*Research Article*

# Unpaired Low-Dose CT Denoising Network Based on Cycle-Consistent Generative Adversarial Network with Prior Image Information

**Chao Tang, Jie Li, Linyuan Wang** (iD)**, Ziheng Li, Lingyun Jiang, Ailong Cai, Wenkun Zhang, Ningning Liang, Lei Li, and Bin Yan** (iD)

*PLA Strategy Support Force Information Engineering University, Zhengzhou, Henan Province 450001, China*

Correspondence should be addressed to Bin Yan; ybspace@hotmail.com

The widespread application of X-ray computed tomography (CT) in clinical diagnosis has led to increasing public concern regarding excessive radiation dose administered to patients. However, reducing the radiation dose will inevitably cause server noise and affect radiologists' judgment and confidence. Hence, progressive low-dose CT (LDCT) image reconstruction methods must be developed to improve image quality. Over the past two years, deep learning-based approaches have shown impressive performance in noise reduction for LDCT images. Most existing deep learning-based approaches usually require the paired training dataset which the LDCT images correspond to the normal-dose CT (NDCT) images one-to-one, but the acquisition of well-paired datasets requires multiple scans, resulting the increase of radiation dose. Therefore, well-paired datasets are not readily available. To resolve this problem, this paper proposes an unpaired LDCT image denoising network based on cycle generative adversarial networks (CycleGAN) with prior image information which does not require a one-to-one training dataset. In this method, cyclic loss, an important trick in unpaired image-to-image translation, promises to map the distribution from LDCT to NDCT by using unpaired training data. Furthermore, to guarantee the accurate correspondence of the image content between the output and NDCT, the prior information obtained from the result preprocessed using the LDCT image is integrated into the network to supervise the generation of content. Given the map of distribution through the cyclic loss and the supervision of content through the prior image loss, our proposed method can not only reduce the image noise but also retain critical information. Real-data experiments were carried out to test the performance of the proposed method. The peak signal-to-noise ratio (PSNR) improves by more than 3 dB, and the structural similarity (SSIM) increases when compared with the original CycleGAN without prior information. The real LDCT data experiment demonstrates the superiority of the proposed method according to both visual inspection and quantitative evaluation.

## 1. Introduction

X-ray computed tomography (CT) is one of the most significant imaging modalities in modern hospitals and clinics. However, the risk of radiation in CT induces genetic, cancerous, and other diseases and has become a critical concern to patients and operators [1–3]]. A common and effective strategy to alleviate the risk is to achieve low-dose CT (LDCT) imaging by reducing the tube current during scanning and consequently decreasing the number of photons received by the detector. The dose reduction increases noise and artifacts in reconstructed CT images, thereby severely degrading the image quality and jeopardizing the clinical diagnosis. To solve this problem, researchers have proposed various noise-reduction strategies, including iterative reconstruction (IR) [4, 5], sinogram domain denoising [6–9], and image domain postprocessing [10–12].

Over the past decades, researchers have focused on developing new iterative algorithms for LDCT image reconstruction. In general, these algorithms optimize an objective function, which incorporates a system model [13, 14],

a statistical noise model, and prior information in the image domain [4, 15, 16]. Well-known image priors consist of total variation (TV) and its variants [17–19], dictionary learning [20, 21], and wavelet frame [22]. These iterative reconstruction algorithms exhibit satisfactory performance in improving image quality, but their computational burden and sensitive parameters limit their practical applications.

Image postprocessing is more computationally efficient compared with iterative reconstruction, which has spawned a lot of simple and effective methods. Nonlocal means (NLM) filtering methods estimate noise components by using multiple patches extracted at different locations in the image [23] and have been widely used for CT [24]. Motivated by compressed sensing methods, an adaptive K-SVD method [25] was proposed to reduce artifacts in CT reconstructions. The block matching 3D (BM3D) method is also an outstanding method for image postprocessing in CT imaging fields [26, 27]; this method exploits similarities in image blocks. These traditional postprocessing methods have improved the quality of CT images; however, the results often undergo edge blurring and/or residual artifacts given the nonuniform distribution of reconstruction noise.

More recently, several supervised machine learning approaches have been proposed for noise reduction in LDCT. These methods usually reveal a relation between the pixel value in the LDCT image and the pixel value at the same location in a corresponding NDCT image by training with paired images. Chen et al. [28] designed a deep convolutional neural network (CNN) to map LDCT images toward its relative normal-dose counterparts in a patch-by-patch manner. Kang et al. [29] used a similar method but adopted CNN to directional wavelet transform of CT images. Then, more complex networks were proposed to handle the LDCT denoising problem such as the residual encoder-decoder convolutional neural network (RED-CNN) in [30], which achieves competitive performance relative to state-of-the-art methods in clinical cases.

Although the abovementioned networks presented impressive denoising results, they all belong to the end-to-end network, which typically utilizes mean squared errors (MSE) between the network output and the ground truth as loss function. However, recent studies [31, 32] indicated that this per-pixel MSE often suffers from oversmoothed edges and loss of details. MSE-based approaches tend to take the mean of high-resolution patches by using Euclidean distance rather than geodesic distance. Given that the medical images usually lie in a highly nonlinear manifold [33], the algorithm is prone to neglect subtle details that are vital for clinical diagnosis when it tries to minimize per-pixel MSE. To overcome the limitations of per-pixel regression in noise reduction, the generative adversarial network (GAN) [34] based on adversarial loss is introduced to medical image reconstruction. In 2017, Wolterink et al. [35] were the first to apply the GAN for cardiac CT image reconstruction. And, Yang et al. [36] utilized a GAN with Wasserstein distance (WGAN). In order to enhance the capability of noise reduction, perceptual loss is simultaneously used to optimize the loss function. Yi and Babyn [37] combined an adversarial trained network and a sharpness detection network to mitigate noise in LDCT and achieved satisfactory performance. Hence, a general framework for estimating generative models by using an adversarial process has shown outstanding performance in medical image reconstruction.

The above-mentioned denoising networks usually require spatially paired counterparts. However, in medical imaging, well-paired counterparts are difficult to obtain. For example, in LDCT imaging, continuously scanning a patient twice in normal and low dose is impossible under normal circumstances. The shortage of paired data has been one of the factors that restrict the wide application of deep learning in low-dose CT reconstruction. Recently, unsupervised variants of GANs, such as CycleGAN [38] and DualGAN [39], have been proposed for mapping different domains without matching data pairs. Motivated by their success in image processing, unpaired GANs have been successfully applied to CS-MRI reconstruction [40] and CT synthesis based on MR images [41, 42]. For LDCT reconstruction, well-paired clinical scans acquired at different dose levels are not readily available. Even if we obtain the same patient data at different dose levels, the data are difficult to match perfectly due to physical activity and the inevitable slight movement of the scanning position, which may affect the denoising ability of the networks.

In this study, we propose an unpaired LDCT denoising network based on CycleGAN with prior image information. In the proposed network, the design of cycle-consistent structure impels the network to learn the mapping relationship between the LDCT image collection and NDCT image collection (Figure 1(a)), rather than an image pair of an LDCT image and NDCT image (Figure 1(b)). Therefore, the proposed network does not need a one-to-one corresponding training dataset and can learn with the unpaired dataset. Meanwhile, the prior image information extracted from the preprocessed image by using LDCT is introduced into the network to supervise the generation of content and ensure correspondence of the image content. The map of image collections through cyclic loss and the supervision of content through prior image loss confer our proposed method to produce results that have not only lower noise but also accurate details.

