ITC 3/54
Information Technology and Control
Vol. 54 / No. 3/ 2025
pp. 937-957
DOI 10.5755/j01.itc.54.3.38668

NAP-CycleGAN: A New Cyclegan-Based CT Images Synthesis Model for Clinical Image Reconstruction Using Brain MR Images

Received 2024/08/31 | Accepted after revision 2025/02/11

HOW TO CITE: Zhang, L., Zhu, L., Cao, G., Wang, C., Qi, Z. (2025). NAP-CycleGAN: A New Cyclegan-Based CT Images Synthesis Model for Clinical Image Reconstruction Using Brain MR Images. *Information Technology and Control*, 54(3), 937-957. https://doi.org/10.5755/j01.itc.54.3.38668

NAP-CycleGAN: A New Cyclegan-Based CT Images Synthesis Model for Clinical Image Reconstruction Using Brain MR Images

Li Zhang, Lin Zhu, Guorui Cao, Cancan Wang
Guangxi Key Laboratory of Special Biomedicine, School of Medicine, Guangxi University, Nanning, 530004, China

Zhongquan Qi
Guangxi Key Laboratory of Special Biomedicine, School of Medicine, Guangxi University, Nanning, 530004, China
Fujian Maternity and Child Health Hospital, 18 Daoshan Road, Fuzhou City, Fujian Province, 350001, China

Corresponding author: qzqsfy@163.com

The intricate structure of the brain often necessitates the combined use of magnetic resonance (MR) and computed tomography (CT) imaging for comprehensive diagnostics in clinical care. However, certain patients cannot be exposed to radiation-intensive CT scans, leading to data scarcity and affecting subsequent treatment. In this regard, this paper proposes a new model noise-attention-pix2pix-CycleGAN (NAP-CycleGAN), replacing the generator with pix2pixHD utilizing multi-scale strategies and context-aware modules. By integrating channel attention, the model effectively extracts relevant image features, allowing adaptive weight assignment and handling of long-range dependencies. Additionally, Gaussian noise is introduced to the discriminator to counteract adversarial sample attacks and prevent gradient vanishing. Furthermore, feature matching loss and cycle consistency loss are integrated to reduce image detail distortion. To verify the model validity, it is compared with seven state-of-the-art methods. The experimental results on the public brain dataset brain01 show that the proposed model outperforms these methods, it yields the best, and the synthetic CTs of the proposed model are closest to the original CT images.

KEYWORDS: CycleGAN, Pix2pixHD, Channel attention, Noise discriminator, MR-CT synthesis

1. Introduction

The structure of the human brain is sophisticated and complex, and brain imaging technology has therefore become an indispensable tool in scientific research and clinical work. Currently, mainstream medical imaging technologies include X-ray, magnetic resonance (MR), computed tomography (CT) so on [31, 3]. At the time of application, MR and CT have been widely used [32], MR provides better soft-tissue contrast, but CT is particularly suitable for bone imaging with better electron density information [20, 44, 8, 33]. It suggests that single-modality medical imaging sometimes struggles to provide a full range of critical information, so a combination of multimodal images is needed to enable physicians to obtain information from different perspectives. To reduce the risk of ionizing radiation, existing MR can be used to generate CT to enable the use of multimodal images as an aid in clinical diagnosis and treatment [25].

To date, many methods have been proposed for MR to CT such as density-based [28], atlas-based [2], and machine learning-based [19] methods. However these methods have limited accuracy [24], lack robustness, and are time-consuming [22]. To solve these problems, researchers have tried to use deep learning techniques for image synthesis [4, 34, 39]. Among them, generative adversarial network(GAN) [16] is one of the most widely used, which is efficient and consumes less manpower to achieve better results. Many GAN networks [12, 49, 48, 40, 36] have been used for image synthesis. Chen et al. [6] proposed the NICEGAN method by reusing the discriminator, one of the two main structures of GAN, to obtain a more impactful and efficient image translation architecture. Aljohani et al. [1] proposed a deep pix2pixGAN method based on deep convolutional neural networks as well as GAN to synthesise better quality CT images from MR maps as a way to synthesise new medical images from different images. Zhao et al. [50] used conditional GAN(CGAN) to synthesize images of multi-centred regions of the pelvic region by applying the feature enhancement module and feature reconstruction loss to GAN and composited a high-quality image with more detailed texture features. Jin et al. [21] proposed MR-GAN based on a dual cycle-consistent structure and combines dual cycle-consistent loss to generate MR from CT using paired and unpaired data to address the context-misalignment problem of unpaired training. Xin et al. [37] used multi-channel multi-path CGAN of multi-parameter MRI with the channel-based independent feature extraction network to generate pseudo-CT and reduce the risk of omission of feature extraction in single-channel images. However, the results achieved by these methods are still far from adequate and limited in the scope of enhancement. For structured medical images, more detailed texture features of medical images need to be extracted. Based on Cycle-Consistent GAN(CycleGAN)'s ability to bi-directionally capture potential mapping relationships of images to achieve preservation of medical structural information, Wang et al. [52] used CycleGAN for deformation-invariant cross-domain information fusion for synthetic CT, which achieved better image alignment. Jelmer et al. [43] used Deep MR to CT based on CycleGAN, using convolutional neural networks (CNNs) instead of GAN for generator and discriminator architecture to generate CT using an unpaired dataset and generating higher quality CT images and preserve key details in the MR images as well as good structural alignment for efficient medical image translation. However, The limited CNN structure of the network and the unpaired sample data still leave ambiguity in the details sometime and clinical features of the generated CTs. New methods need to be found for high-quality CTs [41].

Pix2pixHD is emerging in the past few years, proposed by Wang et al. [42], and used to synthesize high-resolution photo-realistic imagery because of its ability to capture image details. Kévin et al. [5] used a pix2pix-HD-based model to synthesize CT from MR using multicentre data, demonstrating to some extent the utility of the framework to bring the study of MR-CT synthesis techniques for clinical applications a step closer. For the task of MR-CT synthesis using GAN, the researchers have iterated and innovated through multiple channels. Yi et al. [47] proposed DualGAN to reconstruct realistic synthetic images by introducing a cycle consistency loss, and the introduction of it helps to maintain the consistency of image semantics and improve the flexibility and applicability of training. Ding et al. [10] utilized feature matching loss with CGAN to improve the quality of generated images and Highly realistic images are obtained. Moreover, Xue

et al. [45] investigated the enhancement of CycleGAN for image translation by combining the channel attention with evolutionary algorithms and obtained the feature that the channel attention allows learning important texture features of real images, which provides a new direction for the enhancement of CycleGAN image translation. In addition, Emami et al. [11] and Kearney et al. [23] introduce spatial attention in SPA-GAN and attention-aware CycleGAN, respectively, to help the generator focus on key regions of the image, such as soft tissue features and cranial detail texture in MR images. However, the accuracy for generating cross-modal CTs using GAN is currently inconsistent and sometimes produces blurred images.

In this paper, to reduce the risk of patient exposure to radiation, and make it possible to clinically obtain MR-CT multimodal image combinations based on the patient's captured MRs alone thus providing an accurate plan for a comprehensive diagnosis and treatment, Considering that CycleGAN flexibly generates better CT images, We wish to improve the transparency and interpretability of the GAN model, as well as to explore the detail capturing ability of GAN in CT synthesis even further, and obtain higher quality CT images. The contributions of this paper are as follows:

1. We propose a new noise-attention-pix2pix CycleGAN model named NAP-CycleGAN with a pix2pix-HD generator and mainly concentrate on MR to CT synthesis, to provide a potential solution of missing MR-CT paired brain datasets in clinical practice.
2. The channel attention modules are introduced in the generator and efficiently focus on local feature extraction thus capturing important detailed textural features of bones and soft tissues in MR images and performing adaptive weight assignment when necessary.
3. The Gaussian Noise is incorporated into the discriminator to improve the robustness of the discriminator against adversarial sample attacks and avoid gradient vanishing.
4. Feature matching loss and cycle consistency loss are added in the synthesis and reconstruction process to reduce the loss of detail and gradient descent loss, thus improving the detail quality and realism of the generated images. And quantitative as well as qualitative analyses are performed on the public brain datasets with other advanced models.

2. Materials and Methods

2.1. NAP-CycleGAN Model

2.1.1. Model Architecture

The brain CT generation method based on the noise-attention-pix2pix-CycleGAN(NAP-CycleGAN) is shown in Figure 1. The (a) in Figure 1 represents the NAP-CycleGAN architecture. As the same as CycleGAN, the NAP-CycleGAN is essentially a ring-shaped network, divided into two parts, forward and reverse cycles [7], and contains two generators and two discriminators [46]. On the input side, the original MR is regarded as the X domain, and the original CT is regarded as the Y domain, the X and Y domains share the generators G_{CT} and G_{MR}, each with a discriminator D_{CT} and D_{MR} [9]. For the X domain, the generator G_{CT}'s task is to convert MR into CT, and for the Y domain, the generator G_{MR}'s task is to convert CT into MR. In the forward loop, the X domain's MR is used as input to generate the Synthetic CT (SCT) via the generator G_{CT}. The SCT is fed into the discriminator D_{CT} along with the original CT for differentiation. At the same time, the output SCT of the generator G_{CT} is used as input to the generator G_{MR} to generate the fake MR. And then the generated pseudo-MR is compared with the original MR to produce a forward cycle consistency loss. In the reverse loop, the CT of the Y domain is used as a network input to generate a Synthetic MR (SMR) via the generator G_{MR}. The generated SMR is fed into the discriminator D_{MR} along with the original MR for differentiation. At the same time, the output SMR of the generator G_{MR} is used as input to the generator G_{CT} to generate the fake CT. The generated pseudo-CTs are compared with the original CTs to produce a reverse cycle consistency loss [6]. In addition, the network generates some feature matching losses during the actual training of the forward and backward loops.

Figure 1(b) represents the block diagram of MR to CT synthesis using NAP-CycleGAN. The process of generating CT from MR using NAP-CycleGAN is as follows: Input the MR image into NAP-CycleGAN, the generator G_{CT} generates a synthetic CT, and then input it into the discriminator G_{CT} to identify whether the image is a real CT or a synthetic CT, and if the result of the identification is fake, return the result and continue with the next generation and discrimination process, and generates the corresponding adversarial

loss. In this process, the generated SCT is fed into the generator G_{MR} for generating the reconstruction MR, and the comparison of the reconstruction MR with the input MR produces a loss of cycle consistency. At the same time, the generated synthetic CT is compared with the paired CT of the input MR producing a feature matching loss.

2.1.2. Pix2pixHD Generator

Although the traditional Cyclegan generator has better global style migration, its performance is sometimes limited in high-resolution image generation tasks, resulting in loss of details and image blurring [13]. Based on this, we replace the generator with pix2pixHD, constituting the pix2pixHD, constituting the generator structure with encoder, residual block, and decoder as main parts. pix2pixHD is designed for high-resolution image generation [30], capturing both global consistency and local details, with an architecture that is capable of handling complex details and large image sizes. This capability is especially important for generating medical images such as CT or

Figure 1

(a) NAP-CycleGAN architecture. G_{CT}: CT generator, from MR to generate CT; G_{MR}: MR generator, from CT to generate MR; D_{CT}: CT discriminator, discriminate whether the SCT is real or fake; D_{MR}: MR discriminator, discriminate whether the SMR is real or fake; (b) The block diagram of MR to CT synthesis using NAP-CycleGAN.

