2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Improvement of Image Quality of Cone-beam CT Images by Three-dimensional Generative Adversarial Network

Takumi Hase[1], *Student Member, IEEE*, Megumi Nakao[1], *Member, IEEE*, Keiho Imanishi[2], Mitsuhiro Nakamura[3], and Tetsuya Matsuda[1], *Member, IEEE*

***Abstract*— Artifacts and defects in Cone-beam Computed Tomography (CBCT) images are a problem in radiotherapy and surgical procedures. Unsupervised learning-based image translation techniques have been studied to improve the image quality of head and neck CBCT images, but there have been few studies on improving the image quality of abdominal CBCT images, which are strongly affected by organ deformation due to posture and breathing. In this study, we propose a method for improving the image quality of abdominal CBCT images by translating the numerical values to the values of corresponding paired CT images using an unsupervised CycleGAN framework. This method preserves anatomical structure through adversarial learning that translates voxel values according to corresponding regions between CBCT and CT images of the same case. The image translation model was trained on 68 CT-CBCT datasets and then applied to 8 test datasets, and the effectiveness of the proposed method for improving the image quality of CBCT images was confirmed.**

I. INTRODUCTION

Cone beam computed tomography (CBCT) is widely used to capture images of a patient's anatomy in a variety of clinical situations. Its compact device size allows 3D imaging during treatment or surgery [1] with lower radiation dose than CT. However, the physical limitation of the CBCT device restricts imaging area, which can cause beam hardening, artifacts and defects in the reconstructed CBCT image. Improving the image quality of CBCT images is a clinically important issue for accurate radiation therapy and surgery. [2][3].

In recent years, various deep learning methods have been explored to improve the quality of medical images and to reduce artifacts[2]. In the clinical field, it is difficult to acquire paired CT-CBCT images from the same patient with perfectly matched anatomical structures, so image translation using unsupervised learning is of more interest than supervised learning such as scattering correction[4][5].

Generative Adversarial Network (GAN)[6] and its extended method, CycleGAN[7], are relatively new unsupervised learning frameworks that have been widely studied. In particular, CycleGAN has been proposed for image translation and is very effective in the field of medical imaging where paired images are difficult to create[2]. Nakao et al. studied the reduction of dental metal artifacts using a 3D generative adversarial network[8][9]. Liang et al. studied the improvement of image quality of CBCT images in the head and neck region[10]. They used a loss function based on CycleGAN for image translation to achieve pixel values close to those of CT images while maintaining the anatomical structure of CBCT images of the head and neck region. To improve the image quality of abdominal CBCT images, Kida et al. used a method based on CycleGAN with a loss function to perform image translation while maintaining the anatomical structure of abdominal CBCT images[11]. However, due to the variety of anatomical structures in abdominal CBCT and CT images, the generators may fail to learn a meaningful mapping and produce totally distorted images. In this study, to improve the image quality of abdominal CBCT images acquired during radiotherapy for prostate cancer, we applied the CycleGAN framework to construct an generative adversarial network that performs translation between a set of CBCT images and a group of CT images. To solve the problem of translating images while preserving the anatomical structure, we propose to use rigid registration to generate training data for unsupervised learning. Within the framework of image translation based on unsupervised learning, we compared our method with conventional methods. Then, we evaluated the performance of the translation from two perspectives: CT value translation and anatomical structure preservation.

II. METHODS

To improve the image quality of CBCT images based on unsupervised learning, we use a pair of CBCT images and CT images acquired from the same patient. We define the set of CBCT images as $X = \{x_i\}(i = 1, 2, ..., N)$ and the set of CT images as $Y = \{y_j\}(j = 1, 2, ..., M)$, we apply the basic framework of CycleGAN[7] to learn the function of image translation from the image sets X to the image sets Y. The following two points are the target to be achieved in the proposed image translation.

1) Reduces artifacts in the CBCT image and translates it to the correct CT value of the tissue to be imaged
2) Preserve the anatomy of the CBCT image during the translation

A. Dataset and preprocessing

In this study, we used CBCT images (512×512 pixels, 48-93 slices) and CT images (512×512 pixels, 134-226 slices)

[1]T. Hase, M. Nakao, and T. Matsuda are with Graduate School of Informatics, Kyoto University, Kyoto, 606-8501, JAPAN. (email: thase@sys.i.kyoto-u.ac.jp, megumi@i.kyoto-u.ac.jp, tetsu@i.kyoto-u.ac.jp)
[2]K. Imanishi is with e-Growth Co., Ltd., Kyoto 604-8006, JAPAN.
[3]M. Nakamura is with Graduate School of Medicine, Kyoto University, Kyoto, 606-8501, JAPAN.