## 2. Methods

*2.1. Noise Reduction Model.* In LDCT imaging, serious noise typically occurs in CT images as the number of photons received by the detector decreases. One of the effective ways to improve the image quality is designing a tailored network to make the input LDCT images as close as possible to the NDCT images. This process can be classified as an image denoising problem, which can be described by the following model:

$$G : x \longrightarrow y, \tag{1}$$

where $x \in R^{N \times N}$ denotes an LDCT image and $y \in R^{N \times N}$ denotes the corresponding NDCT image. The goal of the noise reduction process is to obtain a transformation $G$ that maps $x$ to $y$.
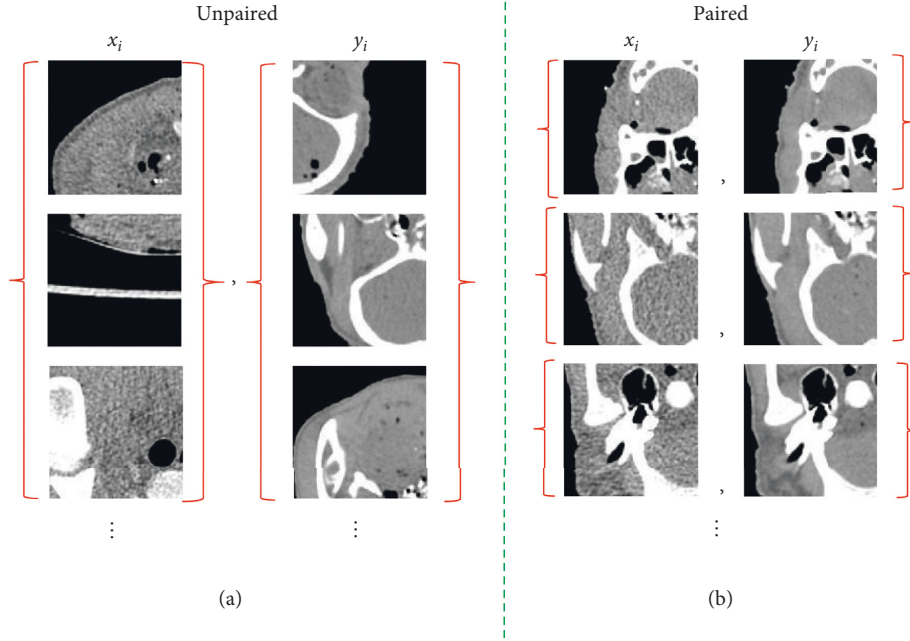
FIGURE 1: Schematic diagram of (a) unpaired dataset and (b) paired dataset.

In this process, $x$ can be seen as a sample from the LDCT distribution $P_l$ and $y$ can be seen as a sample from the NDCT distribution $P_n$. The denoising process transforms $x$ to a certain distribution $P_g$. And, the denoising process aims to determine an optimal $G$ to make $P_g$ close to $P_n$. However, in the reconstructed LDCT images, noise is complicated and uniformly distributed over the whole image; as such, distributions $P_l$ and $P_n$ have no explicit mathematical relationship up to date [36]. The traditional methods usually have difficulty in denoising LDCT images. For deep learning-based methods, the uncertainty of a noise model can be ignored due to the learning capability of high-level features and presentation of data distribution by the CNN. Therefore, designing a tailored CNN is an effective method to suppress noise in LDCT and improve image quality.

*2.2. Introduction of CycleGAN.* In 2017, Zhu et al. [38] proposed an unpaired network named CycleGAN, which has gained extensive attention. This network can capture the special characteristics of one image collection and figure out how these characteristics could be translated into other image collection without using any paired training examples; this network has been successfully utilized in style transfer, object transfiguration, season transfer, and photo enhancement.

Under the assumption that there is some underlying relationship between the source domain $X$ and target domain $Y$, the goal of CycleGAN is to learn mapping $G : X \longrightarrow Y$ so that the distribution of image from $G(x)$ is indistinguishable from the distribution of image from domain $Y$. This network includes two mapping functions, namely, $G : X \longrightarrow Y$ and $F : Y \longrightarrow X$ and also introduces two discriminators, namely, $D_X$ and $D_Y$. Discriminator $D_Y$

aims to distinguish between translated samples $G(x)$ and real samples $y$. Discriminator $D_X$ aims to distinguish between translated samples $F(y)$ and real samples $x$. In theory, adversarial training can identify mappings $G$ and $F$ that produce outputs identically distributed as target domains $Y$ and $X$ [34]. However, with its large sufficient capacity, a network can map the same set of input images to any random permutation of images in the target domain, where any of the learned mappings can induce an output distribution that matches the target distribution. Therefore, adversarial loss alone cannot guarantee that the learned function can map individual input $x$ to desired $y$. To further reduce the space of possible mapping functions, the mapping functions $G$ and $F$ should be cycle consistent. As shown in Figure 2(a), for each image $x$ from domain $X$, the image translation cycle should be able to bring $x$ back to the original image: $x \longrightarrow G(x) \longrightarrow F(G(x)) \approx x$, which is named as forward cycle consistency. Similarly, as demonstrated in Figure 2(b), for each image $y$ from domain $Y$, the image translation cycle should be able to bring $y$ back to the original image: $y \longrightarrow F(y) \longrightarrow G(F(y)) \approx y$, which is named as backward cycle consistency. The abovementioned behavior can be incentivized by cycle-consistency loss.

$$
\begin{aligned}
L_{\text{cyc}}(G, F) = & \, \mathbb{E}_{x \sim P_{\text{data}}(x)}\left[\|F(G(x)) - x\|_1\right] \\
& + \mathbb{E}_{y \sim P_{\text{data}}(y)}\left[\|G(F(y)) - y\|_1\right],
\end{aligned}
\tag{2}
$$

where $P_{\text{data}(x)}$ is the distribution of $x$ and $P_{\text{data}}(y)$ is the distribution of $y$.

In LDCT imaging, although the projection data contain a lot of noise, it is usually complete. Therefore, the reconstructed images still contain useful information that is basically consistent with corresponding NDCT images. This indicates that there is a close relationship between the LDCT
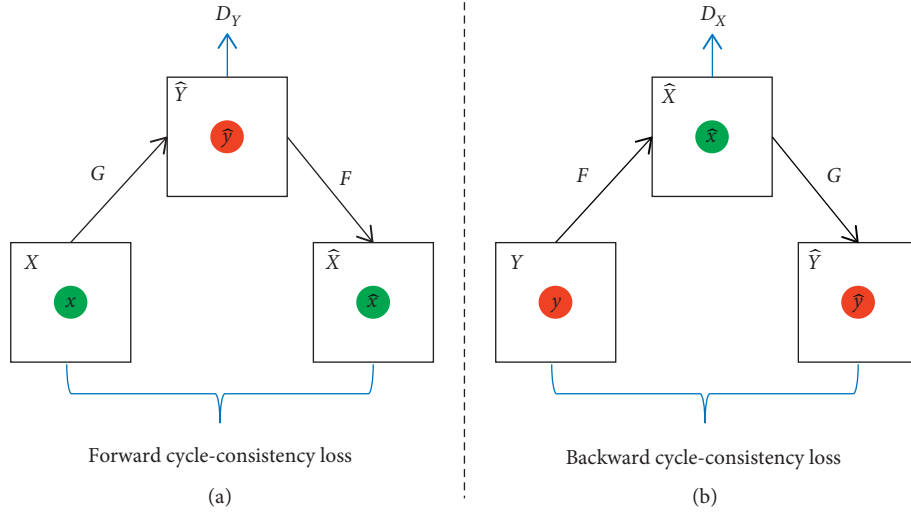
FIGURE 2: Structure diagram of CycleGAN. (a) Forward cycle-consistency loss. (b) Backward cycle-consistency loss.

image and NDCT image and satisfies the basic assumption of CycleGAN. Thus, this study considers using this unpaired network for LDCT image reconstruction.

