



CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy

Yingzi Liu, Yang Lei, Tonghe Wang, and Yabo Fu

Department of Radiation Oncology and Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA

Xiangyang Tang

Department of Radiology and Imaging Sciences and Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA

Walter J. Curran, Tian Liu, Pretesh Patel, and Xiaofeng Yang^{a)}

Department of Radiation Oncology and Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA

(Received 29 August 2019; revised 27 February 2020; accepted for publication 27 February 2020; published xx xxxx xxxx)

Purpose: Current clinical application of cone-beam CT (CBCT) is limited to patient setup. Imaging artifacts and Hounsfield unit (HU) inaccuracy make the process of CBCT-based adaptive planning presently impractical. In this study, we developed a deep-learning-based approach to improve CBCT image quality and HU accuracy for potential extended clinical use in CBCT-guided pancreatic adaptive radiotherapy.

Methods: Thirty patients previously treated with pancreas SBRT were included. The CBCT acquired prior to the first fraction of treatment was registered to the planning CT for training and generation of synthetic CT (sCT). A self-attention cycle generative adversarial network (cycleGAN) was used to generate CBCT-based sCT. For the cohort of 30 patients, the CT-based contours and treatment plans were transferred to the first fraction CBCTs and sCTs for dosimetric comparison.

Results: At the site of abdomen, mean absolute error (MAE) between CT and sCT was 56.89 ± 13.84 HU, comparing to 81.06 ± 15.86 HU between CT and the raw CBCT. No significant differences ($P > 0.05$) were observed in the PTV and OAR dose-volume-histogram (DVH) metrics between the CT- and sCT-based plans, while significant differences ($P < 0.05$) were found between the CT- and the CBCT-based plans.

Conclusions: The image similarity and dosimetric agreement between the CT and sCT-based plans validated the dose calculation accuracy carried by sCT. The CBCT-based sCT approach can potentially increase treatment precision and thus minimize gastrointestinal toxicity. © 2020 American Association of Physicists in Medicine [<https://doi.org/10.1002/mp.14121>]

Key words: self-attention cycle, GANCBCT-based synthetic CT generation, adaptive radiotherapy

1. INTRODUCTION

Pancreatic cancer, being one of the leading causes of cancer-related mortality, continues showing poor prognosis with a 5-year overall survival (OS) of 9%.¹ Although surgery provides the best curative option, a majority of the patients are unresectable because they present with either locally advanced pancreatic cancer (LAPC) or with distant metastasis.² In patients with LAPC, stereotactic body radiotherapy (SBRT) involves high radiation doses within a short course has demonstrated improved locoregional control compared to conventional radiotherapy, which could potentially translate to improvement in survival.³ Pancreatic SBRT, however, presents several challenges, whereby up to 3-cm pancreas motion can be observed during the breathing cycle⁴ and patient weight loss occurs quite common during the treatment course. Caution should be exercised during the planning and delivery process to minimize the suboptimal target coverage and excessive doses to organs-at-risk (OARs) such as stomach and duodenum. Imaging acquisition typically involves respiratory management and the use of daily cone-beam CT (CBCT).

On-board CBCT is currently widely available in most clinics. The incorporation of CBCT allows for target position verification and set-up displacement correction. In the application of pancreatic SBRT, prior to the course of treatment, gold fiducials are usually inserted into the tumor to assist CBCT-based daily setup. This approach takes into consideration of the possible significant inter-fractional movement of the pancreas; however, it does not account for variations in position and shape of the surrounding OARs respect to the pancreas. It is therefore possible that excessive dose would be delivered to the OARs that results in gastrointestinal toxicity. To address this clinical challenge, MRI-guided adaptive planning has been proposed to spare the normal tissue adjacent to the treatment target.⁵ Due to the limitation that MR images cannot provide electron density information, adaptive planning is usually performed on the original planning CT using the deformed MR-based OAR contours. While much better soft-tissue contrast can be obtained by MRI than CBCT, this method has several limitations. First, the MR-CT registration process itself can introduce an uncertainty of 2–5 mm.⁶ In addition, it takes longer to obtain the MR images, which could lead to motion-induced artifact that degrade the precision of tissue delineation. Moreover, the

pancreatic patients can and usually undergo notable body weight/size loss over the course of treatment. Using the deformable registration to register MR images to the original planning CT simply neglect this anatomical change. Another approach of adaptive planning involves deformation of the planning CT to the CBCT to account for anatomical changes.⁷ While this method tends to represent the patient's true anatomy immediately prior to treatment, it still heavily depends on the accuracy of the algorithm that performs deformable registration. It could yield good dosimetry results at disease sites that are more stationary such as head and neck but would perform poorly at the site of abdomen because of the substantial patient motion and different organ water/air filling status.

Obviously, it would be more beneficial to directly use CBCT for adaptive radiotherapy without the additional error brought by the registration process. However, the raw CBCT data is not suitable for dose calculation. It is because of the considerable artifacts associated with CBCT such as streaking, shading, cupping and reduced image contrast caused by scatter contamination.^{8–10} As a result, the CBCT HU values usually fluctuate and differ significantly from the CT-based numbers.¹¹ To overcome the preceding drawbacks that hinder the direct use of CBCT for adaptive planning, CBCT imaging quality and HU fidelity must be improved. There have been a lot of publications that address this task of CBCT quality enhancement. The conventional methods can be broadly categorized into hardware corrections, such as collimator on the source side^{12,13} and anti-scatter grids on the detector side,¹⁴ and model-based correction, such as analytical methods¹⁵ and Monte Carlo simulations.⁷ The limitations of hardware approaches include severe heat loading problems and the degraded image quality because of reduced quantum efficiency; and the concerns found in model-based techniques include the capacity of the physics models, the computational efficiency, and the anatomical differences in CT and CBCT if prior CT information is used.

In light of their huge success in a variety of fields of medical imaging,^{16–27} machine-learning-based methods have also been proposed for the task of CBCT correction, by predicting a CBCT-to-CT model through a training process to learning the comprehensive relationship between CBCT and CT. To make it consistent, the corrected CBCT generated by machine-learning-based methods is called as synthetic CT (sCT) in the rest of the paper. Generally, the training process of machine-learning-based methods could be computationally expensive, but once the training is finished, the prediction of sCT would be within a short time when a new CBCT is fed in. The fast prediction speed makes it desirable in the application of adaptive radiotherapy. While relatively new, there have already been several studies reporting the generation of sCT from CBCT. They can broadly fall into the following two categories in terms of the composition of training dataset: CBCT-corrected CBCT pairs^{28–31} and CBCT-planning CT pairs.^{32–36} In the former category, the corrected CBCT differs from sCT in that it is the projections or reconstructions of the corrected images that are obtained either by analytical methods²⁸ or Monte Carlo simulations.^{29–31}

Although widely varied in the detailed designs and algorithms, deep-learning-based 2D residual U-net networks were used in the former category to generate sCT.^{28–31} In those works, encouraging improvements can be found in the imaging quality and dose calculation accuracy. However, the quantitative imaging and dosimetric comparisons were carried out between the sCT and the reference corrected CBCT, rather than directly comparing to the clinically used planning CT. Although improvement can even be found in the sCT compared to the corrected CBCT,^{28,29} the physics models used for the generation of corrected CBCT could still limit the learning scope of the machine-learning algorithms. CBCT-planning CT-based training dataset, on the other hand, suffers from the limitation that there are different organ filling statuses and air pocket locations between the CBCT and planning CT, which could affect the accuracy of the learned CBCT-to-sCT mapping algorithm. The design of the network must, therefore, be robust enough to mitigate the shortcomings of imperfect training dataset. Kida et al.³² used a 2D U-net deep convolutional neural network (DCNN) and mean absolute error (MAE) as the loss function for the pelvic CBCT-to-sCT generation. Similarly, Li et al.³⁴ used a U-net DCNN with residual convolution unit for the head and neck CBCT-to-sCT generation. The MAE between sCT and CT ranges from 6–27 HU. Lei et al. applied an anatomical feature selection and alternating random forest framework for CBCT artifact correction at the site of both brain and pelvis. The mean MAEs are 12.81 ± 2.04 and 19.94 ± 5.44 HU, respectively. Recently, cycleGAN³⁷ has been proposed to deal with the unpaired training data in multiple applications in medical imaging such as MRI-based sCT generation,^{38–42} organ segmentation,^{43,44} and CBCT-based sCT generation.^{33,36} CycleGAN incorporates an inverse transformation to better constrain the training model toward one-to-one mapping. In the application of CBCT-to-sCT generation, Liang et al.³³ applied cycleGAN to train the CBCT-planning CT dataset without performing deformable registration. The cycleCBCT generated from CT was used to restrain the network. The mean MAE is 29.85 ± 4.94 HU for the head and neck patient cohort. Similarly, our group (Harms et al.)³⁶ published CBCT-to-sCT generation method using cycleGAN with the incorporation of residual blocks and a novel compound loss in the cycle consistency loss function. The mean MAEs at the site of the brain and pelvis in this study are 13.0 ± 2.2 and 16.1 ± 4.5 HU, respectively. In Harms et al.'s work, the authors mentioned that although cycleGAN was initially designed for unpaired mapping,³⁷ in its application in medical imaging, due to the imaging complexity, rigid registration is still suggested to preserve the qualitative values.

