

CT synthesis from MRI using multi-cycle GAN for head-and-neck radiation therapy



Yanxia Liu^a, Anni Chen^a, Hongyu Shi^a, Sijuan Huang^b, Wanjia Zheng^{b,c}, Zhiqiang Liu^a, Qin Zhang^{d,*}, Xin Yang^{b,*}

^a School of Software Engineering, South China University of Technology, Guangzhou, Guangdong, 510006, China

^b Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangzhou, Guangdong, 510060, China

^c Air Force Hospital of Southern Theater of the Chinese People's Liberation Army, Guangzhou, Guangzhou, 510507, China

^d School of Computer Science and Engineering, South China University of Technology, Guangzhou, 510006, China

ARTICLE INFO

Keywords:
Deep learning
MRI
CT synthesis
Cycle GAN

ABSTRACT

Magnetic Resonance Imaging (MRI) guided Radiation Therapy is a hot topic in the current studies of radiotherapy planning, which requires using MRI to generate synthetic Computed Tomography (sCT). Despite recent progress in image-to-image translation, it remains challenging to apply such techniques to generate high-quality medical images. This paper proposes a novel framework named Multi-Cycle GAN, which uses the Pseudo-Cycle Consistent module to control the consistency of generation and the domain control module to provide additional identical constraints. Besides, we design a new generator named Z-Net to improve the accuracy of anatomy details. Extensive experiments show that Multi-Cycle GAN outperforms state-of-the-art CT synthesis methods such as Cycle GAN, which improves MAE to 0.0416, ME to 0.0340, PSNR to 39.1053.

1. Introduction

Computed Tomography (CT) plays a crucial role in the planning of Radiation Therapy (RT), as dose calculation required for the RT plan is directly related to the tissue electron density information obtained from the CT (Han, 2017). However, CT lacks good soft-tissue contrast, especially in the brain, head, neck, and pelvic regions, making it extremely difficult to directly delineate endangering organs and target areas on CT (Rasch et al., 2005). In addition, multiple CT scans expose patients to additional ionizing radiation repeatedly, which brought potentially harmful effects to patients (Han, 2017).

Magnetic Resonance Imaging (MRI) is another modality used for auxiliary diagnosis and treatment. Compared with CT, the structured imaging of MRI has made significant progress in soft-tissue contrast and organ visualization. Besides, MRI-guided Radiation Therapy (MRIGRT) is safer since the patients would not be exposed to ionizing radiation in the image scanning process. It also reduces the cost of treatment because only one set of images is required for scanning (Nyholm and Jonsson, 2014; Owrangi et al., 2018). However, due to the different physical imaging principles of MRI and CT, conventional MRI sequences cannot obtain the patient's bone information (Han, 2017). Besides, MRI cannot

directly provide electron density information (Arabi et al., 2018), making a tremendous challenge for the dose calculation in MRIGRT.

Realizing MRIGRT is very difficult but significantly dedicated to estimating CT values from existing MRI for dose calculation in the radiotherapy planning process. Existing methods can be roughly divided into three categories (Arabi et al., 2018):

(a) **Bulk density assignment methods.** In these methods, the tissue in MRI is carefully segmented, and then a predefined uniform density is set for each segmented area to obtain synthetic CT (sCT). The tissue categories include water, fat, air, and bone (Hsu et al., 2013; Chin et al., 2014; Zheng et al., 2015). However, these methods rely on accurate segmentation results, take a long time, and lack details of heterogeneity within the organization.

(b) **Atlas-based methods.** The atlas mentioned here refers to a manually marked database. These methods first register the MRI in Atlas with the real MRI to obtain the deformation field and then use the deformation field to register the CT in Atlas with the real MRI to obtain sCT corresponding to real MRI (Burgos et al., 2014; Arabi and Zaidi, 2016; Andreasen et al., 2016). However, these methods demand accurate image registration. The maturity of the registration technology will directly affect the reliability of the image translation.

* Corresponding authors.

E-mail addresses: cszhangq@scut.edu.cn (Q. Zhang), yangxin@sysucc.org.cn (X. Yang).

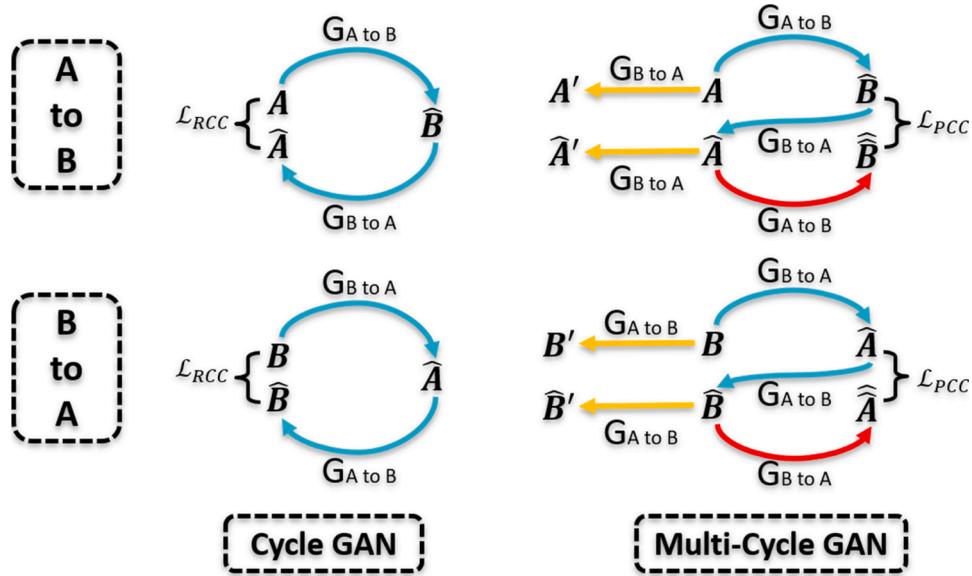


Fig. 1. The difference in the structure between Cycle GAN and Multi-Cycle GAN.

(c) **Learning-based methods.** These methods use self-learning and self-optimizing strategies to find a nonlinear mapping mechanism and estimate the electron density of tissue directly according to the MRI. Recently, many efficient machine learning algorithms have been applied to the generation of sCT, such as Random Forest (Andreasen et al., 2016; Huynh et al., 2016) and Convolutional Neural Networks (CNN) (Han, 2017; Leynes et al., 2017). These algorithms have shown excellent capabilities indirectly generating sCT from MRI.

There has been a growing interest in introducing deep learning to generate CT from MRI recently. The network structure used in the early years is simple, but its performance was much better than the traditional sCT generation methods. Han proposed using 2D Deep Convolutional Neural Networks (DCNN) to generate sCT (Han, 2017), the results were improved compared with traditional atlas-based methods. Nie et al. (Nie et al., 2016) proposed using 3D fully convolutional networks of 3 layers to generate sCT from MRI, which achieved better results than atlas-based methods and Structured Random Forest (SRF) (Huynh et al., 2016).

CT synthesis from MRI can be seen as a style transfer problem from a macro point of view. In 2014, Goodfellow et al. proposed a novel and effective network named Generative Adversarial Nets (GANs) (Goodfellow et al., 2014) to solve the style transfer problem in the field of Computer Version. GANs optimize the generative model by an adversarial step, which trains two models simultaneously: one is the generative model (generators; G) to obtain the data distribution, and the other is the discriminative model (discriminators; D) to estimate the authenticity of the sample. The purpose of G is to maximize the probability of D making mistakes.

Recently, GANs have gained significant attention in the medical field, especially for image-to-image translation tasks. Nie et al. (Nie et al., 2017, 2018) used GANs to apply style transfer to medical images. They first co-registered and normalized the MRI and CT obtained from patients' heads, necks, and pelvis. And then, they took these images as the inputs of GANs during the training phase. Comparing with atlas-based methods and SRF, they improved the results significantly.

Afterward, prior knowledge has been further exploited in image-to-image translation tasks. In order to take full advantage of additional auxiliary conditions, the Conditional GAN is proposed, which was derived from GANs and can be supervised or weakly supervised. In 2019, Zijlstra et al. (2019) proposed a supervised conditional GAN to transform the MRI intensity of the arm, which brought a positive effect.