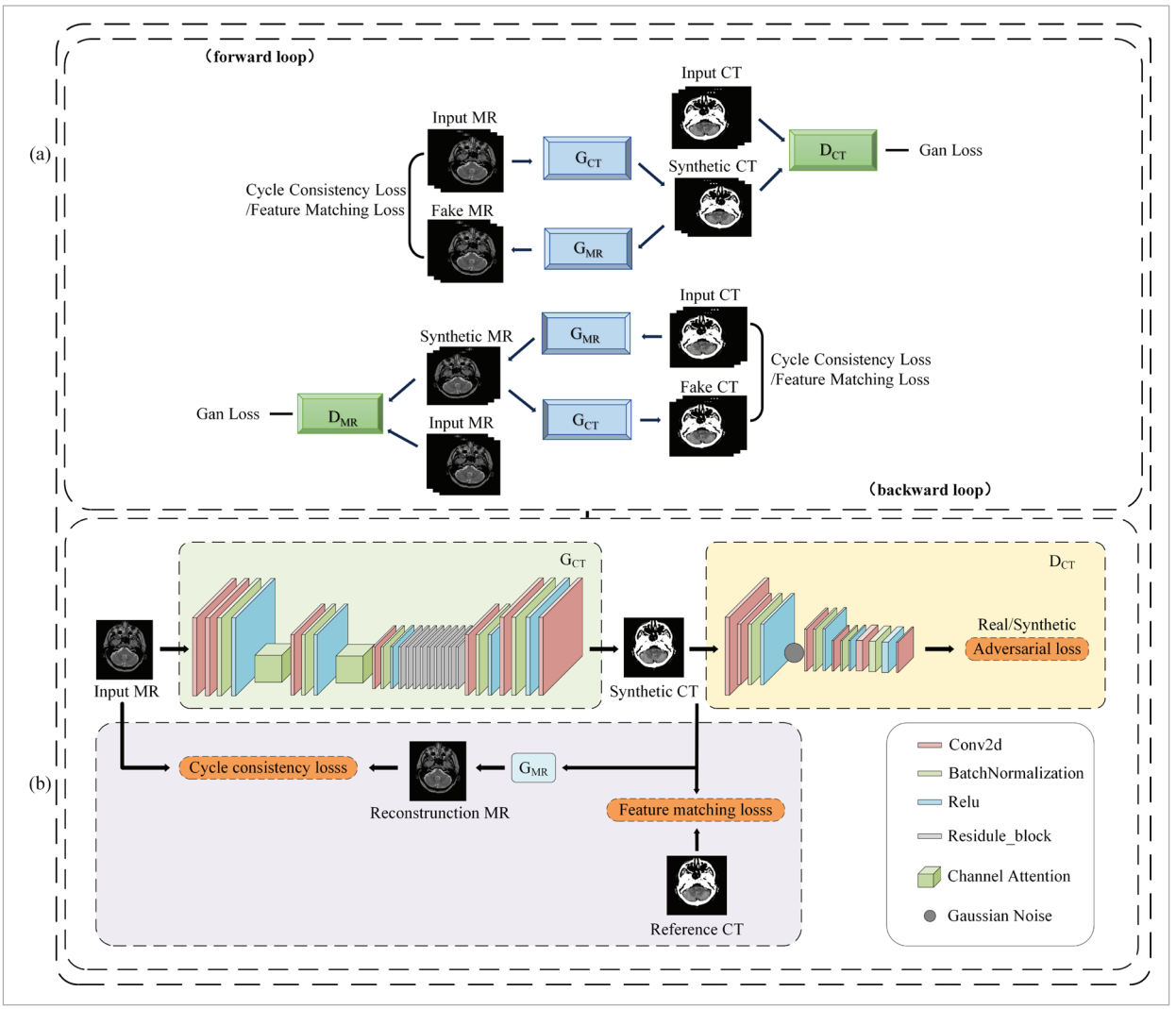
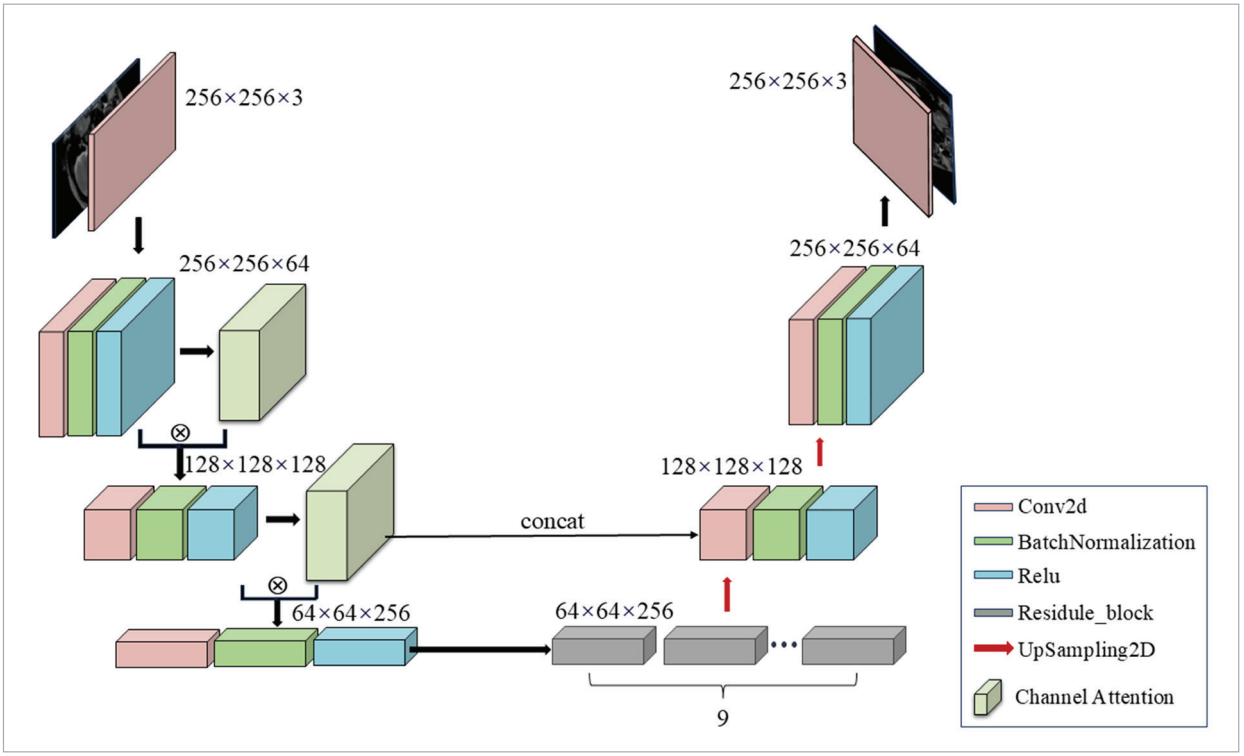


Figure 2
Generator Structure.



images with fine features. In contrast, other common generators such as PatchGAN underperform complex structures and details, and the generated images may be macroscopically incoherent [14]; U-Net generators rely too much on skip connections, which may cause the network to focus more on local information and ignore global features, and are prone to blurring of boundaries [18]. In addition, these generators perform poorly in capturing long-range dependencies and features in detailed regions, which may lead to less realistic quality of the generated images [26, 27]. In contrast, pix2pixHD combines the advantages of both, reasonably introduces the skip connection and up-sampling modules, further enhances the detail expression ability and spatial resolution generation quality, significantly improves the resolution and realism of the generated images, and has superior advantages compared to other generators in the high-resolution image generation task [42], which can better satisfy the needs of the tasks that have high requirements on details and resolution, such as medical imaging.

In the pixpixHD generator, the channel attention is introduced in the encoder, which can help the network learn the important features in the input image more effectively, automatically learn the importance of each channel, and then adjust the weight of the feature map by enhancing or suppressing the response of different channels; the spatial attention mechanism, upsampling layer [35] and skip connection are introduced in decoder, which helps the network pay more attention to the features of different regions in input image, effectively increasing the spatial resolution of generated CTs, thus improving the detail and realism of the image.

Figure 2 shows the generator structure of NAP-CycleGAN, and its workflow is as follows: the input and output of the image in the generator are transmitted using a scale of 1×256×256×3. Among them, the encoder calculates a total of three convolutions, the output scales obtained from each calculation of a convolution module are further calculated and outputted by using channel attention once, and the results obtained from the channel attention are multiplied by

the results obtained from the calculation of the convolution block, and the results are input into the next convolution module. The results obtained from the channel attention are multiplied by the results of the convolution block and fed into the next convolution block for computation, and the results obtained from the last convolution block are directly fed into the residual blocks for computation. In the decoder, two convolution modules are designed, and an up-sampling operation is carried out after every convolution block computation. The result of the first convolutional module is then connected to the corresponding convolutional module in the encoder, and the last convolutional module directly outputs the generated image using the tanh activation function.

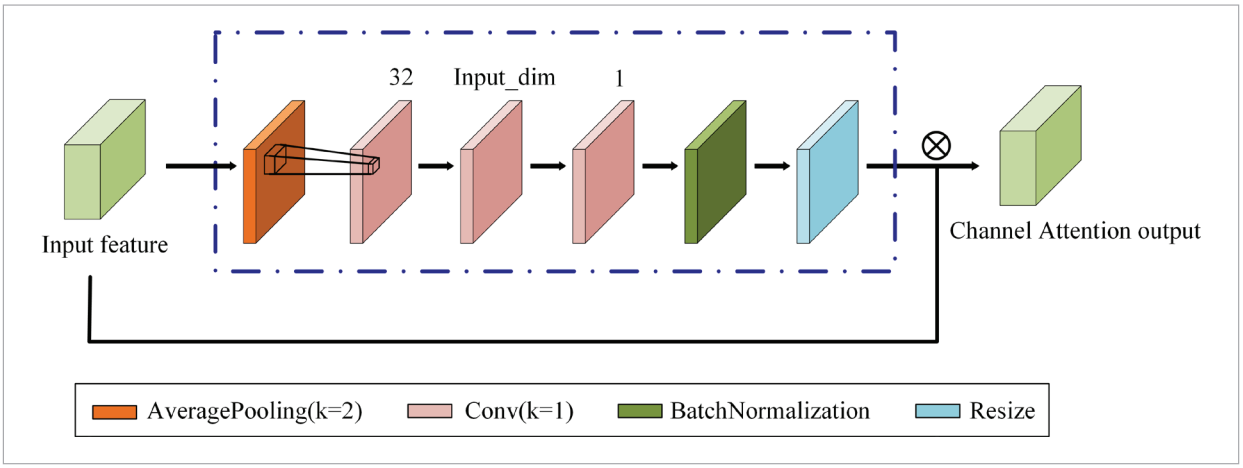
2.1.3. Channel Attention Module

The basic structure of the channel attention mechanism module [51] in the NAP-CycleGAN network is shown in Figure 3. In the channel attention of the embedded encoder, the input feature **F** undergoes a 2×2 average pooling layer and then completes the attention feature extraction by the channel convolutional computation, and completes the detail feature extraction operation of the input image by using three convolutional layers and one normalization layer and one Resize layer. In this process, the size of the convolution kernel (k) is 1×1, the number of channels in the first convolution layer W_1 is set to 32, the number of channels in the second convolution layer W_2 is set to the input scale, the two convolutional layers generates different weights for different input channels. Then the number of channels in the convolution layer W_3 is compressed into a single channel for standardization, which can generate a global single-channel feature map, in which each pixel value represents the overall importance weight of the location, and this weight gathers the Comprehensive weight of the location on all channels, which can reduce the computational complexity as well as enhance the salient features and aggregated the global information, which can increase the computational performance of the model. Next, the output features obtained from a single channel are subjected to the Resize operation with the same shape as input_dim, which makes the output scale identical to the input feature scale **F** and is multiplied to obtain the final output result (M_c) of the channel attention mechanism. By introducing the channel attention mechanism, the encoder can automatically learn the correlation between different channels and adaptively adjust the importance of different channels to better adapt to different types of input images, enabling the encoder to better capture the key features in the input image [15]. This helps to improve the generalization ability of the model and enhance the feature representation ability of the generator, enabling the generator to better process images with different features to produce higher-quality output images. The above process of operating the channel attention mechanism can be represented as follows:

$$M_C=\sigma(W_3(R(W_2(R(W_1(AvgPool(\mathbf{F}))))))\times\mathbf{F}. \quad (1)$$

Figure 3

Schematic diagram of the channel attention.