In this study, we aimed to generate pancreatic sCT that can be used for adaptive radiotherapy. Due to the substantial motion at the abdomen and different bowel air pocket statuses, there are spatial mismatches in internal tumor and organ positions. The CBCT images were paired to the planning CT through deformable registration. The main network was based on the cycleGAN developed by Harms et al.³⁶ to account for the mismatched dataset. A larger 3D patch size ($64 \times 64 \times 64$) was

used to capture more spatial relationships. To address the blurred CBCT image caused by motion artifacts, attention gate (AG)⁴⁵ were incorporated in the generator architecture to learn the structural variations, through which the most relevant semantic contextual information can be captured. A cohort of 30 pancreatic cancer patients who underwent SBRT at a single institute was included in this study. Initial CT-based plans were transferred to CBCT and sCT for dosimetric comparison. The main contributions of this work include: (a) This study applied the use of the deep-learning-based method in CBCT-based adaptive therapy for pancreatic cancer treatment, which is new and would be very helpful for future clinical practice. (b) An attention gate strategy was incorporated in the cycleGAN network to better differentiate organ boundaries and improve the HU prediction accuracy.

2. MATERIALS AND METHODS

2.A. Image acquisition and processing

The study cohort was composed of 30 pancreatic cancer patients who underwent breath-hold SBRT at our clinic from 2017 to 2019. CT scans were acquired on Siemens (Erlangen, Germany) SOMATOM Definition AS with 120 kVp. Contrast was injected in the stomach during the CT scan. mAs ranged from 77 to 378. The CT slice thickness was 1.5 mm with superior-inferior coverage ranging from 291 to 568.5 mm. The voxel size in the left-right and anterior-posterior was 0.977 mm with field of view (FOV) of 500×500 mm². CBCT scans were acquired prior to each treatment fraction on a Varian TrueBeam (Varian Medical Systems, Inc. Palo Alto, USA) with 125 kVp. The exposure current was 80 mA. The CBCT slice thickness was 1.99 mm with superior-inferior coverage of 174.99 mm. The voxel size was 0.91 mm with FOV of 464.91×464.91 mm². The first fraction CBCTs and CTs were used as the training dataset. Image data were extracted retrospectively under an IRB-approved protocol. The images were imported to Velocity AI 3.2.1 (Varian Medical Systems, Inc. Palo Alto, USA) for registration and resampling. First, rigid registration was performed to pair the CBCT to CT by aligning the bony structures such as the vertebral bone. Deformable registration from CBCT to CT was then carried out. In Velocity, the deformable multi-pass was chosen as the deformation algorithm. The deformed CBCT was then resampled so that the data volume and voxel size between the CBCT and CT are identical. After acquiring the paired CBCT-CT dataset, to minimize the influence of different air pocket status that can confuse the machine learning network, the air pockets in the CT were copied to the CBCT; similarly, the air pockets in the CBCT were copied to the CT. During the training and prediction process by our algorithm, leave-one-out cross-validation was used for the 30 CBCT-CT training pairs.

2.B. Self-attention cycleGAN-based CBCT to sCT generation

Due to the different organ shape/locations caused by patient motion, there are inevitable mismatches between the

CBCT and CT images. Therefore, training a CBCT-to-CT transformation model could be highly under-constrained, meaning small data errors could be amplified during this transformation. To cope with the issue of potential erroneous mapping in one direction, our group has employed cycleGAN at the sites of brain and pelvis.³⁶ Interested reader can refer to Harms et al.'s paper for the detailed design of the network, here we only provide a brief description: A typical GAN-based method employs both the generator and discriminator networks competing with each other. Specifically, the generator's training objective is to produce a CT-like image from a CBCT image, called sCT. The discriminator's training objective is to differentiate the sCT from the real CT. Improved performances of both the generator and discriminator networks can be achieved through the training process. A cycleGAN³⁷ doubles the process of GAN by enforcing an inverse transformation, that is, transforming a CT to a CBCT-like image. This doubly constrains the model toward one-to-one mapping attempting to generate more accurate output images. Residual blocks⁴⁶ were employed in the generator network to minimize a residual, or error, image between the CBCT and CT. In Harms et al.'s work, a novel compound loss function was also used, which includes a $l_{1.5}$ -norm loss and a gradient-magnitude distance loss, attempting to solve the problem of blurry tissue boundaries and in the meantime avoid tissue misclassification.

In this work, we aimed to apply the 3D cycleGAN method at the site of abdomen for the potential use in pancreatic adaptive radiotherapy. However, because of patient motion, organ dislocations were still observed and the organs in CBCT can be blurred through the deformable registration process. Compared to the more stationary body sites such as brain and pelvis, direct use of cycleGAN may give suboptimal results, as the mismatch on pancreas site is larger. To deal with this issue, attention gate (AG) was introduced to the generator networks aiming to learn the structural variations. Unlike directly using the feature maps from the encoding path, the feature maps extracted from the coarse-scale were used in gating to disambiguate irrelevant and noisy responses in long skip connections. This was performed immediately prior to the concatenation operation to merge only relevant activations. In addition, AG filtered the neuron activations during both the forward and backward pass. Gradients originating from background regions were down-weighted during the backward pass, which allowed model parameters in shallower layers to be updated mostly based on spatial regions that were relevant to a given task. By using AG, the most salient features from the compression path were highlighted and were passed through the bridge path. Figure 1 shows the schematic workflow of the used methods. The networks can be broken into two stages: training (shown on the left) and prediction (shown on the right). During training, the patches ($64 \times 64 \times 64$) with an overlap of $32 \times 32 \times 32$ between every two neighboring patches from training CBCT and CT images are extracted and go through the generator networks to produce the corresponding synthetic images. The synthetic image patches

then go through the adverse generator to produce the corresponding cycle images. After the training is finished, an incoming CBCT image can be fed into the well-trained model to produce the corresponding sCT image. A whole synthetic CT was obtained via average-based patch fusion of the sCT patches, that is, in patches which overlap upon output, voxel values in the same position were averaged. In the generator architecture, nine residual blocks were used to

learn the differences between CBCT and CT and an AG was used to learn structure variations.

The model was trained and tested on an NVIDIA Tesla V100 GPU with 32 GB memory. The batch size was 20. 10 GB CPU memory and 29 GB GPU memory was used for each batch optimization. The algorithm was implemented by PYTHON 3.7 and Tensorflow 1.14. The training process took around 13 h and the prediction of one sCT set took about 1–2 min.

