

Research Article

Low-Dose CT Image Denoising Based on Improved DD-Net and Local Filtered Mechanism

Hongen Liu ¹, Xin Jin ¹, Ling Liu ¹ and Xin Jin ^{1,2}

¹School of Software, Yunnan University, Kunming 650091, Yunnan, China

²Engineering Research Center of Cyberspace, Yunnan University, Kunming 650000, China

Correspondence should be addressed to Xin Jin; xinxin_jin@163.com

Received 20 March 2022; Revised 2 June 2022; Accepted 20 June 2022; Published 3 August 2022

Academic Editor: Laxmi Lydia

Copyright © 2022 Hongen Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Low-dose CT (LDCT) images can reduce the radiation damage to the patients; however, the unavoidable information loss will influence the clinical diagnosis under low-dose conditions, such as noise, streak artifacts, and smooth details. LDCT image denoising is a significant topic in medical image processing to overcome the above deficits. This work proposes an improved DD-Net (DenseNet and deconvolution-based network) joint local filtered mechanism, the DD-Net is enhanced by introducing improved residual dense block to strengthen the feature representation ability, and the local filtered mechanism and gradient loss are also employed to effectively restore the subtle structures. First, the LDCT image is inputted into the network to obtain the denoised image. The original loss between the denoised image and normal-dose CT (NDCT) image is calculated, and the difference image between the NDCT image and the denoised image is obtained. Second, a mask image is generated by taking a threshold operation to the difference image, and the filtered LDCT and NDCT images are obtained by conducting an elementwise multiplication operation with LDCT and NDCT images using the mask image. Third, the filtered image is inputted into the network to obtain the filtered denoised image, and the correction loss is calculated. At last, the sum of original loss and correction loss of the improved DD-Net is used to optimize the network. Considering that it is insufficient to generate the edge information using the combination of mean square error (MSE) and multiscale structural similarity (MS-SSIM), we introduce the gradient loss that can calculate the loss of the high-frequency portion. The experimental results show that the proposed method can achieve better performance than conventional schemes and most neural networks. Our source code is made available at [https://github.com/LHE-IT/Low-dose-CT-Image-Denoising/tree/main/Local Filtered Mechanism](https://github.com/LHE-IT/Low-dose-CT-Image-Denoising/tree/main/Local%20Filtered%20Mechanism).

1. Introduction

Computed tomography (CT) is crucial in medical diagnosis and illness analysis [1–5]. As excessive CT scan probably causes a series of acute potential cancers and cases, the medical instrument usually adopts a CT radiation dose as low as possible to avert the damage to health. However, it is certain that reducing the radiation dose will cause information loss of human tissue, and a large amount of noise in the image may influence the accuracy of the diagnosis. Thus, how to reduce the noise in low-dose CT (LDCT) images and preserve the image information is one of the critical issues in medical image processing.

As reducing the dose of CT radiation can produce the projection data with a low signal-to-noise ratio, some

methods utilize the nonlinear filter [6, 7] or the statistical characteristic of noise [8, 9] to reduce the noise in the projection data. In addition, some methods attempted to remove the noise and streak artifacts in LDCT images directly. To remove the streak artifacts fused in the tissue structure, approaches including nonlinear diffusion filter [10], sparse representation, and dictionary learning [11–13] were proposed. Due to the excellent performance, some natural image denoising algorithms [14, 15] were applied to remove the noise in LDCH images. These methods belong to post-processing methods and aim to preserve the detailed information and remove the noise and artifacts simultaneously.

Due to the remarkable expressive capacity of neural networks, many researchers try to reduce the noise in the LDCT images based on convolution neural network (CNN)

and generative adversarial network (GAN). Some researchers utilized structures, such as autoencoder, residual block, and dense block, into CNN for LDCT image denoising [16–18]. Some researchers used CNN to learn image features and provide prior knowledge for traditional CT reconstruction schemes such as analytic reconstruction and iterative reconstruction [19, 20]. However, it is hard to generate realistic and diverse details based on CNN. To remedy this, some schemes based on GAN [21–28] were proposed. In these GAN-based solutions, the denoised images are generated by the generator and evaluated by the discriminator. Through competition and self-optimization, the generator can generate realistic normal-dose CT (NDCT) images. Besides, some novel structures, such as leap structure [29] and sharpness detector [23], were developed to enhance the performance of neural networks. Although these networks can perfectly remove the noise and recover the image structures, there are still some ambiguous and incorrect subtle structures in the denoised results. Confronting the problem, the frequency-separation-based networks [30, 31] were proposed to separate the LDCT image into a low-frequency portion and a high-frequency portion for generating each portion, respectively. However, the incorrect subtle structures are usually the areas that are contaminated severely by noise and streak artifacts instead of the edge information in high-frequency domain, as shown in Figure 1. On the one side, frequency separation usually produces information loss during transformation. On the other side, there is an essential correlation between global structures and detailed information, which is useful for subtle structure restoration.

To address the above problems, we use the difference image between the predicted image and the NDCT image to make the mask image. Compared with frequency-separation-based methods, the difference image can reflect the areas that contain incorrect subtle structures accurately. Using the mask image obtained by taking a threshold operation to the difference image, the local filtered mechanism filters the high-quality areas in the LDCT image and NDCT image and preserves the low-quality areas, which will be optimized by the network especially. To utilize the correlation between global structures and detailed information, our model learns the global structure between LDCT image and NDCT image in the first step. Then, the model learns the detailed information between the filtered LDCT image and NDCT image in the second step. By the network optimization in the second step, the ability in subtle structure restoration is enhanced significantly. As the global structures and detailed information are learned in the same network, our model can provide the implicit global features for subtle structure restoration and avoid the transformation loss during frequency separation. However, learning two tasks in one network requires a deeper network structure. Therefore, we propose an improved DD-Net [32] that has mighty ability in feature extraction and denoising. In our network, we replace the dense block by an improved residual dense block [33] to deepen the neural network and improve the performance. Although this mechanism solves the subtle structure restoration problem, the combination of MSE and MS-SSIM as a loss function is detrimental for the edge

information restoration. Thus, we introduce the gradient loss [34] to calculate the loss in the high-frequency portion, which improves the performance reduction caused by the local filtered mechanism.

In this work, we evaluate the performance of our proposed network by comparing with other typical schemes. Experimental results demonstrate that the proposed scheme can restore the ambiguous subtle structures brilliantly and gets higher performance in objective metrics than most comparative methods. The contributions of the paper can be listed as follows:

- (i) We propose an improved network based on DD-Net and a novel local filtered mechanism. Through this mechanism, the network can generate the subtle structures accurately with the global context and accomplish the balance between network generalization and subtle structure restoration.
- (ii) We introduce the gradient loss to enhance the ability in edge information restoration and improve the performance reduction caused by the abatement of generalization significantly.
- (iii) Experiments on low-dose chest image and brain image denoising prove that our network outperforms the conventional schemes and most neural networks in both evaluation metric and visual appearance.

2. Related Work

Since we propose an improved neural network for LDCT image denoising, some significant work will be reviewed optionally in this section.

As neural networks have achieved brilliant performance in image processing, LDCT denoising methods based on the neural network have been presented in the past decades. These methods achieved outstanding results in both objective metrics and visual appearance. Although analytic reconstruction and iterative reconstruction are still the mainstream in commerce scenes, it is certain that neural networks will be applied in commercial CT equipment in the future.

2.1. CNN Methods. Due to the powerful ability of feature extraction and mapping, some researchers attempted to reduce the noise in the LDCT images based on CNN. The deep residual network and cascade network [16–18] were the early applications to improve the performance of LDCT denoising. Zhang et al. [32] combined the dense block and deconvolution structure to build a lightweight network that can reuse the features effectively. Some methods [19, 20] combined neural network with analytic reconstruction or iterative reconstruction and improved the quality of LDCT images in the projection domain. Based on the residual block and dense block, the residual dense block (RDB) [33] achieved excellent performance in superresolution through contiguous memory (CM) mechanism and local feature fusion (LFF). By introducing the feature attention and enhancement attention modules (EAM), the real image denoising network (RIDNet) [35] can denoise real noise

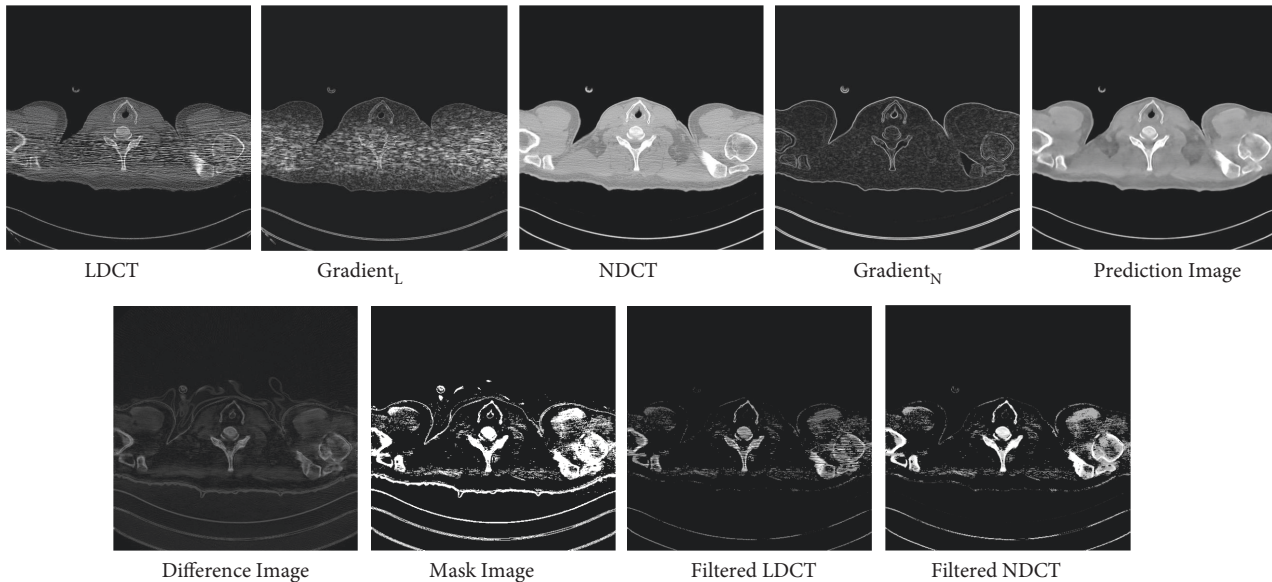


FIGURE 1: The comparison between the frequency-separation-based methods and local filtered mechanism. Gradient_L denotes the gradient image of the LDCT image. Gradient_N denotes the gradient image of the NDCT image. The gradient images were generated by the Sobel filter. The display window is [100, 300] Hu for better visualization.

images efficiently. Inspired by the above study, this work introduces an improved residual dense block based on the DD-Net for achieving further enhancement of the feature representation and denoising.

2.2. GAN Methods. While methods based on CNN can greatly improve the denoising performance, they can only generate the image structures based on prior knowledge, which causes the restriction in subtle structure restoration. Therefore, the schemes based on GAN were presented. Yang et al. [21] applied the Wasserstein distance and perceptual loss to train the GAN network. Wolterink et al. [22] introduced the voxelwise loss to improve the performance of GAN. Ge et al. [24] developed a conditional GAN to generate the thin thickness slices from thick LDCH images. To compete with commerce algorithms, Shan et al. [25] proposed a modularized adaptive processing neural network (MAP-NN). Choi et al. [26] presented the semisupervised GAN, including denoising network and classification network, to reduce the dependence on NDCT images. Taking the consecutive low-dose projections as the input, the comprehensive domain generator [27] with three-dimension was presented to learn the redundant information among slices and generate the subtle structures. To capture structural details, You et al. [29] introduced the leap connection and network in network. To describe the uncertainty of the denoised image, Huang et al. [36] used the CutMix technique and U-Net-based discriminator to provide radiologists with a confidence map. However, although the GAN-based methods can perfectly preserve the texture information in LDCT images, they performed poorly in subtle structure restoration as well. In contrast, our network preserves the areas in low quality by the mask image and strengthens the ability in subtle structure restoration by optimizing the areas in low quality especially.

