the viscous fluid registration method) from CT to CBCT was performed, and the results of the deformation image registration (DIR) were reviewed by an advanced radiation oncologist to ensure their quality. The deformed CT image was further cropped by removing the information outside the body region. Afterward, a reference CT (rCT) was used as a benchmark for the synthetic CTs (sCTs) during training and testing. The image workflow is illustrated in Figure 1.

The cGAN architecture is shown in Figure 2. As a 2-dimensional (2D) model, it generates sCT in a slice-by-slice fashion, and the model learns a mapping from a CBCT image to a CT image with a generator and a discriminator. The generator G is trained to produce sCT images that cannot be distinguished from CT images by an adversarial trained discriminator D , which is trained to detect the sCT images generated by the generator. Both the generator and discriminator of the cGAN observe the input CBCT image. In this study, we implement a full convolutional neural network (CNN) architecture²⁴ as the generator. In this CNN architecture, a U-Net backbone architecture is utilized; it consists of an encoding path and a decoding path. The encoding path and decoding path are combined by using a skipping connection to concatenate multilevel features and take advantage of both low-level and high-level information. To further improve the performance of the U-Net generator, the building blocks of the original U-Net are replaced with residual blocks to achieve consistent training as the depth of the network increases. The detailed network architecture is shown in Figure 2. A U-Net backbone architecture with residual blocks is added into the generator of the cGAN framework to obtain the combination of high- and low-level features. A convolutional PatchGAN classifier²⁵ is invoked as the discriminator; it only penalizes structure at the scale of image patches. This discriminator tries to determine whether each patch in an image is real or fake. The convolution responses of all the patches across the image are averaged to provide the ultimate output of the discriminator.

The loss function is an association between the cGAN loss and the L1 loss. The cGAN loss can be expressed as

$$L_{cGAN} = E[\log D(\text{CBCT}, CT)] + E[\log(1 - D(\text{CBCT}, G(\text{CBCT})))] \quad (1)$$

The discriminator, D , seeks to maximize this cGAN loss while the generator G tries to minimize this term. E means to sum over all the samples. The generator's job is to generate sCT images that are close to the acute CT images. Fooling the discriminator

is just one way to learn to complete the task of the generator. Another approach is adding the L1 distance to the loss function. We used the L1 distance rather than L2 as L1 encourages less blurring. The L1 loss is shown below:

$$L_{L1} = E[||CT - G(\text{CBCT})||_1]. \quad (2)$$

Then, the final objective is finding an optimal generator of the loss function

$$L_{\text{final}} = L_{cGAN} + \lambda L_{L1}, \quad (3)$$

where λ balances the contributions of the 2 terms.

Three different networks (cGAN, CycleGAN, and U-Net) were trained with the same training and validation datasets in the same environment. A mini-batch size of 6 was used on a Tesla P100 (NVIDIA) graphical processing unit (GPU). All weights were initialized from a random normal initialization with a mean of 0 and a standard deviation of 0.02. The learning rate was set to 0.0001, and the Adam optimizer was used. The loss function of U-Net was a combination of the L1 loss and the structural similarity index (SSIM). Compared with cGAN, the loss function of CycleGAN had an additional cycle-consistent loss term. The cGAN, CycleGAN, and U-Net were trained for 200, 200, and 100 epochs, respectively, and the models with the smallest validation loss were selected as the optimal models.

The sCT evaluation strategy was used to evaluate the performance of different synthetic image generators using the pixel-wise HU accuracy, noise level, and structural similarity. The evaluation strategy takes advantage of 4 criteria, including the mean absolute error (MAE), root-mean-square error (RMSE), SSIM, and peak signal-to-noise ratio (PSNR).

The MAE is the magnitude of the mean absolute difference of all the pixel values between 2 images. The MAE formula is illustrated in Equation 4. A smaller difference is indicated by a lower value.

$$\text{MAE} = \frac{1}{n_x n_y} \sum_{i,j} |sCT(i,j) - CT(i,j)|, \quad (4)$$

where $sCT(i,j)$ is the value of the pixel at (i,j) in the sCT or CBCT, $CT(i,j)$ is the value of the pixel at (i,j) in the reference CT, and n_x, n_y represent the total number of pixels in one slice.