Based on its structure, CycleGAN mainly focuses on the map of distributions. This network is better at the overall conversion of images and may overlook the correspondence of details. However, for LDCT noise reduction, the outputs should not only appear similar to an NDCT image but also retain details as much as possible; more importantly, the output must not contain false information, which may cause misdiagnosis. These issues require additional supervision and restraint to the network in the training process to make it more suitable for LDCT reconstruction. In this study, we consider incorporating prior image information into the network to guarantee content correspondence and prevent the generation of fake details during denoising process.

### 2.3. Unpaired Denoising Network Based on CycleGAN with Prior Information.

Figure 3 shows an overview of our proposed method. The network contains forward and backward cycles and two generators and discriminators. In order to more clearly illustrate the training mechanism of the proposed network, we randomly selected an LDCT image from LDCT image collection and an NDCT image from the NDCT image collection as the input-ground truth pair. Note that the LDCT image is not corresponding to the NDCT image, as shown in Figure 4.

In the forward cycle, generator $G_N$ is trained to generate images that are as close to corresponding NDCT images as possible (Figure 3(a)). Generator $F_L$ is trained to translate the resulting image $G_N(x)$ back to the corresponding LDCT image. In the backward cycle, generator $F_L$ is trained to generate images that are as close to corresponding LDCT images as possible (Figure 3(b)). Generator $G_N$ is trained to translate the resulting image $F_L(y)$ back to the NDCT image. In the training process of network, the discriminators $D_N$ and $D_L$ are used to estimate the probability that the sample is

from the real image rather than generating image. At the same time, the generators $G_N$ and $F_L$ attempt to generate images that are not easily distinguishable by the discriminators. This paper utilizes the adversarial loss [43] as the objective function to train the "game" process:

$$
\begin{aligned}
L_{\mathrm{GAN}_N}(G_N, D_N, X, Y) &= \mathbb{E}_{y \sim p_{\mathrm{data}}(y)}[\log D_N(y)] \\
&\quad + \mathbb{E}_{x \sim p_{\mathrm{data}}(x)}[\log(1 - D_N(G_N(x)))], \\
L_{\mathrm{GAN}_L}(F_L, D_L, X, Y) &= \mathbb{E}_{x \sim p_{\mathrm{data}}(x)}[\log D_L(x)] \\
&\quad + \mathbb{E}_{y \sim p_{\mathrm{data}}(y)}[\log(1 - D_L(F_L(y)))].
\end{aligned}
\tag{3}
$$

In order to reduce the feasible domain space of the mapping functions, the cycle consistency is introduced to further constrain the training process of the network so that the network can be trained under unpaired data:

$$
\begin{aligned}
L_{\mathrm{cyc}}(G_N, F_L) &= \mathbb{E}_{x \sim p_{\mathrm{data}}(x)}\left[\left\|F_L(G_N(x)) - x\right\|_1\right] \\
&\quad + \mathbb{E}_{y \sim p_{\mathrm{data}}(y)}\left[\left\|G_N(F_L(y)) - y\right\|_1\right].
\end{aligned}
\tag{4}
$$

For this cycle consistency network, the performance of a generator requires the indirect supervision through the results of another generator. For example, in the forward cycle, in addition to discriminator $D_N$, the performance of generator $G_N$ also needs to be supervised by the results $F_L(G_N(x))$ of another generator $F_L$. This mechanism does not guarantee the accuracy of the final input and may lead to the occurrence of fake details. For LDCT image reconstruction, the accuracy of the results is critical, and if false information is generated, it may cause misdiagnosis, leading to serious consequences. These circumstances require direct constraints to the generators, especially to generator $G_N$, which directly produces the desired outcome. Therefore, this paper incorporates the prior image information into the network to directly supervise the generator $G_N$, as shown in Figure 3(a). For fear of changing the unpaired property of the network, the resulting images
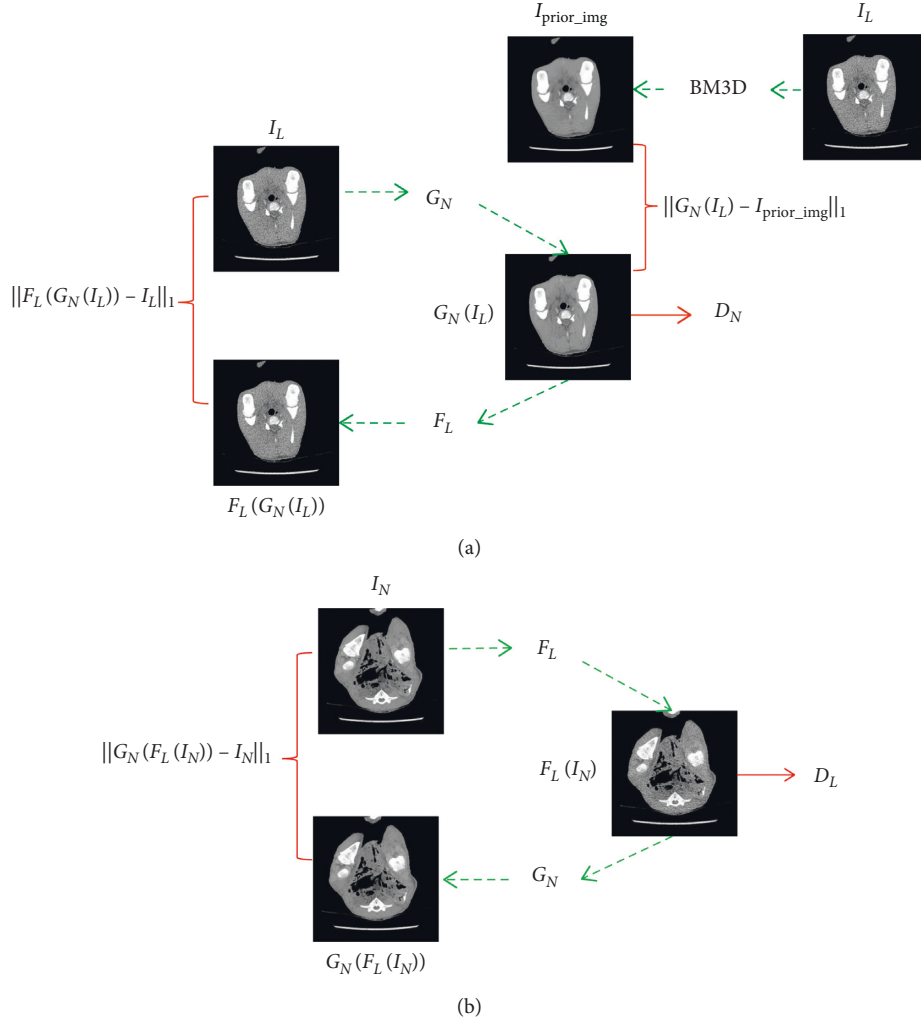
(a)