Most existing methods apply registered images belonging to two or

more domains to train. However, it is not easy to obtain two images with completely identical content in the real world but with different styles. Although GANs can be used for unsupervised image generation, their generation performance is poor, and they cannot preserve details during the conversion process. To solve this problem, in 2017, Zhu et al. proposed Cycle GAN (Zhu et al., 2017), which was improved based on GANs. Cycle GAN is essentially two mirror-symmetric GANs, forming a ring network. Two GANs share two generators, and each GAN has a discriminator, so there are two generators and two discriminators in total. The traditional GAN is generated unidirectionally while Cycle GAN is generated mutually; thus, the images of the two domains can be transferred to each other and do not need to be registered. In the same year, Yi et al. (2017) proposed Dual GAN, which was similar to Cycle GAN, achieving comparable performance with Cycle GAN. Recently, some researchers introduced the Cycle GAN into the medical community to generate sCT. Wolterink et al. (2017) utilized Cycle GAN, with unregistered MRI and CT images, to successfully generate higher quality head and neck sCT. Yang et al. (2018) also verified the excellent performance of Cycle GAN on unregistered MRI and CT datasets. Hiasa et al. (2018) modified the loss function of Cycle GAN to improve the accuracy of the boundary by adding Gradient Consistency (GC) loss (Penney et al., 1998) and this method achieves good results in MRI-to-CT synthesis of the pelvic region. Zeng and Zheng (2019) introduced a weak supervision mechanism when using unregistered images. This method first inputs the MRI into the 2D Cycle GAN to obtain the corresponding sCT (weak label), then fed the real MRI into the 3D generator to generate sCT, finally subtracted the sCT generated by the 2D and 3D generator as a part of the loss. Because Cycle GAN has no direct constraints between the input image and synthetic image, it does not guarantee structural consistency between these two images. Yang et al. (2020) defined an extra structure-consistency loss based on the modality independent neighborhood descriptor. This method produces better synthetic CT images in both accuracy and visual quality compared to Cycle GAN.

The purpose of this paper is to use MRI to generate synthetic CT (sCT). Some image translation algorithms, such as DCNN, GANs, Conditional GAN, and Cycle GAN, have been applied to this field recently. However, they failed to obtain high-precision results and retain anatomy details simultaneously. Therefore, we propose a novel framework named Multi-Cycle GAN based on Cycle GAN. Specifically, we introduce Pseudo-Cycle Consistent module to control the consistency of generation, and add a domain control module to provide additional identical constraints. The former makes full use of the intermediate output and

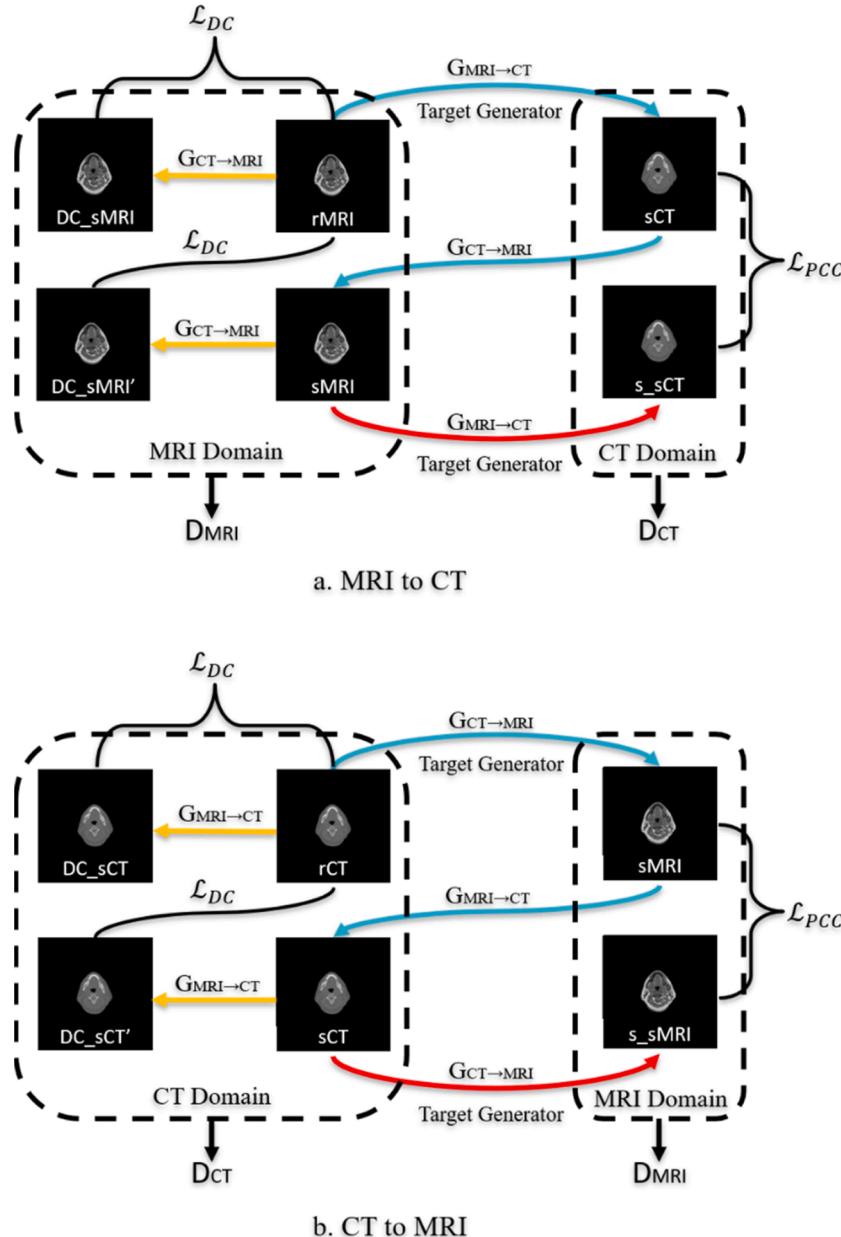


Fig. 2. The concrete realization of Multi-Cycle GAN. The left part is sCT generated by rMRI, and the right part is sMRI generated by rCT.

adds the difference between the target object generated by the first target generator and the object generated by the second target generator to the loss to enhance training, thereby improving the generation effect of the target generator. The latter uses the distance between the input and the output to prevent images in the expected domain from being incorrectly translated to the other domain. Furthermore, we propose a new generator named Z-Net, which adds regular skip connections between the traditional convolutional layers and dramatically improves the accuracy of detail conversion. Extensive experiments show that our method outperforms state-of-the-art CT synthesis methods such as Cycle GAN, which improves MAE to 0.0416, ME to 0.0340, PSNR to 39.1053.

2. Methods

2.1. Multi-Cycle GAN

The proposed Multi-Cycle GAN is an image synthesis model based on Cycle GAN. Compared to Cycle GAN, Multi-Cycle GAN has two

innovations, (i) adding Pseudo-Cycle Consistent module (the red line on the right of Fig. 1) and (ii) introducing domain control module (the yellow line on the right of Fig. 1). The architecture significantly improves the performance and robustness of generators in adversarial learning.

2.1.1. Pseudo-Cycle Consistent module

Real-Cycle Consistent Loss (RCCL) is the loss used by Cycle GAN, which prevents the learned mapping $G_{A \rightarrow B}$ and $G_{B \rightarrow A}$ from contradicting each other. Taking A to B (Fig. 1) as an example, RCCL ensures the image quality of the \hat{A} . However, only using RCCL (single cycle) has no constraint conditions to control the generation effect of $G_{A \rightarrow B}$ in the whole cycle. As a result, \hat{B} fails to get consistency constraints and does not make full use of the intermediate results (\hat{B}) during training. Therefore, we introduce the Pseudo-Cycle Consistent module (PCCM) into Cycle GAN (the red line on the right of Fig. 1), which adds a target generator for each cycle and Pseudo-Cycle Consistent Loss (PCCL) as the consistency constraint for the target generator. After adding PCCL, \hat{B} can

make a difference with \hat{B} to control the generation consistency of $G_A \rightarrow B$ and improve the generation effect of the target generator. Moreover, since \hat{A} is the pseudo data generated by circulation, its imaging effect would be reduced compared with A , so this operation is equivalent to adding noise to the input data of the generator, enhancing the data, and improving the robustness of the generator.