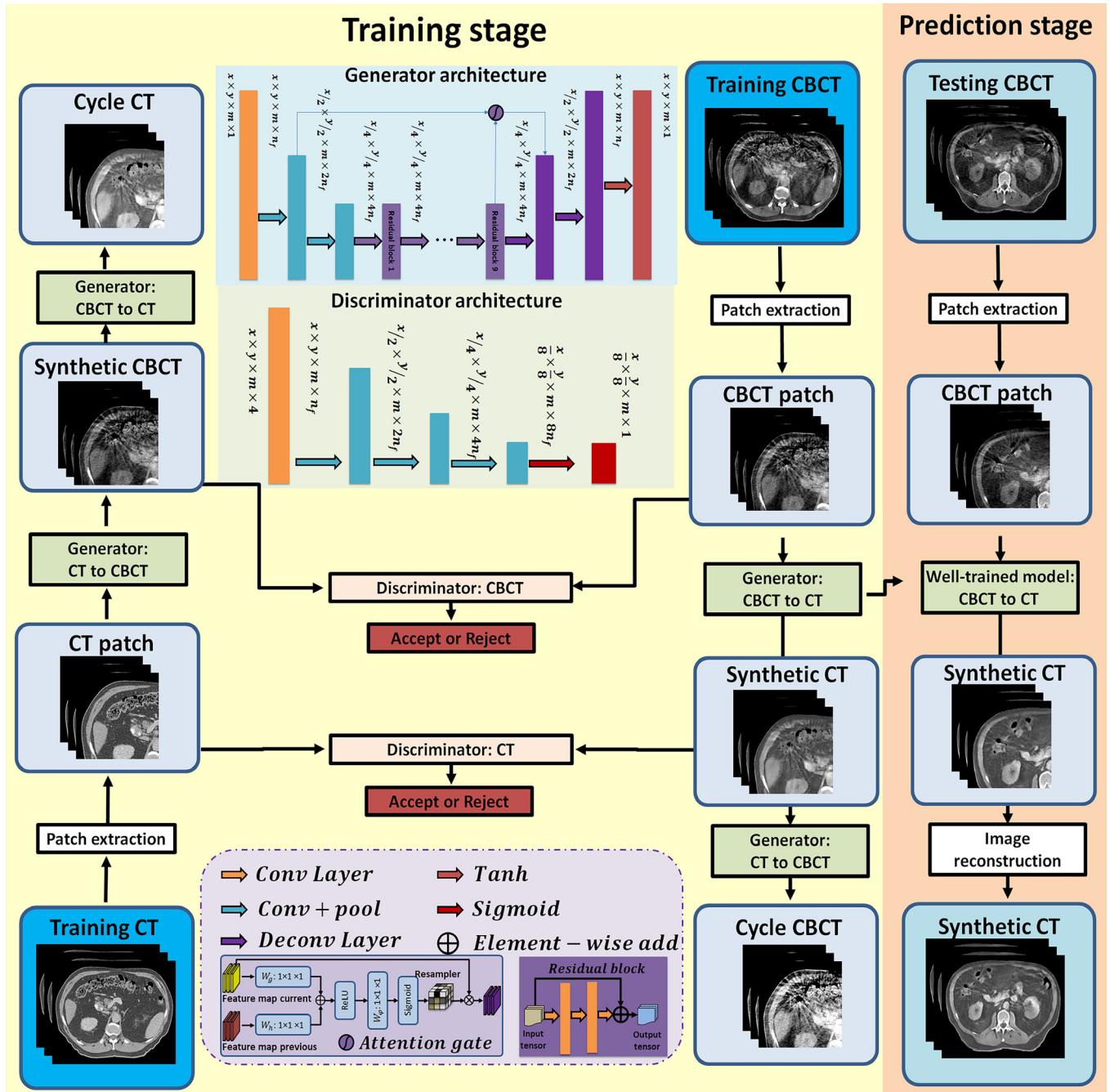


FIG. 1. Schematic flowchart of the deep-attention cycleGAN network for sCT generation. The networks can be broken into two stages: training (shown on the left) and prediction (shown on the right). During training, the patches from training CBCT and CT images are extracted and go through the generator networks to produce the corresponding synthetic images. The synthetic images patches then go through the adverse generator to produce the corresponding cycle images. After the training finished, an incoming CBCT image is fed into the well-trained model to produce the corresponding sCT image. CBCT, cone-beam CT; sCT, synthetic CT

2.C. CBCT-based sCT quality evaluation

The sCT evaluation strategy can be broken into imaging quality and dose calculation capacity. Mean absolute error (MAE), spatial nonuniformity (SNU), peak signal to noise ratio (PSNR), normalized cross-correlation (NCC), and structural similarity index (SSIM) were used as the imaging end-points. They are defined as follows:

$$MAE = \frac{1}{n_i n_j n_k} \sum_{x,y,z} |CT(x,y,z) - sCT(x,y,z)|$$

$$SNU = \frac{\overline{HU}_{max} - \overline{HU}_{min}}{1000} \times 100\%$$

$$PSNR = 10 \times \log_{10} \left(\frac{MAX^2}{\sum_{x,y,z}^{n_i n_j n_k} |CT(x,y,z) - sCT(x,y,z)|^2 / n_i n_j n_k} \right)$$

$$NCC = \frac{1}{n_i n_j n_k} \sum_{x,y,z} \frac{(CT(x,y,z) - \mu_{CT})(sCT(x,y,z) - \mu_{sCT})}{\sigma_{CT} \sigma_{sCT}}$$

$$SSIM(CT, sCT) = \frac{(2\mu_{CT}\mu_{sCT} + c_1)(2\sigma_{CTsCT} + c_2)}{(\mu_{CT}^2 + \mu_{sCT}^2 + c_1)(\sigma_{CT}^2 + \sigma_{sCT}^2 + c_2)}$$

where $CT(x,y,z)$ and $sCT(x,y,z)$ are the value of pixel (x,y,z) in the planning CT and sCT respectively. $n_i n_j n_k$ is the total number of pixels. \overline{HU}_{max} and \overline{HU}_{min} are the averaged maximum and the minimum intensity of region of interests (ROIs) selected for each patient's data. MAX is the maximum intensity in CT and SCT. μ_{CT} and μ_{sCT} are the mean of CT and sCT image. σ_{CT} and σ_{sCT} are the standard deviation of CT and sCT image. MAE is the magnitude of the voxel-based Hounsfield unit (HU) difference between the original CT and the sCT. SNU and PSNR measure if the predicted sCT intensity is evenly or sparsely distributed. NCC is a measure of similarity between CT and sCT as a function of the displacement. In this study, the metrics were calculated within the body mask of the sCT and seven ROIs were selected for each patient at the sites that CBCT-related artifacts occur to calculate SNU.

To prove that the sCT can carry the dose calculation with comparable accuracy to the planning CT, CT-based SBRT treatment plan was transferred to the corresponding sCT. PTV and OAR contours were transferred directly from the planning CT. The relevant DVH metrics for PTV and OARs including D_{10} , D_{50} , D_{95} , D_{min} , D_{mean} , and D_{max} were calculated.

The same evaluation processes were also carried out on the raw CBCT to enable the comparison between the raw CBCT and sCT.

3. RESULTS

3.A. Efficacy of self-attention

The Figure 2 shows the sCT images with and without the use of AG. While both cycleGAN-based networks can obviously improve the image quality and restore the organ

boundaries, the network with the use of AG slightly outperforms the other. Sharper organ boundaries can be found with the use of AG, which are essential for the further development of CBCT-based segmentation. Furthermore, as shown in the second row, the sCT generated with AG shows better intensity uniformity in the same organ. The third and fourth rows provide the checkerboard images to compare the differences between CT, CBCT, and sCTs. The differences between CT and sCTs generated by cycleGAN with attention gate are smaller than the ones generated without using attention gate. There are some structure differences that were originated from the structural differences in CT and CBCT. Although deformable registration was performed, it was still difficult to match the organ structure in CBCT to be the same as CT.

3.B. Image quality

The average MAE between sCT and CT for the patient cohort 56.89 ± 13.84 HU, comparing to 81.06 ± 15.86 HU between CBCT and CT. The average SNU of the sCT was 0.12 ± 0.085 , comparing to 0.25 ± 0.097 for the CBCT. Table I tabulates the comparison between CBCT, sCT generated by 3D U-Net, with and without AG. The original CT and CBCT images as well as the sCT images of a representative patient were compared side-by-side (Fig. 3). Our method can nearly restore finely detailed images structures seen on the CT whereas the CBCT tends to blur these structures. The HU differences between CT and sCT are much smaller compared to between CT and CBCT. The scatter and streaking artifacts near the air pockets and the gold fiducial were significantly reduced in the sCT images. The HU histogram plots of the whole 3D image sets shown in Fig. 3 to illustrate the improvement of HU fidelity corresponding to CT. As can be seen in the HU histogram plot, the sCTs are in good agreement with the CT at the peak around -100 HU. There are some discrepancies at the higher HU peak around 30 – 40 HU that may be due to the contrast injection in the stomach during the CT scan.

Large differences can be found at contrast-injected stomach due to the time intervals between CT and CBCT acquisition. Some discrepancies in the coronal and sagittal views were due to the different organ sizes and locations in CT and CBCT/sCT. These can be considered as one of the limitations of the ground truth (CT-CBCT pairs) that may affect the results of MAE. But on the other hand, it also demonstrated that our algorithm has effectively dealt with the limitations of the training dataset by showing reasonable HU values at the stomach. This method is also robust enough to show the correct organ size and position that are the same as the CBCT, rather than forcing the organ to be alike the one in CT.