(b)

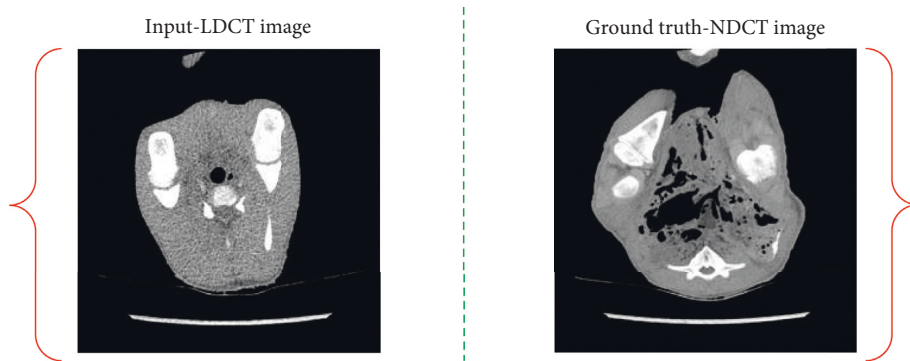FIGURE 3: Overview of the proposed method. (a) Forward cycle. (b) Backward cycle.



FIGURE 4: One input-ground truth pair of the proposed unpaired network.

processed by BM3D method utilizing LDCT images are regarded as prior images. And, the mean absolute error (MAE) between the prior image and the generated image is introduced into the loss function to constrain the training of the network. The adversarial loss of the forward cycle can be written as

$$
\begin{aligned}
L_{\mathrm{GAN}_N}^p \left(G_N, D_N, X, Y\right) = {} & \mathbb{E}_{y \sim p_{\mathrm{data}}(y)}\left[\log D_N\left(y\right)\right] \\
& + \mathbb{E}_{x \sim p_{\mathrm{data}}(x)}\Big[\log\left(1 - D_N\left(x\right)\right) \\
& + \alpha\left\|G_N\left(x\right) - I_{\mathrm{prior\_img}}\right\|_1\Big],
\end{aligned}
\tag{5}
$$

where $\alpha$ is the weight of the MAE. In the training process, the generator $G_N$ tries to minimize the objective function (5) while the discriminator $D_N$ tries to maximize it, that is, $\min_{G_N} \max_{D_N} L^p_{\mathrm{GAN}_N}(G_N, D_N, X, Y)$. We denote the proposed network as CycleGAN-BM3D.

Through the above analysis, the total loss function of our proposed method is

$$
\begin{aligned}
L(G_N, F_L, D_N, D_L) = {} & L^p_{\mathrm{GAN}_N}(G_N, D_N, X, Y) \\
& + L_{\mathrm{GAN}_L}(F_L, D_L, Y, X) \qquad (6) \\
& + \lambda L_{\mathrm{cyc}}(G, F),
\end{aligned}
$$

where $\lambda$ is a nonnegative parameter used to balance the weight of the cycle consistency loss. The total objective function of the proposed CycleGAN-BM3D is

$$
G_N^*, F_L^* = \arg \min_{G_N, F_L} \max_{D_N, D_L} L(G_N, F_L, D_N, D_L). \qquad (7)
$$

*2.4. Network Architecture.* In the proposed method, the network of the generator, as shown in Figure 5(a), includes three submodules: encoder, convertor, and decoder. The encoder extracts the features from the input image utilizing CNN. The network of the encoder includes one $7 \times 7$ Convolution-InstanceNorm-ReLU layer with 64 filters and stride 1 denoted as c7s1-64, two $3 \times 3$ Convolution-InstanceNorm-ReLU layers with $k$ filters and stride 2 in which $k$ equals to 128 and 256, respectively. We denote these two layers as d128 and d256. The convertor, as shown in Figure 5(b), is used to convert feature vectors extracted from the source domain $X$ to the target domain $Y$. The convertor contains six residual network (Resnet) blocks [44], and each block contains two $3 \times 3$ convolutional layers with 256 filters on both layers. The decoder includes three layers. The first two layers are $3 \times 3$ fractional-strided-Convolution-InstanceNorm-ReLU layers with stride 1/2 and 64 and 32 filters, respectively. We denote the two layers as u64 and u32. The third layer is a $7 \times 7$ Convolution-InstanceNorm-ReLU with 3 filters and stride 1 denoted as c7s1-3. The network of the discriminator, as shown in Figure 5(c), consists of five convolution layers. The first four layers are $4 \times 4$ Convolution-InstanceNorm-LeakyReLU layers with stride 2 and 64, 128, 256, and 512 filters, respectively. We denote them as C64, C128, C256, and C512. We use leaky ReLUs with slope 0.2. In the last layer, a $4 \times 4$ convolution layer with 1 filter and stride 1 is utilized to produce a one-dimensional output.

## 3. Experiments and Results

*3.1. Experimental Dataset and Performance Evaluations.* The CT images of a deceased piglet were selected as the experimental dataset to verify the performance of the proposed network. The images were scanned by a 64-slice multidetector GE Healthcare scanner (Discovery CT750 HD) by using $100 \, \mathrm{kV}$ and $0.625 \, \mathrm{mm}$ slice thickness. Five different tube currents were set to yield CT images with different dose levels. The specific scanned parameters and effective dose of different tube currents are listed in Table 1.

In each dose level, 906 images with size $512 \times 512$ were acquired. As shown in Figure 6, we partitioned the slices by taking one's data and then skipping 10 slices. We finally obtained 360 images reconstructed by FBP using the projection data of 5% dose as the noisy dataset, that is, source collection $X$. And, 180 images were obtained for testing. The normal dose dataset utilized consists of 360 images, which are corresponding to the noisy dataset and obtained from the NDCT images constructed by FBP. The normal dose dataset is the target collection $Y$. Given the unpaired property of the proposed network, the training images and the ground truth do not require a one-to-one correspondence. The number of two image datasets can also be different. In the training stage, we performed an unpaired operation for the inputs and labels of the network. In addition, each image was divided into sixteen $128 \times 128$ images to enlarge the training dataset to 5760 images.

For comparison, we selected some representative traditional methods and other networks.

*3.1.1. BM3D.* This method exhibits outstanding performance in noise reduction over other traditional image denoising methods.

*3.1.2. Original CycleGAN.* This network is trained without the constraint of prior image to test the supervised effect of the prior image. The input dataset and ground truth dataset are the same as the proposed method CyclGAN-BM3D, which are not one-to-one correspondence.

Structural similarity (SSIM) [45], peak signal-to-noise ratio (PSNR), and normalized mean absolute distance (NMAD) are selected as measures of reconstruction quality for the quantitative assessment of the proposed network and abovementioned contrast algorithms. Specifically, PSNR and NMAD are defined as follows:

$$
\begin{aligned}
\mathrm{PSNR} &= 10 \log_{10}\left( \frac{\mathrm{MAX}^2(f)}{1/N\sum_{i=1}^{N}\left|f(i) - f_0(i)\right|^2} \right), \\
\mathrm{NMAD} &= \frac{\sum_{i=1}^{N}\left|f(i) - f_0(i)\right|}{\sum_{i=1}^{N}\left|f(i)\right|},
\end{aligned} \qquad (8)
$$

where $f$ and $f_0$ represent the denoising image and ideal image, respectively, $i$ is the pixel in the image, and $N$ is the total number of pixels in the image. A higher PSNR indicates that the image is of higher quality. The NMAD value close to 0 indicates small differences between the ideal image and the reconstructed results. In general, $\mathrm{SSIM} \leq 1$ and $\mathrm{SSIM} = 1$ indicate the exact theoretical reconstruction.