2.3. Subtle Structure Restoration. Currently, although CNN and GAN have remarkably improved the performance in image denoising, it is still hard to restore the subtle structures. Yin and Babyn [23] designed a sharpness detector based on cGAN to preserve more edge information. Wang et al. [30] applied the shearlet transformation to generate the high-frequency information and low-frequency information separately. Fritsche et al. [34] utilized the low pass filter for frequency separation and adopted the GAN loss for subtle structure restoration. Yang et al. [31] designed two sub-networks based on U-Net in the generator for LDCT image denoising in the spatial domain and high-frequency domain. Recent work restores subtle structures through frequency separation, which lacks global information and may bring information loss during transformation. In contrast, our network learns the global structures and detailed information in one network and can provide the implicit global context for subtle structure restoration. Meanwhile, our model can specially optimize the areas with low quality through the mask image and avoid information loss during transformation.

3. The Proposed Scheme

The critical content of LDCT image denoising is to restore the subtle structures while removing the noise. To enhance the ability in feature representation and denoising, we propose an improved DD-Net. Besides, we present a new local filter mechanism and introduce a novel gradient loss to restore subtle structures accurately.

3.1. Network Structure. The proposed neural network employs a similar structure with DD-Net [32], which achieves brilliant performance in medical image denoising. However, as the global structures and detailed information are learned

in one network, our model requires a deeper network structure and more powerful feature representation ability. Thus, the improved residual dense block (IRDB) [33] is introduced. The improved residual dense block is composed of the dense connected [37] block and the enhanced residual block [35]. Considering the scale variation in the max-pooling layers [38], we remove the batch normalization in the dense connected block. The detailed structure of the improved residual dense block is shown in Figure 2.

The detailed network structure is represented in Figure 3. It includes 1 convolution layer, 4 max-pooling layers, 4 improved residual dense blocks, 4 upsampling layers, and 8 deconvolution layers, followed by Relu and batch normalization.

The input and output of the network are $512 \times 512 \times 1$ medical images. We adopt a 7×7 convolution layer after the input layer. After the 7×7 convolution layer, there are 4 encoder modules and 4 decoder modules. The encoder modules employ the max-pooling layer and the improved residual dense block [33] to extract multiscale features. The decoder modules employ the upsampling layers and 2 deconvolution layers followed by the Relu and batch normalization to restore the image information. The layers with the same feature shape in encoder modules and decoder modules will be concatenated. The kernel size of the last deconvolution layer is 1×1 to generate the denoised image.

3.2. Local Filtered Mechanism. As the subtle structure restoration is a significant content in LDCT image denoising, most schemes attempt to restore subtle structure by frequency separation. However, high-frequency information cannot reflect the subtle structures accurately, and approaches that process high-frequency information and low-frequency information separately may cause information loss during the transformation. To solve the problem, we propose a local filter mechanism to enhance the ability in subtle structure restoration. The mechanism restores the unclear subtle structures by two steps. In the first step, the LDCT image is inputted into the network to get the denoised image I_1 . In the second step, the difference image D_1 between the NDCT image and I_1 is obtained. Then, the mask image is generated by taking the threshold operation to D_1 . Using the mask image to conduct elementwise multiplication operation with LDCT and NDCT images, the areas with high quality are filtered, and the areas with poor quality are preserved. At last, the filtered LDCT image is inputted into the network to get the filtered denoised image I_2 . The original loss between the denoised image I_1 and NDCT image N_1 is calculated to enhance the ability in global structure restoration. The correction loss between the filtered denoised images I_2 and filtered NDCT image N_2 is calculated to enhance the ability in unclear subtle structure restoration. The detailed process is shown in Figure 4 and Algorithm 1.

Compared with the frequency-separation-based methods, the proposed network can achieve feature sharing between global structure and detailed information and can provide more context information for detail restoration. Thus, it can generate more realistic and precise subtle structures, as shown in Figures 5–8. In fact, the above

mechanism can be considered as a confrontation between network generalization and subtle structure restoration. The neural network tends to discard some detailed information of a specific image and make the subtle structures over-smoothed, which is beneficial for the enhancement of generalization and the reduction of overall error. To alleviate this phenomenon, this mechanism is designed to filter the areas with high quality and drive the network to specially optimize the areas with low quality that contain detailed information and are harder to optimize. Finally, the network can achieve the balance between the image denoising and subtle structure restoration through the sum of original loss and correction loss of the improved DD-Net.

Table 1 and Figure 9 demonstrate the results of different threshold values and training strategies. Although the model which was trained without a local filter mechanism can achieve higher PSNR and SSIM results, it drops some subtle structures to improve the generalization ability and cannot achieve further enhancement of subtle structure quality. The models which were trained with a local filter mechanism for 160 epochs can restore more appealing subtle structures but achieve poor performance in low-frequency portion denoising. Therefore, all networks without a pretrained model obtain lower PSNR and SSIM results than a pretrained model. The reason is that the global structures and detailed information are learned separately when adopting a local filtered mechanism at the beginning of the training phase. Without the prior knowledge of global structures, the network cannot utilize the global structure information learned in the first step to restore the subtle structures in the second step efficiently. Models with threshold values 0.01 and 0.04, which were trained with a pretrained model and local filter mechanism for 160 epochs, obtain higher PSNR and SSIM results than a pretrained model, but the improvement is slower than the model which was trained without a local filter mechanism for 160 epochs. The reason is that the subtle structures are more difficult to optimize. As is shown in Figure 9, it can be seen that all models trained with pretrained model and local filter mechanism pay more attention to subtle structure restoration.

Furthermore, how to set the threshold value is a crucial problem. When we use the low threshold value, the filtered images contain a large amount of low-frequency information which hinders the network from especially optimizing subtle structures. When we use the high threshold value, most subtle structures are filtered, and large black areas in the filtered images will influence the global structure restoration as there is a large difference between the filtered LDCT image and real LDCT image. Here, we trained the model without the local filtered mechanism for 80 epochs at first and visualized the filtered results using different threshold values. The visualization results are shown in Figure 10. As the threshold value of 0.04 can filter most low-frequency areas and save enough subtle structures, we choose 0.04 as our training threshold value in the chest dataset. Adopting the above parameter selection strategy, we choose 0.004 as our training threshold value in the brain dataset. Table 1 shows that both larger threshold value and lower threshold value will reduce the

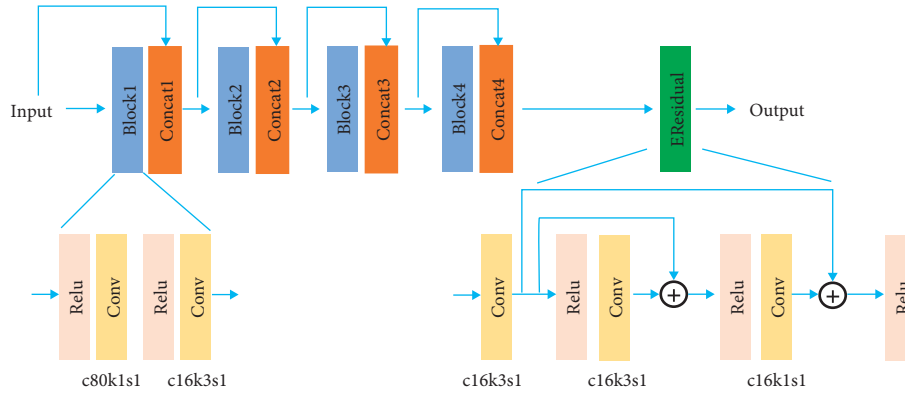


FIGURE 2: The structure of improved residual dense block (IRDB). c80k1s1 denotes that the channel, kernel size, and stride are 80, 1×1 , and 1, respectively. c16k3s1 denotes that the channel, kernel size, and stride are 16, 3×3 , and 1, respectively. c16k1s1 denotes that the channel, kernel size, and stride are 16, 1×1 , and 1, respectively. EResidual denotes the enhanced residual block.

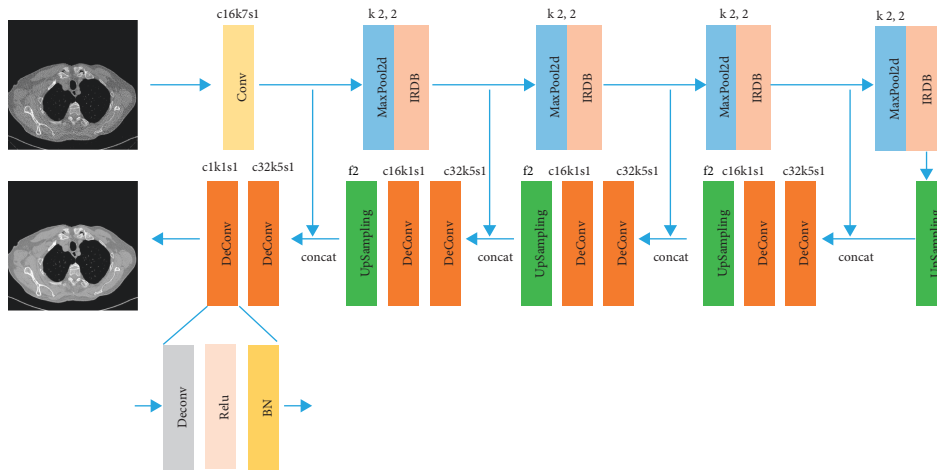


FIGURE 3: An overview of our network structure. The c16k7s1 denotes that the channel, kernel size, and stride are 16, 7×7 , and 1, respectively. $k, 2$ denotes that the kernel size of max-pooling $2d$ is 2×2 . $f2$ denotes that the upscale factor of upsampling is 2. c32k5s1 denotes that the channel, kernel size, and stride are 32, 5×5 , and 1. c16k1s1 denotes that the channel, kernel size, and stride are 16, 1×1 , and 1. c1k1s1 denotes that the channel, kernel size, and stride are 1, 1×1 , and 1.

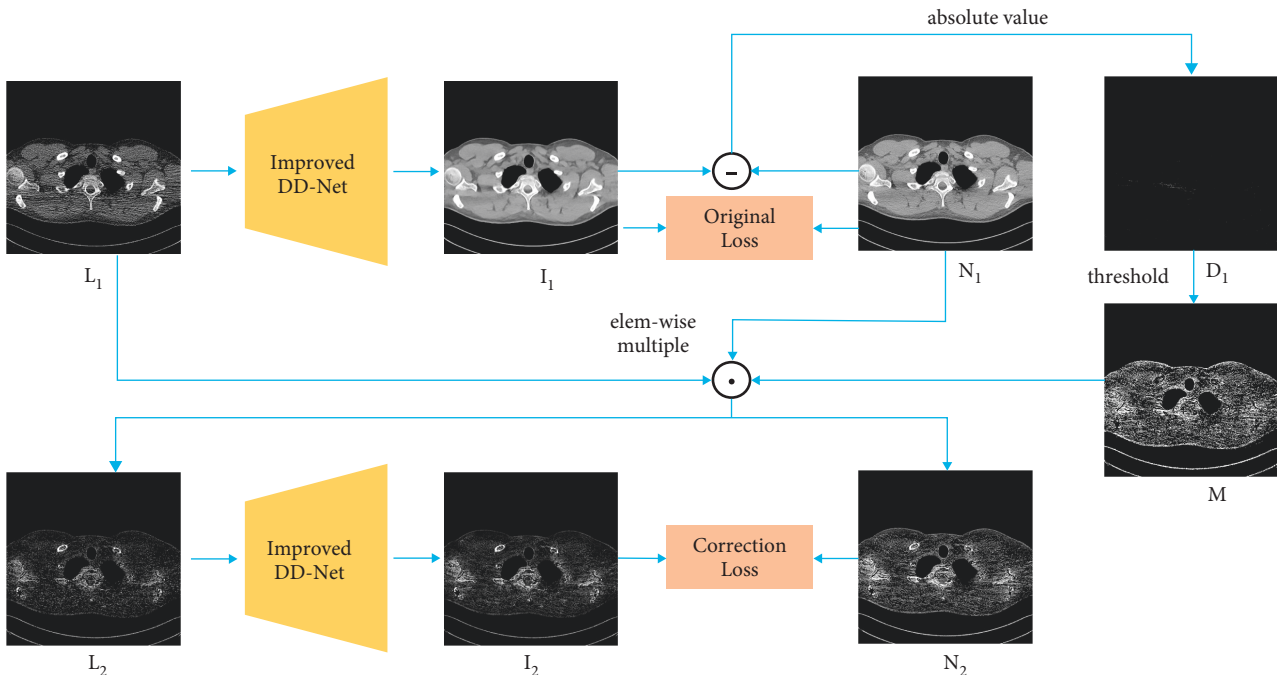


FIGURE 4: The framework of local filtered mechanism.