3.C. Dose calculation capacity

Figure 4 shows the dose difference map of a representative patient. Minimal dose differences can be found between the CT- and sCT-based plans compared to the CT- and CBCT-

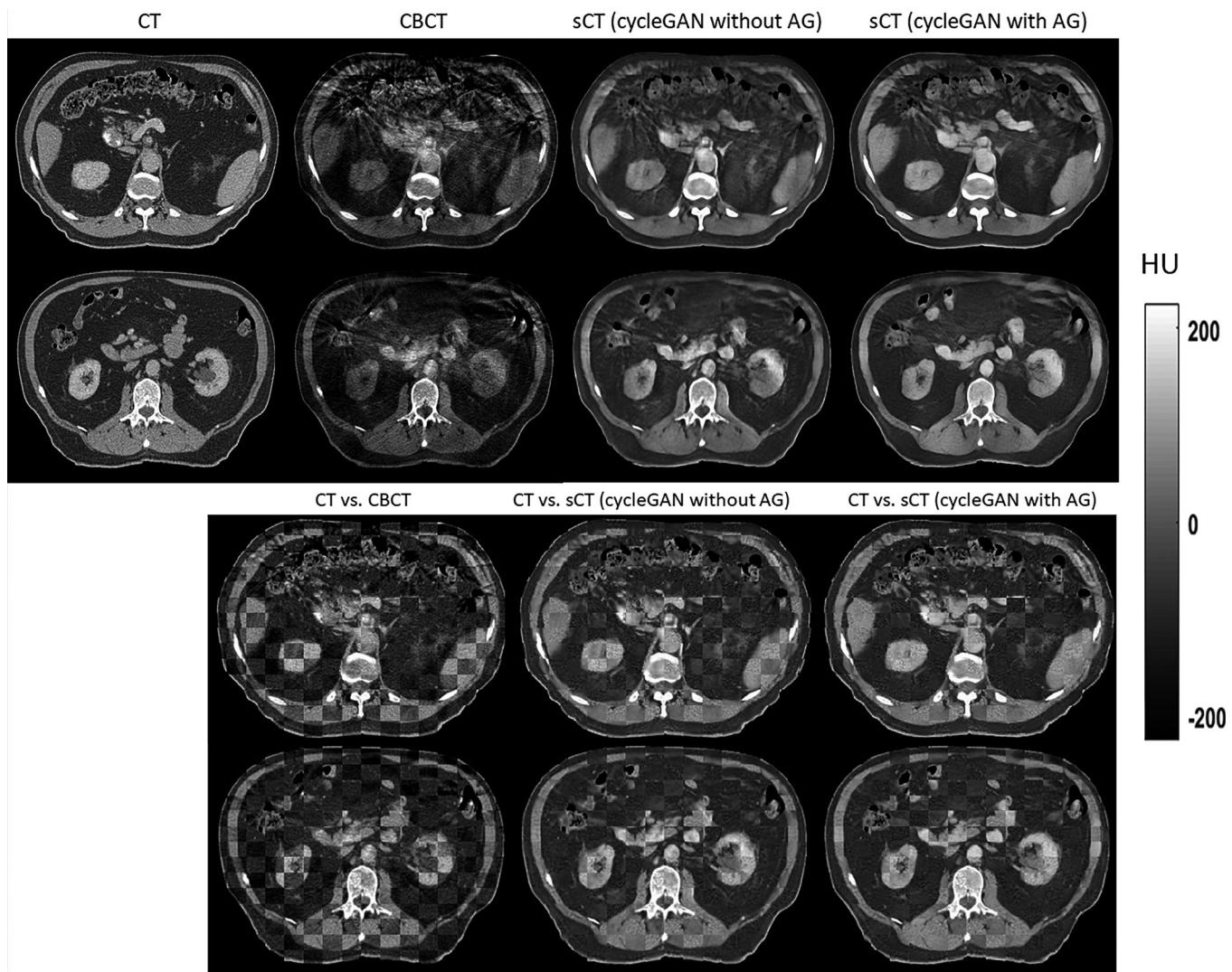


FIG. 2. Visual comparison with and without the use of AG in the cycleGAN networks

TABLE I. Numerical comparison among CBCT, sCT generated by U-Net, with and without AG

	CBCT	U-Net	CycleGAN	CycleGAN with AG
MAE (HU)	81.06 ± 15.86	66.71 ± 15.82	63.72 ± 10.55	56.89 ± 13.84
SNU	0.25 ± 0.097	0.19 ± 0.088	0.15 ± 0.096	0.12 ± 0.085
PSNR (dB)	22.89 ± 2.89	28.17 ± 3.34	25.69 ± 3.40	28.80 ± 2.46
NCC	0.85 ± 0.081	0.87 ± 0.033	0.94 ± 0.040	0.96 ± 0.022
SSIM	0.60 ± 0.063	0.56 ± 0.085	0.70 ± 0.044	0.71 ± 0.032

AG, attention gate; CBCT, cone-beam CT; MAE, mean absolute error; PSNR, peak signal to noise ratio; SNU, spatial nonuniformity; sCT, synthetic CT.

based plans. The OARs include bowel, duodenum, liver, kidney and stomach. Figure 5 shows the box plot of dose difference between CT- and sCT-based plans in DVH metrics for PTV and OARs. All the DVH dose differences were less than 1 Gy. Table II shows that no significant differences ($P > 0.05$) were observed in the PTV and OAR DVH metrics including D_{95} , D_{mean} , and D_{max} between the CT- and sCT-based plans, while significant differences ($P < 0.05$) were found between the CT- and CBCT-based plans such as in

PTV DVHs. These results indicate that our sCTs can provide reliable dose calculation results compared to the planning CT.

4. DISCUSSION

This work generated pancreatic sCT from routine CBCT images based on our self-attention cycleGAN method. The image quality of the sCT has been greatly improved by showing clearer organ boundaries and less scattering artifacts. The