*3.2. Implementation Details.* In the training process, the negative log likelihood objective of the first two items in $L_{\mathrm{GAN}_N}$ and $L_{\mathrm{GAN}_L}$ is replaced by least squares loss to stabilize the proposed model [46]. After the replacement, we train $G_N$ to minimize
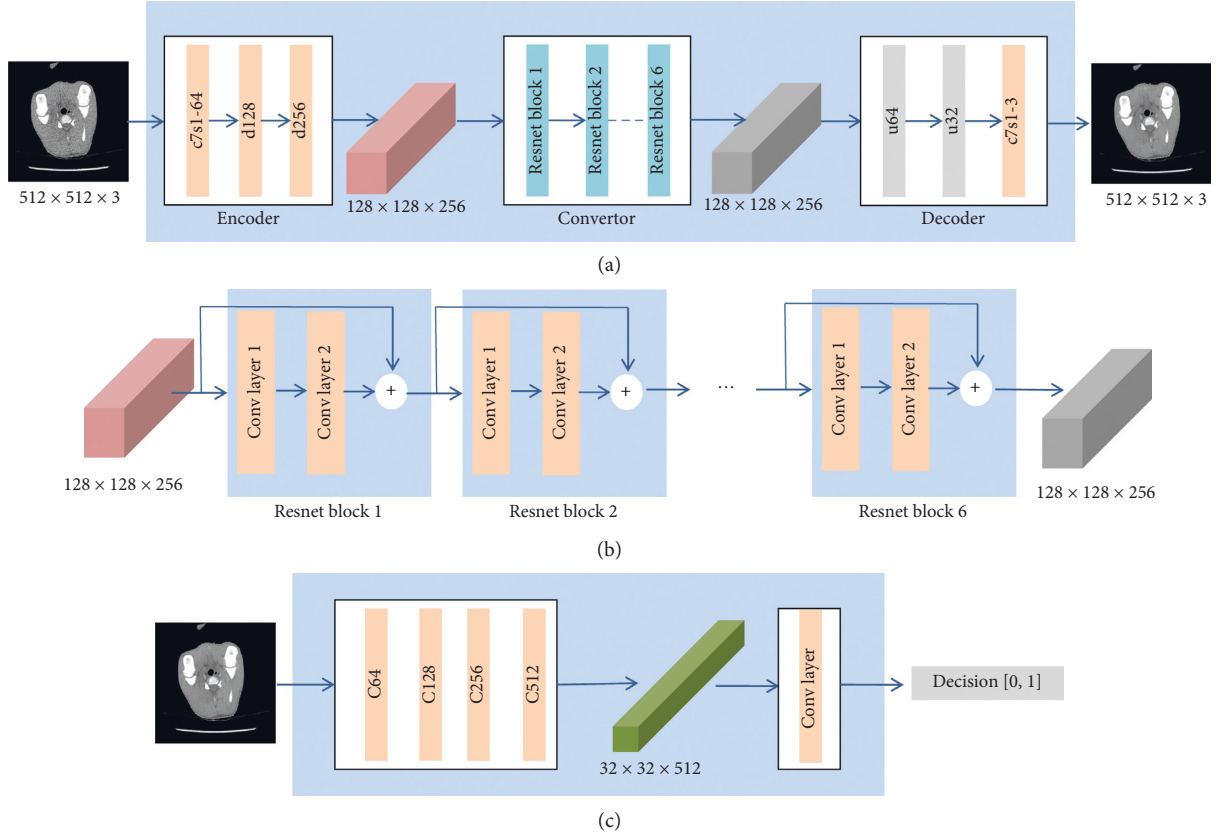
FIGURE 5: Network structure. (a) Diagram of generator's network structure. (b) Diagram of convertor's network structure. (c) Diagram of discriminator's network structure.

TABLE 1: CT scanning protocol for the experiment.

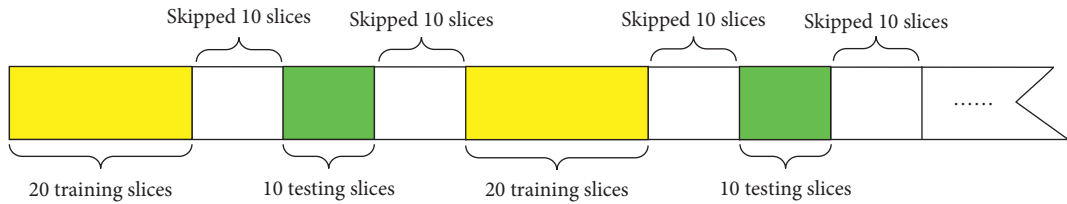|  | Normal dose | 50% dose | 25% dose | 10% dose | 5% dose |
| --- | --- | --- | --- | --- | --- |
| Tube current (mAs) | 300 | 150 | 75 | 30 | 15 |
| Effective dose (mSv) | 14.14 | 7.07 | 3.54 | 1.41 | 0.71 |



FIGURE 6: Schematic diagram of data preparation. The yellow rectangular blocks represent the training data, the green rectangular blocks represent the test data, and the white rectangular blocks represent the skipped slices.

$$\mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\left(D_N\left(G_N\left(x\right)\right)-1\right)^2 + \alpha\left\|G_N\left(x\right)-I_{\text{prior\_img}}\right\|_1\right].$$

$$(9)$$

We also train $F_L$ to minimize

$$\min \mathbb{E}_{y\sim p_{\text{data}}(y)}\left[\left(D_L\left(F_L\left(y\right)\right)-1\right)^2\right].$$

$$(10)$$

For discriminators $D_N$ and $D_L$, the objectives are

$$\min \mathbb{E}_{y\sim p_{\text{data}}(y)}\left[\left(D_N\left(y\right)-1\right)^2\right] + \mathbb{E}_{x\sim p_{\text{data}}(x)}\left[D_N\left(G(x)^2\right)\right],$$

$$\min \mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\left(D_L\left(x\right)-1\right)^2\right] + \mathbb{E}_{y\sim p_{\text{data}}(y)}\left[D_L\left(F_L\left(y\right)^2\right)\right].$$

$$(11)$$

For the setting of parameters, $\lambda$ is set as 10. Adam solver with a batch size of 1 is selected as the optimizer to optimize the networks. We keep the same learning rate for the first

10,000 epochs and linearly decay the rate to zero over the next 10,000 epochs. The proposed network in this paper trained 40,000 epochs. $\alpha$ is an important parameter for controlling the weight of prior information during training. When the value of $\alpha$ is too small, the effect of prior information will be negligible and cause minimal improvement in image quality. By contrast, a very large $\alpha$ will overemphasize the role of prior information and, to some extent, limit the learning ability of the network itself. In this paper, a series of networks was trained by setting different values of $\alpha$ to determine a suitable value. For the sake of fairness, each network has the same parameters' setting, except $\alpha$. We randomly selected 10 LDCT images to test the performance of different networks. The effect of $\alpha$ was quantitatively determined by plotting the average SSIM and NMAD of the denoising images in Figure 7.