In Equation (1), M_C represents the attention weight of each channel i.e. the channel attention output result, and **F** is the input feature map, such as the dimension of 256×256×64, which represents the height, width, and number of channels. AvgPool(**F**) is the average pooling operation, W_1, W_2, and W_3 are convolution layers, R is the activation function Relu of W_1 and W_2, and σ is the sigmoid activation function of W_3.

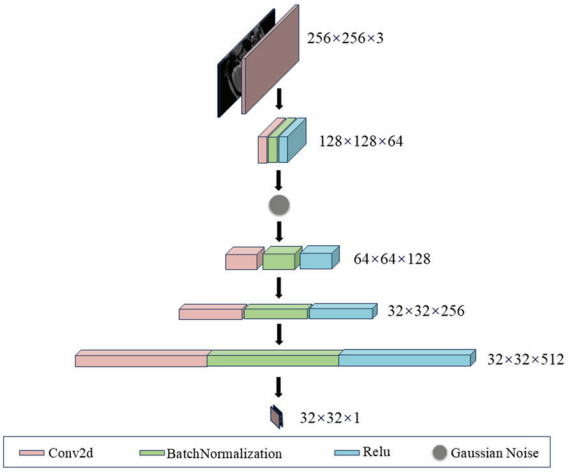
2.1.4. Discriminator Structure

For the discriminator, on top of the traditional CycleGAN network discriminator, Gaussian Noise is proposed to be embedded before the second convolutional module after the first one. Gaussian

Noise can be regarded as a regularisation technique, and the introduction of Gaussian Noise to form the noise discriminator increases the robustness and generalization of the model and improves the model's ability to fight against adversarial attacks, making it easier for the model to learn the true features of real data. As shown in Figure 4, in the discriminator, there are a total of four convolutional modules consisting of convolutional, standard, and activation layers for stacking computation.

Figure 4

Discriminator structure.



2.2. Loss Function

In the training process of the NAP-CycleGAN network, the generator objective function used to compute the loss contains the traditional adversarial loss, the cycle consistency loss, and the feature matching loss, and the loss function used for the discriminator has only the adversarial loss. The main definition of the generator objective function is as follows:

$$L_{\text{NAP-CycleGAN}} = L_{\text{adv}} + \beta L_{\text{fml}} + \gamma L_{\text{ccl}}. \quad (2)$$

In Equation (2), L_{adv} stands for the adversarial loss, L_{fml} stands for the feature matching loss, L_{ccl} stands for the cycle consistency loss, and β and γ stand for the corresponding loss weights, where the value of β is set to 0.1 and γ is set to 10.

2.2.1. Adversarial Loss

The total adversarial loss in training combines the adversarial loss in the x and y domains. For the calculation of the adversarial loss, a biparadigm of squared L2 norms is used to measure the difference between the samples produced by the generator and the true samples. The formula for this is as follows:

$$L_{\text{adv}} = E_{x\sim pdata(x)}[\log D(x)] + E_{z\sim p(z)}[\log(1-D(G(z)))]. \quad (3)$$

In Equation (3), $D(x)$ is the output of the discriminator on the true data x, $G(z)$ is the output of the generator on the random noise z, and $D(G(z))$ is the output of the discriminator on the generated sample $G(z)$, $E_{x\sim pdata(x)}$ denotes the expectation of the true data, where pdata(x) is the distribution of the true data, $E_{z\sim p(z)}$ denotes the expectation of the random noise, i.e., the prediction of the discriminator D on the fake data $G(z)$ generated by the generator G, where p(z) is the distribution of the random noise.

2.2.2. Feature Matching Loss

The goal of the feature matching loss function is to facilitate feature matching between the generator and the discriminator by minimizing the distance between the real data and the generated data in the feature map output by the discriminator, thus improving the quality and realism of the generated image. Feature matching loss is a technique in GAN that aims to improve the training stability of the generator and the quality of the generated images. The feature matching loss can be seen as a further development and optimization of the traditional perceptual loss as it also involves the use of intermediate features of the convolutional neural network, but it focuses more on the training strategy of the generative adversarial net-

work rather than simply as a choice of loss function. The difference between the features is measured using Euclidean distance and the sum of the mean values is calculated as the feature matching loss with the following formula:

$$L_{fml}=\frac{1}{N}\sum_{i=1}^{N}(\|E(x)_i-E(G(x))_i\|_2^2+\|D(y)_i-D(F(y))_i\|_2^2). \quad (4)$$

Equation (4) calculates the discriminator's feature similarity identification between the real reference image and the fake image, where N represents the number of pixels of the feature image, G and F are generators, E and D are discriminators, $E(x)$ denotes the feature image of real data x at the output of generator E, $G(x)$ denotes the generated data $G(x)$, $D(y)$ denotes the feature image of real data y at the output of discriminator D, and $F(y)$ denotes the generated data $F(y)$.

2.2.3. Cycle Consistency Loss

The cycle consistency loss is used to promote consistency between the generated image and the original image. The basic idea is to transform an image to another domain through a generator, and then reverse the transformation back to the original domain through another generator, and as close as possible to the original image. By minimizing this difference, the performance of the generator can be improved by making the generated image closer to the original input image. Its calculation formula is as follows:

$$L_{ccl}=\frac{1}{\gamma}\left(\frac{1}{N}\sum_{i=1}^{N}|F(G(x_i))-x_i|+\frac{1}{M}\sum_{i=1}^{M}|G(F(\gamma_i))-\gamma_i|\right). \quad (5)$$

Equation (5) calculates the similarity between the fake input image synthesised from the generated image and the reference input image by calculating the L1 distance between the x-domain and the y-domain, N and M represent the number of samples sampled from domains x and y, respectively. G is the generator from domain x to domain y, and F is the generator from domain y to domain x. $F(G(x))$ represents the fake image x generated from the synthetic image y generated from the real image x by the generator F, while $G(F(y))$ represents the fake image y obtained from the synthetic image x generated from the real image y by the generator G. γ represents the weight of the Cycle consistency loss over the total loss.

2.3. Experimental Setup

2.3.1. Data Acquisition and Pre-processing

The model uses the publicly available dataset brain01 from the GitHub website(https://github.com/ChengBinJin/MRI-to-CT-DCNN-TensorFlow) and has been introduced in [17]. The dataset includes 367 pairs of 2D brain MR-CT raw paired images, each MR and CT image size is 512 × 256 pixels. We finish the benchmark by rectifying N4 bias correction, histogram matching, and head mask generation. Then we adjust images by left-right rotation at different angles. To ensure the accuracy of the data, we cropped all MR-CT separated images after alignment to 256*256 pixels, and 290 pairs of MR-CT paired images were randomly selected for training and 70 pairs of images for testing. Our processed dataset is divided into two training sets and two test sets named trainA, trainB, and testA, testB respectively. There are 290 images in both folders of the training set, and 77 images in the two folders of the test set, where A stands for MR, and B stands for CT.

2.3.2. Evaluation Metrics

For the comparison of model generation effects, the structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), mean absolute error (MAE), Visual Information Fidelity (VIF), and Fréchet Inception Distance (FID) are used to complete the comparative evaluation of different models.

The SSIM is a kind of index used to measure the degree of similarity between two images, which takes into account the three aspects of brightness, contrast, and structure, and ranges between [-1, 1]. SSIM is essential for preserving anatomical structures, particularly in disease detection and segmentation tasks. Higher SSIM indicates better retention of structural information, which is important for accurate medical diagnosis and anatomical structure analysis, especially in image conversion. The formula is as follows:

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+c_1)(2\mu_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)}. \quad (6)$$

In Equation (6), x and y are the two images to be compared, μ_x and μ_y are the mean values, σ_x^2 and σ_y^2 are the variances, σ_{xy} is the covariance, and c_1 and c_2 are small constants added for stability.

PSNR is a measure of image distortion that evaluates image quality by comparing the mean square error in decibels (dB) between different images. Increased PSNR values mean less distortion in the generated images, so doctors can rely on higher-quality generated images for decision-making without worrying about misdiagnosis due to image distortion. The formula is as follows:

$$\mathrm{PSNR}(x,y)=10\cdot\log_{10}\left(\frac{MAX^2}{\mathrm{MSE}}\right). \tag{7}$$

MAX is the maximum possible value of the pixel value and MSE is the mean square error, the sum of the squares of the average differences in each loudness between the two images divided by the number of pixels, it is calculated as follows:

$$\mathrm{MSE}(x,y)=\frac{1}{N}\sum_{i=1}^{N}(x_i-y_i)^2 . \tag{8}$$

MAE is a measure of the average absolute difference between two images, ensuring the overall quality of the synthesized image, Lower MAE values indicate better retention of detail in the resulting image, which helps the physician to more accurately identify and diagnose the lesion area and is calculated as Equation (9):

$$\mathrm{MAE}(x,y)=\frac{1}{N}\sum_{i=1}^{N}|x_i-y_i| . \tag{9}$$

In Equation (9), x and y are image pixel values and N represents the number of pixels.

VIF measures image quality by evaluating the information fidelity of image pairs. VIF measures the fidelity of visual information in an image. A high VIF value means that the resulting image maintains good visual quality in the presence of noise, which is particularly important for medical imaging diagnosis. Defined as:

$$\mathrm{VIF}(x,y)=\frac{\sum_{k=1}^{K}\sum_{l=1}^{L_K} I(x_{kl};\hat{x}_{kl}|\sigma^2_{nkl})}{\sum_{k=1}^{K}\sum_{l=1}^{L_K} I(x_{kl};n_{kl}|\sigma^2_{nkl})} . \tag{10}$$

In Equation (10), x_{kl} is the original image in the kth scale and lth subband, and $\hat{x}_{kl}$ is the output image in the corresponding scale and subband. $I(x_{kl};\hat{x}_{kl}|\sigma^2_{nkl})$ is the information between the input image x_{kl} and the output image $\hat{x}_{kl}$ under the condition of noise variance σ^2_{nkl}; $I(x_{kl};n_{kl}|\sigma^2_{nkl})$ is the information between the input image x_{kl} and the output image n_{kl} under the condition of noise variance σ^2_{nkl}.

FID is used to assess the quality of the generated image by comparing the statistical properties of the generated image with the real image. Lower FID values indicate that the generated image is closer to the real image in terms of statistical properties and the generated image is more clinically usable, especially in disease diagnosis where high similarity between statistical properties and real image means more reliable judgement. That can be computed as:

$$\mathrm{FID}=\left\|\boldsymbol{\mu}_1-\boldsymbol{\mu}_2\right\|_2^2+\mathrm{Tr}(\Sigma_1+\Sigma_2-2\sqrt{\Sigma_1\Sigma_2}) . \tag{11}$$

In Equation (11), $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the mean vectors of the real and generated image features, respectively, Σ_1 and Σ_2 are the covariance matrices of the real and generated image features, respectively, $||\boldsymbol{\mu}_1-\boldsymbol{\mu}_2||_2^2$ is the Euclidean distance between the two mean vectors, and $\mathrm{Tr}(\cdot)$ denotes the trace (the sum of the diagonal elements of the matrix), and $\sqrt{\Sigma_1\Sigma_2}$ is the square root of the product of the two covariance matrices.