The RMSE is frequently used for measuring differences between 2 images by computing the RMSE of all the pixel values, as shown in Eq. 5. Lowering the RMSE improves the image quality.

$$\text{RMSE} = \sqrt{\frac{1}{n_x n_y} \sum_{i,j} (sCT(i,j) - CT(i,j))^2}, \quad (5)$$

where $sCT(i,j)$ is the value of the pixel at (i,j) in the sCT or CBCT, $CT(i,j)$ is the value of the pixel at (i,j) in the reference CT, and n_x, n_y represent the total number of pixels in one slice. The SSIM is a perceptual metric that quantifies the image quality using that of the reference image. The formula of the SSIM is shown in Equation 6. Structural information is

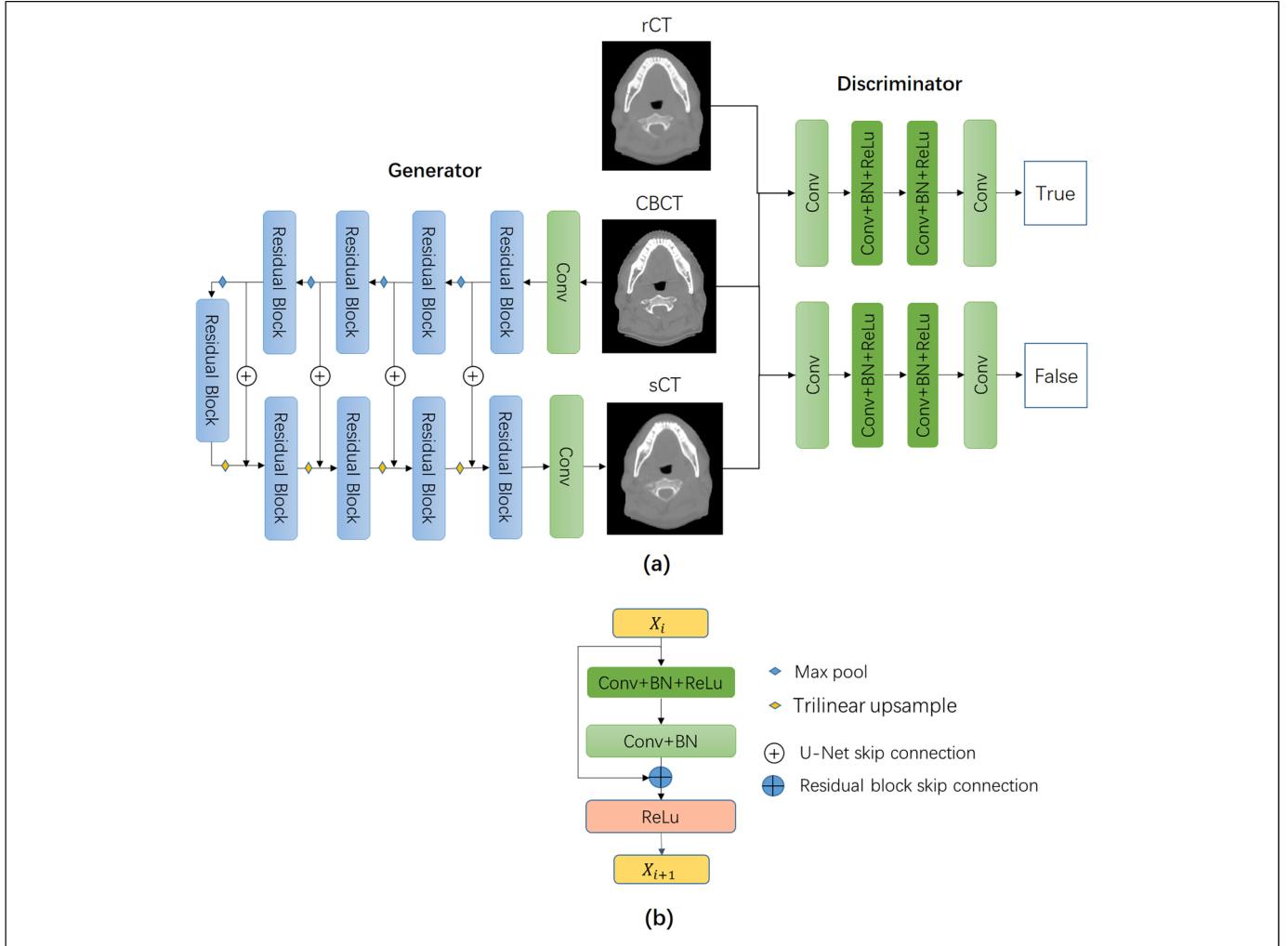


Figure 2. Proposed conditional GAN architecture used to map a cone-beam computed tomography (CBCT) image to a CT image. The U-Net backbone with residual blocks is integrated into the generator. (A) The overall architecture diagram. (B) The detailed architecture of a residual block.

essentially considered. The SSIM value ranges from 0 to 1, and increasing the SSIM improves the image quality.

$$\text{SSIM} = \frac{(2\mu_{sCT}\mu_{CT} + c_1)(2\sigma_{sCTCT} + c_2)}{(\mu_{sCT}^2 + \mu_{CT}^2 + c_1)(\sigma_{sCT}^2 + \sigma_{CT}^2 + c_2)}, \quad (6)$$