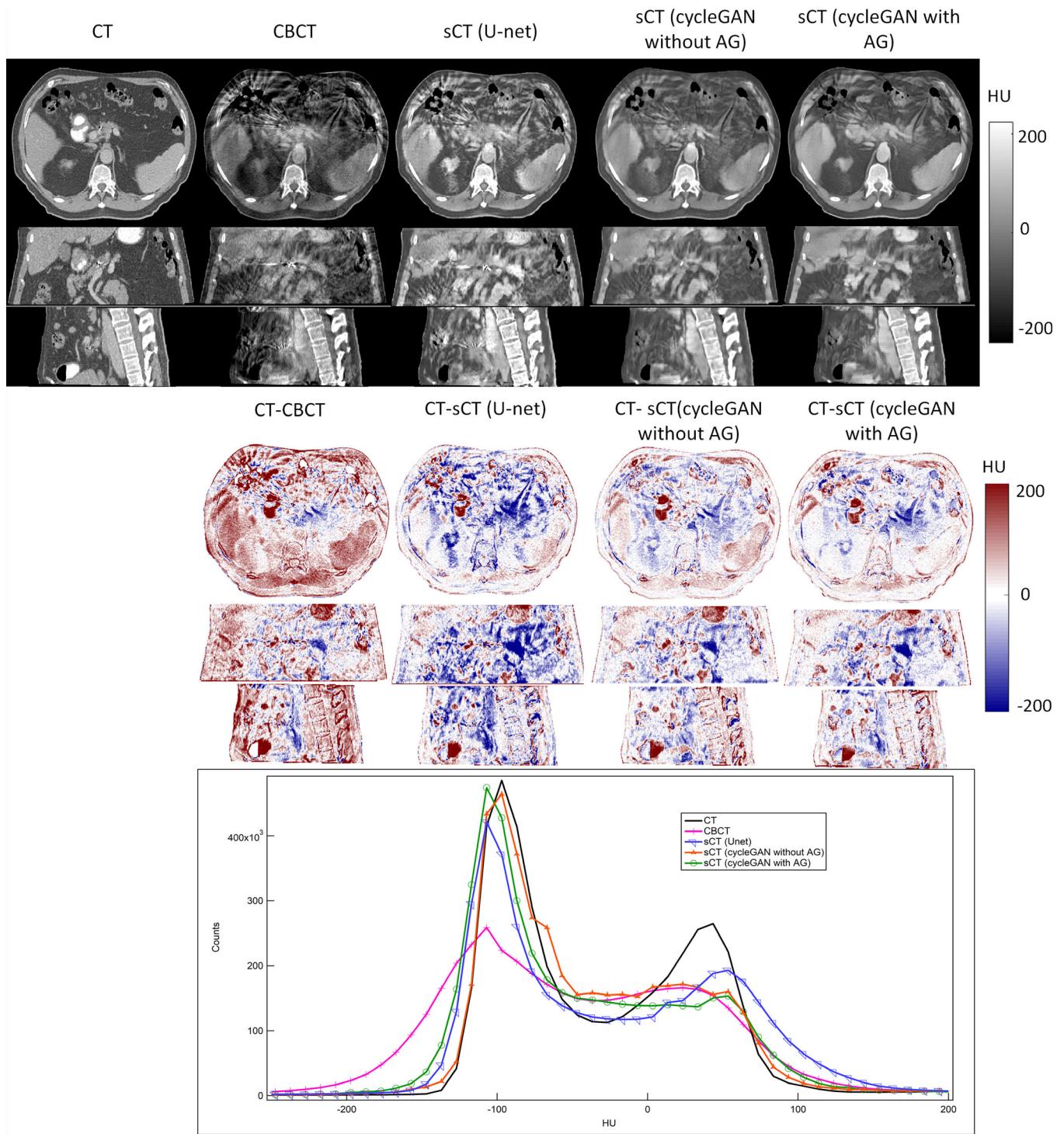


FIG. 3. Summary of CBCT to sCT results in a representative patient. CBCT, cone-beam CT; sCT, synthetic CT

average MAE between CT and sCT was 56.89 ± 13.84 HU, comparing to 81.06 ± 15.86 HU between CT and the original CBCT. An overall improvement around 24 HU was thus found between CT-CBCT and CT-sCT. The mean MAE achieved in this study (56.89 ± 13.84 HU) appears to be larger than the other CBCT to sCT studies that already described in the introduction. There are several reasons for this discrepancy: the main reason could be that different body

parts were used to generate the corresponding sCTs. Both brain and pelvis were evaluated in Harms et al.'s study and Lei et al.'s study, the average MAE between CT and raw CBCT were 27.8 ± 6.6 HU (Harms et al.)/ 23.8 ± 5.1 (Lei et al.) HU for brain and 56.3 ± 19.7 HU (Harms et al.)/ 45.5 ± 12.3 (Lei et al.) HU for pelvis. In contrast, the average MAE between CT and CBCT was 81.06 ± 15.86 HU at the site of abdomen. This MAE discrepancy partly lies

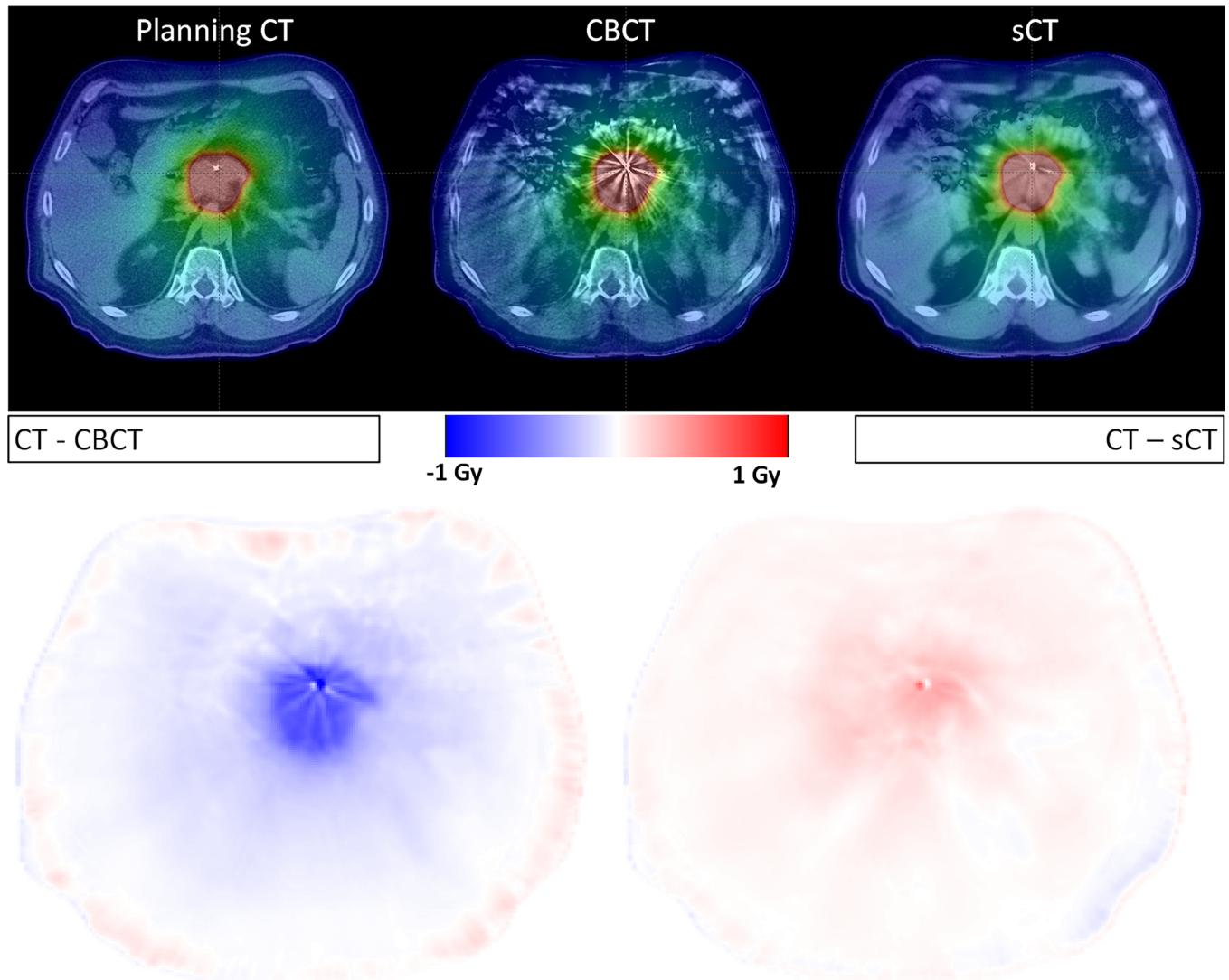


FIG. 4. Dose comparison between CT- and CBCT-, as well as between CT- and sCT-based plans. CBCT, cone-beam CT; sCT, synthetic CT

in the HU inaccuracy in the CBCT dataset and partly is due to the different organ shapes and locations in CT and CBCT. For the latter source of discrepancy, it can be traced back as the limitation of the training dataset but not from the proposed deep-learning-based algorithm. Brain could be considered a stationary site with little organ dislocations, followed by pelvis when the different bladder filling status and bowel movement could contribute to a higher MAE, and the highest MAE is found in abdomen because of the substantial motion caused by patient respiration and internal organ movement. It is worth noting for the purpose of adaptive radiotherapy, the deep learning algorithm should not force the predicted organ locations in sCT to be the same as those in the planning CT, but to focus on the CBCT-to-CT intensity mapping algorithm that is capable of mitigating the artifacts and enhancing the HU accuracy for the tissues in CBCT. Therefore, while the calculation of MAE still serves as one of the main approaches to evaluate the quality of sCT, the same dataset or at least same body

site needs to be used to make comparison between different algorithms.