From the curves of Figure 7, when $\alpha = 10$, the SSIM reaches the maximum, indicating that the denoising images are most similar to the NDCT images in structure, that is, under this circumstance, the network has the greatest ability to retain the details of the LDCT images. NMAD reflects the accuracy of denoising to some extent. Smaller NMAD indicates that the noise in the LDCT images is removed more completely. Considering the detail retention and noise reduction of the network, this paper set $\alpha$ as 10.

To analyze the potential denoising capability of selected algorithms and networks, two representative slices and the corresponding zoomed regions of interest (ROIs) are shown in Figures 8–10, respectively.

As shown in Figure 8, when using traditional methods (BM3D) or deep learning-based methods, the noise that appears in the LDCT image is suppressed to varying degrees. The classic BM3D method has an outstanding noise suppression effect but makes the processed image oversmoothed so that some vital details disappear. As indicated by the red arrows in ROI, the results of BM3D lose some information of the bone. Although the result of the original CycleGAN is not oversmoothed, it shows fake details, connecting the unconnected bones in the NDCT image. The CycleGAN-BM3D, which introduces prior image information, does not have redundant details and retains the information that should be retained.

Figure 9 shows the overall second slice of different methods and the corresponding absolute difference images between NDCT images and the resulting images. In the difference images, the darker the color is, the smaller the error will be. It can be clearly observed that the result of CycleGAN-BM3D has the smallest difference.

For further analysis of image details, two regions were selected as ROIs, which are shown in Figure 10. In ROI I, the tissue pointed by red squares is smeared out in the BM3D images but is easily identifiable in the CycleGAN and CylceGAN-BM3D images. As marked by the yellow ellipses in ROI II, the three black holes in the results of BM3D and CycleGAN are blurred and inseparable but are recognizable in the result of CycleGAN-BM3D. Furthermore, the smooth area below the ROI II of CycleGAN-BM3D is the most similar to the NDCT image.

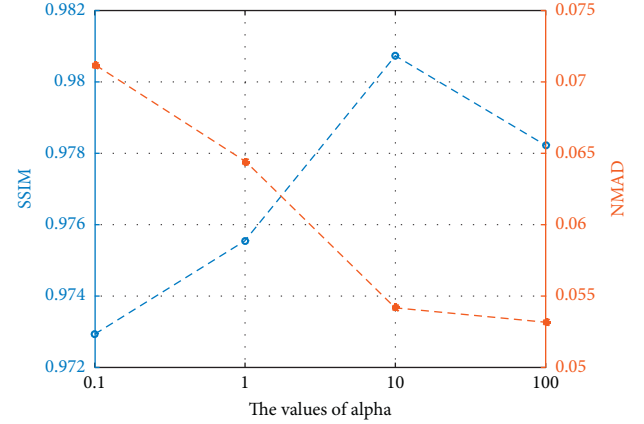Based on the visual effect, the proposed network CycleGAN-BM3D can not only better suppress noise but



FIGURE 7: Average SSIM and NMAD of 10 images in different values of $\alpha$. The blue line indicates the SSIM curve and the orange line represents the NMAD curve.

also retain more details than the other networks. More importantly, after adding the constraint of prior information, CycleGAN-BM3D can effectively prevent the generation of fake details compared with the original CycleGAN without prior information.

For quantitative analysis, the average of PSNR, SSIM, and NMAD was calculated for 180 slices in the test dataset to measure performance of the proposed method and the other compared methods. (Table 2).

In each evaluation item, the results with the best performance are marked black. CycleGAN-BM3D ranks first in terms of SSIM even with the unpaired training dataset. As such, the results of CycleGAN-BM3D are the most structurally similar to the NDCT images. In terms of PSNR and NMAD, CycleGAN-BM3D also exhibits satisfactory performance, indicating that the noise removal is relatively clean. Compared with the other algorithms, CycleGAN shows the worst performance because it mainly focuses on mapping the data distribution from the LDCT to NDCT, and cyclic loss may not be enough to supervise the generation of details and suppression of noise. This method needs to add supervision during training to improve image quality. The introduction of prior information in CycleGAN-BM3D enhances the constraints to the image content. That is, in the proposed CycleGAN-BM3D method, cyclic loss plays a role in distribution mapping and prior information loss is used to guarantee the relevance of the content. Therefore, the proposed method demonstrates good performance in noise suppression and detail preservation. We note that the numerical results of BM3D are in the front rank. From the visual effect, the noise removal of BM3D is complete and the main information of the image is basically retained, as much, BM3D has a high quantitative evaluation result. That finding is the reason why we choose the result of BM3D as the prior information.

## 4. Discussion and Conclusion

In the modern CT imaging field, the hidden risk of radiation dose has increased the demand for LDCT. However, LDCT
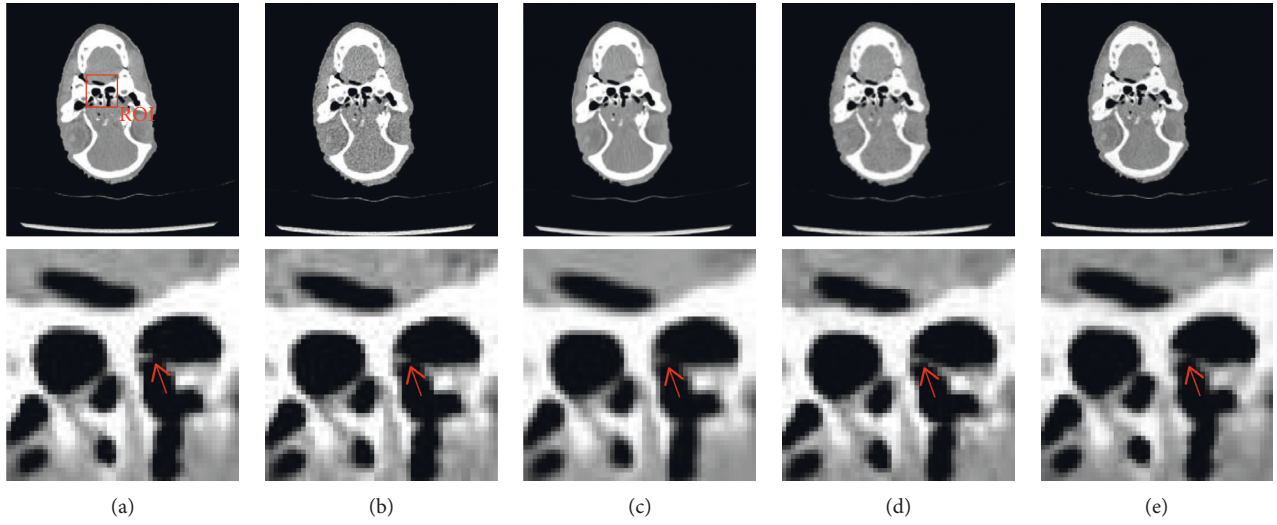
Figure 8: Results of (a) NDCT, (b) LDCT, (c) BM3D, (d) CycleGAN, and (e) CycleGAN-BM3D, respectively. The display widow is [800, 1300].
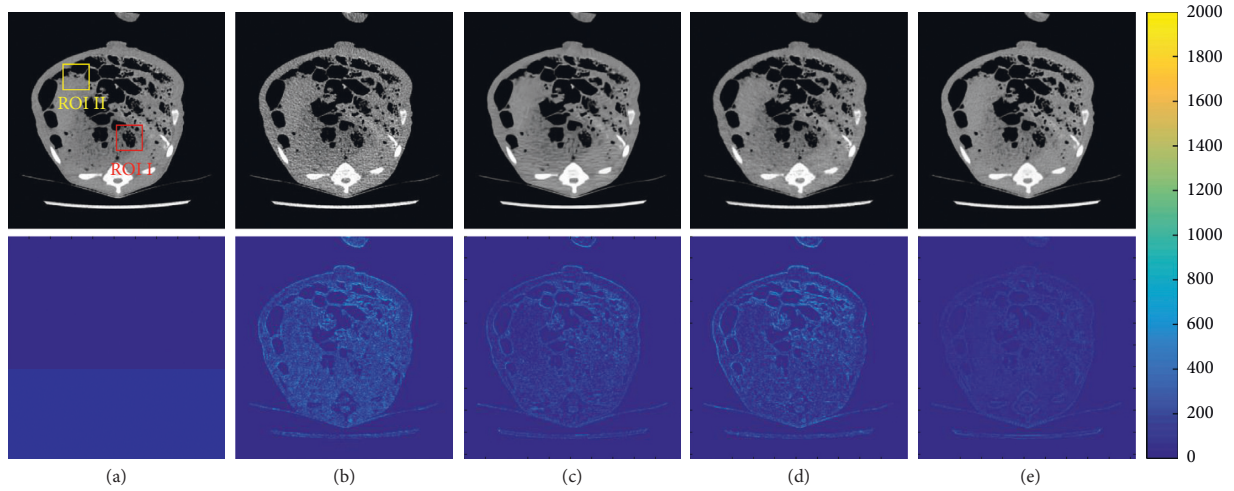