where μ_{sCT} is the mean of the pixel values of the sCT or CBCT image, μ_{CT} is the mean of the pixel values of the reference CT image, σ_{sCT} is the standard deviation of sCT or CBCT, σ_{CT} is the standard deviation of the CT image, σ_{sCTCT} is the covariance of 2 images sCT and CT, and c_1 and c_2 are 2 variables that stabilize division with a weak denominator.

The PSNR is frequently used to evaluate image quality, especially for noise reduction. It is a combination of the mean squared error and the maximum intensity value of the reference image. The formula of the PSNR is shown in Eq. 7. Increasing

the PSNR improves the image quality.

$$\text{PSNR} = 10 \times \log_{10} \left(\frac{\text{MAX}^2}{\frac{1}{n_x n_y} \sum_{i,j} (sCT(i,j) - CT(i,j))^2} \right), \quad (7)$$

where $sCT(i,j)$ is the value of the pixel at (i,j) in the sCT or CBCT, $CT(i,j)$ is the value of the pixel at (i,j) in the reference CT, n_x, n_y represent the total number of pixels in one slice, and MAX is the maximum intensity in the sCT or CBCT.

Statistical Analysis

Results are presented for an independent testing set consisting of the data of 30 head and neck cancer patients with 2790 slices. After removing the all-black slices, a total of 2670 slices were included in the testing set. A paired t-test was