Deformable registration was performed on the CBCT to pair it to the corresponding CT. The reasons to perform deformable registration were to: (a) The OAR shapes and locations, as well as body contour sizes can be different between CT and CBCT images, which could confuse the training networks because it is not one-to-one mapping. For example, the anterior body part in CT may be paired with air in the corresponding CBCT due to patient weight loss. (b) To prove that the HU values from the synthetic CT are accurate, we compared the HU differences and doses calculated on synthetic CT and planning CT. To exclude the potential influences from the different body size, it is necessary to use the synthetic CT generated from the deformed CBCT. To compare the performance of training data using rigid registration and deformable registration, the network was re-trained with rigidly registered CT-CBCT pairs. The newly generated sCTs were then imported to Velocity to be deformably registered to

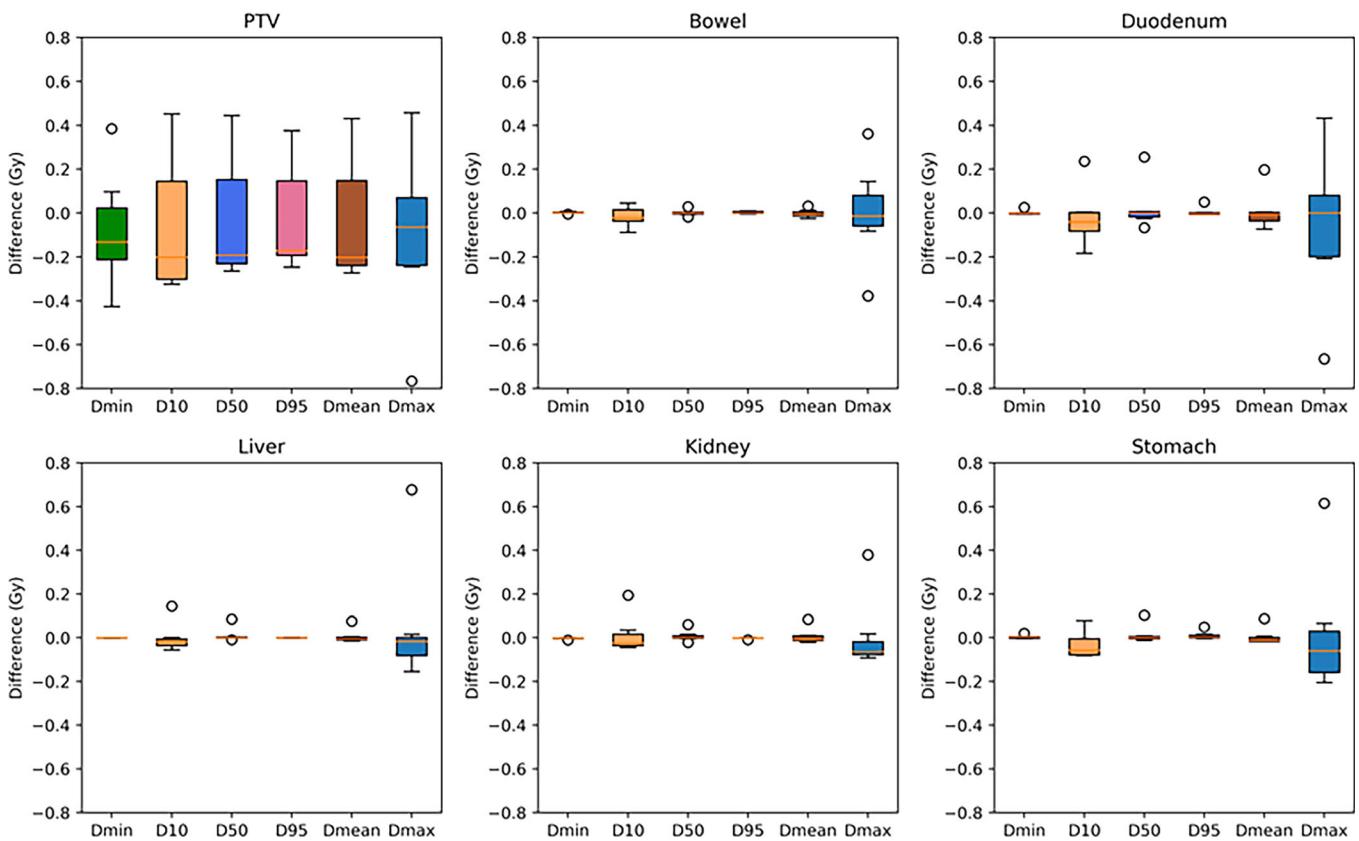


FIG. 5. Box plot of the DVH dose difference between CT- and sCT-based plans in PTV and OARs. The central orange line indicates the median value, and the borders of the box represent the 25th and 75th percentiles. The outliers are plotted by the black 'O' marker. OARs, organs-at-risk; sCT, synthetic CT

TABLE II. *P*-value comparison between CT- and CBCT-based plans, as well as CT- and sCT-based plans

	PTV	Bowel	Duodenum	Liver	Kidney	Stomach
CT vs. CBCT	D95	0.008	0.016	0.530	0.192	0.079
	Dmean	0.010	0.027	0.110	0.344	0.212
	Dmax	0.011	0.033	0.082	0.100	0.055
CT vs. sCT	D95	0.845	0.193	0.581	0.534	0.186
	Dmean	0.769	0.769	0.842	0.619	0.623
	Dmax	0.509	0.981	0.630	0.665	0.902

CBCT, cone-beam CT; sCT, synthetic CT.

the corresponding CT. The mean MAE was 58.45 ± 13.88 , comparing to 56.89 ± 13.84 HU achieved by the model trained with deformably registered CT-CBCT pairs. Although there was no significant difference ($P = 0.21$) between the rigidly registered trained and deformably registered trained networks, the side-by-side comparison between the sCT generated by the two methods demonstrated some disagreement, in which the sCTs generated by the model trained by deformably registered CT-CBCT pair appeared to be less noisy and have better organ boundaries than the sCT generated by the model trained by rigidly registered CT-CBCT pairs (Fig. 6).

To evaluate whether the sCT generated by our method is capable for accurate dose calculation, both the CBCT- and

sCT-based treatment plans were compared to the CT-based plans, showing no significant differences between sCT- and CT-based plans, while significant differences were found between CBCT- and CT-based plans. It is important to clarify that the doses were calculated on the deformed CBCT and sCT images with the same air pockets status to the planning CT to exclude the potential influences from the different body size and air pocket status. Therefore, the evaluation is focused on the HU accuracy of CBCT and sCT. One may argue that the organ size and location could be notably different between CT and CBCT/sCT, which makes the PTV and OAR DVH comparison questionable when a direct transferring of contours from CT to CBCT/sCT were performed. While it is true as the mismatch between CT and CBCT/sCT serves as the major limitation of all the machine learning-based algorithms, it is worth pointing out that other than bone and air, the HU differences among the other organs are relatively small in which the dosimetric impact would be minimal when different locations of those organs were found. Moreover, the SBRT treatment plans use the volumetric body radiotherapy (VMAT) technique that employs multiple entry points that will ultimately mitigate the impact of different organ size and locations. To summarize, the dose comparison conducted in this study only served the purpose to show that the sCT generated by our method is capable to provide relatively accurate HU numbers that shows little dosimetric

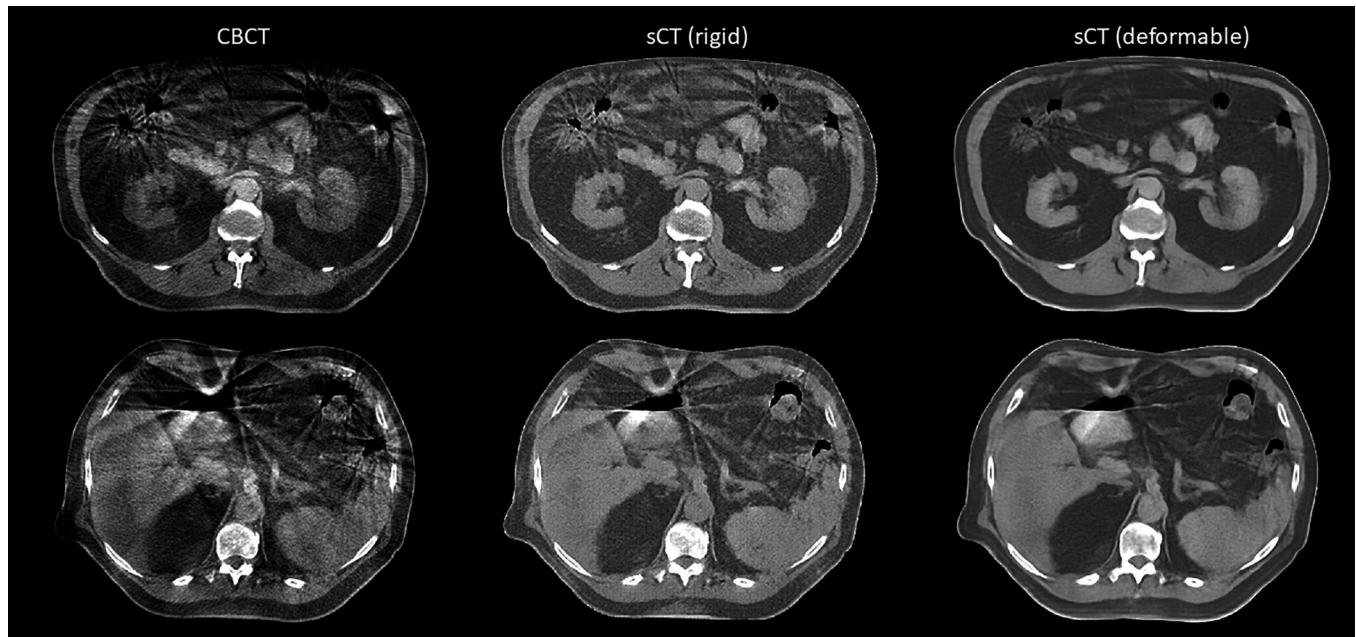


FIG. 6. Comparisons between the networks trained by rigid and deformable registration