Our goal is to learn mapping function between MRI and CT, so we set two domains: $data_{MRI}$ and $data_{CT}$, which are respectively denoted by X and Y , where $x_i \in X$ represents the $data_{MRI}$ domain, and $y_i \in Y$ represents the $data_{CT}$ domain. And we use $x \sim p_{data}(x)$ and $y \sim p_{data}(y)$ to represent the data distribution of MRI and CT respectively. To achieve the style transfer from MRI to CT, we use two generators $G_{MRI \rightarrow CT} : MRI \rightarrow CT$ and $G_{CT \rightarrow MRI} : CT \rightarrow MRI$ simultaneously. In addition, we use two discriminators D_{MRI} and D_{CT} , where D_{MRI} is to judge the authenticity of x and $G_{CT \rightarrow MRI}(y)$, and D_{CT} is to judge the authenticity of y and $G_{MRI \rightarrow CT}(x)$.

From Fig. 2, we can also determine that each independent cycle structure (Fig. 2a and Fig. 2b) has its own training goal. Taking Fig. 2a as an example, the input is real MRI (rMRI), and the goal is getting the realistic-looking synthetic CT (sCT), so $G_{MRI \rightarrow CT}$ is the target generator of this part. Similarly, the target generator of Fig. 2b is $G_{CT \rightarrow MRI}$.

We use Real-Cycle Consistent Loss (\mathcal{L}_{RCC}) and Pseudo-Cycle Consistent Loss (\mathcal{L}_{PCC}) to control the consistency of generator output. \mathcal{L}_{RCC} is proposed by Cycle GAN, which uses the difference between the synthetic image generated by a single cycle and the real image to ensure the consistency of the cycle of $G_{MRI \rightarrow CT}$ and $G_{CT \rightarrow MRI}$. The purpose of \mathcal{L}_{RCC} is to make $G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(x)) \approx x$. Similarly, we need to satisfy backward cycle consistency: $G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(y)) \approx y$. The specific equation is as follows:

$$\begin{aligned} \mathcal{L}_{RCC} = & \mathbb{E}_{x \sim p_{data}(x)} [\|G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(x)) - x\|_1] \\ & + \mathbb{E}_{y \sim p_{data}(y)} [\|G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(y)) - y\|_1] \end{aligned} \quad (1)$$

$$\|z\|_1 = \sum_i^n |z_i| \quad (2)$$

where \mathbb{E} stands for Expectation, and L1 norm (Eq. (2)) is used to calculate the difference between the output and the real data.

\mathcal{L}_{PCC} adds a target generator to generate target domain objects based on the single cycle. Then, the translation effect of the target generator is enhanced by comparing the target domain objects generated by the two target generators. When \mathcal{L}_{PCC} is not included, the sCT generated by $G_{MRI \rightarrow CT}$ is not constrained by the loss of consistency. Thus the quality of $G_{MRI \rightarrow CT}$ generation cannot be controlled. Therefore, we put forward the concept of Multi-cycle, which adds a target generator based on Real-Cycle, and further strengthens the target generator. When generating CT from MRI, \mathcal{L}_{PCC} ensures $G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(x))) \approx G_{MRI \rightarrow CT}(x)$, while generating MRI from CT, $G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(y))) \approx G_{CT \rightarrow MRI}(y)$ is also satisfied. The equation is described as follows:

$$\begin{aligned} \mathcal{L}_{PCC} = & \mathbb{E}_{x \sim p_{data}(x)} [\|G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(x))) - G_{MRI \rightarrow CT}(x)\|_1] \\ & + \mathbb{E}_{y \sim p_{data}(y)} [\|G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(y))) - G_{CT \rightarrow MRI}(y)\|_1] \end{aligned} \quad (3)$$

The purpose of the generator is to generate objects sufficient to deceive the discriminator, and the goal of the discriminator is to accurately distinguish the authenticity of the image. \mathcal{L}_{RCC} and \mathcal{L}_{PCC} are the losses set for the generation effect of the generator, without any constraint on the discriminator. Therefore, it needs to be introduced adversarial loss to achieve the adversarial training of the entire model. The realization is described as follows:

$$\begin{aligned} \mathcal{L}_{adv_G} = & \mathbb{E}_{x \sim p_{data}(x)} [(1 - D_{ct}(G_{MRI \rightarrow CT}(x)))^2] \\ & + \mathbb{E}_{x \sim p_{data}(x)} [(1 - D_{ct}(G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(x))))^2] \end{aligned}$$

$$\begin{aligned} & + \mathbb{E}_{y \sim p_{data}(y)} [(1 - D_{MRI}(G_{CT \rightarrow MRI}(y)))^2] \\ & + \mathbb{E}_{y \sim p_{data}(y)} [(1 - D_{MRI}(G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(y))))^2] \end{aligned} \quad (4)$$

$$\begin{aligned} \mathcal{L}_{adv_D} = & \mathbb{E}_{x \sim p_{data}(x)} [(D_{ct}(G_{MRI \rightarrow CT}(x)))^2] + \mathbb{E}_{y \sim p_{data}(y)} [(1 - D_{ct}(y))^2] \\ & + \mathbb{E}_{y \sim p_{data}(y)} [(D_{MRI}(G_{CT \rightarrow MRI}(y)))^2] + \mathbb{E}_{x \sim p_{data}(x)} [(1 - D_{MRI}(x))^2] \end{aligned} \quad (5)$$

where \mathcal{L}_{adv_G} (Eq. (4)) works on the generator in cycles, \mathcal{L}_{adv_D} (Eq. (5)) works on the discriminator in cycles. The goal of D_{MRI} is to distinguish between real MRI and synthetic MRI. While the output value tends to 1, it means that the generator has a better generation effect and achieves the cycle consistency, and the image is considered real. On the contrary, if the output value tends to 0, it is considered fake. Similarly, the goal of D_{CT} is to distinguish between real CT and synthetic CT.

Compared with Cycle GAN, the adversarial losses we used is consist of the discrimination of synthetic sCT (s_sCT) and synthetic sMRI (s_sMRI) (Fig. 2) while removing the discrimination terms of real objects ($D_{ct}(y)$ and $D_{MRI}(x)$). This change aims to deceive the discriminator to the greatest extent without increasing the amount of calculation to improve the performance of the generator.

2.1.2. Domain control module

The Domain Control module (DCM) controls the domain invariance of the synthetic image (Xu et al., 2020), and we use it to provide additional identical constraints. Two target generators are added (the yellow line on the right of Fig. 1) into our framework, and the input and output data fields of these generators belong to the same domain, which is different from the generators in the cycles. The purpose of using different inputs into the same generator is to prevent images already in the expected domain from being incorrectly translated to the other domain. The Domain Control module ensures the domain correctness of the image synthesis and improves the stability of the generator.

The loss function of the domain control module can be divided into two parts: Domain Control Loss (\mathcal{L}_{DC}) and Domain Control adversarial loss (\mathcal{L}_{DC_adv}). Take Fig. 2a as an example. Using \mathcal{L}_{DC} make sure $G_{CT \rightarrow MRI}(x) \approx x$ and $G_{CT \rightarrow MRI}(G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(x))) \approx x$. Similarly, when the input data is rCT (Fig. 2b), it also satisfies $G_{MRI \rightarrow CT}(y) \approx y$ and $G_{MRI \rightarrow CT}(G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(y))) \approx y$. The specific equation of \mathcal{L}_{DC} is shown in Eq. (6).