Figure 9: Results of (a) NDCT, (b) LDCT, (c) BM3D, (d) CycleGAN, and (e) CycleGAN-BM3D, respectively. The display widow is [800, 1300].
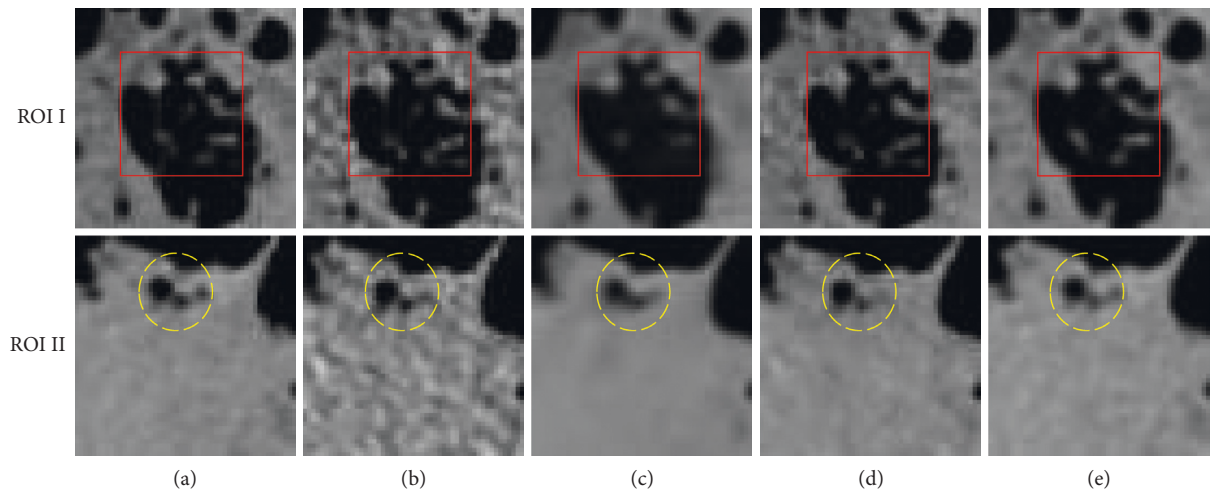


Figure 10: Zoomed-in ROIs of slice 2. The first column is the NDCT (a). The following columns are the results from (b) LDCT, (c) BM3D, (d) CycleGAN, and (e) CycleGAN-BM3D, respectively. The display widow is [800, 1300].

Table 2: Quantitative evaluation of results by different algorithms on 180 slices in the test dataset.

|  | PSNR | SSIM | NMAD |
|---|---|---|---|
| LDCT | $27.6514 \pm 1.3862$ | $0.8698 \pm 0.0375$ | $0.1124 \pm 0.0174$ |
| BM3D | $32.2185 \pm 0.9599$ | $0.9271 \pm 0.0245$ | $0.0681 \pm 0.0087$ |
| CycleGAN | $29.0833 \pm 0.8661$ | $0.9059 \pm 0.0263$ | $0.0890 \pm 0.0096$ |
| CycleGAN-BM3D | $\mathbf{34.3577 \pm 0.1475}$ | $\mathbf{0.9798 \pm 0.0041}$ | $\mathbf{0.0583 \pm 0.0012}$ |

images often suffer from serious noise, which degrades image quality and troubles clinical diagnosis. In the past two years, deep neural network provides a new idea for LDCT noise reduction. Most of the existing neural networks for LDCT reconstruction usually require well-matched datasets for network training. However, the well-matched CT images of different dose levels are difficult to obtain. This may affect the performance of networks and lead to blurred details or fake information in the resulting images.

To improve the quality of LDCT image and broaden the application of neural networks in LDCT noise reduction, this paper proposed an unpaired network based on CycleGAN with prior image information. In contrast to existing denoising networks, the proposed network can be trained by unpaired datasets, thereby alleviating the limitation of paired dataset requirement. Most GANs used to reduce noise mainly focus on the distribution mapping from LDCT to NDCT, and this process may overlook the accurate content correspondence. To enhance the constraint to the content and prevent producing fake details, we incorporate the prior image processed by BM3D into CycleGAN to supervise the generation of image content. In the experiment of real data, visual inspection demonstrated that the proposed method can suppress noise in the LDCT image and prevent the generation of fake details. The result of quantitative evaluations indicated that, after incorporating prior information, the PSNR improved more than 3 dB and SSIM also increased compared with the original CycleGAN without prior information. The results of qualitative and quantitative evaluations indicated that the proposed method exhibits reasonable performance and outperforms the original CycleGAN when applied to LDCT reconstruction.

The validity of prior information affects the performance of the proposed method. In this work, the LDCT images processed by traditional methods are obtained as prior information, which is a simple and efficient way. In the future, we intend to explore other representative shared features between LDCT and NDCT images as prior information to further improve the performance of the proposed network, such as the sparsity or sharpness information.

## Data Availability

The data used in the study can be obtained from http://homepage.usask.ca/~xiy525/publication/sagan/.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Chao Tang and Jie Li contributed equally to this work.

## References

[1] D. J. Brenner and E. J. Hall, "Cancer risks from CT scans: now we have data, what next?," *Radiology*, vol. 265, no. 2, pp. 330-331, 2012.

[2] O. W. Linton and F. A. Mettler, "National conference on dose reduction in CT, with an emphasis on pediatric patients," *American Journal of Roentgenology*, vol. 181, no. 2, pp. 321–329, 2003.

[3] D. J. Brenner and E. J. Hall, "Computed tomography—an increasing source of radiation exposure," *New England Journal of Medicine*, vol. 357, no. 22, pp. 2277–2284, 2007.