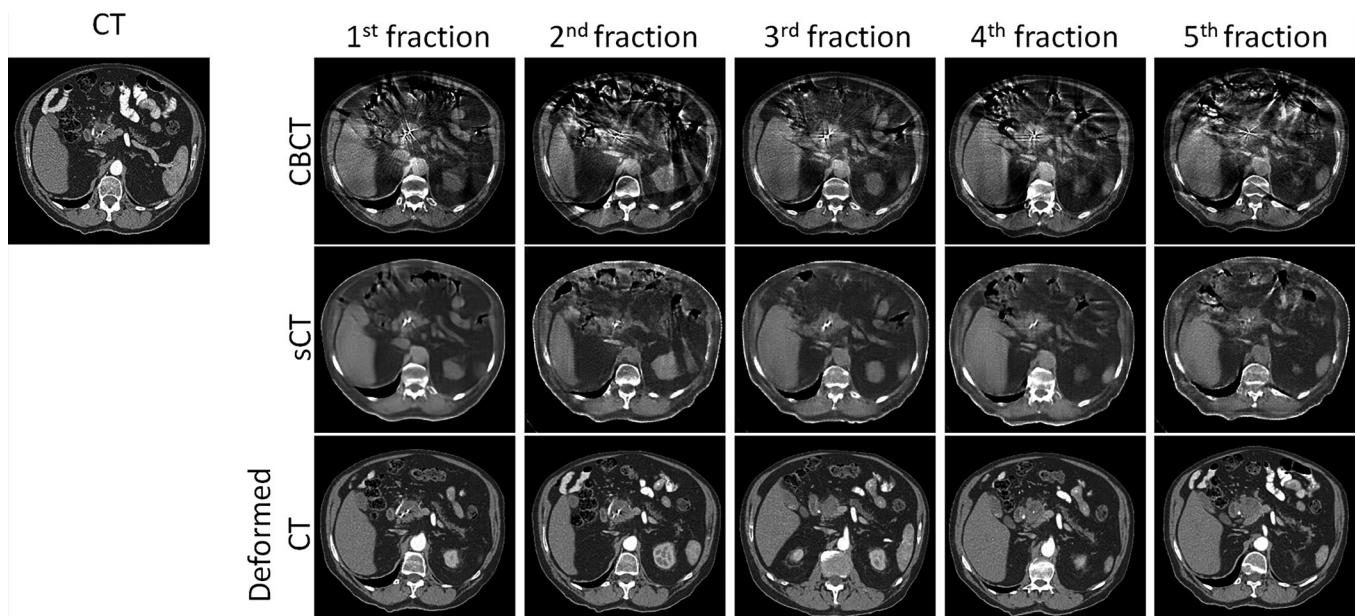


FIG. 7. Comparisons of the shapes and locations of PTV and OARs between DIR- and sCT-based images. DIR, deformable image registration; OARs, organs-at-risk; sCT, synthetic CT

difference compared to the planning CT. As we have proved that the sCT is capable for accurate dose calculation in this study, future work will focus on the impact of patient size and motion on the PTV and OAR DVHs with the use of the true size CBCT/sCT and the online registration that was performed prior to each treatment with the updated contours showing the real organ locations. More specifically, the sCT images will be generated directly based on the newly acquired CBCT images and the fully trained CBCT-to-sCT model. Then, automatic multi-organ segmentation can be performed on the sCT images to generate the OAR contours reflecting

the real shapes and locations when the patient is on the treatment couch. A quick dose calculation should subsequently be performed showing the DVHs of PTV and OARs. The physician can then determine whether the treatment can be delivered or modification is required for the treatment plan in response to the changes in PTV and OAR DVHs.

To compare deformable image registration (DIR)-based images and sCT, we deformed the planning CT to the 5 fraction CBCTs. The sCTs were generated directly by the well-trained model. As shown in Fig. 7 below, although DIR-based method can provide better HU values, it cannot reflect

the real daily anatomy of the patient due to the registration error, especially for the complex mixture of air and soft tissue. The purpose of pancreatic adaptive RT is to evaluate and adjust the plan based on the accurate anatomic information on the treatment position, which can better be achieved through CBCT-to-sCT-based method since the sCT have the same geometry as CBCT. Therefore, even though DIR may have better dose calculation results, its unmatched geometry may introduce errors in re-contouring. The error in contouring can lead to under- and overdose at the level of prescription dose in target and OARs, respectively, which is much larger than the error <1 Gy of sCT in this study.

As can be seen from the last paragraph, CBCT image quality and dose calculation capacity are not the only hindrance to clinical implementation of on-line adaptive radiotherapy. Accurate and rapid multi-organ segmentation and dose calculation are the other key factors. CT-based abdominal multi-organ segmentation has been reported through atlas-based method^{47,48} as well as learning-based method.^{49,50} However, CBCT-based segmentation has not yet been studied. It might due to the severe imaging artifact seen in the original CBCT image. Since we have proved the imaging quality enhancement through our CBCT-to-CT algorithm by showing much sharper organ boundaries, the future of CBCT-based multi-organ segmentation is promising. Regarding the rapid dose calculation, besides the fast development of GPU-based dose calculation engine,⁵¹ learning-based automatically dose distribution generation has appeared as a hot topic.^{52,53} However, the inherent limitation of CBCT must be understood. For example, the acquired images have less field-of-view compared to the CT images, and the longer acquisition time may induce larger motion artifacts. Since co-planar plan is usually used for pancreatic SBRT, the authors do not anticipate that the limited field-of-view would be a big factor that impacts the dose calculation accuracy. But full validation needs to be carried out. Better motion management during the CBCT acquisition is also required.

5. CONCLUSIONS

In this study, we have developed a self-attention cycle-GAN method for CBCT to sCT generation. The sCT images could provide accurate dose calculation that is comparable to the planning CT. Much less artifacts and sharper organ boundaries were found in the sCT compared to the CBCT, which could facilitate the future development of CBCT-based multi-organ segmentation. This approach can potentially increase treatment precision and thus minimize gastrointestinal toxicity.

ACKNOWLEDGMENTS

This work is supported in part by the National Institutes of Health under Award Number R01CA215718, and the Emory Winship Cancer Institute pilot grant.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

^{a)}Author to whom correspondence should be addressed. Electronic mail: xiaofeng.yang@emory.edu; Tel: (404)-778-8622; Fax: (404)-778-4139