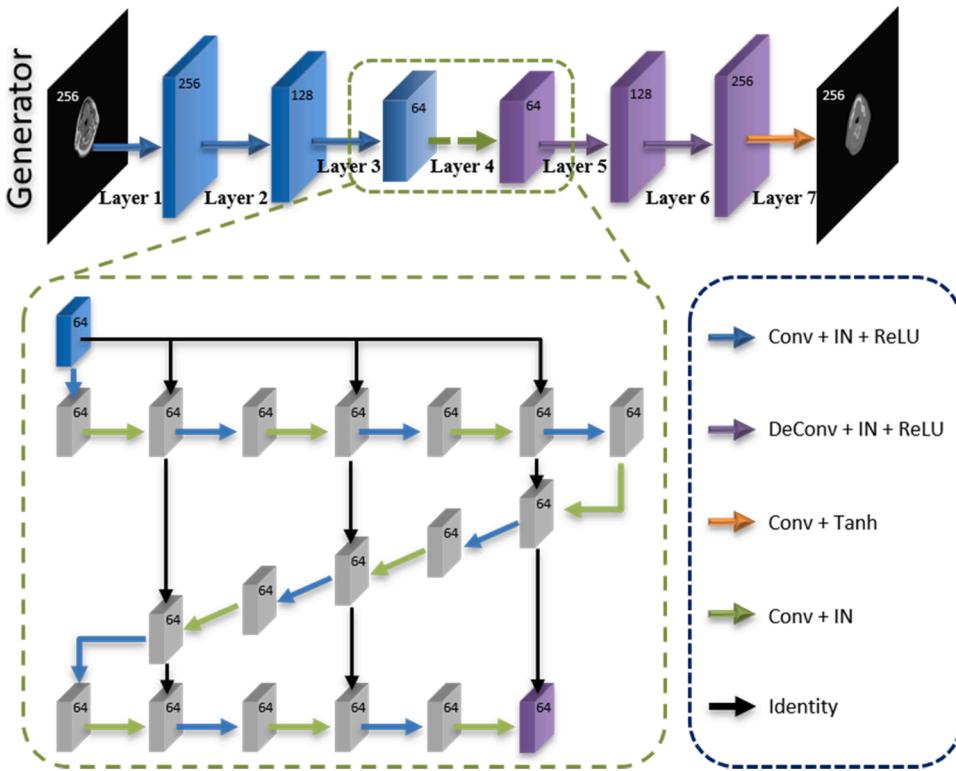
$$\begin{aligned} \mathcal{L}_{DC} = & \mathbb{E}_{x \sim p_{data}(x)} [\|G_{CT \rightarrow MRI}(x) - x\|_1] + \mathbb{E}_{y \sim p_{data}(y)} [\|G_{MRI \rightarrow CT}(y) - y\|_1] \\ & + \mathbb{E}_{x \sim p_{data}(x)} [\|G_{CT \rightarrow MRI}(G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(x))) - x\|_1] \\ & + \mathbb{E}_{y \sim p_{data}(y)} [\|G_{MRI \rightarrow CT}(G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(y))) - y\|_1] \end{aligned} \quad (6)$$

Although the loss function involves discriminator terms, it is only used for the optimization of the generator and not for the optimization of the discriminator. \mathcal{L}_{DC_adv} is to distinguish between the real images and the synthetic ones after adding the domain control module. The detail is described by Eq. (7).

$$\begin{aligned} \mathcal{L}_{DC_adv} = & \mathbb{E}_{x \sim p_{data}(x)} [(1 - D_{MRI}(G_{CT \rightarrow MRI}(x)))^2] \\ & + \mathbb{E}_{x \sim p_{data}(x)} [(1 - D_{MRI}(G_{CT \rightarrow MRI}(G_{CT \rightarrow MRI}(G_{MRI \rightarrow CT}(x))))^2] \\ & + \mathbb{E}_{y \sim p_{data}(y)} [(1 - D_{CT}(G_{MRI \rightarrow CT}(y)))^2] \\ & + \mathbb{E}_{y \sim p_{data}(y)} [(1 - D_{CT}(G_{MRI \rightarrow CT}(G_{MRI \rightarrow CT}(G_{CT \rightarrow MRI}(y))))^2] \end{aligned} \quad (7)$$

2.2. Z-Net Generator

Existing Cycle GAN-based medical image generation models typically use a simple series of convolutional layers (Nie et al., 2016) or ResNet (He et al., 2016) as generators, which do not consider the detail



retention of the synthetic images. To address this problem, we propose a novel generator named Z-Net (shown in Fig. 3). Z-Net is composed of three parts, including down-sampling layers (Layer 1~Layer 3), feature extraction module (Layer 4), and up-sampling layers (Layer 5~Layer 7), with the operations of Conv, Instance Normalization (IN) (Ulyanov et al., 2016) and Rectified Linear Unit (ReLU) activation function (Glorot et al., 2011). Low-level layers usually detect some fundamental patterns (e.g., edge and corner), which might be helpful for the detailed conversion of images. Thus, we design a novel feature extraction module, which attaches nine skip connections to different positions. In this way, we effectively extract low-level image features and fuse them to high-level features, which significantly improves the transformation of small objects.

Z-Net not only improves the accuracy of the small object generation but also has robust scalability. It can be widely applied to the generator proposed in this paper and can be applied to different style transfer or image synthesis models for better performance, as explained in section 3.2.

2.3. Full objective

Multi-Cycle GAN is composed of multiple generators and discriminators (the detail of discriminator can be found in Discriminator Architecture of the Appendix), and it involves two parts of loss. One is the Consistent Loss, and the other is the Adversarial Loss. In summary, the total loss of the generator is as follows:

$$\mathcal{L}_{\text{Generator}}(G, D, X, Y) = \alpha \mathcal{L}_{\text{RCC}} + \beta \mathcal{L}_{\text{PCC}} + \delta \mathcal{L}_{\text{DC}} + \gamma \mathcal{L}_{\text{adv_G}} + \varepsilon \mathcal{L}_{\text{DC_adv}} \quad (8)$$

where $\alpha, \beta, \delta, \gamma$ and ε are adjustable parameters to adjust the proportion of each part of the loss in the training process.

$\mathcal{L}_{\text{adv_D}}$ is the loss of training discriminator, expressed as follows:

$$\mathcal{L}_{\text{discriminator}}(G, D, X, Y) = \mathcal{L}_{\text{adv_D}} \quad (9)$$

The adversarial learning between the generator and the

Fig. 3. The architecture of the Z-Net. The generator is composed of 7 layers, where the operations represented by the blue line is Conv, IN, and ReLU, the purple line is DeConv, IN, and ReLU, the yellow line is Conv and tanh. The specific structure in the green dashed box is given below the figure, where the solid green lines indicate Conv, IN, and the black line represents the element-wise addition of two matrices (for interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

discriminator helps to generate high-quality objects.

3. Experiments and results

3.1. Data collection and preprocessing

Both CT and MR images for each patient were acquired on the same day with the same setup procedure (patient-specific polyurethane foam immobilization devices and head&neck&shoulder immobilization mask). The data of each patient include CT images and MR images, as well as the radiotherapy plan. The institutional review board approved our study and the ethical review office from the institution, and the data had been submitted to a public Research Data Deposit (RDD) platform (www.researchdata.org.cn), with approval RDD number as RDAD2021001910. See data information in the Appendix for details.

Since the MRI and CT data we collected have no one-to-one correspondence in quantity and spatial location, we need to perform coarse registration operations before training. Because the medical images are sampled from the patient's head and neck, and the probability of tissue deformation in this area is tiny, the Mutual Information Rigid Registration algorithm (Han, 2017) can be used to perform rigid registration on each patient's unpaired MRI and CT images.

After that, we use the N4 Bias Field Correction algorithm (Sled et al., 1998) to correct each MRI slice to solve varying intensity inside the MRI slice. And then, the Histogram Matched method proposed by Cox et al. (1995) is applied to match all MRIs with randomly selected templates to standardize the image intensity of different patients. Non-human parts such as bed boards and fixed frames would appear on the original CT image. In order to remove these useless structures, we need to obtain the mask from MRI and multiply the mask of MRI to denoise the CT image. The specific data preprocessing flow is shown in Fig. 4. And then, resample the image size from 512 * 512 to 256 * 256 and perform the normalization operation before training to reduce the amount of calculation. Finally, the preprocessed data can be input into the network for training.