2.3.3. Statistical Analysis

For the analysis of statistical differences in model performance metrics, we performed the paired samples Wilcoxon signed rank test using GraphPad Prim 8, The criterion for the significant difference in results was $p < 0.05$. When $p < 0.05$, it indicates that there is a difference between the models.

2.3.4. Implementation Details

The task of generating pseudo-CT from MR for the NAP-CycleGAN network is completed using Python as the programming language, with a comparative analysis of the evaluation metrics and a statistical analysis of the generated results with seven state-of-art models, namely NICEGAN [6], Deep MR to CT [43], DualGAN [47], CGAN [42], CycleGAN [52], pix2pix [38] and VAE [29]. After completing training using a training set of 290 pairs of MR-CT paired images, we used a test set of 77 pairs of MR-CT paired images as evaluation samples to compare model performance.

During the training process, the training parameters of all models are kept the same as NAP-CycleGAN, and the Adam optimizer is used for gradient descent and parameter updating of the models. The parameters of NAP-CycleGAN are set to λ with 10 and β with 0.5, the *batch-size* is set to 1, the initial *learning rate* is 0.0002, the number of epochs to keep the learning

rate constant during training, *epoch_step* is 100, the visualization results are saved every 100 steps and sets the random number seed to the same number as the training set images, 290. And we set the weights of feature matching loss and cycle consistency loss to 0.1 and 10 respectively. A total of 200 *epochs* of training are performed with a training duration of 12 hours and a testing duration of 1 second per new image.

The experimental platform is relying on the NVIDIA 2080Ti GPU graphics card Platform, based on python3.8, Tensorflow2.5, and cuda11.2.

3. Results

3.1. Evaluation of SCTs by Different Methods

In order to verify the effectiveness of the proposed network, the seven MR-CT synthesis novel methods of NICEGAN, Deep MR to CT, DualGAN, CGAN, CycleGAN, pix2pix, VAE, and NAP-CycleGAN are comparatively, respectively with the same dataset brain01.

The Synthetic CTs(SCTs) generated by these state-of-the-art methods and their difference maps are shown in Figure 5, showing the training synthetic

Figure 5

Comparison of Synthetic CTs and difference maps from different models using the whole testing dataset including 77 paired MR-CT images. (a) Original MR and the Reference CT, (b) NICEGAN SCT and difference map, (c) Deep MR to CT SCT and difference map, (d) DualGAN SCT and difference map, (e) CGAN SCT and difference map, (f) CycleGAN SCT and difference map, (g) pix2pix SCT and difference map, (h) VAE SCT and difference map, and (i) NAP-CycleGAN SCT and difference map.

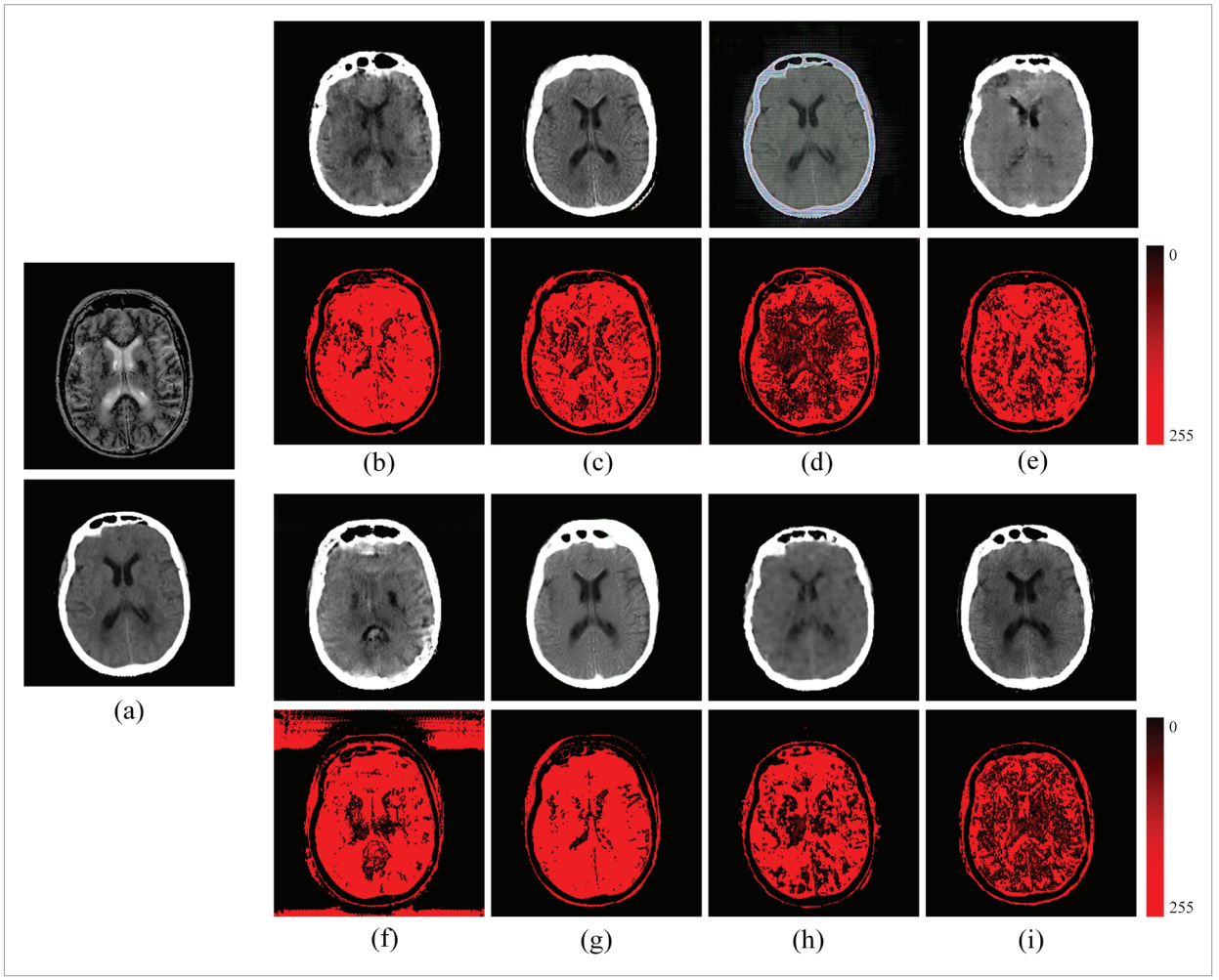


Figure 6

Box & whiskers of the whole test dataset's assessment metrics(MR to CT synthesis) between NAP-CycleGAN and other methods. The title indicates the currently used comparison assessment metric, the horizontal coordinate represents the method name, and the vertical coordinate represents the set of values for this assessment metric for all synthetic CTs of the method on the 77-pair MR-CT paired image test set.

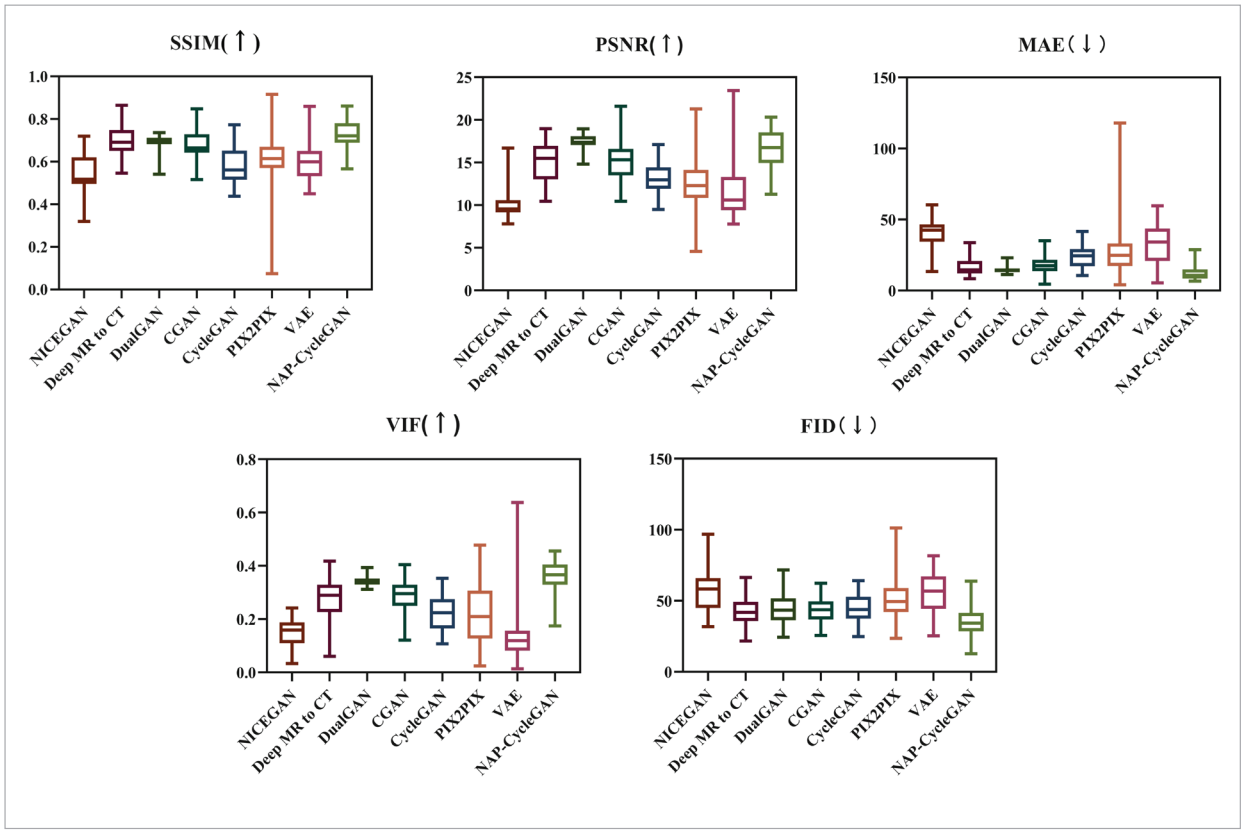


Table 1

Comparison of image average assessment metrics between NAP-CycleGAN and other methods using the whole testing dataset including 77 paired MR-CT images. (The paired samples Wilcoxon signed rank test results: One star represents $p<0.05$, two stars represent $p<0.01$).