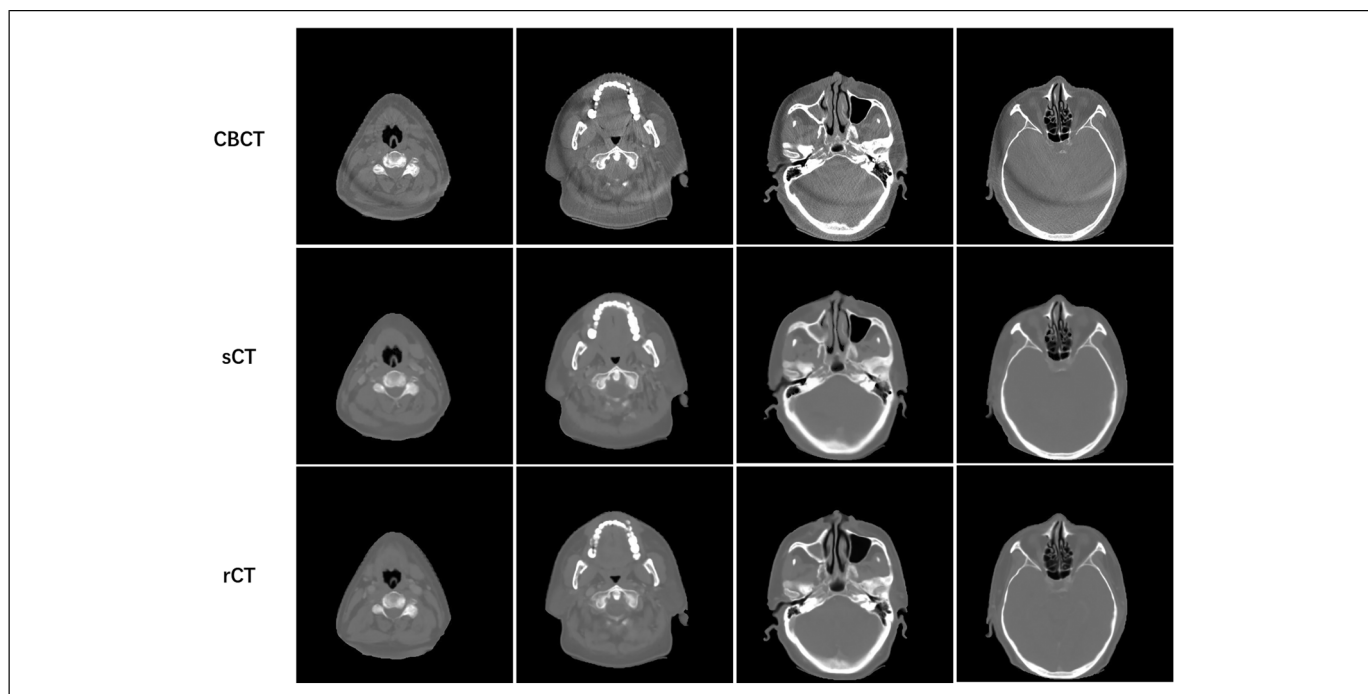


Figure 3. Sample results of the predicted sCT using the proposed method compared to the rCT with the corresponding cone-beam computed tomography (CBCT) images. These images are placed at different anatomical locations of the scans from 4 patients. ($W = 1000$ and $L = 100$ for all images).

used for statistical analyses of different methods, and a P value $\leq .05$ was considered statistically significant.

Results

Figure 3 displays the CBCT images, reference CT images, and sCT images generated by the cGAN for the visual evaluation of images selected randomly from the test dataset. The ring artifacts, as well as the noise, found in the CBCT images were greatly reduced in all the sCT images, while the anatomy is kept the same as in the CBCT images. In contrast, the sCT images have the same brightness distribution as the CT images, and the noise is negligible.

The sCT images generated from 3 different models (cGAN, CycleGAN, and U-Net) are shown in Figure 4. Maps in pseudo-color represent the differences between the ground truth rCT and each sCT image. Ring artifacts are reduced on the sCT image generated by the CycleGAN, but some residual ring artifacts were still present in these regions; they are marked by the white and black arrows in Figure 4D and H. Fewer ring artifacts are found in the sCT image from U-Net, and almost no artifacts are found in the sCT image from the cGAN, but the soft tissue around the nasal cavity became more blurred in sCT images from U-Net compared with the cGAN, as shown by the small red arrows in Figure 4C and E.

To further analyze the details of the sCT images generated from diverse models, the local image details of one test sample are shown in Figure 5. The sCT image generated by U-Net is blurry at the soft tissue boundary, and the edges of

bone are not as sharp as the edges in the CT image. Conversely, the image quality of the sCT images generated by the cGAN or CycleGAN had no visual degradation. Soft tissues, such as the thyroid gland, can be identified. HU profile lines through the trachea and cervical spine are provided in Figure 5B and C. It can be seen that the cGAN and U-Net on line 1 are closer to the results of rCT, while the U-Net on line 2 differs the most from the results of rCT.

The MAE, RMSE, SSIM, and PSNR were used to quantitatively evaluate the HU accuracy of all testing sCT images and CBCT images compared with CT images (within the patient's body); the results are presented in Table 1. Compared with CycleGAN, the MAE and RMSE of the cGAN decreased by 3.91 HU and 8.38 HU, respectively. The SSIM and PSNR of the cGAN increased by 0.02 and 1.35, respectively. The differences between the U-Net method and the cGAN method are relatively small for those parameters, but statistical analysis shows that the cGAN is significantly better than U-Net.

We also measured the time cost for predicting one sCT from CBCT. The predicting time was 0.62 s, 0.89 s, 0.21 s per CBCT slice for cGAN, CycleGAN, U-Net, respectively.