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA: A Cancer J Clin.* 2019;69:7–34.
2. Brunner TB, Nestle U, Grosu A-L, Partridge M. SBRT in pancreatic cancer: What is the therapeutic window? *Radiother Oncol.* 2015;114:109–116.
3. Ng SP, Herman JM. Stereotactic radiotherapy and particle therapy for pancreatic cancer. *Cancers (Basel).* 2018;10:75.
4. Langen KM, Jones DTL. Organ motion and its management. *Int J Radiat Oncol*Biol*Phys.* 2001;50:265–278.
5. Luterstein E, Cao M, Lamb J, et al. Stereotactic MRI-guided Adaptive Radiation Therapy (SMART) for Locally Advanced Pancreatic Cancer: A Promising Approach. *Cureus.* 2018;10:e2324–e2324.
6. Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. *Radiat Oncol (London, England).* 2017;12:28–28.
7. Bootsma GJ, Verhaegen F, Jaffray DA. Efficient scatter distribution estimation and correction in CBCT using concurrent Monte Carlo fitting. *Med Phys.* 2015;42:54–68.
8. Schulze R, Heil U, Gross D, et al. Artifacts in CBCT: a review. *Dentomaxillofac Radiol.* 2011;40:265–273.
9. Cho PS, Johnson RH, Griffin TW. Cone-beam CT for radiotherapy applications. *Phys Med Biol.* 1995;40:1863–1883.
10. Barrett JF, Keat N. Artifacts in CT: recognition and avoidance. *RadioGraphics.* 2004;24:1679–1691.
11. Abe T, Tateoka K, Saito Y, et al. Method for converting cone-beam CT values into Hounsfield units for radiation treatment planning. *Int J Med Phys, Clin Eng Rad Oncol.* 2017;6:361.
12. Lai C-J, Chen L, Zhang H, et al. Reduction in x-ray scatter and radiation dose for volume-of-interest (VOI) cone-beam breast CT—a phantom study. *Phys Med Biol.* 2009;54:6691–6709.
13. Schmidt TG, Fahrig R, Pelc NJ, Solomon EG. An inverse-geometry volumetric CT system with a large-area scanned source: a feasibility study. *Med Phys.* 2004;31:2623–2627.
14. Vogtmeier G, Dorschied R, Engel KJ, et al. *Two-dimensional anti-scatter grids for computed tomography detectors.* Vol 6913: SPIE; 2008.
15. Park Y-K, Sharp GC, Phillips J, Winey BA. Proton dose calculation on scatter-corrected CBCT image: feasibility study for adaptive proton therapy. *Med Phys.* 2015;42:4449–4459.
16. Dong X, Lei Y, Tian S, et al. Air, bone and soft-tissue Segmentation on 3D brain MRI Using Semantic Classification Random Forest with Auto-Context Model; 2019.
17. Dong X, Lei Y, Wang T, et al. Automatic multiorgan segmentation in thorax CT images using U-net-GAN. *Med Phys.* 2019;46:2157–2168.
18. Lei Y, Dong X, Tian Z, et al. CT prostate segmentation based on synthetic MRI-aided deep attention fully convolution network. *Med Phys.* 2019;47:530–540.
19. Lei Y, Liu Y, Dong X, et al. Automatic multi-organ segmentation in thorax CT images using U-Net-GAN. Paper presented at: Medical Imaging; 2019.
20. Wang T, Lei Y, Tian S, et al. Learning-based automatic segmentation of arteriovenous malformations on contrast CT images in brain stereotactic radiosurgery. *Med Phys.* 2019;46:3133–3141.
21. Liu Y, Lei Y, Wang Y, et al. Evaluation of a deep learning-based pelvic synthetic CT generation technique for MRI-based prostate proton treatment planning. *Phys Med Biol.* 2019;64:205022.
22. Wang B, Lei Y, Tian S, et al. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. *Med Phys.* 2019;46:1707–1718.
23. Lei Y, Wang T, Tian S, et al. Male pelvic multi-organ segmentation aided by CBCT-based synthetic MRI. *Phys Med Biol.* 2019;65:035013.

24. Lei Y, Wang T, Wang B, et al. Ultrasound prostate segmentation based on 3D V-Net with deep supervision. Paper presented at: Medical Imaging; 2019.
25. Wang T, Lei Y, Tian Z, et al. Deep learning-based image quality improvement for low-dose computed tomography simulation in radiation therapy. *J Med Imaging* (Bellingham). 2019;6:043504.
26. Wang T, Lei Y, Tang H, et al. A learning-based automatic segmentation and quantification method on left ventricle in gated myocardial perfusion SPECT imaging: a feasibility study. *J Nucl Cardiol*. 2019. <https://doi.org/10.1007/s12350-019-01594-2> (in press).
27. Lei Y, Wang T, Harms J, et al. Image quality improvement in cone-beam CT using deep learning. Paper presented at: Medical Imaging; 2019.
28. Hansen DC, Landry G, Kamp F, et al. ScatterNet: a convolutional neural network for cone-beam CT intensity correction. *Med Phys*. 2018;45:4916–4926.
29. Nomura Y, Xu Q, Shirato H, Shimizu S, Xing L. Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network. *Med Phys*. 2019;46:3142–3155.
30. Maier J, Berker Y, Sawall S, Kachelrieß M. *Deep scatter estimation (DSE): feasibility of using a deep convolutional neural network for real-time x-ray scatter prediction in cone-beam CT*. Vol 10573: SPIE; 2018.
31. Maier J, Eulig E, Vöth T, et al. Real-time scatter estimation for medical CT using the deep scatter estimation: Method and robustness analysis with respect to different anatomies, dose levels, tube voltages, and data truncation. *Med Phys* 2019;46:238–249.
32. Kida S, Nakamoto T, Nakano M, et al. Cone beam computed tomography image quality improvement using a deep convolutional neural network. *Cureus*. 2018;10:e2548–e2548.
33. 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;64:125002.
34. 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;64:145010.
35. Lei Y, Tang X, Higgins K, et al. Learning-based CBCT correction using alternating random forest based on auto-context model. *Med Phys*. 2019;46:601–618.
36. Harms J, Lei Y, Wang T, et al. Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography. *Med Phys*. 2019;46:3998–4009.
37. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. Paper presented at: Proceedings of the IEEE international conference on computer vision; 2017.
38. Lei Y, Harms J, Wang T, et al. MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks. *Med Phys*. 2019;46:3565–3581.
39. Liu Y, Lei Y, Wang Y, et al. MRI-based treatment planning for proton radiotherapy: dosimetric validation of a deep learning-based liver synthetic CT generation method. *Phys Med Biol*. 2019;64:145015.
40. Wolterink JM, Dinkla AM, Savenije MH, Seevinck PR, van den Berg CA, Işgum I. Deep MR to CT synthesis using unpaired data. Paper presented at: International Workshop on Simulation and Synthesis in Medical Imaging; 2017.
41. Hiasa Y, Otake Y, Takao M, et al. Cross-Modality Image Synthesis from Unpaired Data Using CycleGAN. Paper presented at: Simulation and Synthesis in Medical Imaging; 2018//, 2018; Cham.
42. Wang C, Macnaught G, Papanastasiou G, MacGillivray T, Newby D. Unsupervised learning for cross-domain medical image synthesis using deformation invariant cycle consistency networks. Paper presented at: International Workshop on Simulation and Synthesis in Medical Imaging; 2018.
43. Liu Y, Khosravan N, Liu Y, et al. Cross-Modality Knowledge Transfer for Prostate Segmentation from CT Scans. Paper presented at: Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data; 2019//, 2019; Cham.
44. Dong X, Lei Y, Tian S, et al. Synthetic MRI-aided multi-organ segmentation on male pelvic CT using cycle consistent deep attention network. *Radiother Oncol*. 2019;141:192–199.
45. Oktay O, Schlemper J, Folgoc LL, et al. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:180403999. 2018.
46. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Paper presented at: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
47. Ki K, Oda M, Kitasaka T, et al. Multi-atlas pancreas segmentation: atlas selection based on vessel structure. *Med Image Anal*. 2017;39:18–28.
48. Wolz R, Chu C, Misawa K, Fujiwara M, Mori K, Rueckert D. Automated abdominal multi-organ segmentation with subject-specific atlas generation. *IEEE Trans Med Imaging*. 2013;32:1723–1730.
49. Wang Y, Zhou Y, Shen W, Park S, Fishman EK, Yuille AL. Abdominal multi-organ segmentation with organ-attention networks and statistical fusion. *Med Image Anal*. 2019;55:88–102.
50. Okada T, Linguraru MG, Hori M, Summers RM, Tomiyama N, Sato Y. Abdominal multi-organ segmentation from CT images using conditional shape–location and unsupervised intensity priors. *Med Image Anal*. 2015;26:1–18.
51. Tian Z, Li Y, Hassan-Rezaeian N, Jiang SB, Jia X. Moving GPU-OpenCL-based Monte Carlo dose calculation toward clinical use: automatic beam commissioning and source sampling for treatment plan dose calculation. *J Appl Clin Med Phys*. 2017;18:69–84.
52. Chen X, Men K, Li Y, Yi J, Dai J. A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning. *Med Phys*. 2019;46:56–64.
53. Liu Z, Fan J, Li M, et al. A deep-learning method for prediction of three-dimensional dose distribution of helical tomotherapy. *Med Phys*. 2019;46:1972–1983.