Require: L_1 (The low-dose CT image)

Require: N_1 (The normal-dose CT image)

Ensure: L_{total} (The total loss to take the back propagation and optimize the network)

NOTE: D_1 is the difference image which can reflect the difference between I_1 and N_1 . M is the mask image to filter the areas with high quality and save the areas with low quality. P_{xy} is the gray value of D_1 at pixel (x, y) . M_{xy} is the gray value of the mask image M at pixel (x, y) . *Improve dD DNet* is an improved convolution neural network based on DD-Net [32]. *thresh* is the threshold value. \odot is the elementwise multiplication operation. N_2 and L_2 are the filtered normal-dose image and low-dose image. Loss is the loss function of our network. L_{original} is the loss between I_1 and N_1 . L_{correct} is the loss between I_2 and N_2 .

- (1) $I_1 \leftarrow \text{ImprovedDDNet}(L_1)$
- (2) $D_1 \leftarrow |N_1 - I_1|$
- (3) **for** P_{xy} in D_1 **do**
- (4) **if** $P_{xy} < \text{thresh}$ **then**
- (5) $0 \leftarrow M_{xy}$
- (6) **else**
- (7) $1 \leftarrow M_{xy}$
- (8) **end if**
- (9) **end for**
- (10) $N_2 \leftarrow M \odot N_1$
- (11) $L_2 \leftarrow M \odot L_1$
- (12) $I_2 \leftarrow \text{ImprovedDDNet}(L_2)$
- (13) $L_{\text{original}} \leftarrow \text{Loss}(I_1, N_1)$
- (14) $L_{\text{correct}} \leftarrow \text{Loss}(I_2, N_2)$
- (15) $L_{\text{total}} \leftarrow L_{\text{original}} + L_{\text{correct}}$

ALGORITHM 1: The local filtered mechanism.

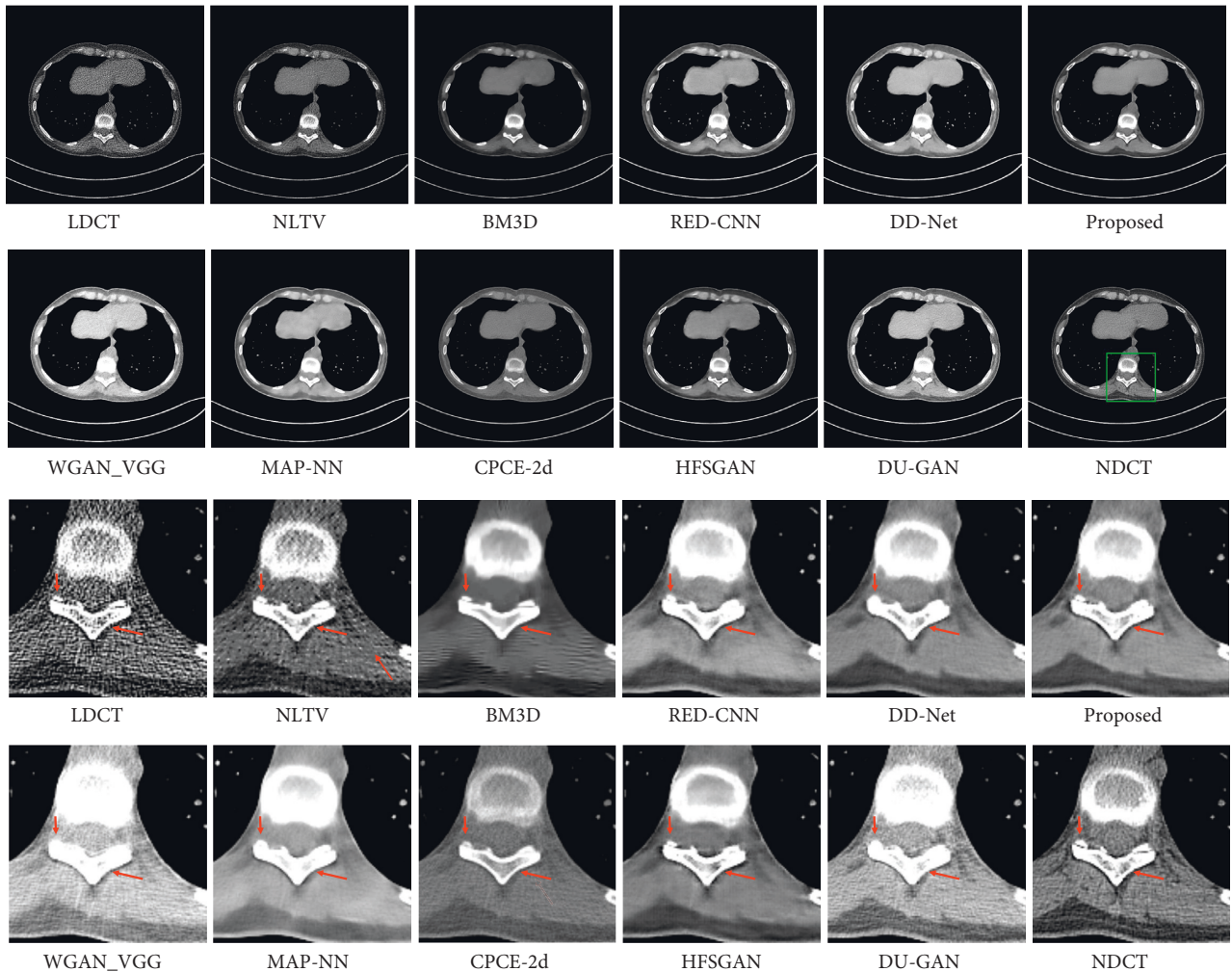


FIGURE 5: The denoised result for low-dose CT image in group 1. The loss function of DD-Net is the combination of MSE and MS-SSIM. The display window is [150, 250] Hu for better visualization.

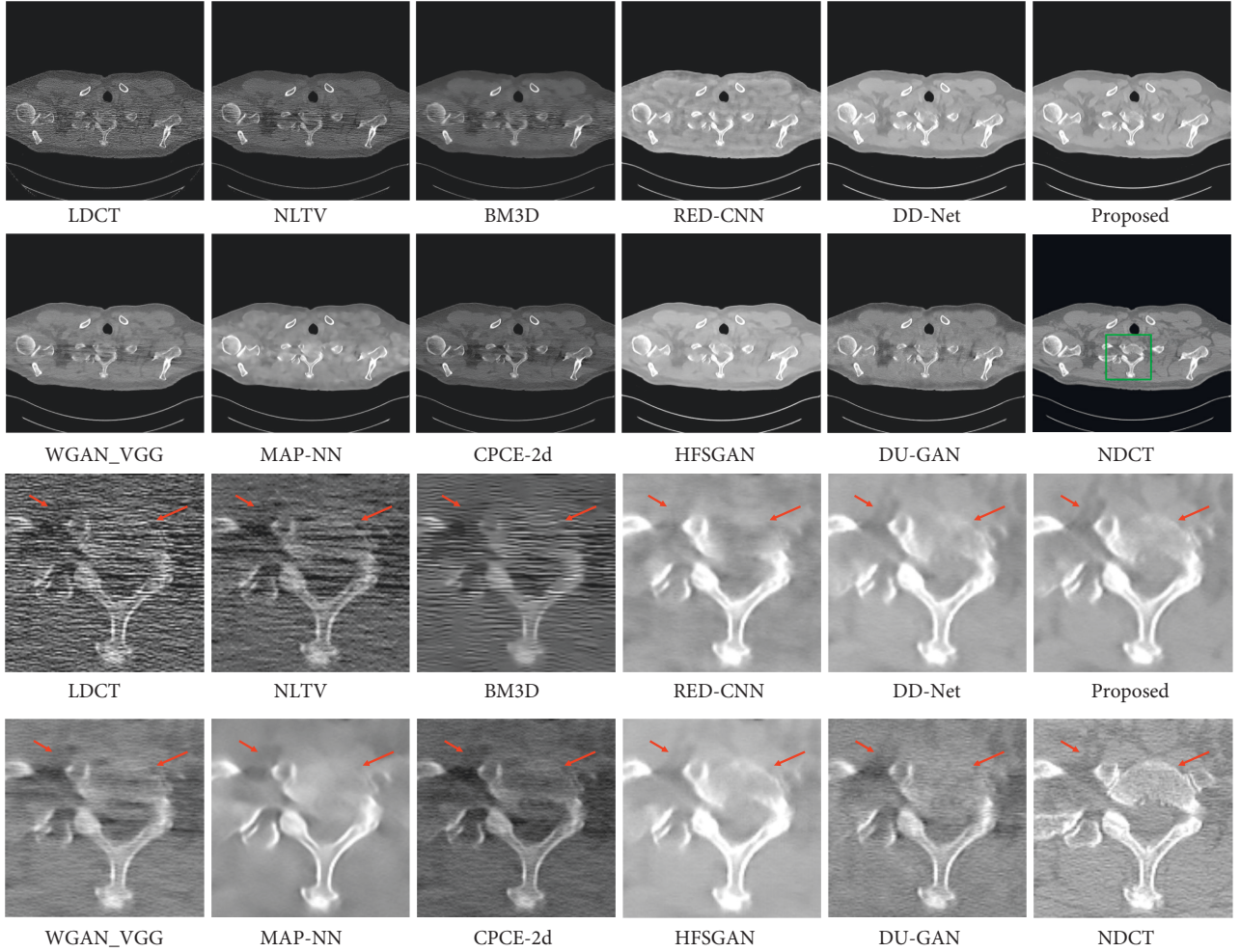


FIGURE 6: The denoised result for low-dose CT image in group 2. The display window is [100, 300] Hu for better visualization.

quality of the denoised results. The model whose threshold value is 0.04 obtains the highest PSNR and SSIM results in the chest dataset regardless of whether we use pretrained model or not. Furthermore, Figure 9 illustrates that an appropriate threshold can obtain a better visual appearance of subtle structures. Experiment verification indicates that the above parameter selection strategy can obtain a relatively proper threshold value.

3.3. Loss Functions. In our experiments, we use the weighted sum of mean square error, multiscale structural similarity [39], and gradient loss as the final loss function. To implement the local filter mechanism, it is required to calculate the original loss between the denoised image I_1 and NDCT image N_1 when the LDCT image L_1 is used as the input, as well as the correction loss between the denoised image I_2 and filtered NDCT image N_2 when the filtered LDCT image L_2 is used as the input. Thus, the total loss function is the sum of original loss and correction loss of the improved DD-Net, as equations (1) and (2).

$$L = \lambda_1 L_{MSE} + \lambda_2 L_{MS-SSIM} + \lambda_3 L_{grad} \quad (1)$$

$$L_{total} = L_{original} + L_{correct} \quad (2)$$

where λ_1 , λ_2 , and λ_3 are the hyperparameters.