Network	SSIM(↑)	PSNR(↑)	MAE(↓)	VIF(↑)	FID(↓)
NICEGAN [6]	0.54±0.09**	10.16±1.69**	40.21±9.49**	0.14±0.04**	57.14±15.10**
Deep MR to CT [43]	0.69±0.06**	15.08±2.25**	16.67±6.26**	0.27±0.07**	42.58±9.61**
DualGAN [47]	0.69±0.03*	17.39±0.75*	14.58±1.92**	0.33±0.01*	44.54±10.42**
CGAN [42]	0.67±0.06**	15.08±2.19**	18.53±5.78**	0.28±0.06**	43.75±8.83**
CycleGAN [52]	0.57±0.07**	13.06±1.50**	23.92±6.90**	0.22±0.05**	44.93±9.69**
Pix2pix [38]	0.61±0.11**	12.61±3.04**	28.43±17.40**	0.21±0.11**	50.98±12.98**
VAE [29]	0.60±0.08**	11.40±2.72**	32.94±13.18**	0.13±0.09**	56.52±14.45**
NAP-CycleGAN	**0.72±0.06**	**16.85±2.20**	**12.49±4.85**	**0.35±0.06**	**35.01±10.97**

SCTs and difference maps of the different methods. To make the differences more obvious as well as to facilitate visualization, we adjusted the difference maps to a black-to-red mode, where the red color in the figure represents the difference between the SCT maps of each method and the reference CT maps. When there are more red parts in the difference maps, it indicates that the SCTs of the model are more different from the RCTs. It is obvious from Figure 5 that the SCTs generated using Deep MR to CT, CGAN, CycleGAN, pix2pix and VAE are more different from the RCTs, and the textures in the SCTs are not the same, with more red parts in the difference maps, while the SCTs generated by DualGAN method and NAP- CycleGAN are closer to the RCT, with the least red parts in the difference maps. After further visual analysis we found that the SCTs of NAP-CycleGAN are visually closer to the RCTs compared to DualGAN, The SCT of NAP-CycleGAN produces clearer and more realistic texture features compared to DualGAN, almost the same as the RCTs, and the difference map has less red parts than it, which have more clarity and homogeneity. Based on the consideration of assessing the quality of model-generated images from various aspects, we separately assessed the quality of SCTs generated by all the above methods.

The comparison results are shown in Figure 6 and Table 1. Figure 6 shows the set of five assessment metrics values for 77 SCTs generated by testing NAP-CycleGAN with seven state-of-the-art methods on the test set containing 77 pairs of MR-CT paired images, and in horizontal coordinates they are arranged in order from left to right: NICEGAN, MR to CT, DualGAN, CGAN, CycleGAN, Pix2pix, VAE. By looking at the data distribution of Figure 6, we can initially find that Nap-CylceGAN performs best, with better values of the assessment metrics for the 77 SCTs than the other methods in general. The distribution of the FID values as well as the MAE values are significantly lower than the values of the other methods, while the lower MAE values indicate that the details of the generated images are better preserved, which helps the doctors to identify and diagnose the lesion areas more accurately. For medical image generation, it indicates that the generated image is close to the real image in terms of high-level features, which improves the credibility of the generated image in the clinic. It shows that our method generates images with more clinical applications.

In order to further confirm the differences of the NAP-CycleGAN compared to other methods, we calculated the mean value of the five evaluation metrics of these SCTs respectively, as shown in Table 1, where the worst metrics values are the NICEGAN, the pix2pix, and VAE with SSIM values of 0.54, 0.61 and 0.60, PSNR values of 10.16, 12.61 and 11.40, MAE values of 40.21, 28.43 and 32.94, respectively. Compared to the other models' results, although not all indicators are at the lowest level, most of the metrics have worse result values than those of the other methods, which suggests that the other models are more effective. It is obvious that the SCTs generated by the NAP-CycleGAN method in the table have an average SSIM value of 0.72 HU, an average PSNR value of 16.85, an average MAE value of 12.49, VIF of 0.35, and FID of 35.01, which gives a more stable quality assessment results compared to the other models that generate better results. It can be seen that NAP-CycleGAN shows better performance in SCT map generation, both in terms of the visual sensation of the generated images and the image evaluation metrics. Higher SSIM indicates that NAP-CycleGAN exhibits stronger structural fidelity compared to other methods, higher PSNR indicates that the details of the generated images are better preserved, which helps the doctor to better observe small lesions when processing high-resolution images, and lower MAE indicates that the SCT is closer with a small pixel gap to the reference CT. We can see this from the SCT as well as the disparity map in Figure 5, where the SCT of NAP-CycleGAN is closer to the reference CT in terms of structural similarity, pixels and realism, containing skeletal contours as well as muscular tissue structural shapes. After using statistical analysis we have also found that the paired samples Wilcoxon signed rank test results between the five evaluation metrics and the models of NAP-CycleGAN with brain01 dataset are all $p<0.05$, so there is a significant difference in the evaluation metrics between our models. We also have compared their training computation efficiency, and we find that the training duration of CGAN, CycleGAN and VAE is the shortest, with 11 hours,11 hours and 10 hours, the NAP-CycleGAN is about 12 hours, but the Deep MR to CT cost longest duration with almost 1.5

hours per epoch, with the other models have run for about the same 20 hours.

3.2. Ablation Experiments and Sensitivity Analysis

For the purpose of using the pix2pixHD generator structure, the embedded Gaussian noise attention used over the spatial attention, and the effectiveness of the loss function as proposed in this paper, an ablation study is conducted. This ablation mainly focuses on the spatial attention, the channel attention, the noise discriminator, and the discriminator, the advantages of the channel generator structure, excluding the feature loss function and the cycle consistency function, and all the ablation experiments use the same dataset for training and testing with the same settings. Figure 7 shows the results of the ablation experiments, from the generated SCTs, it can be seen that the SCTs generated by NAP-CycleGAN are the closest to the RCTs, the detailed texture is also clearer and closer to the RCTs, and the red part is the least in the difference maps, it can be found that no matter which module is embedded into the model alone, it can not get better SCTs, only by combining

Figure 7

Comparison of Synthetic CTs and difference maps from ablation experiments of different components using the whole testing dataset including 77 paired MR-CT images. (a) Original MR and reference CT(No maps), (b) Noise-Cycle: CycleGAN using noise discriminator, (c) Spatten-Cycle: CycleGAN using spatial attention, (d) Channel-Cycle: CycleGAN using channel attention, (e) Noise-pix: pix2pixHD generator with noise discriminator, (f) Spatten-pix: pix2pixHD with the spatial attention, (g) Chaatten-pix: pix2pixHD with the channel attention, (h) NAP-Nonoise: NAP-CycleGAN with the spatial attention and the channel attention, (i) NAP-NoChaatten: NAP-CycleGAN without channel attention, (j) NAP-Spatten: NAP-CycleGAN with the spatial attention, and (k) NAP-CycleGAN.

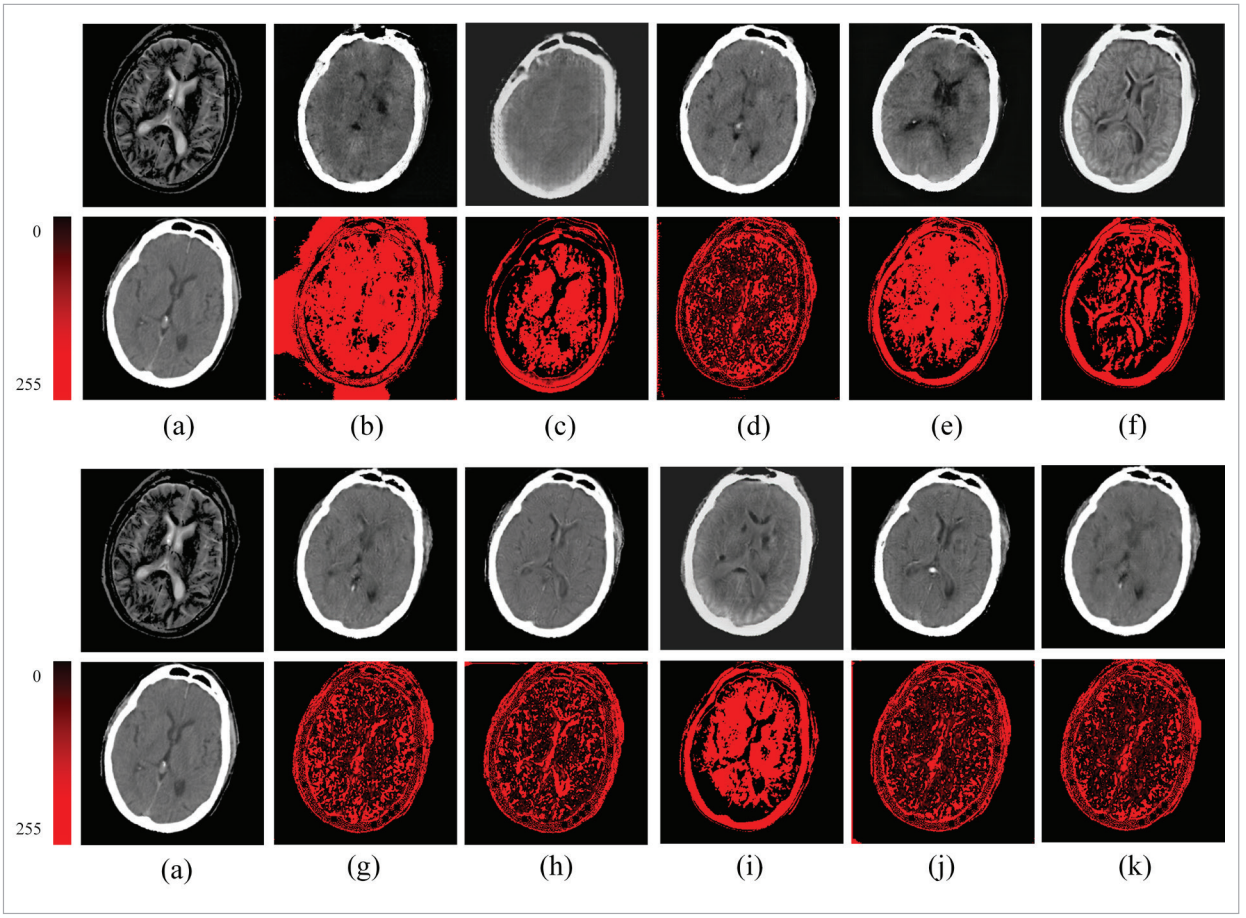


Figure 8

Box & whiskers of the test dataset's assessment metrics(MR to CT synthesis) from ablation experiments of different components. The title indicates the currently used comparison assessment metric, the horizontal coordinate represents the method name, and the vertical coordinate represents the set of values for this assessment metric for all synthetic CTs of the method on the 77-pair MR-CT paired image test set.