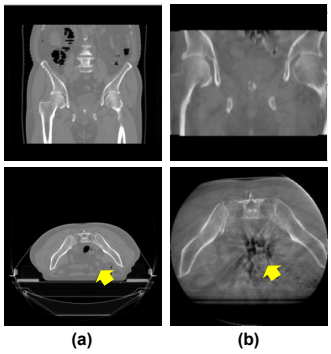

Fig. 1. Visual comparison of CT and CBCT images; CBCT images contain artifacts and have a smaller imaging range than CT images. (a) CT image taken for treatment planning, (b) CBCT image taken on the day of treatment.

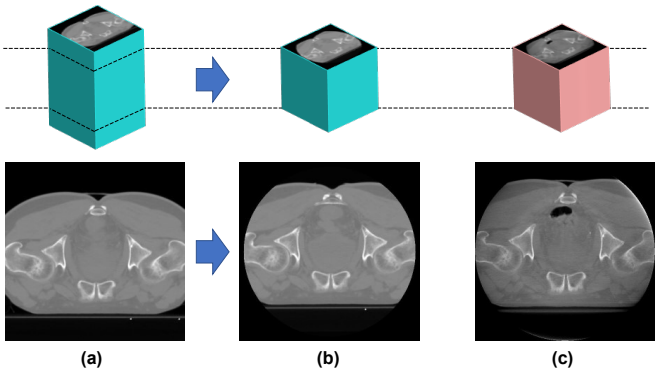

Fig. 2. Rigid registration as a pre-processing of the dataset to align the imaging ranges of CT and CBCT images. (a) CT image (source), (b) CT image after registration, (c) CBCT image (target).

of 76 prostate cancer patients who underwent radiotherapy at Kyoto University Hospital are used as image group X and image group Y. Figure 1 (a) shows the CT image obtained before radiotherapy planning, and (b) shows the CBCT image during radiotherapy of the day. The yellow arrows show the difference of appearance between CBCT and CT images of a same patient, and only the CBCT images contain artifacts and defects. In addition, the CBCT has a narrower field of view than the CT image, and therefore it can visualize a circular area of 25.5 cm in diameter centered on the prostate in a slice. CBCT images consist of 48-93 slices, and around ± 7 cm are visualized in an axial direction. On the other hand, CT images have a wider imaging field of view and imaging area than CBCT images.

Conventional approaches based on the framework of unsupervised learning may not be able to learn appropriate translations for each region sufficiently due to the variety of organ shapes and bone structures in abdominal CBCT and CT images[11][12]. To solve this problem, in this study, we argue that it is effective to select and input slice images whose imaging regions are corresponded between CBCT and CT images. The datasets in this study are CBCT volume images and CT volume images of the each patient, and we can find spatially corresponded regions between CBCT image and CT image. For learning, it is enough to prepare slice images with corresponding imaging regions, and there is no need to involve deformation. Therefore as a preprocessing step, by rigid registration using the CT image as the source and the CBCT image as the target of the same patient, two images were three-dimensionall corresponded. Figure 2 shows an example of the registration result. By rigid registration, the imaging range and the size per voxel can be matched between the CBCT image and the CT image of the same patient. Of the 76 datasets created by this preprocessing, 8 datasets randomly selected were used as test data, and the 68 datasets were used as training data.

B. CycleGAN

In a CycleGAN, a network is constructed using two GANs. From two image groups X and Y, slice image x and y are used as input data, and the image generator G and discriminator D pairs (G_Y, D_Y) and (G_X, D_X) are trained adversarially. The basic bjective of CycleGAN can be described as

$$\mathcal{L}_{cgan} = \mathcal{L}_{adv}(G_Y, D_Y, X, Y) + \mathcal{L}_{adv}(G_X, D_X, X, Y) + \lambda_{cyc}\mathcal{L}_{cyc}(G_Y, G_X) \quad (1)$$

Here, $\mathcal{L}_{adv}$ refers to adversarial loss, which is the original loss function of GAN. G_Y tries to generate image $G_Y(x)$ that is similar to image x in X from image y in Y, whereas D_Y tries to discriminate whether the image is the real image y or generated image $G_Y(x)$. The same is true for G_X and D_X. Adversarial loss measures the performance of D. $\mathcal{L}_{cyc}$ refers to cycle consistency loss, where weight λ_{cyc} controls the relative strength of the adversarial loss and cycle consistency loss. By adding the cycle consistency loss to the loss function, the model is trained to generate a reconstructed image that corresponds to the real image. Therefore, we can request the model to generate an image that retains the real image features.