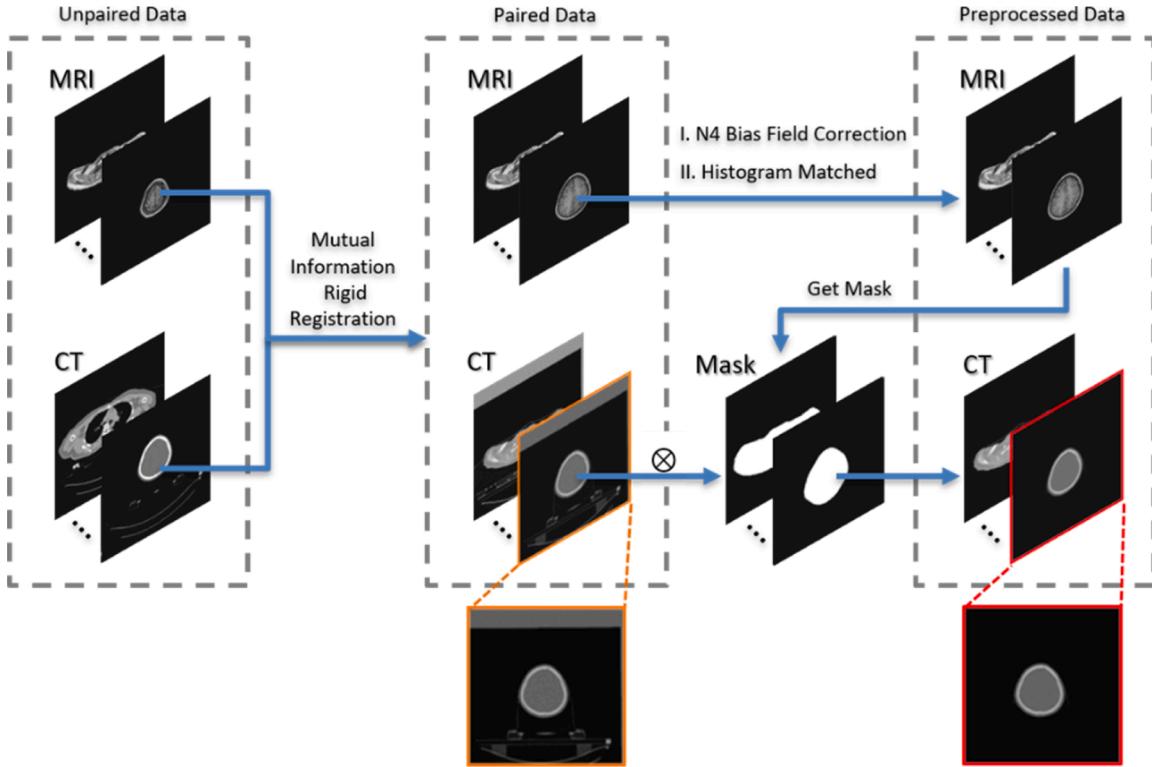


Fig. 4. Data preprocessing. The MRI and CT of the same patient are registered by Mutual Information Rigid Registration algorithm, then N4 Bias Field Correction and Histogram Matched are used to adjust the intensity to get the final MRI. CT and the mask obtained from MRI are multiplied to obtain a denoised CT.

Table 1
The results for different methods.

Method	ME	MAE	PSNR
DCNN	0.4155	1.5630	29.4489
GANs	0.0491	0.0700	31.8132
Conditional GAN	0.0480	0.0527	35.3122
Cycle GAN	0.0409	0.0465	37.1023
Ours	0.0340	0.0416	39.1053

Table 2
Ablation study: the results for each module outperforms.

Backbone	Z-Net	Multi-Cycle GAN		ME	MAE	PSNR
		PCCM	DCM			
✓				0.0409	0.0465	37.1023
✓	✓			0.0379	0.0443	37.3005
✓	✓	✓		0.0348	0.0421	38.1984
✓	✓	✓	✓	0.0340	0.0416	39.1053

Table 3
The results of the model with or without Z-Net.

Model	ME	MAE	PSNR
GANs	0.0491	0.0700	31.8132
GANs with Z-Net	0.0429	0.0677	32.2647
Conditional GAN	0.0480	0.0527	35.3122
Conditional GAN with Z-Net	0.0453	0.0501	35.9557
Cycle GAN	0.0409	0.0465	37.1023
Cycle GAN with Z-Net	0.0379	0.0443	37.3005

3.2. Experimental details

We implement our methods in TensorFlow, and all the computation during the training process runs on the GPU. The batch size is one, and the learning rate is linearly decayed from 2×10^{-4} to 2×10^{-6} , and the input and output data sizes are 256*256. The loss weight term in Eq. (8), α , β , γ , δ , and ϵ are set to 30, 30, 1, 10, and 10, respectively. Other compared methods are implemented based on the author-provided codes and the implementation detail can be found in the Appendix of baseline models.

It is necessary to obtain a generator with a clear generation goal in the early training stage and improve its accuracy and stability in the later stage. Since the training purpose of the pre- and post-training period is slightly different, this paper designs a three-stage training method. The epoch from 0_{th} to 20_{th} of training is the first stage, the model uses the structure of Cycle GAN, that is, \mathcal{L}_{RCC} and \mathcal{L}_{adv_D} are used as the loss function of the optimized generator; and from 20_{th} to 40_{th} epoch is the second stage, the model starts to use the structure of Multi-Cycle GAN without domain control module. It uses \mathcal{L}_{RCC} , \mathcal{L}_{PCC} , and \mathcal{L}_{adv_G} as the loss function of the optimized generator until the 40_{th} epoch. And in the last stage (after 40_{th} epoch), we use the Multi-Cycle GAN as the model, and the total loss is used as the loss function of the generator until the end of training.

3.3. Evaluation

We use Mean Absolute Error (MAE), Mean Error (ME), and Peak Signal to Noise Ratio (PSNR) to evaluate the difference between rCT and sCT.

The definition of MAE is:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - G_{MRI-CT}(x_i)| \quad (10)$$

where n is the total number of voxels, y_i is rCT, x_i is rMRI, x_i and y_i are

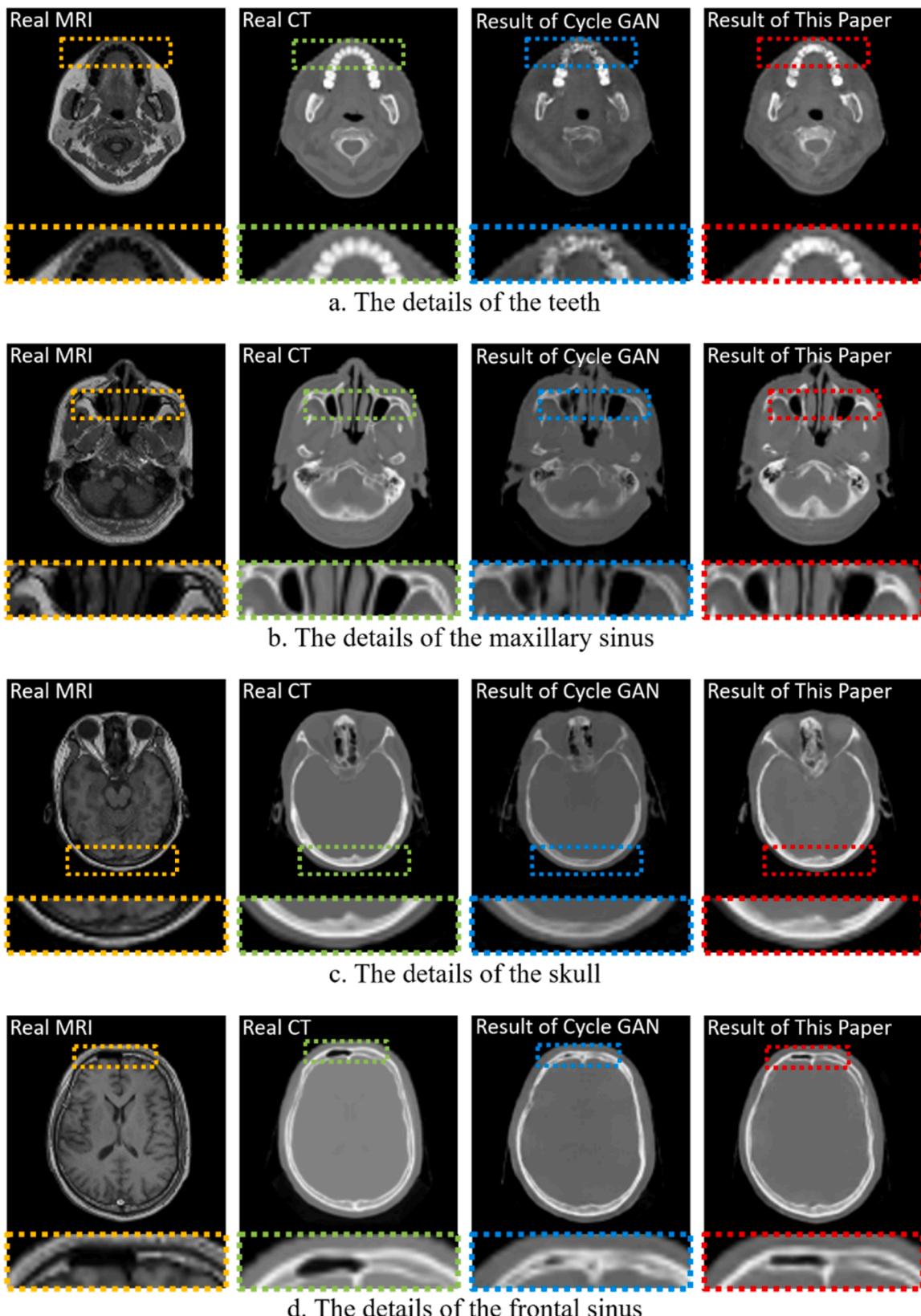


Fig. 5. Experimental results details display. There are four groups: a, b, c, d in the figure, each group from left to right is rMRI, rCT, Cycle GAN result, and our model.