Discussion

In this study, a novel deep learning network based on the cGAN that integrates a U-Net backbone architecture with residual blocks into a GAN framework is proposed to synthesize high-quality CT-like images. A total of 120 NPC patients were used to train, validate, and test this network. The resulting

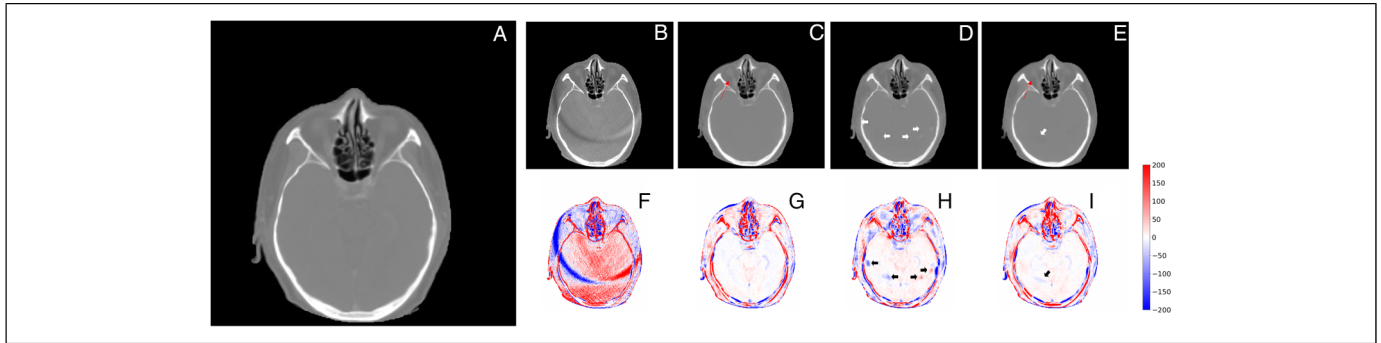


Figure 4. An example of difference maps between rCT and other images. (A) rCT image as the reference; (B) cone-beam computed tomography (CBCT) image; (C) sCT image by conditional generative adversarial network (cGAN); (D) sCT image by CycleGAN; (E) sCT image by U-Net; (F-I) HU difference pseudo-color maps between rCT and the corresponding sCT images above.