C. Loss function with regularization

In order to achieve the preservation of anatomical structure during the translation, which is one of the tasks of this study, we added Intensity Loss[9], which is defined by the following equation (2).

$$\mathcal{L}_{int} = \mathbb{E}_x||G_Y(x) - x||_1 + \mathbb{E}_y||G_X(y) - y||_1 \quad (2)$$

The first term in Equation (2) is a loss function that acts as a regularization term that penalizes the error before and after passing through the generator G_Y, that is, the difference between the original CBCT image x and the translated CT image $G_Y(x)$. This prevents the CT value of the entire image from changing substantially before and after the translation, and as a result, it is considered effective in preserving the anatomical structure.

The above equations (1) and (2) are summarized in equation (3), which is defined as the final objective function of this study. This is the proposed objective function.

$$\mathcal{L} = \mathcal{L}_{cgan} + \lambda_{int}\mathcal{L}_{int} \quad (3)$$

λ_{int} is a weight that controls the strength of the regularization term. The following equation (4) gives the learned model G_Y^* to be obtained.

$$G_Y^*, G_X^* = \arg \min_{G_Y, G_X} \max_{D_Y, D_X} \mathcal{L} \tag{4}$$

III. EXPERIMENTS AND RESULTS

In this study, we divided the training data of 68 subjects into the CBCT image group and the CT image group, and performed unsupervised learning using the proposed method. Throughout this study we use $\lambda_{cyc} = 10$ according to the previous experiment . Then we quantitatively evaluated the translation performance of the trained model using the proposed method, and compared conventional methods.

A. Training and evaluation methods

In order to investigate the image translation performance of the proposed method, two experiments are conducted: one is to investigate the relationship between the weight parameters of the regularization term and the translation performance, and the other is to compare the proposed method quantitatively with the conventional method. Each experiment is divided into a training phase and an evaluation phase. First, in the training phase, CBCT and CT image groups of 68 patients were used as training data, and adversarially learned using the proposed method. Next, in the evaluation phase, we translated the test CBCT images of 8 patients using the trained model, and after checking the translated images, we quantitatively evaluated the translation performance. For the quantitative evaluation, we need a reference image that can be used as a ground truth. The reference image must satisfy the following two conditions.

1) The anatomical structure should match the CBCT image of the test data.
2) Artifacts should be reduced and the CT values of the tissue to be imaged should be representative.

In this study, a reference image was created by aligning a pair of CBCT and CT images of the same patient. Specifically, in the test data, we deformed the CT image by non-rigid registration with the CT image as the source and the CBCT image as the target of the same patient. After visually confirming that the registration was highly accurate near the prostate, the deformed image was defined as the reference image. In this study, the evaluation range was limited to a local region near the prostate to evaluate the translation performance. Figure 3 shows that 256×256 pixels and 20 slices near the prostate were extracted from the three-dimension volume images and used for evaluation. The range of CT values for each image was set to be small so that finer differences in CT values could be evaluated. Specifically, CT values below -300 were set to -300 and CT values above 150 were set to 150, resulting in a final range of CT values of [-300, 150]. In the evaluation, histograms are obtained for the CBCT image, the translated image, and the reference image of the local volume. The mean and standard deviation(SD) of the CT values of each image are then calculated. The correlation coefficient between the histogram of the translated image and the histogram of the reference image is also calculated. These evaluation indices were used to quantitatively evaluate the translation performance.

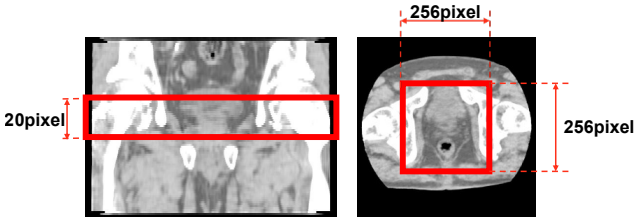

Fig. 3. Region of interest for evaluation. 256x256pixels, 20 slices near the prostate were extracted and used for evaluation.