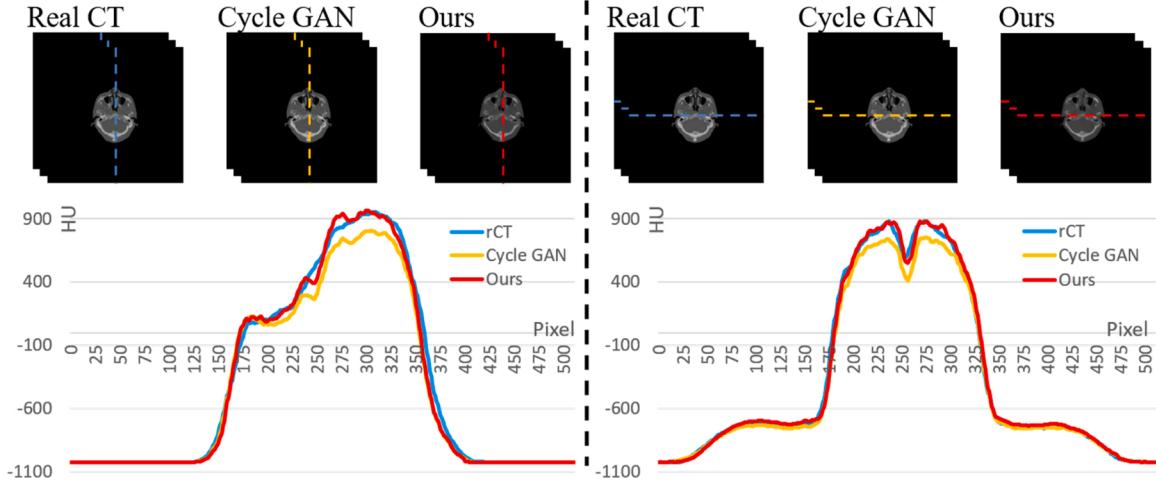


Fig. 6. Comparison chart of HU values. The HU mean values of real CT, Cycle GAN result, and ours are shown in the figure, where a and b refer to longitudinal and transverse comparisons between different methods, respectively.

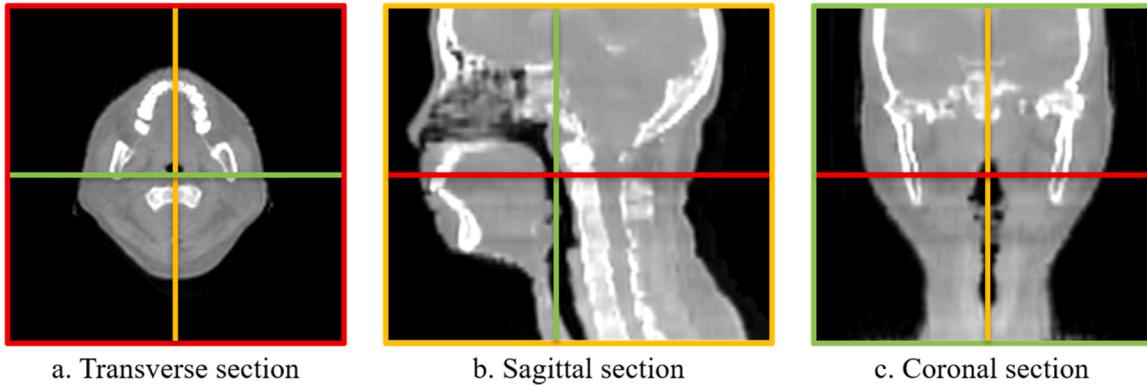


Fig. 7. 3D (three-dimensional) reconstruction of our method.

already registered, the difference between rCT and sCT generated by rMRI through generator $G_{MRI \rightarrow CT}$ is compared. In addition, the definition of ME is:

$$ME = \frac{1}{n} \sum_{i=1}^n (y_i - G_{MRI \rightarrow CT}(x_i))^2 \quad (11)$$

The PSNR is defined as follows:

$$PSNR = 20 \log_{10} \frac{MAX_{rCT}}{MSE} \quad (12)$$

where MAX_{rCT} represents the maximum value of rCT, the MSE equation is as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - G_{MRI \rightarrow CT}(x_i))^2 \quad (13)$$

In the above evaluation metrics, the smaller the value of MAE and ME, the smaller the difference between rCT and sCT, and the larger the PSNR, the higher the similarity between rCT and sCT.

3.4. Result

In order to verify the effectiveness of the method proposed in this paper, we compare with state-of-the-art methods used for medical image synthesis tasks, including DCNN (Han, 2017), GANs (Goodfellow et al., 2014), Conditional GAN (Mirza and Osindero, 2014) and Cycle GAN (Zhu et al., 2017). The results are shown in Table 1, and our method

obtains the best MAE, ME, and PSNR. It can be observed that our Multi-Cycle GAN based method with Z-Net as the generator outperforms Cycle GAN.

To further demonstrate the effectiveness of each module, we now report the ablation studies. As shown in Table 2, all the models are trained with the same training set introduced before. The backbone is Cycle GAN. We can observe that the Z-Net generator can help improve the model's performance, and PCCM and DCM have an outstanding contribution to improving the quality of model generation. Besides, we can note that the comprehensive improvement effect of adding PCCM is more significant than the other two optimizations.

In this experiment, we investigate the effect of the Z-Net generator. We choose three popular CT synthesis models as backbones, i.e., GANs, Conditional GAN, and Cycle GAN. We show the experimental results in Table 3. From this table, we can observe that the performance of different models has improved after using the Z-Net generator, which proves that the Z-Net generator is helpful to improve the accuracy of CT synthesis and has good scalability. Therefore, Z-Net can be used in a variety of CT synthesis models.

We randomly select several organs, including teeth, maxillary sinus, skull, and frontal sinus, and visualize the synthetic results in Fig. 5 for qualitative analysis. The figure includes four groups a-d, and the conversion details of rMRI, rCT, Cycle GAN, and our method are shown from left to right in each group. The first row of each group shows the original images, and the second row is listed as a detailed magnification of the four images. The results of our method are closer to real CT, which shows that our method outperforms Cycle GAN in detail conversion.

Fig. 6 shows the different HU values of the rCT, Cycle GAN result, and ours. All figures have been resampled and converted from the output of 256*256 to 512*512, which is the size of the origin CT. The left of **Fig. 6** is the longitudinal mean of all test data, and the right of **Fig. 6** is the transverse mean, where the blue line represents the HU mean values of the rCT, the yellow line represents Cycle GAN, and the red line represents the sCT HU mean values generated by our model. As can be seen from the line chart, the HU mean values of the sCT generated by our method are closer to the rCT than Cycle GAN, and the trend is the same.

Besides, we also pay attention to the continuity of the sCT 3D (three-dimensional) reconstructed image. **Fig. 7** shows the sCT reconstructed result generated by our method where a, b, and c represent transverse section, sagittal section, and coronal section, respectively. It can be seen that there is continuity between slices.

4. Discussion

In this work, we proposed a Multi-Cycle GAN method for unpair head and neck MRI-to-CT synthesis. The experiment results in **Table 1** demonstrate our superior performance. Especially, our model performs much better than Cycle GAN, where MAE improves from 0.0465 to 0.0416, ME from 0.0409 to 0.0340, PSNR from 37.1023 to 39.1053. Our method uses a simple yet effective module named Pseudo-Cycle Consistent module to control the synthetic consistency within the cycle, improving the utilization of intermediate output and enhancing the effect of the target generator. Therefore, the difference between real CT and synthetic CT is more negligible. Besides, our method uses the domain control module, which guarantees that the output result has the same domain no matter which domain the generator input data is. The Domain Control module ensures the domain correctness of the image synthesis and improves the stability of the generator.