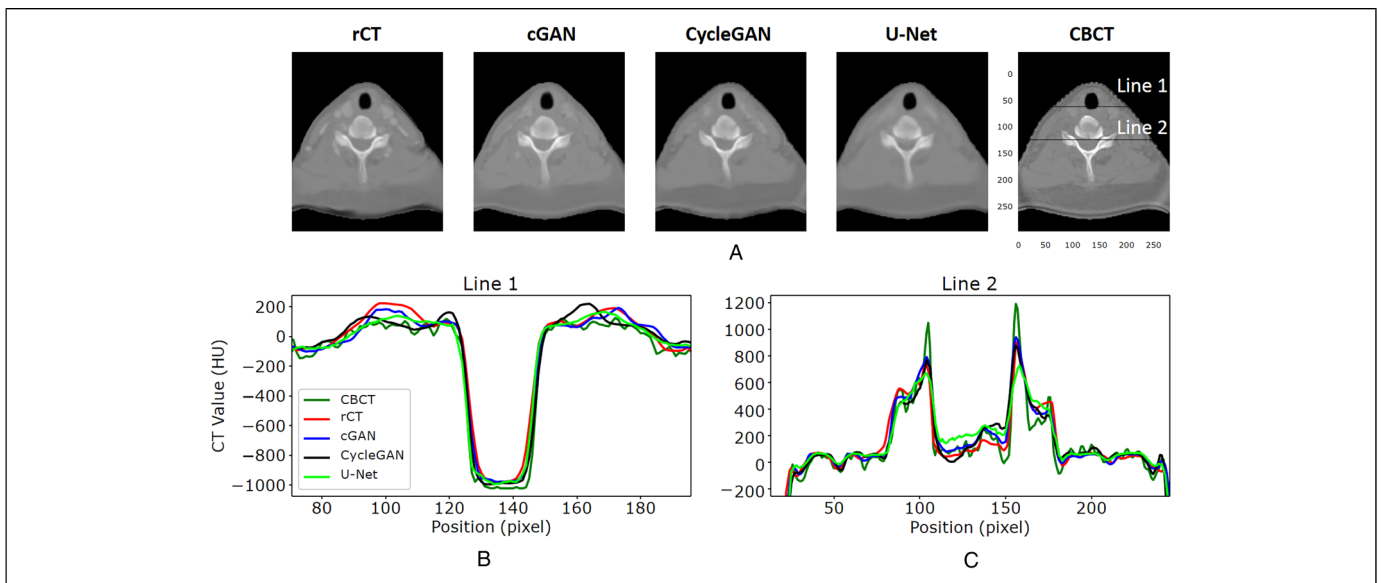


Figure 5. An example of the HU line profiles of rCT, sCT, and cone-beam computed tomography (CBCT), and sCT generated by conditional generative adversarial network (cGAN), CycleGAN, and U-Net. (A) Traversal of rCT, 3 sCTs, and CBCT. (B) HU profile of line 1. (C) HU profile of line 2.

sCT images are visually similar to planning CT images; thus, this method can be directly used for quantitative applications such as dose calculation and adaptive treatment planning.

In terms of image quality, the sCT image generated by the cGAN achieves a better image quality for soft tissues and contains fewer artifacts due to the design of this network, which retains the advantages of a GAN and incorporates the features of U-Net. The deformable registration between paired CBCT images and planning CT images is difficult to make accurate for the soft tissue boundary since the CBCT image is not clear in this region. As the U-Net method is meant to generate sCT images that are close to the registered CT images, the synthesized image is distorted in small areas and soft tissue boundaries may be unclear. Fortunately, the cGAN and CycleGAN can construct a nonlinear mapping relationship between 2 image domains. Therefore, the soft tissue of the rCT image can be more clearly reflected in the CT images synthesized by the GAN than those

from U-Net. However, the ring artifacts, as well as the noise, are obvious in the CBCT image, and they are, to some extent, inevitably retained in images synthesized by CycleGAN. This is because this method requires the synthesized images to keep strong characteristics of CBCT images to restore CBCT images from synthesized images as well as possible. Liang et al. also illustrated the problem that CycleGAN cannot suppress metal and dental artifacts, although the image quality can be improved.¹⁵ By integrating the U-Net architecture into the generator of the cGAN, the discriminator acts as an automatically learned loss function to guide the generator to exploit both low-level and high-level features to reduce artifacts in CBCT images. Furthermore, the discriminator was also designed to take the CBCT image as input, thus ensuring that the synthesized CT image would have the same anatomy structures as the original CBCT image. Consequently, through the cGAN, anatomy structures are kept and almost all of the artifacts in CBCT are removed.

Table 1. Quantitative Comparison Result of 4 Different Images.^a

	MAE(HU)	RMSE(HU)	SSIM	PSNR(dB)
CBCT	36.23 ± 20.24	104.60 ± 41.05	0.83 ± 0.08	25.34 ± 3.19
cGAN	16.75 ± 11.07	58.15 ± 28.64	0.92 ± 0.04	30.58 ± 3.86
CycleGAN	20.66 ± 12.15	66.53 ± 29.73	0.90 ± 0.05	29.29 ± 3.49
U-Net	16.82 ± 10.99	58.68 ± 28.34	0.92 ± 0.04	30.48 ± 3.83

Abbreviations: CBCT, cone-beam computed tomography; cGAN, conditional generative adversarial network; MAE, mean absolute error; RMSE, root-mean-square error; SSIM, structural similarity index; PSNR, peak signal-to-noise ratio.

^aP(CBCT vs cGAN), P(CycleGAN vs cGAN), and P(U-Net vs cGAN) in MAE, RMSE, SSIM, and PSNR < 0.001.