3.3.1. Mean Square Error (MSE). The mean square error (MSE) is used to generate the objective information of images and calculate the pixel difference between denoised images and NDCT images. The MSE is described in equations (3) and (4).

$$I(i, j) = \text{ImprovedDDNet}(L(i, j)). \quad (3)$$

$$L_{MSE} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (N(i, j) - I(i, j))^2, \quad (4)$$

where $I(i, j)$ denotes the gray value of denoised images at pixel (x, y) and $N(i, j)$ denotes the gray value of NDCT images at pixel (x, y) .

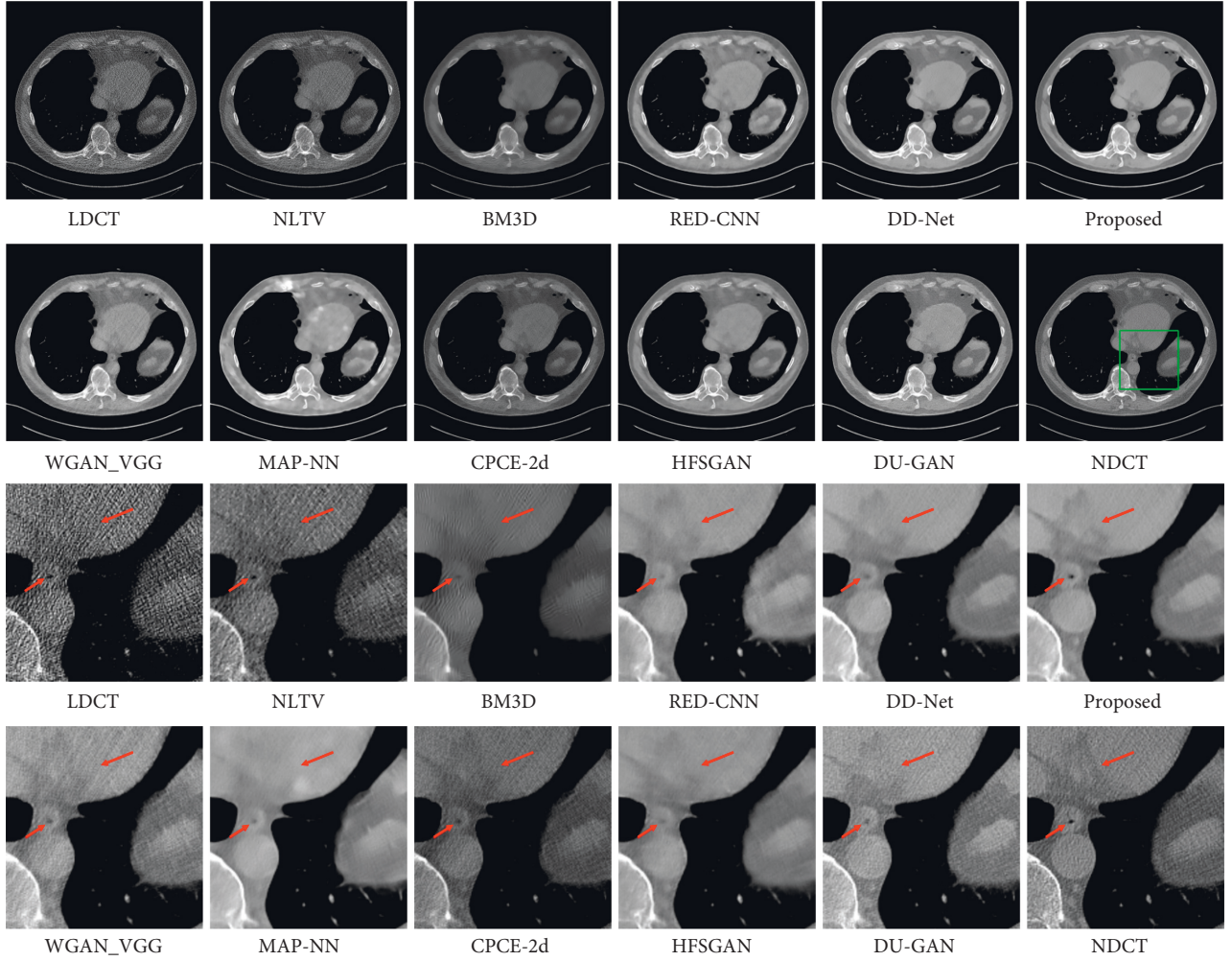


FIGURE 7: The denoised result for low-dose CT image in group 3. The display window is [100, 300] Hu for better visualization.

3.3.2. *Multiscale Structural Similarity (MS-SSIM)*. The structural similarity (SSIM) is a common metric for evaluating the perceptual loss in a single scale. As optimization of SSIM, the multiscale structural similarity (MS-SSIM) [39] is conducted over different resolutions and has a better performance. The MS-SSIM is described in equations (5)–(9).

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1},$$

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2},$$

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y}.$$

$$\text{MS-SSIM}(x, y) = [l_M(x, y)]^{\alpha_M} \cdot \prod_{j=1}^M [c_j(x, y)]^{\beta_j} \cdot [s_j(x, y)]^{\gamma_j}. \quad (6)$$

$$L_{\text{MS-SSIM}} = 1 - \text{MS-SSIM}, \quad (7)$$

$$\alpha_j = \beta_j = \gamma_j \quad (8)$$

$$\sum_{j=1}^M \gamma_j = 1, \quad (9)$$

where $I(x, y)$ is used to evaluate luminance, $c(x, y)$ is used to evaluate the contrast, and $s(x, y)$ is used to evaluate the structural similarity. α, β, γ are constants. In general, $\alpha_1 = \beta_1 = 0.0448$, $\alpha_2 = \beta_2 = 0.2856$, $\alpha_3 = \beta_3 = 0.3001$, $\alpha_4 = \beta_4 = 0.2363$, and $\alpha_5 = \beta_5 = 0.1333$. C_1, C_2, C_3 are constants. $C_1 = (0.01 \cdot (2^B - 1))^2$, $C_2 = (0.03 \cdot (2^B - 1))^2$, $C_3 = (C_2/2)$, and B denotes the bit depth of the image.

3.3.3. *Gradient Loss*. The gradient loss [34] is used to generate the edge information of images. In the first step, we take the convolution operation with denoised image and NDCT image, respectively, using a high-pass filter kernel K instead of the Sobel operator. In the second step, the gradient loss can be obtained by calculating the mean square error between the denoised image G_{predict} and NDCT image G_{normal} . The gradient loss is described in equations (10)–(13).

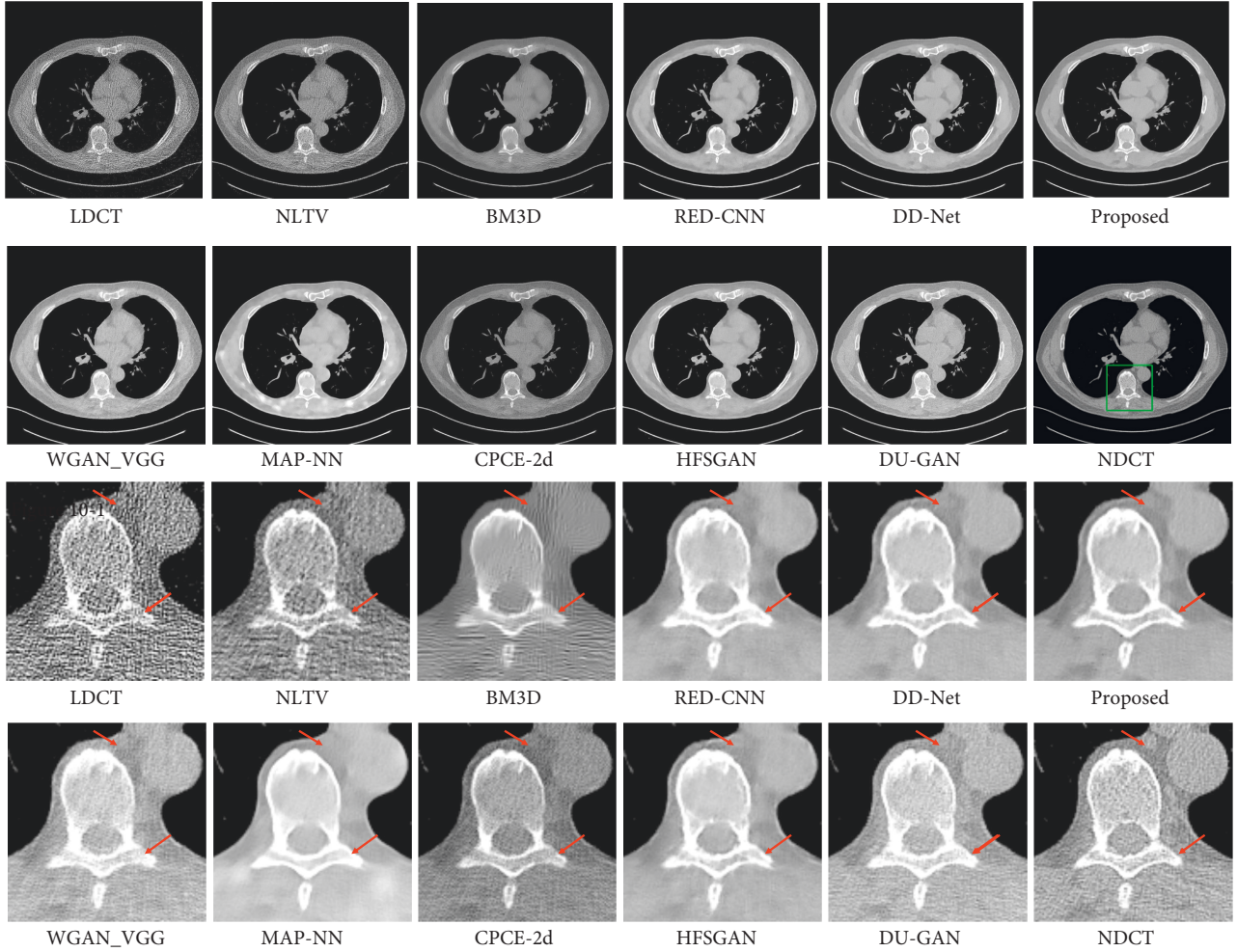


FIGURE 8: The denoised result for low-dose CT image in group 4. The display window is [100, 300] Hu for better visualization.

TABLE 1: The PSNR and SSIM results of chest datasets using different threshold values and training strategies. The w/o mask denotes that the model was trained without local filter mechanism. The w/o pretrained 80 denotes that the model was trained with local filter mechanism for 160 epochs. Pretrained 80 denotes that the model was trained without local filtered mechanism for 80 epochs at first and was trained with local filter mechanism for 160 epochs.

Epoch	Threshold	Method	PSNR	SSIM
80	—	W/o mask	31.7071	0.7601
160	—	W/o mask	31.9231	0.7636
160	0.01	W/o pretrained 80	30.9619	0.7423
160	0.04	W/o pretrained 80	31.3512	0.7511
160	0.07	W/o pretrained 80	31.1152	0.7464
160	0.10	W/o pretrained 80	30.5216	0.7367
160	0.01	Pretrained 80	31.8175	0.7614
160	0.04	Pretrained 80	31.8401	0.7613
160	0.07	Pretrained 80	31.6834	0.7589
160	0.10	Pretrained 80	31.5151	0.7583

$$K = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad (10)$$

$$G_{\text{predict}} = I \otimes K, \quad (11)$$

$$G_{\text{normal}} = H \otimes K, \quad (12)$$

$$L_{\text{grad}} = \text{MSE}(G_{\text{predict}}, G_{\text{normal}}), \quad (13)$$

where K denotes the high-pass filter kernel and \otimes denotes the convolution operation.