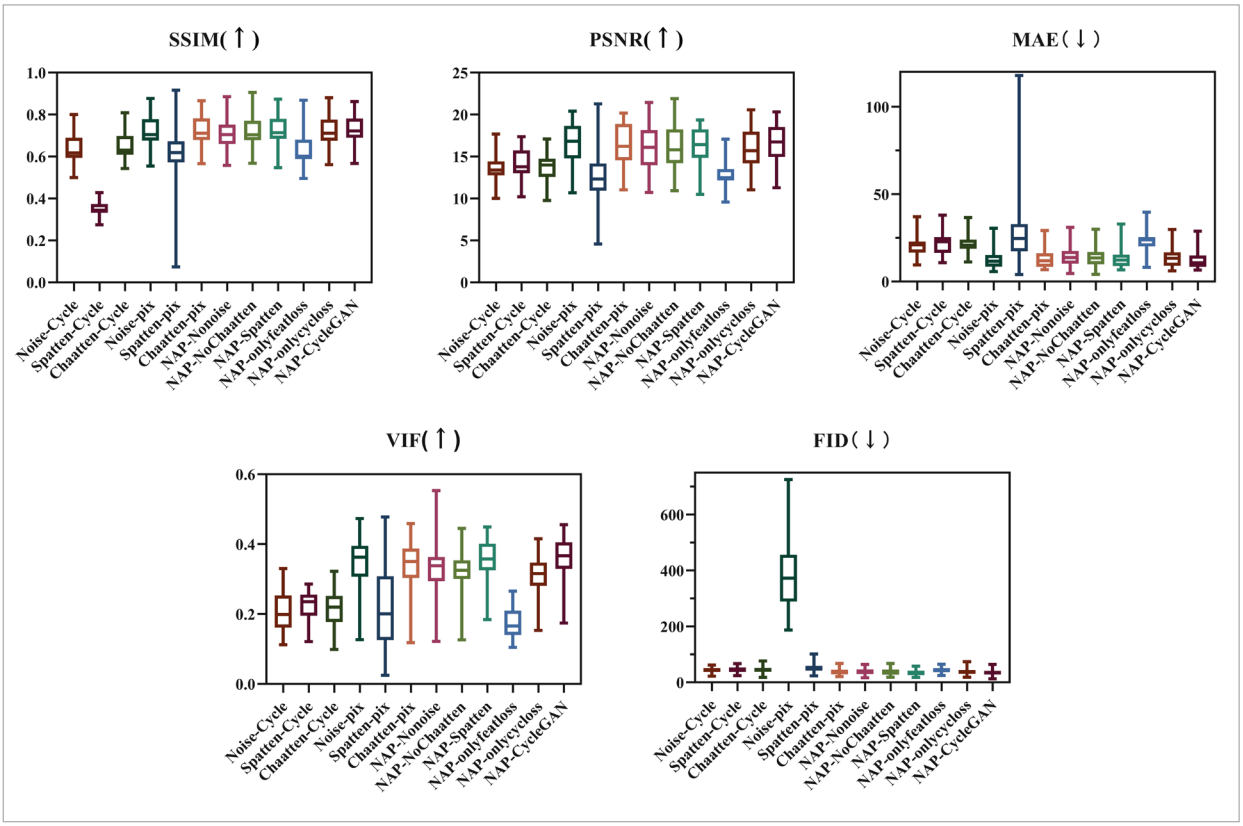


Table 2

Comparison of image average assessment metrics for ablation experiments using the whole testing dataset including 77 paired MR-CT images. (The paired samples Wilcoxon signed rank test results: One star represents $p<0.05$, two stars represent $p<0.01$).

Network	SSIM(↑)	PSNR(↑)	MAE(↓)	VIF(↑)	FID(↓)
Noise-Cycle	0.63±0.06**	13.44±1.28**	20.92±5.21**	0.21±0.05**	44.08±9.33**
Spatten-Cycle	0.34±0.03**	14.05±1.65**	21.81±5.76**	0.22±0.03**	45.77±10.22**
Chaatten-Cycle	0.64±0.05**	13.63±1.50**	21.92±4.66**	0.21±0.04**	45.19±10.67**
Noise-pix	0.71±0.07**	16.52±2.38*	12.90±5.32*	0.34±0.07*	37.64±10.86**
Spatten-pix	0.61±0.11**	12.66±3.02**	28.13±17.31**	0.21±0.11**	51.06±13.04**
Chaatten-pix	0.71±0.06**	16.42±2.39*	13.05±5.14**	0.33±0.07**	39.09±10.77**
NAP-Nonoise	0.70±0.06**	15.96±2.41**	14.67±5.31**	0.33±0.08**	38.67±11.49*
NAP-NoChaatten	0.71±0.06**	16.17±2.35**	14.20±5.20**	0.31±0.05**	37.10±10.81*
NAP-Spatten	0.72±0.06*	16.25±2.06**	13.23±5.28**	0.35±0.06*	34.86±10.04
NAP-CycleGAN	**0.72±0.06**	**16.85±2.20**	**12.49±4.85**	**0.35±0.06**	35.01±10.97

all the modules and the loss function can we get the SCTs with the best quality. The best quality SCT can only be obtained by combining all modules and loss functions without spatial attention.

In addition, the quality of the generated images is also evaluated for the ablation experiments and the results are shown in Figure 8 and Table 2. Firstly, By comparing the specific value distributions of the five metrics for the 77 synthetic samples in the test set in Figure 8, we find that pix2pixHD with the spatial attention (Spatten-pix) exhibits poor data distributions in the boxplots, with the lowest values of SSIM/PSNR/VIF, and the highest distributions of MAE values. FID distribution is also higher. In contrast, the distributions of the five metrics for the NAP-CycleGAN and other ablation component test results performed better in the boxplots, all showing the correct tendencies as they should. In order to further obtain the best experimental combination, we compared the mean and standard deviation of the metric values of the 77 sample test results, as shown in Table 2. It can be seen from Table 2 that in the case of not changing the loss function, just changing the other module models is not as effective as adding a new loss function. The introduction of Gaussian noise can effectively change the quality of generated image evaluation, and the introduction of only the noise discriminator can achieve a better quality of image generation compared to embedding only the spatial attention. If the pix2pixHD structure is introduced into the generator structure alone, the image quality of the SCT is worse than if the module is not introduced together with other modules. It can be seen that in the case of introducing each module alone, only a small change in the quality of the generated image can be achieved, and if the added modules are combined two by two, it may also result in poor SCT generation, especially the use of spatial attention. But if all the modules are embedded and added with the feature matching loss and the cycle consistency loss function, i.e. NAP-CycleGAN, the generated SCT image quality of the CycleGAN model can be greatly improved and a better image performance can be obtained. Besides, the statistical analysis results that the paired samples Wilcoxon signed rank test results between the five evaluation metrics and the models of NAP-CycleGAN with brain01 dataset almost all $p<0.05$, so there is a significant difference in the evaluation metrics between our models. The training computation efficiency is similar, which is during 12-13 hours.

In order to investigate the impact of a truly single loss function on model performance, we do the abla-

Figure 9

Comparison of Synthetic CTs and difference maps from ablation experiments of different losses using the whole testing dataset including 77 paired MR-CT images. (a) Original MR and reference CT(No maps), (b) Noise-onlyfeatloss: NAP-CycleGAN without cycle consistency loss, (c) Noise-onlycycloss: CycleGAN without feature matching loss, and (d) NAP-CycleGAN: feature matching loss with cycle consistency loss and adversarial loss.

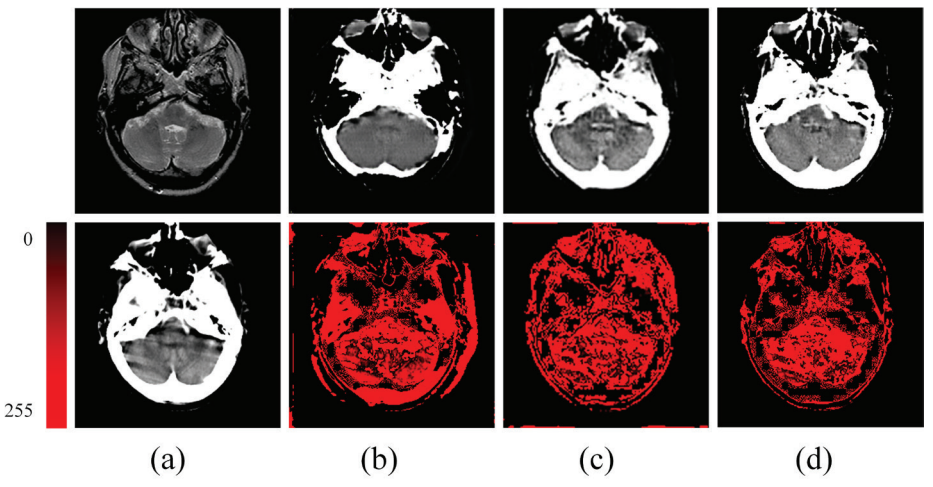


Figure 10

The sensitivity analysis results of MR to CT synthesis in the training process for β(β increased from 0.1 to 1 per 0.1 with other settings the same). The scatter plots from left to right and from top to bottom are: average SSIM values comparison, average PSNR values comparison, average MAE values comparison, average VIF values comparison, and average FID values comparison.

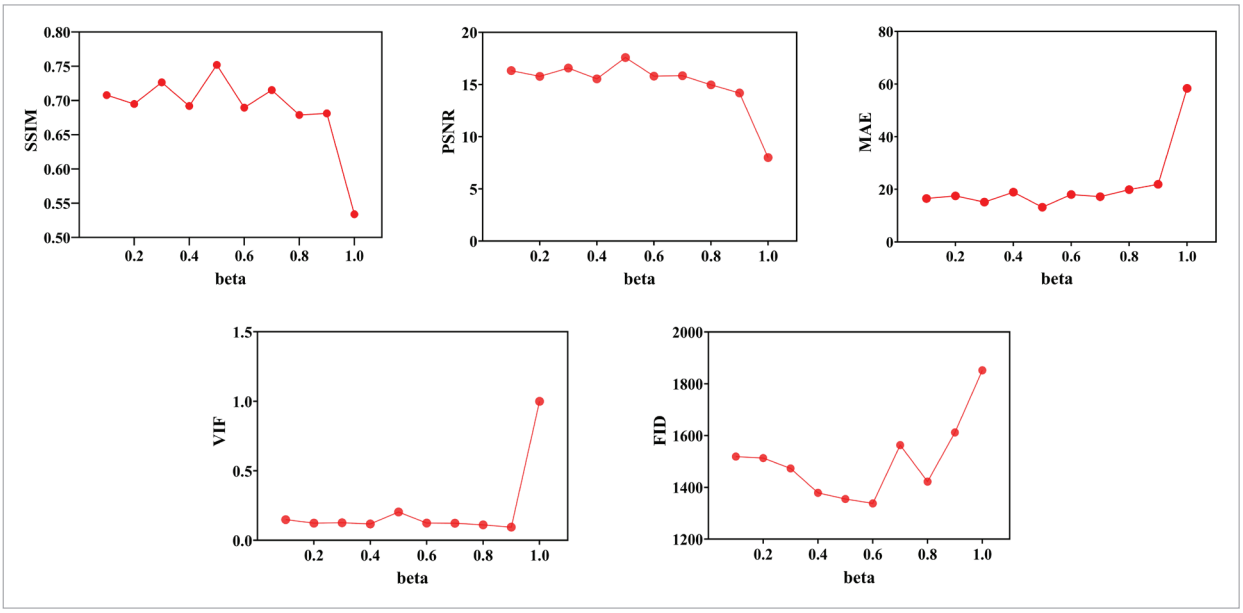


Table 3

Comparison of image average assessment metrics for loss function ablation experiments using the whole testing dataset including 77 paired MR-CT images. (The paired samples Wilcoxon signed rank test results: One star represents p<0.05, two stars represent p<0.01).

Network(ganloss)	SSIM(↑)	PSNR(↑)	MAE(↓)	VIF(↑)	FID(↓)
NAP-onlyfeatloss	0.62±0.06**	12.66±1.28**	23.75±5.54**	0.17±0.04**	43.71±10.41**
NAP-onlycycloss	0.71±0.06**	15.89±2.24**	13.84±5.25**	0.30±0.05**	38.09±11.58*
NAP-CycleGAN	**0.72±0.06**	**16.85±2.20**	**12.49±4.85**	**0.35±0.06**	**35.01±10.97**

tion experiments for the cycle consistency loss and the feature matching loss separately. As is shown in Figures 8-9 and Table 3 are SCTs and difference maps of different losses. From the results, it can be found that on the surface, the SCT of NAP-CycleGAN is not much different from the other generated SCT, but the difference map shows that the red part in the difference map of NAP-CycleGAN is less than the other two components, which means that NAP-CycleGAN is closer to the reference CT. Finally, in order to further validate the model performance of these three methods, we calculated five metrics for each of the three methods: SSIM, PSNR, MAE, VIF, and FID. From Table 3, we can also find that the methods of integrated feature matching loss and cycle consistency loss produce a better performance of the metrics and get better SCT which is closest to RCT compared to others. This is a good indication that the best results are achieved when two loss functions are combined.