TABLE I

AVERAGE OF 8 CASES OF THE EVALUATION VALUES (MEAN, STANDARD DEVIATION(SD), CORRELATION COEFFICIENT) OF THE ORIGINAL CBCT IMAGE AND THE TRANSLATED CT IMAGE LEARNED BY VARYING THE WEIGHT COEFFICIENTS λ_{int} OF THE REGULARIZATION TERM.

	original	0	5	10	15	20	25	reference
Mean	-67	0.97	7	8.6	7.1	-62	-65	16
SD	100	83	81	77	76	95	97	79
Correlation	0.45	0.92	0.89	0.92	0.9	0.47	0.43	1

B. Parameter Tuning

The purpose of this experiment was to tune the weight parameters of the regularization term and to verify the effectiveness of the loss function with regularization in the proposed method. After adversarial learning with six different weight coefficients of the regularization term ($\lambda_{int} = 0, 5, 10, 15, 20, 25$), the performance of each translation was quantitatively evaluated. Average of 8 cases of the evaluation values results of each weight are shown in Table I. It can be confirmed that the proposed method can translation images close to the reference image up to $\lambda_{int} = 15$. For $\lambda_{int} = 20, 25$, the regularization term was too strong and the translation failed, outputting an image that was almost the same as the original image. Overall, the results show that the conversion performance is high when $\lambda_{int} = 10$.

C. Comparison with conventional methods

The purpose of this experiment is to verify the effectiveness of the proposed method, and to do so, we compare it with the conventional method, CycleGAN[7]. We trained training data for each of the proposed method and CycleGAN, and performed quantitative evaluation and comparison using the test data. In this experiment, we set the regularization term weight factor $\lambda_{int} = 10$ for the proposed method based on the results of the previous experiment. First, we visually compared the translated images of each method. An example of the translation results is shown in Figure 4 (a). Comparing the yellow arrows in case A and the green arrows in Case B, we can see that the proposed method can retain the anatomical structure of the original image. Next, the histogram of the evaluation range for each image for case A is shown in Figure 4 (b). Comparing the

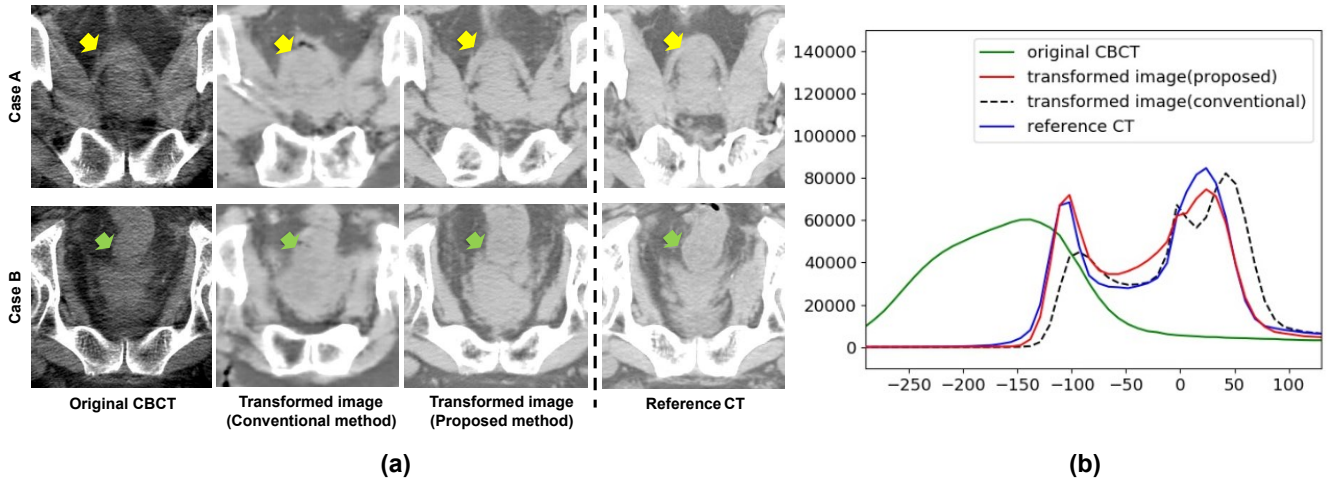

Fig. 4. Comparison of two examples of translation results between a model using the conventional method and a model using the proposed method. (a) The first column shows the reference image of each case. The second column shows the original CBCT images of each case. The third column shows the images translated from the CBCT images of each case by the conventional method. The fourth column shows the images translated from the CBCT images of each case by the proposed method, (b) Histogram of each image in case A