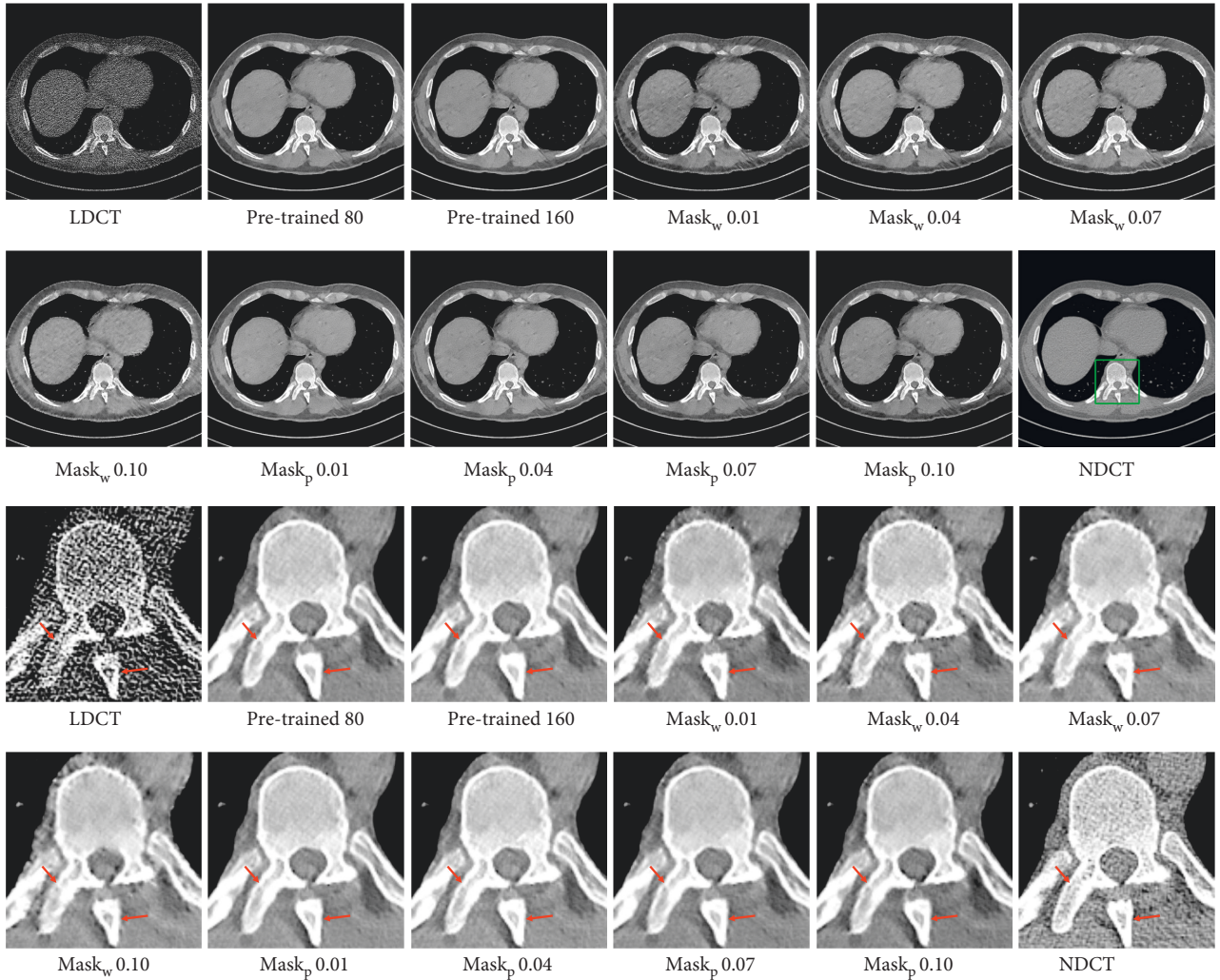


FIGURE 9: The denoised result for low-dose CT image with different threshold values and training strategies. The display window is [100, 300] Hu for better visualization. Pretrained 80 denotes the image generated by the model trained without local filtered mechanism for 80 epochs. Pretrained 160 denotes the image generated by the model trained without local filtered mechanism for 160 epochs. Both Mask_w and Mask_p denote the image generated by the model trained with local filtered mechanism. Mask_w denotes the image generated by the model trained without the pretrained model. Mask_p denotes the image generated by the model trained with the pretrained model.

4. Experiments

The experimental details and result analysis are introduced in this section to validate the effectiveness of our proposed scheme.

4.1. Dataset and Implement Details

4.1.1. Dataset. In our experiments, the dataset comes from the “2016 NIH-AAPM Mayo Clinic Low Dose CT Grand Challenge.” It comprises 100 chest scans with 10% of the routine dose, 99 head scans with 25% of the routine dose, and 100 abdomen scans with 25% of the routine dose. In our study, 10% chest scan and 25% head scan are used as the chest dataset and brain dataset. For chest dataset, we randomly select 25 cases for training and 10 cases for testing. For brain dataset, we randomly select 40 cases for training and 10 cases for testing. There is no overlap

between training and testing. The medical image dataset can be obtained from the Cancer Imaging Archive (TCIA) website at <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=52758026>.

4.1.2. Details. We trained and tested the proposed model in the chest dataset and brain dataset, respectively. The network was trained for 80 epochs without local filtered mechanism to obtain the pretrained model. Based on the pretrained model, the model was trained for 160 epochs with a local filtered mechanism to obtain the final model. The batch size was 8. The optimizer was Adam with β_1 as 0.9 and β_2 as 0.999. The initial learning rate was set to 1×10^{-4} which reduced to 5×10^{-5} at the 130-th epoch. The convolution and deconvolution layers were initialized with the Gaussian function whose mean square was 0, and variance was 0.01. We, respectively, set λ_1 , λ_2 , and λ_3 to 1, 0.15, and 0.8 as the hyperparameters of loss function. All models were trained

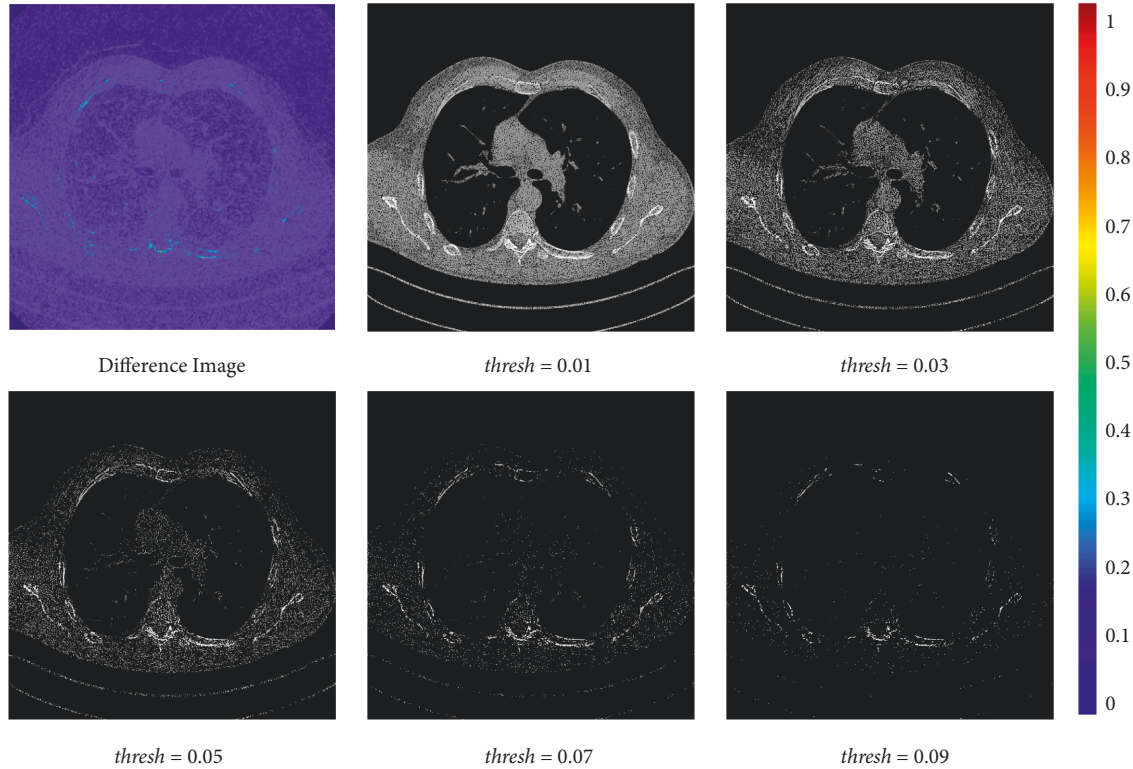


FIGURE 10: The difference image and filtered normal-dose CT images using different threshold values. The difference image was processed by pseudocolorization technology for better visualization. The display window is $[0, 400]$ Hu for filtered normal-dose CT images.

and tested on the NVIDIA RTX 2080 Ti GPU. Besides, the training curves in Figures 11 and 12 demonstrate that our method is easier to achieve higher performance than DD-Net using MS-SSIM. Although the model which was trained without local filtered mechanism can get higher PSNR and SSIM results, it will drop some subtle structures to improve the generalization ability, which is shown in Figure 9. The PSNR and SSIM results on the chest testing dataset, respectively, stabilize at 31.80 and 0.76 when the epoch is 120. The PSNR and SSIM results on the brain dataset, respectively, stabilize at 53.70 and 0.99 when the epoch is 140.

To verify the effectiveness of our scheme in subtle structure restoration, we selected some competitive methods for evaluation. Nonlocal total variation (NLTV) [40] is known as a statistical iterative method based on the compressed sensing (CS) technique. It adopts the global search and nonuniform weight penalization to improve the denoised image quality. Block-matching and 3d filtering (BM3D) [15] is the postprocessing method. It has the advantages of nonlocal methods and transform methods and removes the noise by searching and matching the similar blocks. Residual encoder-decoder convolution neural network (RED-CNN) [16] is a CNN-based method. It utilizes the residual structure and encoder-decoder structure and takes MSE as a loss function. Different from RED-CNN, DD-Net [32] applies the dense block in the encoder-decoder structure and uses the combination of MSE and SSIM as a loss function. Generative adversarial network with Wasserstein distance and perceptual loss (WGAN-VGG) [21] is a GAN-based method. It takes the combination of perceptual

similarity calculated by the VGG-19 network and Wasserstein distance as a loss function. Adopting the GAN structure as well, MAP-NN [25] applies the multiple conveying path-based convolution encoder-decoder (CPCE) modules in the generator. It takes the combination of MSE, Wasserstein distance, and edge incoherence calculated by the Sobel operator as a loss function. 2D conveying path-based convolution encoder-decoder network (CPCE-2d) [28] uses single CPCE module for LDCT denoising and takes the combination of adversarial loss and perceptual loss as the loss function. High-frequency sensitive generative adversarial network (HFSGAN) [31] is a frequency-separation-based method. It obtains high-frequency portion and low-frequency portion by guided filter and applies two U-Net to process the high-frequency portion and whole image separately. Generative adversarial networks with dual-domain U-Net-based discriminators (DU-GAN) [36] apply two U-Net-based discriminators to evaluate the difference in image domain and gradient domain and adopt the CutMix technique to describe the uncertainty of the denoised image.