Another main disadvantage of CBCT is that HU tends to be less accurate than the reference CT images. Three methods have the ability to rectify the CT values of CBCT: among them, the cGAN gives the best results, U-Net gives the second-best results, and CycleGAN gives the worst results (see Table 1). The MAE and RSME are used to evaluate the ability to rectify CT values. Compared with previous studies, the MAE of synthesized CT images based on CycleGAN and U-Net in this study is lower. For example, the results of Liang showed that the MAE decreases from 69.29 HU down to 29.85 HU¹⁵ for the CycleGAN method. Maspero found that the results based on CycleGAN can be decreased from 195 HU to 51 HU,¹⁶ and the MAE in the pelvic region for pseudo-CT images based on an improved 3D CycleGAN is above 50 HU.¹⁹ Chen illustrated that the MAE of synthesized CT images based on U-Net was 18.98 HU,⁸ and Li explained that the range of the MAE for U-Net improved to (6, 27) HU.¹³ The CycleGAN requires the fake image to keep all the information in the original image; as a result, its CT correction ability was reduced to some degree. The U-Net methods use paired data for training, which allows them to obtain more information about the corresponding CT, including HU values. The cGAN combines the benefits of U-Net and GAN, so it can further improve the CT accuracy.

The results of the quantitative analysis of image quality show that the SSIM values of the cGAN and U-Net are higher than those of CycleGAN, as compared with 0.8911 from a study by Chen,⁸ 0.85 from a study by Liang,¹⁵ and 0.81 from a study by Sun.¹⁹ The opposite result was found in Liu's study,²⁰ where SSIM values based on U-Net were much lower than those of the CycleGAN method (0.56 vs 0.70). Some researchers have put forward this use of 3D CycleGAN as a way to improve the image quality of the pseudo-CT. The results of Sun showed that the SSIM increases from 0.81 to 0.86 using the 3D CycleGAN method for cervical cancer patients.¹⁹ Liu found that the SSIM of synthesized CT images improved to 0.71 from using the 3D CycleGAN rather than the 2D CycleGAN.²⁰

Ideally, a reference replanning CT image with an anatomical structure that is identical to CBCT is required to train the model. However, such an ideal CBCT-CT pair is not readily available,

as they are not acquired simultaneously. Thus, the anatomical structure is not identical between the CT and CBCT: a certain amount of deformation was observed. 3D deformable registration between CBCT and CT was then performed for optimal performance. The synthesizing performance to some extent depends on the performance of the deformable registration. The displacement caused by the deformable registered deviation directly affects parameters such as the MAE. An evaluation of the registration accuracy and how the accuracy influences the synthesis performance should be performed in a future study.

The present study had several limitations. First, it is also unknown whether cGAN will improve the accuracy of dose calculation since doses calculated based on sCT were not analyzed in the current work. The difference between the HU values of the images obtained by both cGAN and u-NET networks in this study is very small, which most likely has almost no effect on the dose calculation, while the time to generate pseudo-CT images from CBCT is about 3 times longer than that of U-NET (0.62 s vs 0.21 s per slice). Using the synthetic CT of CBCT for quantitative applications such as dose calculation and adaptive treatment planning will be the focus of our future studies. Second, we only concentrated on patients with head and neck cancer in this work, which contains less anatomical variations caused by organ motion or filling than the tumor in thorax, abdomen, or pelvis sites. The applicability of our proposed model to other disease sites will be considered in future studies. In addition, since this is a retrospective analysis, more prospective large randomized clinical trials and multicenter clinical samples are needed to further validate the reliability of our study findings.

Conclusions

In conclusion, a novel cGAN was used to generate sCT from paired CBCT and CT scans. Experiments demonstrated that the sCT images had sharp clarity and were comparable to CT images. In general, the cGAN can generate sCT images with accurate HU values and anatomy structures. It is a promising method for facilitating adaptive radiotherapy treatments.

Authors' Note

The data used and analyzed during the current study are available from the corresponding author. This study was approved by the Medical Ethics Committee of Jiangxi Cancer Hospital, and the approval number is 2022ky012.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by grants from National Cancer Center Climbing Fund of China (No.NCC201814B040 for Xiao-chang Gong).