When evaluating the impact of Z-Net in **Table 3**, Z-Net is used to replace the generators in GANs, Conditional GAN, and Cycle GAN, respectively, and it shows that using Z-Net as the generator in different models can improve their performance. This is because Z-Net uses skip connections to combine the low-level features with high-level features. It can prevent the explosion of non-convexity that occurs when networks get deeper and retain more high-resolution details information in high-level features, thereby improving the accuracy of image synthesis.

To examine the effectiveness of each module, we compare the Cycle GAN with each module. As shown in **Table 2**, by replacing the Z-Net with the regular generator or replacing the Cycle GAN to Multi-Cycle GAN, the synthetic accuracy would be improved. Moreover, using the domain control module can also improve the quality of image synthesis because it ensures that the synthetic domain is specific no matter the input domain, which can improve the synthetic stability of the generator.

It is worth mentioning that our method achieves excellent performance without consuming additional computing resources and time. In the training phase, the model we proposed requires about 100 epochs to converge on average, which takes about 140 h on average. In comparison, Cycle GAN requires about 150 epochs to converge on average, and it takes about 145 h on average. In the testing phase, the generation speed of the method we proposed generates one case every 4 s (about 85

slices per case), which is the same speed as Cycle GAN.

The proposed Multi-Cycle GAN has performed well in image translation since it would significantly affect the clinical diagnosis of the doctors. The qualitative results of Cycle GAN and our model are presented in **Fig. 6**. Overall, the detailed structures of synthetic CT generated by Cycle GAN are often lost or smoothed. For example, the teeth and maxillary sinus generated by Cycle GAN are blurred, as shown in **Fig. 5a** and **Fig. 5b**, and the skull and frontal sinus are incomplete, as shown in **Fig. 5c** and **Fig. 5d**. On the contrary, both bones and soft tissues information are complete and clear in the synthetic CT generated by our Multi-Cycle GAN.

5. Conclusion

In this study, we use Multi-Cycle GAN to transfer MRI to sCT and find out that Pseudo-Cycle Consistent module can enhance the performance of the target generator, and the domain control module can improve the stability of the generator and the quality of the synthetic image. Additionally, a new generator named Z-Net is proposed to improve the accuracy of detail conversion. Extensive experiments show that our method outperforms state-of-the-art CT synthesis methods such as Cycle GAN. In future research, we consider adding the dose calculation as an evaluation metric to judge the results and segment each tissue in MRI and sCT to make a more detailed contrast. Besides, we will also use multi-sequence MRI images as complementary inputs to generate sCT for better performance.

Funding

This work was supported by Key Realm R&D Program of Guangzhou (202007030007), Natural Science Foundation of Guangdong Province (2017A030310217) and Pearl River S&T Nova Program of Guangzhou (201710010162).

CRediT authorship contribution statement

Yanxia LIU, Hongyu SHI and Xin YANG: designed the analytical framework of this study and drafted the manuscript. **Anni CHEN, Zhi-qiang LIU and Qin ZHANG:** provided methodological advices and made major revisions of the manuscript. **Sijuan HUANG and Wanjia ZHENG:** assisted collecting and processing data and conducted analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Many thanks to Mr. Zesen CHENG, Mr. Tianyu ZENG, Mr. Yin HONG, Mr. Rui LI, Ms. Wenqi WANG, Ms. Suizhu YANG, Ms. Jiali GUO, for their assistant on this paper.

Appendix A

Data information

The CT simulator in our department, a Brilliance Big Bore CT scanner (Philips Medical Systems, Cleveland, OH, USA), had monthly/quarterly/yearly QA/QC based on the American College of Radiology (ACR) CT accreditation phantom before scanning the patients. Furthermore, these CT simulator images showed that image quality was within ACR guidelines for all tested scanning protocols.

A pre-defined set of scanning protocols was performed on Nasopharyngeal Cancer (NPC) patients using the CT simulator, and the image data were reconstructed with a 3 mm thickness. These images were generated using the more common clinical protocols of 3 mm slices at 3 mm increments with

Table A1
Data distribution.

	Total(n = 164)	Train(n = 117)	Validate (n = 18)	Test(n = 29)
Male	122(74.4 %)	89(76.1 %)	13(72.2 %)	20(69.0 %)
Female	42(25.6 %)	28(23.9 %)	5(27.8 %)	9(31.0 %)
<18y	0(0.0 %)	0(0.0 %)	0(0.0 %)	0(0.0 %)
18–45y	71(43.3 %)	52(44.4 %)	7(38.9 %)	12(41.4 %)
>45y	93(56.7 %)	65(55.6 %)	11(61.1 %)	17(58.6 %)

the scanning protocols most commonly used in clinical practice. The detailed parameters for these protocols were given as follows: Voltage 120kVp, Exposure 300mAs, Slice Thickness 3 mm, Increment 3 mm, Collimation 16×0.75 (mm), Display FOV 600 mm, Scan FOV 600 mm, Reconstruction Filter Type UB/B, and Pitch 0.567.

We retrospectively collected MRI studies of the nasopharynx from patients with histologically proven and radiation therapy-naïve NPC between October 1, 2018, and June 30, 2019, from a single institute (Sun Yat-sen University Cancer Center, SYSUCC).

The MRIs were acquired on a 3.0 T MR simulator (Philips Medical Systems, Ingenia 3.0, Netherlands). Four sequences (T1, T2, T1C, and T1DixonC-water) of MR images were collected in our center to facilitate physicians to diagnose lesions and delineate target and normal organ structures.

The four MR sequences and corresponding scanning parameters are stated as follows:

T1-weighted 3D fast field echo (FFE) sequence (2.4 ms echo time, 4.9 ms repetition time, 12° flip angle, $0.95 \times 0.95 \times 3$ mm 3 acquisition resolution, $0.69 \times 0.69 \times 3$ mm 3 reconstructed voxel, 528×295 acquisition matrix, 720×720 reconstruction matrix, 3.20 min average acquisition duration).

T2-weighted 2D turbo spin-echo (TSE) sequence (110 ms echo time, 5712 ms repetition time, 90° flip angle, $0.90 \times 1.09 \times 3$ mm 3 acquisition resolution, $0.57 \times 0.57 \times 3$ mm 3 reconstructed voxel, 556×256 acquisition matrix, 880×880 reconstruction matrix, 5 min average acquisition duration).

After T1-weighted and T2-weighted images were acquired, gadodiamide was injected intravenously with 0.2 mL/kg, T1C images were then collected with the same parameters as T1.

Afterward, dual-echo Dixon 3D fast field echo (FFE) sequence (1.53 and 2.8 ms echo time, 5.1 ms repetition time, 15° flip angle, $1.03 \times 1.04 \times 3$ mm 3 acquisition resolution, $0.69 \times 0.69 \times 3$ mm 3 reconstructed voxel, 484×273 acquisition matrix, 720×720 reconstruction matrix, and 2.55 min average acquisition duration) was used and to reconstruct T1DixonC-water images.

The final data set is comprised of 164 patients with CT (a total of 17,348 slices) and T1-weighted (a total of 14,882 slices), which are randomly assigned to three cohorts:

- (a) A training cohort of 117 patients for GAN construction;
- (b) A validation cohort of 18 patients for optimization of the GAN hyperparameters;
- (c) A testing cohort of 29 patients to test the performance of the sCT generator tool.

Table A1 shows the gender ratio and the age distribution of the data set.

Discriminator Architecture

Since the discrimination task in this paper is relatively simple and can be regarded as a binary classification task, there is no need to use a highly complex discriminator to occupy a large memory and consume additional computing resources to improve the accuracy of the discrimination. The discriminator consists of a 4-layer network structure, as shown in Fig. A1. The red arrow represents the combined operation of Conv and ReLU, the blue arrow represents the combination of Conv, IN, and ReLU, and the gray arrows represent Conv operations. Furthermore, the number in the upper left corner of each image or feature in the figure represents its size.