In the NLTV method, we set the number of iterations, time step (dt), gradient regularization (ϵ), and fidelity term (λ) to 5, 0.2, 1×10^{-6} , and 1.2, respectively. In the BM3D method, the parameter σ was 25, and the hard threshold value was $2.7 \times \sigma$. β in Kaiser filter was 2. In the basic estimation, we set the match threshold value, maximum number of group matched blocks, block size, block stride, search step, and search window size to 2500, 16, 8, 3, 3, and 39, respectively. In the final estimation, we set the match threshold value, maximum number of group matched

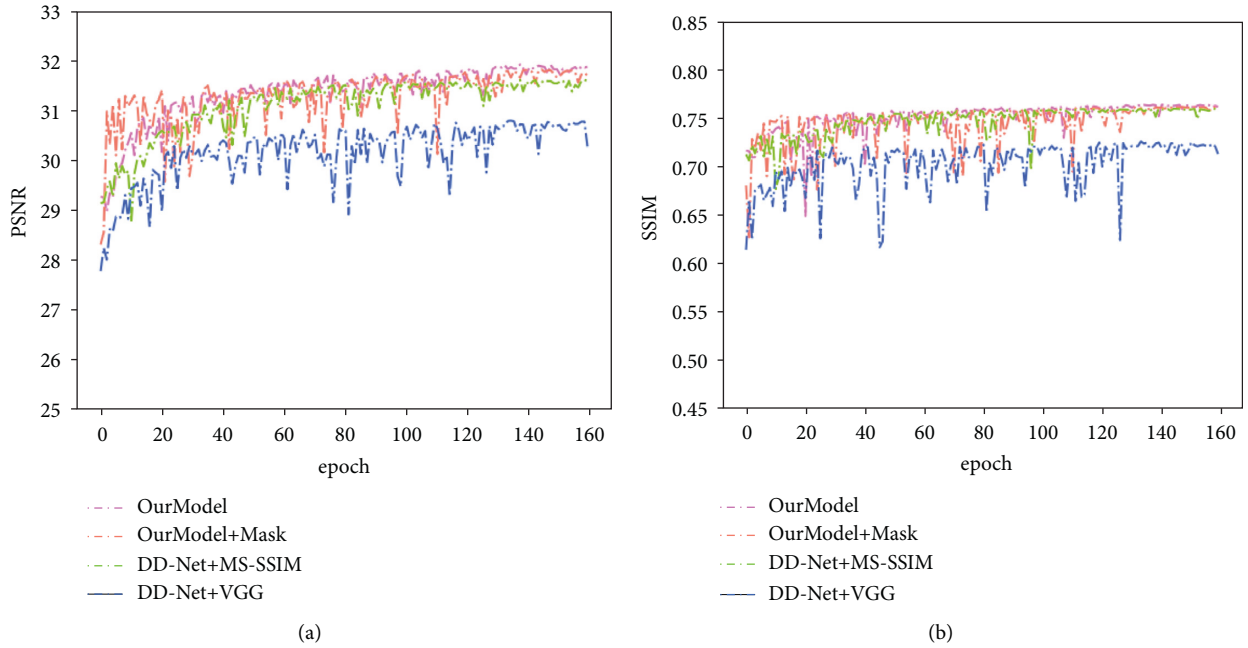


FIGURE 11: The curves of PSNR and SSIM on the chest testing dataset during the network training.

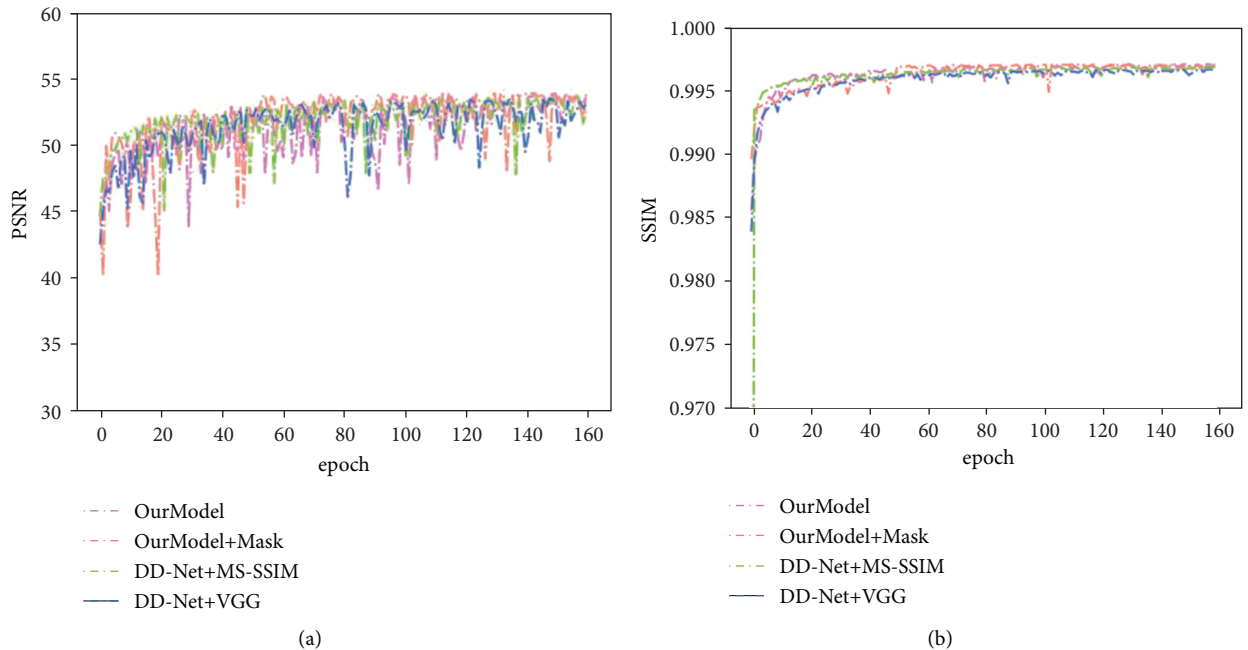


FIGURE 12: The curves of PSNR and SSIM on the brain testing dataset during the network training.

blocks, block size, block stride, search step, and search window size to 400, 32, 8, 3, 3, and 39, respectively. In the RED-CNN method, the LDCT images were cropped into 64×64 patches as the input, and the batch size was 160. The network was trained for 100 epochs using Adam with default parameters; namely, β_1 was 0.9 and β_2 was 0.999. The initial learning rate was set to 1×10^{-5} and decreased by 0.5 every 30 epochs. In the DD-Net method, it took the 512×512 images as the input, and the batch size was 8. The network

was trained for 160 epochs using Adam with default parameters. The initial learning rate was set to 1×10^{-4} and slowly decreased to 1×10^{-5} . In the WGAN-VGG method, the LDCT images were cropped into 64×64 patches as the input, and the batch size was 128. The network was trained for 200k iterations using Adam with β_1 as 0.5 and β_2 as 0.9, respectively. The learning rate was set to 1×10^{-5} , and λ_1 in the loss function was set to 0.1. λ_p for the gradient penalty was 10. In the MAP-NN method, the LDCT images were

cropped into 64×64 patches as the inputs, and the batch size was 128. The network was trained for 80 epochs using Adam with default parameters. The initial learning rate was set to 1×10^{-4} and decreased by $1/\sqrt{t}$ after the t -th epoch. λ_m and λ_e in the loss function were set to 50 and 50, respectively. λ_p for the gradient penalty was 10. The number of conveying-link-oriented network encoder-decoders (CLONE) was 5. In the CPCE-2d method, the LDCT images were cropped into 64×64 patches as the inputs, and the batch size was 128. The network was trained for 40 epochs using Adam with default parameters. The initial learning rate was set to 1×10^{-4} and decreased by $1/t$ after the t -th epoch. λ_p for perceptual loss was set to 0.1. In the HFSGAN method, it took the 512×512 images as the input, and the batch size was 12. The network was trained for 200 epochs using Adam with β_1 as 0.5 and β_2 as 0.999, respectively. The initial learning rate was set to 2×10^{-4} . λ_1 and λ_2 in the loss function were set to 100 and 50, respectively. In the DU-GAN method, the LDCT images were cropped into 64×64 patches as the inputs, and the batch size was 64. The network was trained for 100000 iterations using Adam with default parameters. The initial learning rate was 1×10^{-4} . λ_{adv} , λ_{img} , and λ_{grd} in the loss function were set to 0.1, 1, and 20, respectively. The code of NLTV was downloaded at http://math.sjtu.edu.cn/faculty/xqzhang/NLIP_v1.zip. The code of BM3D was implemented using Python. Other deep learning-based methods were implemented using PyTorch according to the official codes.

4.2. Evaluation Metrics. Appropriate evaluation metrics are crucial for the evaluation of LDCT image denoising because medical images contain more subtle structures and fewer channels than natural images. We select the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) as the evaluation metrics. These evaluation metrics are widely used in image processing tasks, such as image superresolution and image inpainting.

4.2.1. Peak Signal-to-Noise Ratio (PSNR). PSNR is an objective metric to measure the error of image pixels, and it is usually used in the images which are sensitive to error, as equations (14)-(15). In general, the higher the PSNR is, the lower distortion the image has.

$$MSE = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (x(i, j) - y(i, j))^2, \quad (14)$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE} \right), \quad (15)$$

where x, y, N denote the denoised image, normal-dose image, and the width or height of the image, respectively, and MAX denotes the maximum of gray value.

4.2.2. Structural Similarity (SSIM). As the PSNR cannot completely reflect the subjective visual difference, the SSIM is used as a supplement to measure the visual appearance of images, as equation (16). In general, the higher the SSIM is,

the more abundant and appealing visual appearance the image has.

$$MS-SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x + \sigma_y + c_2)}, \quad (16)$$

where μ denotes the mean value, σ denotes the variance, σ_{xy} denotes the covariance of the denoised image and normal-dose image, and c_1 and c_2 are constants. $c_1 = (0.01 \cdot MAX)^2$, and $c_2 = (0.03 \cdot MAX)^2$.

4.3. Experimental Results and Analysis

4.3.1. Qualitative Evaluations. The LDCT denoised images are shown in Figures 5–8. The regions of interest (ROIs) are marked by green rectangles, and the red pointing arrows denote the difference between our method and comparative methods.

From the enlarged view under the denoised results, it can be seen that the BM3D and NLTV reduce the noise to some extent but fail to remove the artifacts, which is shown in Figure 6. The denoised images generated by RED-CNN and MAP-NN are slightly oversmoothed due to adopting MSE as a loss function. As the perceptual loss and adversarial loss cannot reflect the pixel level difference accurately, the WGAN-VGG and CPCE-2d preserve the texture information in LDCT images but cannot remove the streak artifacts completely. Using the MS-SSIM as a loss function, the DD-Net can remove the noise and streak artifacts effectively but cannot restore some subtle structures accurately, which can be shown in Figure 7. As the frequency-separation-based method, the HFSGAN enhanced the structural similarity between the denoised image and NDCT image in high-frequency portion and can restore the subtle structures clearly. However, it blurred the image information in the low-frequency portion. As is shown in Figure 6, the DU-GAN can generate the image with realistic texture information but cannot restore the subtle structure when the LDCT image is contaminated severely by noise and streak artifacts. Overall, as is shown in Figures 6 and 7, our method can restore subtle structures effectively and keep the color and structure consistent with the NDCT image.

4.3.2. Quantitative Evaluations. To validate the performance of our neural network quantitatively, we adopted peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) as objective metrics. Table 2 shows the PSNR and SSIM results of the chest dataset and brain dataset using different methods. The NLTV and BM3D are traditional denoising methods, the RED-CNN, DD-Net, and our model are CNN-based denoising methods, and the WGAN-VGG, MAP-NN, CPCE-2D, HFSGAN, and DU-GAN are GAN-based denoising methods.

First, as the NLTV and BM3D remove the noise and artifacts through the information of single LDCT image solely, they perform poorly in image information restoration and get the lower PSNR and SSIM results than deep learning-based methods. Second, as the perceptual loss and

TABLE 2: The PSNR and SSIM results of chest dataset and brain dataset using different methods. The chest-10% denotes that the dose of low-dose chest images is 10% of the routine dose. The brain-25% denotes that the dose of low-dose brain images is 25% of the routine dose.