ORCID iD

Yun Zhang  <https://orcid.org/0000-0001-5075-727X>

References

1. Boda-Heggemann J, Lohr F, Wenz F, et al. KV cone-beam CT-based IGRT: a clinical review. *Strahlenther Onkol.* 2011;187(5):284-291.
2. Sorcini B, Tilikidis A. Clinical application of image-guided radiotherapy, IGRT (on the Varian OBI platform). *Cancer Radiother.* 2006;10(5):252-257.
3. Schulze R, Heil U, Gross D, et al. Artefacts in CBCT: a review. *Dentomaxillofac Radiol.* 2011;40(5):265-273.
4. Wang H, Dong L, Lii MF, et al. Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy. *Int J Radiat Oncol Biol Phys.* 2005;61(3):725-735.
5. Ramella S, Fiore M, Silipigni S, et al. Local control and toxicity of adaptive radiotherapy using weekly CT imaging: results from the LARTIA trial in stage III NSCLC. *J Thorac Oncol.* 2017;12(7):1122-1130.
6. Nomura Y, Xu Q, Shirato H, et al. Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network. *Med Phys.* 2019;46(7):3142-3155.
7. Ronneberger O, Fischer P, Brox T. *U-Net: Convolutional Networks for Biomedical Image Segmentation.* Springer; 2015.
8. Chen L, Liang X, Shen C, et al. Synthetic CT generation from CBCT images via deep learning. *Med Phys.* 2020;47(3):1115-1125.
9. Hansen DC, Landry G, Kamp F, et al. Scatternet: a convolutional neural network for cone-beam CT intensity correction. *Med Phys.* 2018;45(11):4916-4926.
10. Lalonde A, Winey B, Verburg J, et al. Evaluation of CBCT scatter correction using deep convolutional neural networks for head and neck adaptive proton therapy. *Phys Med Biol.* 2020;65(24):245022.
11. Minnema J, van Eijnatten M, Hendriksen AA, et al. Segmentation of dental cone-beam CT scans affected by metal artifacts using a mixed-scale dense convolutional neural network. *Med Phys.* 2019;46(11):5027-5035.
12. Yuan N, Dyer B, Rao S, et al. Convolutional neural network enhancement of fast-scan low-dose cone-beam CT images for head and neck radiotherapy. *Phys Med Biol.* 2020;65(3):035003.
13. Li Y, Zhu J, Liu Z, et al. A preliminary study of using a deep convolution neural network to generate synthesized CT images based on CBCT for adaptive radiotherapy of nasopharyngeal carcinoma. *Phys Med Biol.* 2019;64(14):145010.
14. Jin KH, McCann MT, Froustey E, Unser M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Process.* 2017;26(9):4509-4522.
15. Liang X, Chen L, Nguyen D, et al. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys Med Biol.* 2019;64(12):125002.
16. Maspero M, Houweling AC, Savenije MHF, et al. A single neural network for cone-beam computed tomography-based radiotherapy of head-and-neck, lung and breast cancer. *Phys Imaging Radiat Oncol.* 2020;14:24-31.
17. Harms J, Lei Y, Wang T, et al. Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. *Med Phys.* 2019;46(9):3998-4009.
18. Kurz C, Maspero M, Savenije MHF, et al. CBCT Correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation. *Phys Med Biol.* 2019;64(22):225004.
19. Sun H, Fan R, Li C, et al. Imaging study of pseudo-CT synthesized from cone-beam CT based on 3D CycleGAN in radiotherapy. *Front Oncol.* 2021;11:603844.
20. Liu Y, Lei Y, Wang T, et al. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. *Med Phys.* 2020;47(6):2472-2483.
21. Maspero M, Savenije MHF, Dinkla AM, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Phys Med Biol.* 2018;63(18):185001.
22. Li Z, Wu M, Zheng J, et al. Perceptual adversarial networks with a feature pyramid for image translation. *IEEE Comput Graph Appl.* 2019;39(4):68-77.
23. Zhang Y, Yue N, Su MY, et al. Improving CBCT quality to CT level using deep learning with generative adversarial network. *Med Phys.* 2021;48(6):2816-2826.
24. Liu Z, Liu F, Chen W, et al. Automatic segmentation of clinical target volumes for post-modified radical mastectomy radiotherapy using convolutional neural networks. *Front Oncol.* 2021;10:581347.
25. Isola P, Zhu J, Zhou T, et al. Image-to-Image translation with conditional adversarial networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2017:5967-5976.