Moreover, We have also conducted the sensitivity experiment of NAP-CycleGAN to β. In order to achieve this purpose, we vary β from 0.1 to 1 and increase per 0.1 with other settings that are the same. As the results are shown in Figure 10, The scatter plots from left to right and from top to bottom are: The scatter plots from left to right and from top to bottom is: average SSIM values comparison, average

PSNR values comparison, average MAE values comparison, average VIF values comparison, average FID values comparison. From the scatter plots of 5 metrics, we can find that the NAP-CycleGAN performs best when β is set to 0.5.

4. Discussion

During the above works, we have used NAP-CycleGAN for the task of generating CT from MR, and have good results. NAP-CycleGAN integrates the collection of Channel attention and multiple losses to form a new combination of losses on top of the CycleGAN for the purpose of generating a real CT image that is adversarial and robust. To get the best-performing model, we have done some ablation experiments for analysis and comparison.

The analysis of those shown from Figures 7-10, Table 2, and Table 3 indicates that the performance of NAP-CycleGAN far exceeds the other components. Specifically, the introduction of Gaussian noise reduces the image detail loss, this can be seen from the comparison of metrics values between noise-Cycle and CycleGAN in Tables 1 and 2, where five metrics are significantly improved. By comparing the Spatten-Cycle, Chaatten-Cycle and Noise-Cycle rows of Table 2 and Figure 8, it can be found that the Chaatten-Cycle performs best with the mean value of the SSIM of 0.64, PSNR of 13.63, the VIF of 0.21, and Noise-Cycle has better values of 20.92 and 44.08 on MAE and FID. This indicates that Noise-Cycle performs better in SCT's accuracy and features, but overall Chaatten-Cycle is better in image structure similarity, and detail texture, which can also be found in Figure 7. From here we can find the importance and advantages of the channel attention. In order to explore the optimal combination of pix2pixHD generator with other components, we performed it with six combinations, and the results are shown in Figures 7-8 and Table 2. From noise-pix to Nap-Spatten, the SCT of Spatten-pix performs worst in SSIM/PSNR/MAE/VIF, and the SCT of Noise-pix has the worst FID, which indicates the poor structural similarity for the SCT of Spatten-pix, insufficient fidelity for visual information, and relatively fewer clinical features in Noise-pix. From Figure 7, it can also be found that the SCT of Spatten-pix has too much darkness resulting in blurring of some of the textures, and the SCT of Noise-pix has features that are too general. Comparatively, the SCTs of NAP-Spatten and NAP-CycleGAN have obvious boundaries, the features are obvious with less distortion, and the evaluation metrics are also found to be optimal in Figure 8 and Table 2, especially for NAP-CycleGAN. Moreover, from Figures 9-10 and Table 3, we have found that single loss function does not improve the performance, whereas when we combine them, we get the best SCTs with the most comprehensive clinical information features. Also, some parameters during the experiment also affect the results. As is shown in Figure 10, when the beta is set to 0.5, relatively good metrics are obtained, which means it has an important impact on the model generation. These results suggest that the combination of those components plays a crucial role in the synthesis of SCTs.

To validate the effectiveness of our model, we conducted a comparative study of our approach with several of the more classical methods. From Figure 5, it is obvious that NAP-CycleGAN's SCT with distinct skeletal boundaries has accurate soft tissue structures and boundaries that are closest to the reference CT compared to other models, with clear and accurate tissue texture features at the very boundaries of the skeleton, and the least amount of distortion in the images. In contrast, the SCT of NICEGAN, CGAN, CycleGAN, and VAE was severely distorted, with more clinical details lost, and the difference maps also showed more red parts, the SCT skeletal boundary of pix2pix was too thick, the SCT skeletal boundary of Deep MR to CT showed errors, and the picture was more distorted, and the SCT of DualGAN had darker tones. In Figure 6 and Table 1, the image evaluation metrics SSIM and VIF are higher than those of the other models, with a lower MAE and FID value, even though its PSNR metrics are lower than DualGAN, overall NAP-CycleGAN's metrics outperform other state-of-the-art methods. Since SSIM mainly measures the structural similarity of clinical images, PSNR is used to quantify the image distortion, MAE reflects the image accuracy, VIF reflects the visual information fidelity and clinical texture features of the images, and FID emphasises the degree of clinical features of the generated images, which is in complete agreement with the above-mentioned features of the SCT image perfor-

mance of NAP-CycleGAN, indicating that NAP-CycleGAN's SCT is superior to other methods both in terms of visual analysis and performance analysis.

In the research, the dataset we used is paired data, but the SCTs of other methods still have problems like blurring of details and mismatch of detailed features, while the feature extraction of the NAP-CycleGAN is more comprehensive. This further demonstrates the superiority of NAP-CycleGAN for clinical medical image translation.

5. Conclusions

In this research, the NAP-CycleGAN is proposed to finish the process of MR-to-CT synthesis according to the brain MR-CT data. It is a novel model that replaced the CycleGAN generator with pix2pixHD, introduces the Channel attention into generator incorporates the Gaussian Noise to integrate the noise discriminator, and uses the cycle consistency loss and the feature matching loss, which combines the advantages of detail feature extraction of Channel attention and the advantages of loss function loss computation and reduction of gradient vanishing solves the problems of detail distortion and instability of generated images in previous CycleGAN network generated images.

Experimental results show that NAP-CycleGAN has obtained the best evaluation results in terms of the evaluation metrics and difference maps, and generates better CT images in terms of image quality and realism compared to state-of-the-art methods.

The limitation of the NAP-CycleGAN is that it is trained using images from public datasets and lacks more practical and clinical validation. Our future work focuses on researching deep learning models that can perform multimodality not limited to MR-CT, integrating with clinics as much as possible, exploring the possibility of obtaining more actual clinical medical datasets for research, as well as whether it is possible to perform 3D MR-CT image conversion and completing more relevant statistical analyses to refine the model performance such as more clinically-oriented metrics or expert evaluations.

Acknowledgement

This work was financially supported by the National Natural Science Foundation of China [grant number 81971505].

References

1. Abeer, A., Nawaf, A. Generating Synthetic Images for Healthcare with Novel Deep Pix2Pix GAN. Electronics, 2022, 11(21), 3470. https://doi.org/10.3390/electronics11213470
2. Arabi H., Dowling J.A., Burgos N., Han X., Greer P.B., Koutsouvelis N., Zaidi H. Comparative Study of Algorithms for Synthetic CT Generation from MRI: Consequences for MRI-Guided Radiation Planning in the Pelvic Region. Medical Physics, 2018, 45(11), 5218-5233. https://doi.org/10.1002/mp.13187
3. Armanious K., Jiang C., Fischer M., Küstner T., Hepp T., Nikolaou K., Gatidis S., Yang B. MedGAN: Medical Image Translation Using GANs. Computerized Medical Imaging and Graphics, 2020, 79, 101684. https://doi.org/10.1016/j.compmedimag.2019.101684
4. Boulanger M., Nunes J.-C., Chourak H., Largent A., Tahri S., Acosta O., De Crevoisier R., Lafond C., Barateau A. Deep Learning Methods to Generate Synthetic CT from MRI in Radiotherapy: A Literature Review. Physica Medica, 2021, 89, 265-281. https://doi.org/10.1016/j.ejmp.2021.07.027
5. Brou Boni K. N. D., Klein J., Vanquin L., Wagner A., Lacornerie T., Pasquier D., Reynaert N. MR to CT Synthesis with Multicenter Data in the Pelvic Area Using a Conditional Generative Adversarial Network. Physics in Medicine and Biology, 2020, 65(7), 075002. https://doi.org/10.1088/1361-6560/ab7633
6. Chen, R., Huang, W., Huang, B., Sun, F., Fang, B. Reusing Discriminators for Encoding: Towards Unsupervised Image-to-Image Translation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, 8168-8177. https://doi.org/10.1109/CVPR42600.2020.00819
7. Chen, X., Cao, Y., Zhang, K., Wang, Z., Xie, X., Wang, Y., Men, K., Dai, J. Technical Note: A Method to Synthesize Magnetic Resonance Images in Different Patient Rotation Angles with Deep Learning for Gantry-Free Radiotherapy. Medical Physics, 2023, 50(3), 1746-

1755. https://doi.org/10.1002/mp.15981

8. Chen, Y., Yang, X.-H., Wei, Z., Heidari, A. A., Zheng, N., Li, Z., Chen, H., Hu, H., Zhou, Q., Guan, Q. Generative Adversarial Networks in Medical Image Augmentation: A Review. Computers in Biology and Medicine, 2022, 144, 105382. https://doi.org/10.1016/j.compbiomed.2022.105382
9. Deng, L., Ji, Y., Huang, S., Yang, X., Wang, J. Synthetic CT Generation from CBCT Using Double-Chain-CycleGAN. Computers in Biology and Medicine, 2023, 161, 106889. https://doi.org/10.1016/j.compbiomed.2023.106889
10. Ding, S., Zheng, J., Liu, Z., Zheng, Y., Chen, Y., Xu, X., Lu, J., Xie, J. High-Resolution Dermoscopy Image Synthesis with Conditional Generative Adversarial Networks. Biomedical Signal Processing and Control, 2021, 64, 102224. https://doi.org/10.1016/j.bspc.2020.102224
11. Emami, H., Aliabadi, M. M., Dong, M., Chinnam, R. B. SPA-GAN: Spatial Attention GAN for Image-to-Image Translation. IEEE Transactions on Multimedia, 2020, 23, 391-401. https://doi.org/10.1109/TMM.2020.2975961
12. Emami, H., Dong, M., Nejad-Davarani, S. P., Glide-Hurst, C. K. Generating Synthetic CTs from Magnetic Resonance Images Using Generative Adversarial Networks. Medical Physics, 2018, 45(8), 3627-3636. https://doi.org/10.1002/mp.13047
13. Estakhraji, S. I. Z., Pirasteh, A., Bradshaw, T., McMillan, A. On the Effect of Training Database Size for MR-Based Synthetic CT Generation in the Head. Computerized Medical Imaging and Graphics, 2023, 107, 102227. https://doi.org/10.1016/j.compmedimag.2023.102227
14. Fan, C., Lin, H., Qiu, Y. U-Patch GAN: A Medical Image Fusion Method Based on GAN. Journal of Digital Imaging, 2023, 36(1), 339-355. https://doi.org/10.1007/s10278-022-00696-7
15. Gong, C., Huang, Y., Luo, M., Cao, S., Gong, X., Ding, S., Yuan, X., Zheng, W., Zhang, Y. Channel-Wise Attention Enhanced and Structural Similarity Constrained CycleGAN for Effective Synthetic CT Generation from Head and Neck MRI Images. Radiation Oncology, 2024, 19(1), 37. https://doi.org/10.1186/s13014-024-02429-2
16. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. Generative Adversarial Nets. Communications of the ACM, 2020, 63(11), 139-144. https://doi.org/10.1145/3422622
17. Han, X. MR-Based Synthetic CT Generation Using a Deep Convolutional Neural Network Method. Medical Physics, 2017, 44(4), 1408-1419. https://doi.org/10.1002/mp.12155
18. Huang, Y., Bian, S., Li, H., Wang, C., Li, K. DS-UNet: A Dual Streams UNet for Refined Image Forgery Localization. Information Sciences, 2022, 610, 73-89. https://doi.org/10.1016/j.ins.2022.08.005
19. Huynh, T., Gao, Y., Kang, J., Wang, L., Zhang, P., Lian, J. Estimating CT Image from MRI Data Using Structured Random Forest and Auto-Context Model. IEEE Transactions on Medical Imaging, 2016, 35(1), 174-183. https://doi.org/10.1109/TMI.2015.2461533
20. Jiang, J., Hu, Y.-C., Tyagi, N., Zhang, P., Rimner, A., Deasy, J.O., Veeraraghavan, H. Cross-Modality (CT-MRI) Prior Augmented Deep Learning for Robust Lung Tumor Segmentation from Small MR Datasets. Medical Physics, 2019, 46(10), 4392-4404. https://doi.org/10.1002/mp.13695
21. Jin, C.-B., Kim, H., Jung, W., Joo, S., Park, E., Saem, A.Y., Han, I.H., Lee, J.I., Cui, X. Deep CT to MR Synthesis Using Paired and Unpaired Data. Sensors (Basel), 2019, 19(10), 2361. https://doi.org/10.3390/s19102361
22. Johnstone, E., Wyatt, J.J., Henry, A.M., Short, S.C., Sebag-Montefiore, D., Murray, L., Kelly, C.G., McCallum, H.M., Speight, R. Systematic Review of Synthetic Computed Tomography Generation Methodologies for Use in Magnetic Resonance Imaging-Only Radiation Therapy. International Journal of Radiation Oncology, Biology, Physics, 2018, 100(1), 199-217. https://doi.org/10.1016/j.ijrobp.2017.08.043
23. Kearney, V., Zeimer, B.P., Perry, A., Wang, T., Chan, J., Yom, S.S., Solberg, T.D. Spatial Attention Gated Variational Autoencoder Enhanced Cycle-Consistent Generative Adversarial Networks for MRI to CT Translation. International Journal of Radiation Oncology, Biology, Physics, 2019, 105(1), E720-E721. https://doi.org/10.1016/j.ijrobp.2019.06.897
24. Keereman, V., Fierens, Y., Broux, T., De Deene, Y., Lonneux, M., Vandenberghe, S. MRI-Based Attenuation Correction for PET/MRI Using Ultrashort Echo Time Sequences. Journal of Nuclear Medicine, 2010, 51(5), 812-818. https://doi.org/10.2967/jnumed.109.065425
25. Kim, T., Cha, M., Kim, H., Lee, J. K., Kim, J. Learning to Discover Cross-Domain Relations with Generative Adversarial Networks. International Conference on Machine Learning: PMLR, 2017, p. 1857-1865.