Method	Chest-10%		Brain-25%	
	PSNR	SSIM	PSNR	SSIM
LDCT	23.1989	0.4487	51.4490	0.9938
NLTV [40]	25.6894	0.6236	35.1604	0.8905
BM3D [15]	26.9452	0.6748	40.9885	0.7938
RED-CNN [16]	30.9656	0.7518	53.1678	0.9961
DD-Net + MS-SSIM [32]	31.6241	0.7595	53.6189	0.9967
DD-Net + VGG	30.8235	0.7240	53.4942	0.9966
Proposed	31.8401	0.7613	53.8237	0.9970
WGAN-VGG [21]	29.2275	0.6932	48.0044	0.9897
MAP-NN [25]	30.6552	0.7567	52.0912	0.9952
CPCE-2d [28]	27.5885	0.6889	50.7368	0.9935
HFSGAN [31]	31.0291	0.7468	38.5513	0.9162
DU-GAN [36]	29.9207	0.7012	54.0428	0.9968

adversarial loss reflect the style difference, the WGAN-VGG and CPCE-2d cannot remove the streak artifacts completely. Therefore, they obtain lower PSNR and SSIM results than methods using MSE as one of the loss functions, including RED-CNN, DD-Net, MAP-NN, HFSGAN, DU-GAN, and our method in the chest dataset. Meanwhile, as the RED-CNN was trained with MSE loss solely, it produces over-smoothed denoised results in Figure 6. Due to the MS-SSIM loss, the denoised results generated by DD-Net keep the high structural similarity with NDCT images. Therefore, the DD-Net obtains the second-best PSNR and SSIM results in the chest dataset and the third-best PSNR and SSIM results in the brain dataset. As the HFSGAN processes the high-frequency portion especially, it can accurately restore the subtle structures. However, it blurs the image information in the low-frequency portion and obtains lower objective metrics results than DD-Net and our model in the chest dataset. In addition, due to the information loss during frequency separation, the HFSGAN performs poorly when the difference between LDCT images and NDCT images is relatively small. Therefore, the HFSGAN obtains the lowest PSNR and SSIM results among deep learning-based methods in the brain dataset. Although the DU-GAN can preserve the texture information in LDCT images, it performs poorly in subtle structure restoration, which can be shown in Figure 6, and cannot keep color consistency with NDCT images. Besides, as there is a relatively large difference in texture information between LDCT images and NDCT images, the DU-GAN obtains lower PSNR and SSIM results than most deep learning-based methods in the chest dataset. However, when the difference between LDCT images and NDCT images is small, the DU-GAN can accurately sense the difference through a confidence map and get the best PSNR result in the brain dataset. Our model achieves high performance in subtle structure restoration and obtains the best objective metrics results in the chest dataset. However, our model obtains the second-best PSNR result and the best SSIM result in the brain dataset because the confidence map and taking the 64×64 patches as the input in the DU-GAN can sense the difference with low confidence score more

accurately than our method. In contrast, our model can achieve higher performance when the LDCT images contain a large amount of noise and streak artifacts.

Because the perceptual loss using VGG-19 [41] is another effective loss function to enhance the visual appearance, it is compared with MS-SSIM. As is shown in Figures 11 and 12, the model using MS-SSIM as a loss function obtains higher PSNR and SSIM results than perceptual loss using VGG-19. The reason is that VGG-19 is trained on natural images and cannot reflect the visual features of medical images perfectly. Moreover, as shown in Figure 13, whether it is DD-Net or our network, the image generated by the model training with VGG-19 contains more noise than the model using MS-SSIM.

4.3.3. Uncertainty Visualization. The uncertainty visualization of different methods is shown in Figure 14. As both BM3D and NLTV reduce the noise in the LDCT image to some extent, they can obtain higher global scores of D_{dec}^{img} than the LDCT image. However, they cannot remove the streak artifacts and restore the image information efficiently, which causes low per-pixel confidence. The RED-CNN can efficiently remove the noise and streak artifacts but over-smoothens the LDCT image simultaneously. Therefore, its global score and per-pixel confidence are lower than other deep learning-based methods. In addition, although the WGAN-VGG and CPCE-2d can avoid the oversmooth problem, they cannot remove streak artifacts completely. Therefore, their global scores are higher than RED-CNN but lower than most deep learning-based methods. The MAP-NN will pay more attention to subtle structure restoration by adopting the gradient loss as the loss function but blurs the low-frequency portion according to per-pixel confidence. Therefore, it obtains a higher global score and per-pixel confidence than WGAN-VGG. The HFSGAN can generate realistic subtle structures but over-smoothens the low-frequency portion in the LDCT image as well according to the confidence map. The DD-Net can remove the noise and streak artifacts efficiently and keep high structural similarity

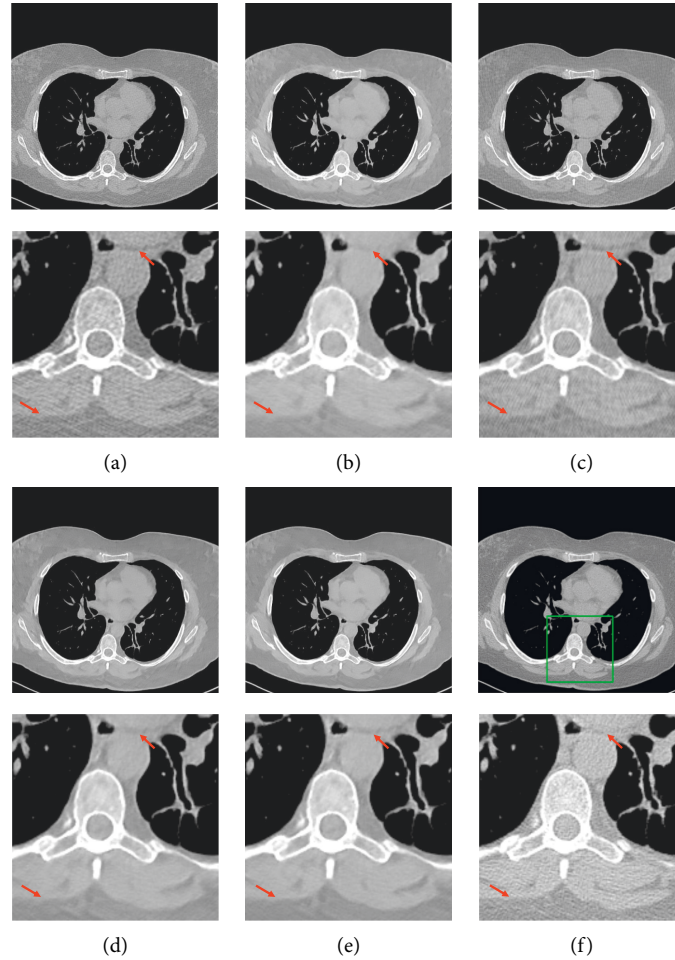


FIGURE 13: The denoised results generated by different loss functions and neural networks. (a) DD-Net + VGG. (b) DD-Net + MS-SSIM. (c) Proposed + VGG. (d) Proposed + MS-SSIM. (e) Proposed + MS-SSIM + local filtered mechanism + gradient loss. (f) NDCT. The display window is [100, 300] for better visualization.

with the NDCT image according to the confidence map. However, compared with our method, it cannot restore some subtle structures accurately and obtains a lower global score than our method. As the DU-GAN is trained with the discriminator and can adjust the image quality pertinently, it can generate the photo-realistic denoised results according to the confidence map and obtains the best global score. The proposed method obtains higher per-pixel confidence than DU-GAN in the subtle structures and obtains the second-best global score, indicating that our method achieves better performance in subtle structure restoration.

4.4. Ablation Study. Since we introduce some novel structures and local filtered mechanism based on DD-Net, it is significant to take a comparative analysis. The detailed results are presented in Table 3, which shows the objective metrics of our modifications.

4.4.1. Use Improved Residual Dense Block. As the global structure and detailed information are learned in one

network, we introduce the improved residual dense block to improve the representation ability of the neural network. As is shown in Table 3, our model obtains higher PSNR and SSIM results than DD-Net when the network uses the improved residual dense block and does not introduce other improvements. Furthermore, when the DD-Net was trained with local filtered mechanism, using the improved residual dense block can solve the performance reduction remarkably, indicating that the representation ability of DD-Net is insufficient for the local filtered mechanism, and introducing the above module can enhance the performance in subtle structure restoration.

4.4.2. Use Local Filtered Mechanism. As the high-frequency portion cannot reflect the incorrect subtle structures accurately and frequency-separation-based methods cannot utilize the correlation between global structures and detailed information, we introduce the local filtered mechanism. Filtering the areas with high quality, this mechanism drives the network to specially optimize the subtle structures in low quality and enhances the ability in subtle structure

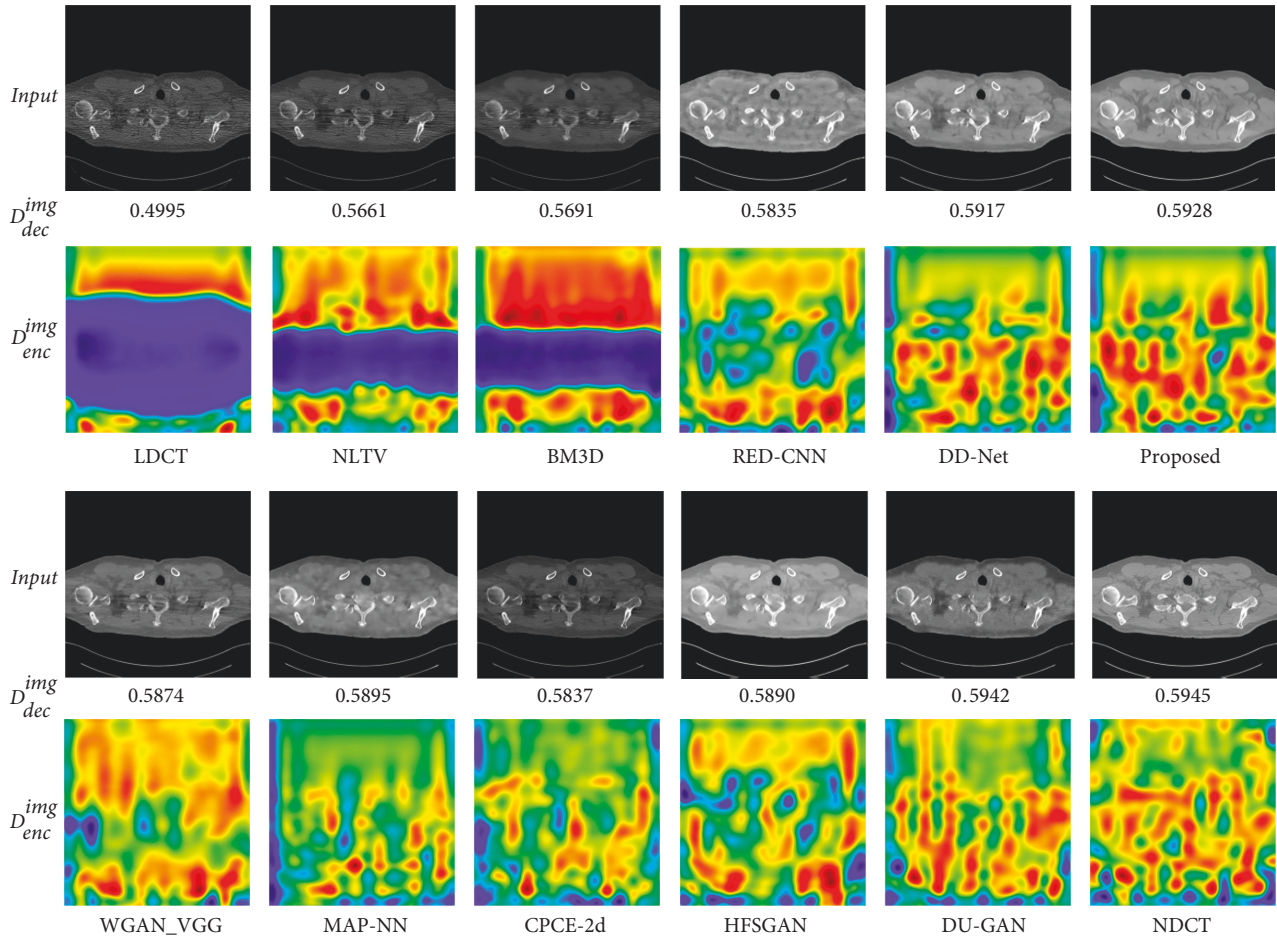


FIGURE 14: Uncertainty visualization of applying the trained discriminator to the outputs of different methods. The display window is [100, 300] Hu for better visualization. Note that the blue color of D_{dec}^{img} indicates a lower confidence score while the red color indicates a high confidence score.