TABLE II
EVALUATION VALUES (MEAN, STANDARD DEVIATION (SD), CORRELATION COEFFICIENT) OF THE ORIGINAL CBCT IMAGE AND THE CT IMAGE AFTER TRANSLATION BY EACH METHOD

	original	conventional	proposed	reference
Mean	-67	-0.69	8.6	16
SD	100	83	77	79
Correlation	0.45	0.88	0.92	1

dashed line and the blue line in the histogram, we can see that the red line is closer to the shape of the blue line. Therefore, the proposed method is considered to have better conversion performance. Then, we compared the conversion performance of the proposed method and the conventional method with each quantitative evaluation index. Average of 8 cases of the evaluation values results of each method are shown in Table II. From these results, it can be confirmed that the proposed method has a higher conversion performance than the conventional method.

IV. CONCLUSIONS

In this study, we used an adversarial generative network that maps the input slice images based on the CycleGAN framework to improve the image quality of CBCT images based on the unsupervised learning framework. As a result of evaluating the translation performance of the learned model by the proposed method, we confirmed that the proposed method is effective for removing artifacts and translating CT values, and that mapping regions contained in slices is effective for preserving anatomical structures.

ACKNOWLEDGMENT

A part of this study was supposed by a JSPS Grant-in-Aid for challenging Exploratory Research (grant number 18K19918). We thank Karl Embleton, PhD, from Edanz Group (https://en-author-services.edanz.com/ac) for editing a draft of this manuscript.

REFERENCES

[1] H. Maekawa, M. Nakao, K. Mineura, T.F. Chen-Yoshikawa, T. Matsuda, Model-based registration for pneumothorax deformation analysis using intraoperative cone-beam CT images, Proc. 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2020.
[2] Y. Xin, W. Ekta, B. Paul, Generative adversarial network in medical imaging: A review, Medical Image Analysis, Vol. 58, 101576, 2019.
[3] M. Nakao, M. Nakamura, T. Mizowaki, T. Matsuda, Statistical deformation reconstruction using multi-organ shape features for pancreatic cancer localization, Medical Image Analysis, Vol. 67, 101829, 2021.
[4] K. Tateoka, Y. Saito, T. Nakazawa, M. Yano, K. Nakata, M. Someya, M. Hori, K. Sakata, Method for Converting Cone-Beam CT Values into Hounsfield Units for Radiation Treatment Planning, International Journal of Medical Physics, Clinical Engineering and Radiation Oncology, No. 6, pp. 361–375, 2017.
[5] C. Zöllner, S. Rit, C. Kurz, G. Vilches-Freixas, et al., Decomposing a prior-CT-based cone-beam CT projection correction algorithm into scatter and beam hardening components, Physics and Imaging in Radiation Oncology, No. 3, pp. 49–52, 2017.
[6] I. Goodfellow, J. Pouget, M. Mirza, B. Xu, D. Warde, S. Ozair, A. Courville, Y. Bengio, Generative Adversarial Nets, Advances in Neural Information Processing Systems 27, pp. 2672–2680, 2014.
[7] J. Zhu, T. Park, P. Isola, A. Efros, Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks, The IEEE International Conference on Computer Vision, pp. 2223–2232, 2017.
[8] M. Nakao, K. Imanishi, N. Ueda, Y. Imai, T. Kirita, T. Matsuda, Regularized three-dimensional generative adversarial nets for unsupervised metal artifact reduction in head and neck CT images, IEEE Access, Vol. 8, pp. 109453-109465, 2020.
[9] M. Nakamura, M. Nakao, K. Imanishi, H. Hirashima, Y. Tsuruta, Geometric and dosimetric impact of 3D generative adversarial network-based metal artifact reduction algorithm on VMAT and IMPT for the head and neck region, Radiation Oncology, 16, 96, 2021.
[10] L. Xiao, C. Liyuan, N. Dan, Z. Zhiguo, G. Xuejun, Y. Ming, W. Jing, J. Steve, Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy, arXiv, 2019.
[11] S. Kida, S. Kaji, Visual enhancement of Cone-beam CT by use of CycleGAN, Medical Physics, Vol.47, No.3, pp. 998–1010, 2020.
[12] T. Hase, M. Nakao, K. Imanishi, M. Nakamura, T. Matsuda, Proposal of 3D Generative Adversarial Network for Improving Image Quality of Cone-beam CT Images, IEICE Technical Report (MI), 120(156), pp.51-56, 2020.