26. Klages, P., Benslimane, I., Riyahi, S., Jiang, J., Hunt, M., Deasy, J.O., Veeraraghavan, H., Tyagi, N. Patch-Based Generative Adversarial Neural Network Models for Head and Neck MR-Only Planning. Medical Physics, 2020, 47(2), 626-642. https://doi.org/10.1002/mp.13927

27. Lan, L., You, L., Zhang, Z., Fan, Z., Zhao, W., Zeng, N., Chen, Y., Zhou, X. Generative Adversarial Networks and Its Applications in Biomedical Informatics. Frontiers in Public Health, 2020, 8, 164. https://doi.org/10.3389/fpubh.2020.00164

28. Largent, A., Barateau, A., Nunes, J.-C., Lafond, C., Greer, P.B., Dowling, J.A., Saint-Jalmes, H., Acosta, O., de Crevoisier, R. Pseudo-CT Generation for MRI-Only Radiation Therapy Treatment Planning: Comparison Among Patch-Based, Atlas-Based, and Bulk Density Methods. International Journal of Radiation Oncology, Biology, Physics, 2019, 103(2), 479-490. https://doi.org/10.1016/j.ijrobp.2018.10.002

29. Li, F., Huang, W., Luo, M., Zhang, P., Zha, Y. A New VAE-GAN Model to Synthesize Arterial Spin Labeling Images from Structural MRI. Displays, 2021, 70, 102079. https://doi.org/10.1016/j.displa.2021.102079

30. Li, S., Zhao, X. High-Resolution Concrete Damage Image Synthesis Using Conditional Generative Adversarial Network. Automation in Construction, 2023, 147, 104739. https://doi.org/10.1016/j.autcon.2022.104739

31. Liu, Z., Wang, S., Dong, D., Wei, J., Fang, C., Zhou, X., Sun, K., Li, L., Li, B., Wang, M., Tian, J. The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. Theranostics, 2019, 9(5), 1303-1322. https://doi.org/10.7150/thno.30309

32. Paudyal, R., Shah, A. D., Akin, O., Do, R. K. G., Konar, A. S., Hatzoglou, V., Mahmood, U., Lee, N., Wong, R. J., Banerjee, S., Shin, J., Veeraraghavan, H., Shukla-Dave, A. Artificial Intelligence in CT and MR Imaging for Oncological Applications. Cancers (Basel), 2023, 15(9), 2573. https://doi.org/10.3390/cancers15092573

33. Reddy, S. Generative AI in Healthcare: An Implementation Science Informed Translational Path on Application, Integration and Governance. Implement Sci, 2024, 19(1), 27. https://doi.org/10.1186/s13012-024-01357-9

34. Spadea, M. F., Pileggi, G., Zaffino, P., Salome, P., Catana, C., Izquierdo-Garcia, D., Amato, F., Seco, J. Deep Convolution Neural Network (DCNN) Multiplane Approach to Synthetic CT Generation from MR Images-Application in Brain Proton Therapy. International Journal of Radiation Oncology, Biology, Physics, 2019, 105(3), 495-503. https://doi.org/10.1016/j.ijrobp.2019.06.2535

35. Szmul, A., Taylor, S., Lim, P., Cantwell, J., Moreira, I., Zhang, Y., D'Souza, D., Moinuddin, S., Gaze, M. N., Gains, J. Deep Learning Based Synthetic CT from Cone Beam CT Generation for Abdominal Paediatric Radiotherapy. Physics in Medicine and Biology, 2023, 68(10), 105006. https://doi.org/10.1088/1361-6560/acc921

36. Tan, Y., Patel, R. V., Wang, Z., Luo, Y., Chen, J., Luo, J., Chen, W., Mao, Z., Huang, R. Y., Wang, H., Bi, W. L., Yao, S. Generation and Applications of Synthetic Computed Tomography Images for Neurosurgical Planning. Journal of Neurosurgery, 2024,141(3),742-751. https://doi.org/10.3171/2024.1.JNS232196

37. Tie, X., Lam, S.-K., Zhang, Y., Lee, K.-H., Au, K.-H., Cai, J. Pseudo-CT Generation from Multi-Parametric MRI Using a Novel Multi-Channel Multi-Path Conditional Generative Adversarial Network for Nasopharyngeal Carcinoma Patients. Medical Physics, 2020, 47(4), 1750-1762. https://doi.org/10.1002/mp.14062

38. Toda, R., Teramoto, A., Kondo, M., Imaizumi, K., Saito, K., Fujita, H. Lung Cancer CT Image Generation from a Free-Form Sketch Using Style-Based Pix2Pix for Data Augmentation. Sci Rep, 2022, 12(1), 12867. https://doi.org/10.1038/s41598-022-16861-5

39. Vukovic, D., Ruvinov, I., Antico, M., Steffens, M., Fontanarosa, D. Automatic GAN-Based MRI Volume Synthesis from US Volumes: A Proof of Concept Investigation. Scientific Reports, 2023, 13(1), 21716. https://doi.org/10.1038/s41598-023-48595-3

40. Wang, C., Yang, G., Papanastasiou, G., Tsaftaris, S. A., Newby, D. E., Gray, C., Macnaught, G., MacGillivray, T. J. DiCyc: GAN-Based Deformation Invariant Cross-Domain Information Fusion for Medical Image Synthesis. Information Fusion, 2021, 67, 147-160. https://doi.org/10.1016/j.inffus.2020.10.015

41. Wang, J., Wu, Q. M. J., Pourpanah, F. DC-CycleGAN: Bidirectional CT-to-MR Synthesis from Unpaired Data. Computerized Medical Imaging and Graphics, 2023, 108, 102249. https://doi.org/10.1016/j.compmedimag.2023.102249

42. Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J., Catanzaro, B. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 8798-8807. https://doi.org/10.1109/CVPR.2018.00917

43. Wolterink, J. M., Dinkla, A. M., Savenije, M. H., Seevinck, P. R., van den Berg, C. A., Išgum, I. Deep MR to CT Synthesis Using Unpaired Data. Simulation and Synthesis in Medical Imaging: Second International Workshop, SASHIMI 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 10, 2017, Proceedings 2: Springer, 2017, 14-23. https://doi.org/10.1007/978-3-319-68127-6_2

44. Xiao, Q., Monfaredi, R., Musa, M., Cleary, K., Chen, Y. MR-Conditional Actuations: A Review. Annals of Biomedical Engineering, 2020, 48(12), 2707-2733. https://doi.org/10.1007/s10439-020-02597-8

45. Xue, Y., Zhang, Y., Neri, F. A Method Based on Evolutionary Algorithms and Channel Attention Mechanism to Enhance Cycle Generative Adversarial Network Performance for Image Translation. International Journal of Neural Systems, 2023, 33(05), 2350026. https://doi.org/10.1142/S0129065723500260

46. Yang, H., Sun, J., Carass, A., Zhao, C., Lee, J., Prince, J. L. Unsupervised MR-to-CT Synthesis Using Structure-Constrained CycleGAN. IEEE Trans Med Imaging, 2020, 39(12), 4249-4261. https://doi.org/10.1109/TMI.2020.3015379

47. Yi, Z., Zhang, H., Tan, P., Gong, M. DualGAN: Unsupervised Dual Learning for Image-to-Image Translation. Proceedings of the IEEE International Conference on Computer Vision, 2017, 2849-2857. https://doi.org/10.1109/ICCV.2017.310

48. Yu, B., Zhou, L., Wang, L., Shi, Y., Fripp, J., Bourgeat, P. EA-GANs: Edge-Aware Generative Adversarial Networks for Cross-Modality MR Image Synthesis. IEEE Transactions on Medical Imaging, 2019, 38(7), 1750-1762. https://doi.org/10.1109/TMI.2019.2895894

49. Zhang, F., Zhao, H., Li, Y., Wu, Y., Sun, X. CBA-GAN: Cartoonization Style Transformation Based on the Convolutional Attention Module. Computers and Electrical Engineering, 2023, 106, 108575. https://doi.org/10.1016/j.compeleceng.2022.108575

50. Zhao, B., Cheng, T., Zhang, X., Wang, J., Zhu, H., Zhao, R., Li, D., Zhang, Z., Yu, G. CT Synthesis from MR in the Pelvic Area Using Residual Transformer Conditional GAN. Computerized Medical Imaging and Graphics, 2023, 103, 102150. https://doi.org/10.1016/j.compmedimag.2022.102150

51. Zhao, S., Geng, C., Guo, C., Tian, F., Tang, X. SARU: A Self-Attention ResUNet to Generate Synthetic CT Images for MR-Only BNCT Treatment Planning. Medical Physics, 2023, 50(1), 117-127. https://doi.org/10.1002/mp.15986

52. Zhu, J.-Y., Park, T., Isola, P., Efros, A. A. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. Proceedings of the IEEE International Conference on Computer Vision, 2017, 2223-2232. https://doi.org/10.1109/ICCV.2017.244