TABLE 3: The ablation study in the chest dataset of our experiments.

Improved residual dense block [33]	Local filtered mechanism	Gradient loss [34]	PSNR	SSIM
×	×	×	31.6241	0.7595
✓	×	×	31.7742	0.7618
×	✓	×	31.6130	0.7569
×	×	✓	31.8317	0.7593
×	✓	✓	31.7400	0.7571
✓	✓	×	31.7363	0.7622
✓	×	✓	31.9231	0.7636
✓	✓	✓	31.8401	0.7613

restoration. However, it causes performance reduction but can make the network generate more precise and realistic subtle structures. The reason is that the areas preserved by this mechanism are more difficult to optimize, and this mechanism hinders the network from dropping some subtle structures for better generalization ability.

4.4.3. Use Gradient Loss. As the combination of MSE and SSIM cannot keep the edge information consistency

between the denoised result and NDCT image, the network will oversmooth some subtle structures, which is shown in Figure 7. To remedy this, we introduce the gradient loss to sense the edge information in the LDCT image. As is shown in Table 3, introducing the gradient loss can solve the performance reduction caused by the local filtered mechanism efficiently as well. In addition, as is shown in Figure 13, the color and edge information of subtle structures are enhanced through the local filtered mechanism and gradient loss, indicating that introducing

the gradient loss is beneficial for subtle structure restoration.

5. Conclusion

In this work, we propose a novel scheme for LDCT denoising based on improved DD-Net and local filtered mechanism. As the incorrect subtle structures are usually the areas that are contaminated severely by noise and streak artifacts instead of the edge information in high-frequency domain, previous studies cannot restore the subtle structures efficiently. Therefore, we present the local filtered mechanism to filter the areas with high quality and make the network optimize the subtle structures especially. Based on the original loss and correction loss of the improved DD-Net, the proposed method can accomplish the balance between network generalization and subtle structure restoration. However, as learning global structures and detailed information in one network requires more powerful feature representation ability and the edge information is significant for subtle structure restoration, we introduce the improved residual dense block and gradient loss to deepen the network structures and keep the edge information consistency between the denoised result and NDCT image, respectively. The ablation study validates the effectiveness of the above components.

The quantitative results show that our scheme can obtain higher scores in objective metrics than conventional schemes and most neural networks. Meanwhile, the visual comparison and uncertainty visualization also show that our scheme can provide a brilliant approach for subtle structure restoration in LDCT image denoising. In addition, the proposed network achieves competitive performance in both chest dataset and brain dataset, even if their radiation doses are quite different, which demonstrates the generalization ability of our network in different scenarios. However, training the network needs to, respectively, calculate the original loss and correction loss of the improved DD-Net for each batch, which introduces more computational cost. Moreover, the effectiveness of our scheme requests further validation in other image processing tasks, such as image superresolution and restoration, which is a significant research direction in the future.

Data Availability

The data supporting this work are from previously reported studies and datasets, which have been cited. The processed data are available at <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=52758026>.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (nos. 62101481 and 62002313), Basic Research Project of Yunnan Province (no.

202201AT070112), Key Laboratory in Software Engineering of Yunnan Province in China (2020SE408), and Industrial Internet Security Situation Awareness Platform of Yunnan Province.

References

- [1] Z. Jing, G. Qiang, H. Fang, L. Zhan-Li, L. Hong-An, and S. Yu, "A novel 3D reconstruction algorithm of motion-blurred CT image," *Computational and Mathematical Methods in Medicine*, vol. 2020, no. 3, pp. 1–13, 2020.
- [2] H. Qi, Z. Chen, and L. Zhou, "CT image reconstruction from sparse projections using adaptive TpV regularization," *Computational and Mathematical Methods in Medicine*, vol. 2015, pp. 1–8, 2015.
- [3] W. Zhao and L. Wang, "Research on 3D reconstruction algorithm of medical CT image based on parallel contour," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 11828–11835, 2020.
- [4] X. Liang, P. Hu, L. Zhang, J. Sun, G. Yin, and G. Yin, "MCFNet: multi-layer concatenation fusion network for medical images fusion," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 7107–7119, 2019.
- [5] R. KumarKhan, A. A. Khan, J. Kumar et al., "Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging," *IEEE Sensors Journal*, vol. 21, no. 14, pp. 16301–16314, 2021.
- [6] L. Yu, J. M. Reinhardt, and J. Pluim, "Sinogram smoothing with bilateral filtering for low-dose CT," in *Medical Imaging 2008: Physics of Medical Imaging*, vol. 6913, p. 8, SPIE, Washington, United States, 2008.
- [7] J. Hsieh, "Adaptive streak artifact reduction in computed tomography resulting from excessive x-ray photon noise," *Medical Physics*, vol. 25, no. 11, pp. 2139–2147, 1998.
- [8] 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.
- [9] 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.
- [10] Y. Chen, Z. Yang, Y. Hu et al., "Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means," *Phys Med Biol*, vol. 57, no. 9, pp. 2667–2688, 2012.
- [11] 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.
- [12] Y. Chen, X. Yin, L. Shi et al., "Improving abdomen tumor low dose ct images using a fast dictionary learning based processing," *Physics in Medicine and Biology*, vol. 58, no. 16, pp. 5803–5820, 2013.
- [13] Y. Hou, C. Zhao, D. Yang, and Y. Cheng, "Comments on "image denoising by sparse 3-D transform-domain collaborative filtering,"" *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 268–270, 2011.
- [14] Z. Li, L. Yu, J. D. Trzasko, J. G. Fletcher, C. H. McCollough, and A. Manduca, "Adaptive non-local means filtering based on local noise level for CT denoising," in *Medical Imaging 2012: Physics of Medical Imaging*, vol. 8313, pp. 447–456, SPIE, Washington, United States, 2016.
- [15] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative

- filtering,” *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [16] H. Chen, Yi Zhang, M. K. Kalra et al., “Low-dose CT with a residual encoder-decoder convolutional neural network,” *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.
- [17] X. Yin, Q. Zhao, J. Liu et al., “Domain progressive 3D residual convolution network to improve low-dose CT imaging,” *IEEE Transactions on Medical Imaging*, vol. 38, no. 12, pp. 2903–2913, 2019.
- [18] D. Wu, K. Kim, G. El Fakhri, and Q. Li, *A Cascaded Convolutional Neural Network for X-ray Low-Dose CT Image Denoising*IEEE, New Jersey, United States, arXiv:1705.04267, 2017.
- [19] D. Wu, K. Kim, G. El Fakhri, and Q. Li, “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.
- [20] E. Kang, J. Yoo, and J. C. Ye, *Deep Convolutional Framelet Denoising for Low-Dose CT via Wavelet Residual Network*IEEE, New Jersey, United States, arXiv:1707.09938, 2017.
- [21] 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, 2018.
- [22] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, “Generative adversarial networks for noise reduction in low-dose CT,” *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.
- [23] X. Yi and P. Babyn, “Sharpness-Aware low-dose CT denoising using conditional generative adversarial network,” *J Digit Imaging*, vol. 31, no. 5, pp. 655–669, 2018.
- [24] R. Ge, G. Yang, C. Xu, Y. Chen, L. Luo, and S. Li, “Stereo-correlation and noise-distribution aware resvox gan for dense slices reconstruction and noise reduction in thick low-dose ct,” in *Proceedings of the Medical Image Computing and Computer Assisted Intervention*, pp. 328–338, MICCAI, Cambridge, UK, September 2019.
- [25] H. Shan, A. Padole, F. Homayounieh et al., “Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction,” *Nat Mach Intell*, vol. 1, no. 6, pp. 269–276, 2019.
- [26] K. Choi, M. Vania, and S. Kim, “Semi-supervised learning for low-dose CT image restoration with hierarchical deep generative adversarial network (HD-GAN),” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2683–2686, EMBC, Berlin, Germany, July 2019.
- [27] Y. Zhang, D. Hu, Q. Zhao et al., “CLEAR: comprehensive learning enabled adversarial reconstruction for subtle structure enhanced low-dose CT imaging,” *IEEE Transactions on Medical Imaging*, vol. 40, no. 11, pp. 3089–3101, 2021.
- [28] H. Shan, Y. Zhang, Q. Yang et al., “3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1522–1534, 2018.
- [29] C. You, L. Yang, Y. Zhang, and G. Wang, “Low-dose CT via deep CNN with skip connection and network in-network,” in *Proceedings of Developments in X-Ray Tomography XII*, vol. 11113, Article ID 111131W, August 2019.
- [30] C. Wang, S. Wang, Z. Xia et al., “Medical image super-resolution via deep residual neural network in the shearlet domain,” *Multimedia Tools and Applications*, vol. 80, no. 10, pp. 1–19, 2021.
- [31] L. Yang, H. Shangguan, X. Zhang, A. Wang, and Z. Han, “High-frequency sensitive generative adversarial network for low-dose CT image denoising,” *IEEE Access*, vol. 8, pp. 930–943, 2020.
- [32] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, “A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1407–1417, 2018.
- [33] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, “Residual dense network for image super-resolution,” in *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2472–2481, IEEE, Salt Lake City, UT, USA, June 2018.
- [34] M. Fritsche, S. Gu, and R. Timofte, “Frequency separation for real-world super-resolution,” in *Proceedings of IEEE/CVF International Conference on Computer Vision Workshop*, pp. 3599–3608, ICCVW, Seoul, Korea, Oct. 2019.
- [35] S. Anwar and N. Barnes, “Real image denoising with feature attention,” in *Proceedings of IEEE/CVF International Conference on Computer Vision*, pp. 3155–3164, ICCV, Seoul, Korea, Oct. 2019.
- [36] Z. Huang, J. Zhang, Y. Zhang, and H. Shan, “DU-GAN: generative adversarial networks with dual-domain U-Net-Based discriminators for low-dose CT denoising,” *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2022.
- [37] G. Huang and Z. Liu, L. Van Der Maaten and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2261–2269, CVPR, Honolulu, HI, USA, July 2017.
- [38] X. Wang, K. Yu, S. Wu et al., “EsrGAN: enhanced super resolution generative adversarial networks,” in *Proceedings of Computer Vision – ECCV 2018 Workshops*, pp. 63–79.
- [39] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, “Loss functions for neural networks for image processing,” *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 1–11, 2015.
- [40] H. Kim, J. Chen, A. Wang, C. Chuang, M. Held, and J. Pouliot, “Non-local total-variation (NLTv) minimization combined with reweighted L1-norm for compressed sensing CT reconstruction,” *Physics in Medicine and Biology*, vol. 61, no. 18, pp. 6878–6891, 2016.
- [41] Y. Gu, Z. Zeng, H. Chen et al., “Medsrgan: medical images super-resolution using generative adversarial networks,” *Multimedia Tools and Applications*, vol. 79, no. 29–30, pp. 21815–21840, 2020.