[4] I. A. Elbakri and J. A. Fessler, "Statistical image reconstruction for polyenergetic X-ray computed tomography," *IEEE Transactions on Medical Imaging*, vol. 21, no. 2, pp. 89–99, 2002.

[5] J. Wang, T. Li, and L. Xing, "Iterative image reconstruction for CBCT using edge-preserving prior," *Medical Physics*, vol. 36, pp. 252–260, 2009.

[6] P. J. L. Riviere, J. Bian, and P. A. Vargas, "Penalized-likelihood sinogram restoration for computed tomography," *IEEE Transactions on Medical Imaging*, vol. 25, no. 8, pp. 1022–1036, 2006.

[7] J. Wang, T. Li, H. Lu, and Z. Liang, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low dose X-ray computed tomography," *IEEE Transactions on Medical Imaging*, vol. 25, no. 10, pp. 1272–1283, 2006.

[8] Y. Zhang, J. Zhang, and H. Lu, "Statistical sinogram smoothing for low-dose CT with segmentation-based adaptive filtering," *IEEE Transactions on Nuclear Science*, vol. 57, no. 5, pp. 2587–2598, 2010.

[9] Q. Xie, D. Zeng, Q. Zhao et al., "Robust low-dose CT sinogram preprocessing via exploiting noise-generating mechanism," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2487–2498, 2017.

[10] A. M. R. Schilham, B. van Ginneken, H. Gietema, and M. Prokop, "Local noise weighted filtering for emphysema scoring of low-dose CT images," *IEEE Transactions on Medical Imaging*, vol. 25, no. 4, pp. 451–463, 2006.

[11] A. Borsdorf, R. Raupach, T. Flohr, and J. Hornegger, "Wavelet based noise reduction in CT-images using correlation analysis," *IEEE Transactions on Medical Imaging*, vol. 27, no. 12, pp. 1685–1703, 2008.

[12] Y. Chen, Z. Yang, Y. Hu et al., "Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means," *Physics in Medicine and Biology*, vol. 57, no. 9, pp. 2667–2688, 2012.

[13] B. D. Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Physics in Medicine and Biology*, vol. 49, no. 11, pp. 2463–2475, 2004.

[14] R. M. Lewitt, "Multidimensional digital image representations using generalized Kaiser-Bessel window functions," *Journal of

the *Optical Society of America A*, vol. 7, no. 10, pp. 1834–1846, 1990.

[15] B. R. Whiting, P. Massoumzadeh, O. A. Earl, J. A. O'Sullivan, D. L. Snyder, and J. F. Williamson, "Properties of pre-processed sinogram data in X-ray computed tomography," *Medical Physics*, vol. 33, no. 9, pp. 3290–3303, 2006.

[16] S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 677–688, 2012.

[17] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in Medicine and Biology*, vol. 53, no. 17, pp. 4777–4807, 2008.

[18] Y. Liu, J. Ma, Y. Fan, and Z. Liang, "Adaptive-weighted total variation minimization for sparse data toward low-dose X-ray computed tomography image reconstruction," *Physics in Medicine and Biology*, vol. 57, no. 23, pp. 7923–7956, 2012.

[19] Z. Tian, X. Jia, K. Yuan, T. Pan, and S. B. Jiang, "Low-dose CT reconstruction via edge-preserving total variation regularization," *Physics in Medicine and Biology*, vol. 56, no. 18, pp. 5949–5967, 2011.

[20] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 31, no. 9, pp. 1682–1697, 2012.

[21] Y. Zhang, X. Mou, G. Wang, and H. Yu, "Tensor-based dictionary learning for spectral CT reconstruction," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 142–154, 2017.

[22] J. K. Choi, B. Dong, and X. Zhang, "Limited tomography reconstruction via tight frame and simultaneous sinogram extrapolation," *Journal of Computational Mathematics*, vol. 34, no. 6, pp. 575–589, 2016.

[23] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[24] Z. Li, L. Yu, J. D. Trzasko et al., "Adaptive nonlocal means filtering based on local noise level for CT denoising," *Medical Physics*, vol. 41, no. 1, Article ID 011908, 2014.

[25] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[26] P. F. Feruglio, C. Vinegoni, J. Gros, A. Sbarbati, and R. Weissleder, "Block matching 3D random noise filtering for absorption optical projection tomography," *Physics in Medicine and Biology*, vol. 55, no. 18, pp. 5401–5415, 2010.

[27] D. Kang, P. Slomka, R. Nakazato et al., "Image denoising of low-radiation dose coronary CT angiography by an adaptive block-matching 3D algorithm," in *Proceedings of the Medical Imaging: Image Processing*, International Society for Optics and Photonics, March 2013.

[28] H. Chen, Y. Zhang, W. Zhang et al., "aLow-dose CT via convolutional neural network," *Biomedical Optics Express*, vol. 8, no. 2, pp. 679–694, Prague, Czech Republic, 2017.

[29] E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *Medical Physics*, vol. 44, no. 10, pp. e360–e375, 2017.

[30] H. Chen, Y. Zhang, M. K. Kalra et al., "Low-dose CT with a residual encoder-decoder convolutional neural network (RED-CNN)," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.

[31] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proceedings of the European Conference on Computer Vision*, pp. 694–711, Springer, Amsterdam, The Netherlands, October 2016.

[32] C. Ledig, L. Theis, F. Huszar et al., "Photo-realistic single image super-resolution using a generative adversarial network," 2016, https://arxiv.org/abs/1609.04802.

[33] D. Wu, K. Kim, F. G. El et al., "Iterative low-dose CT reconstruction with priors trained by artificial neural network," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2479–2486, 2017.

[34] I. Goodfellow, "Nips 2016 tutorial: generative adversarial networks," 2016, https://arxiv.org/abs/1701.00160.

[35] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Isgum, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.

[36] Q. Yang, P. Yan, Y. Zhang et al., "Low dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1348–1357, 2017.

[37] X. Yi and P. Babyn, "Sharpness-aware low-dose CT denoising using conditional generative adversarial network," *Journal of Digital Imaging*, vol. 31, no. 5, pp. 655–669, 2018.

[38] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2242–2251, Venice, Italy, October 2017.

[39] Z. Yi, H. Zhang, P. Tan, and M. Gong, "DualGAN: unsupervised dual learning for image-to-image translation," in *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2868–2876, Venice, Italy, October 2017.

[40] T. M. Quan, T. Nguyen-Duc, and W. K. Jeong, "Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1488–1497, 2018.

[41] J. M. Wolterink, A. M. Dinkla, M. H. F. Savenije et al., "Deep MR to CT synthesis using unpaired data," in *Simulation and Synthesis in Medical Imaging*, pp. 14–23, Springer, Berlin, Germany, 2017.

[42] C.-B. Jin, H. Kim, W. Jung et al., "Deep CT to MR synthesis using paired and unpaired data," 2018, https://arxiv.org/abs/1805.10790.

[43] I. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," in *Proceedings of the Advances in Neural Information Processing Systems 27 (NIPS 2014)*, Montreal, Quebec, Canada, 2014.

[44] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[45] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[46] X. Mao, Q. Li, H. Xie, R. Y. Lau, and Z. Wang, "Multiclass generative adversarial networks with the L2 loss function," 2016, https://arxiv.org/abs/1611.04076v1.