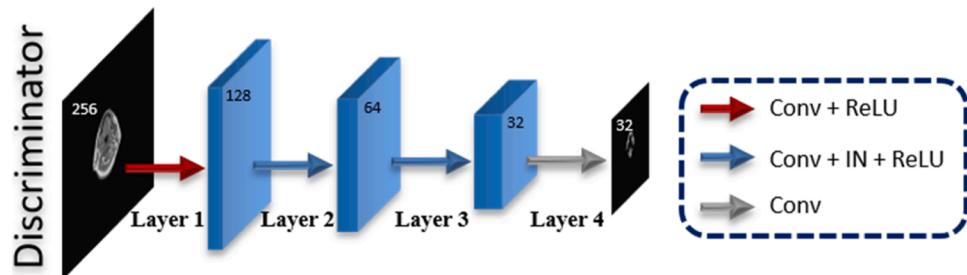


Fig. A1. The architecture of the discriminator. The discriminator is composed of 4 layers, where the operations represented by the yellow line are Conv and ReLU, the blue line is Conv, IN, and ReLU, the gray line is Conv.

Baseline Models

a. DCNN

Deep Convolutional Neural Network (DCNN) (Han, 2017) is improved based on U-Net (Ronneberger et al., 2015) to realize the conversion from MRI to CT. DCNN needs training data paired, and the difference between sCT and rCT is taken as the loss to optimize the whole model.

b. GANs

Generative Adversarial Nets (GANs) (Goodfellow et al., 2014) optimize the generative model through an adversarial step. It trains two models simultaneously: one is the generative model G for obtaining the data distribution. The other is the discriminative model D for discriminating the authenticity of the samples. The training purpose of G is to maximize the probability of D making mistakes. GANs is an unsupervised method. Therefore, it does not need to register the MRI and CT; that is, it can use unpaired data.

c. Conditional GAN

Conditional GAN (Mirza et al., 2014) is derived based on GANs. The main difference is that the method based on GANs is unsupervised, while the method based on Conditional GAN adds additional auxiliary conditions, which can be supervised or weakly supervised.

d. Cycle GAN

Cycle GAN (Zhu et al., 2017) is improved based on GANs. Cycle GAN is essentially two mirror-symmetric GANs, forming a ring network. Unpaired data sets can be applied to Cycle GAN for training. The difference between Cycle GAN and our model is that Pseudo-Cycle Consistent Loss is added in our method, and Z-Net is used as the generator.

In this paper, all the compared methods we mentioned above are implemented based on the author-provided codes with identical hyperparameters setting as ours: the batch size is set to 1, the learning rate is linearly decayed from 2×10^{-4} to 2×10^{-6} , and the input and output data sizes both are 256*256.

References

- Andreasen, D., Van Leemput, K., Edmund, J.M., 2016. A patch-based pseudo-CT approach for MRI-only radiotherapy in the pelvis. *Med. Phys.* 43, 4742–4752.
- Arabi, H., et al., 2018. Comparative study of algorithms for synthetic CT generation from MRI: consequences for MRI-guided radiation planning in the pelvic region. *Med. Phys.* 45 (11), 5218–5233.
- Arabi, H., Zaidi, H., 2016. One registration multi-atlas-based pseudo-CT generation for attenuation correction in PET/MRI. *Eur. J. Nucl. Med. Mol. Imaging* 43 (11), 2021–2035.
- Burgos, N., et al., 2014. Attenuation correction synthesis for hybrid PET-MR scanners: application to brain studies. *IEEE Trans. Med. Imaging* 33 (12), 2332–2341.
- Chin, A.L., Lin, A., Anamalayil, S., Teo, B.K.K., 2014. Feasibility and limitations of bulk density assignment in MRI for head and neck IMRT treatment planning. *J. Appl. Clin. Med. Phys.* 15 (5), 100–111.
- Cox, I.J., Roy, S., Hingorani, S.L., 1995. Dynamic histogram warping of image pairs for constant image brightness. 2, Conference Proceedings 366–369.
- Glorot, X., Bordes, A., Bengio, Y., 2011. Deep sparse rectifier neural networks. *J. Mach. Learn. Res.* 15, 315–323.
- Goodfellow, I., et al., 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems*, pp. 2672–2680.
- Han, X., 2017. MR-based synthetic CT generation using a deep convolutional neural network method. *Med. Phys.* 44 (4), 1408–1419.
- He, K., Zhang, X., Ren, S., et al., 2016. Deep residual learning for image recognition. In: *IEEE Conference on Computer Vision & Pattern Recognition*. IEEE Computer Society.
- Hiasa, Y., et al., 2018. Cross-modality image synthesis from unpaired data using CycleGAN. In: *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer, pp. 31–41.
- Hsu, S.-H., Cao, Y., Huang, K., Feng, M., Balter, J.M., 2013. Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy. *Phys. Med. Biol.* 58, 8419–8435.
- Huynh, T., et al., 2016. Estimating CT image from MRI data using structured random forest and auto-context model. *IEEE Trans. Med. Imaging* 35 (1), 174–183.
- Leynes, A.P., et al., 2017. Direct PseudoCT generation for pelvis PET/MRI attenuation correction using deep convolutional neural networks with multi-parametric MRI: zero echo-time and Dixon Deep pseudoCT (ZeDD-CT). *J. Nucl. Med.* <https://doi.org/10.2967/jnumed.117.198051>.
- Mirza, Mehdi, Osindero, Simon, 2014. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.
- Nie, D., Cao, X., Gao, Y., Wang, L., Shen, D., 2016. Estimating CT image from MRI data using 3D fully convolutional networks. *Deep Learning and Data Labeling for Medical Applications*. Springer, pp. 170–178.
- Nie, D., et al., 2017. Medical image synthesis with context-aware generative adversarial networks. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 417–425.
- Nie, D., et al., 2018. Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans. Med. Imaging* 65 (12), 2720–2730.
- Nyholm, T., Jonsson, J., 2014. Counterpoint: opportunities and challenges of a magnetic resonance imaging-only radiotherapy work flow. *Semin. Radiat. Oncol.* 24 (3), 175–180. Elsevier.
- Owrangi, A.M., Greer, P.B., Glide-Hurst, C.K., 2018. MRI-only treatment planning: benefits and challenges. *Phys. Med. Biol.* 63 (5), 05TR01.
- Penney, G.P., Weese, J., Little, J.A., Desmedt, P., Hill, D.L.G., Hawkes, D.J., 1998. A comparison of similarity measures for use in 2-D-3-D medical image registration. *IEEE Trans. Med. Imaging* 17 (4), 586–595.
- Rasch, C., Steenbakkers, R., van Herk, M., 2005. Target definition in prostate, head, and neck. *Semin. Radiat. Oncol.* 15 (3), 136–145. Elsevier.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 234–241.
- Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* 17, 87–97.
- Ulyanov, D., Vedaldi, A., Lempitsky, V.S., 2016. Instance normalization: the missing ingredient for fast stylization. *CoRR abs/1607.08022*.
- Wolterink, J.M., Dinkla, A.M., Savenije, M.H., Seevinck, P.R., van den Berg, C.A., Igum, I., 2017. Deep MR to CT synthesis using unpaired data. In: *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer, pp. 14–23.
- Xu, Zhe, et al., 2020. Adversarial uni-and multi-modal stream networks for multimodal image registration. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham.
- Yang, H., et al., 2018. Unpaired brain mr-to-ct synthesis using a structure-constrained cyclegan. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, pp. 174–182.
- Yang, H., et al., 2020. Unsupervised MR-to-CT synthesis using structure-constrained CycleGAN. *IEEE Transactions on Medical Imaging* PP 99, 1–1.
- Yi, Z., Zhang, H., Tan, P., Gong, M., 2017. Dualgan: unsupervised dual learning for image-to-image translation. *Proceedings of the IEEE International Conference on Computer Vision* 2849–2857.
- Zeng, G., Zheng, G., 2019. Hybrid generative adversarial networks for deep MR to CT synthesis using unpaired data. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 759–767.
- Zheng, W., Kim, J.P., Kadbi, M., Movsas, B., Chetty, I.J., Glide-Hurst, C.K., 2015. Magnetic resonance-based automatic air segmentation for generation of synthetic computed tomography scans in the head region. *Int. J. Radiat. Oncol. Biol. Phys.* 93, 497–506.
- Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the IEEE International Conference on Computer Vision* 2223–2232.
- Zijlstra, F., et al., 2019. CT synthesis from MR images for orthopedic applications in the lower arm using a conditional generative adversarial network. *Medical Imaging 2019: Image Processing* vol. 10949, p. 109491J. International Society for Optics and Photonics.