ISSN 1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20220304044850
Preliminary communication

A Low-Dose CT Image Denoising Method Combining Multistage Network and Edge Protection

Zhitao GUO*, Feng ZHOU, Yuqing CHEN, Jinli YUAN

Abstract: Low-dose CT is an effective method to reduce the amount of radiation and the potential impact of radiation on patients. However, the reduction of radiation dose will lead to a large amount of noise in the reconstructed image, which will blur the edge details of the internal tissues and the organizational structure, thus causing confusion to the quantitative judgment of doctors in the diagnosis process. For this reason, a multi-stage denoising convolutional neural network emphasizing edge protection (MEP-Net) is proposed in this paper. Firstly, to overcome the problem of fixed input feature scale in the single-stage network, a multi-stage network structure is constructed which can extract multi-scale image noise features. Secondly, in order to protect the edge information of the image, the edge protection module is designed, which is combined with the convolutional neural network to extract the edge features of different stages in the multichannel. Finally, the cross-stage feature fusion mechanism is adopted to restore the edge details of the noisy image. In addition, this paper introduces a compound loss function consisting of Charbonnier loss and edge loss suitable for multi-stage network structure as the loss function to guide the multistage network for training. Experiments are carried out on AAPM2016 public dataset to verify that the edge protection module can effectively protect image edge details. Compared with other advanced algorithms, the proposed multilevel denoising algorithm is optimal in both subjective and objective evaluation indexes.

Keywords: edge protection; image processing; low-dose CT denoising; multistage

1 INTRODUCTION

Computed tomography (CT) technology is widely used for medical diagnosis [1], image-guided surgery [2], and radiotherapy [3] due to its ability to scan specific parts of the human body rapidly and accurately. But ionizing radiation generated during CT scan can harm the human body and increase the risk of cancer. Currently, the radiation dose of a single CT scan in actual clinical practice ranges from 2-15 mSv, such as approximately 8 mSv for a chest CT scan and 10 mSv for an abdominal CT scan, while the medical recommendation is to receive no more than 20 mSv of radiation per year [4]. In clinical practice, low-dose CT imaging with only 1/4-1/5 of the normal dose is usually used to reduce the risk of CT radiation to humans [5]. Today, low dose has gradually become an important reference index for the development of CT equipment. However, the reduction in radiation dose often generates a large amount of noise, leading to a serious decrease in CT imaging [6] and difficulties in identifying internal tissue structures.

There are three main types of methods available to improve the quality of low-dose CT images (LDCT): projection domain filtering algorithms, iterative reconstruction algorithms, and post-processing methods [7-9]. Different from the previous two methods, the post-processing method can directly process reconstructed CT images without relying on projected data. Various algorithms currently used have excellent denoising effects. For example, the context-based block Matching 3D (BM3D) algorithm was adopted by Chen et al. [10] to achieve higher quality denoising effects. The methods of dictionary learning and sparse representation are also widely used for LDCT image recovery. Based on sparse representation theory, Lu et al. [11] introduced a fast dictionary learning method to process low-dose CT images, which effectively suppressed speckle noise and streak artifacts in LDCT images.

In recent years, with the rapid development of neural networks and deep learning in the field of image processing, Chen et al. [12] applied the first applied convolutional neural network in the postprocessing of low-dose CT. Compared to the traditional method, it has both advantages in visual effect and evaluation index. The RED-CNN network with residual encoder-decoder structure proposed by Chen et al. [13], achieved good results in denoising results. Wu et al. [14] proposed a cascaded CNN denoising algorithm that obtained higher quality LDCT denoised images compared to conventional CNN algorithms. Wolterink et al. [15] first applied the Generative Adversarial Network (GAN) to denoise the LDCT image denoising. Yang et al. [16] used the Wasserstein distance in the GAN network to form a WGAN network for low-dose CT image denoising. Yi et al. [17] used a sharpness detection network to guide the network. The resulting denoised images had very small resolution loss and improved the imaging quality of CT images. Yin et al. [18] proposed DP-ResNet to denoise LDCT denoising based on an asymptotic 3D residual convolutional network.

Taking advantage of the development of convolutional neural networks, many functional structures can be used for image recovery, such as recursive residual learning [19], attention mechanism [20, 21], encoder-decoder structure [22], etc. In the field of medical image denoising, existing denoising models are based on single-level models, such as the encoder-decoder structure [13] and the image generation model [15, 16]. Some work has been done on multistage design. However, most of the models in these works are based on single-stage designs. In contrast to advanced vision tasks, the performance of multistage models has been shown to be more effective than single-stage models. The work of Zamir et al. [25] demonstrated that multistage design which can reconstruct the image progressively in a progressive manner through subnets at each stage is feasible. With the development of technology, new advances have been made in low-dose CT denoising.

Although the above algorithms have achieved good results in LDCT denoising, the following problems still exist:

(1) Most existing algorithms are based on a single-level network structure, which is not conducive to the extraction of large-scale image features or the maintenance of high-level context features of images. In addition, due to

the single network structure, it is impossible to realize the cooperative reconstruction of multi-part LDCT images.

(2) Although existing algorithms have achieved good results in image denoising, they still have the problem of insufficient protection of edge details. In the medical diagnosis, the image details after denoising are too smooth, which is not conducive to measuring internal tissues such as nodules and lesion areas.

(3) Individual root mean square error (MSE) as a traditional objective function is often used in many low-dose CT denoising algorithms to achieve a higher signal-to-noise ratio. However, this pixel-by-pixel comparison method is found to be prone to loss of detailed information and thus blurred recovery results.

Taking into account the special characteristics of medical image processing, a complex and clever balance between the pixel-by-pixel denoising effect and spatial details needs to be maintained in denoising. We propose a denoising method that can combine multistage network and edge protection (MEP-Net).

This method can extract multi-scale information of LDCT images as well as multichannel features to facilitate feature fusion in different dimensions. Aiming at maintenance of the spatial details in LDCT, a portable Edge Protection Module (EPM) was constructed to extract the edge information of the input image and generate multi-channel edge features, which could be added to the multi-level network structure to realize the protection of edge details. At the same time, a loss function that combines the Charbonnier loss of the output of different stages and the edge loss that can constrain the edge features in the network is introduced to train the model, so as to obtain better output results. Finally, the AAPM dataset is used to conduct sufficient experiments to show the effectiveness of the network structure designed in this paper.

Our main contributions can be summarized as follows.

(1) In order to obtain better quality denoised images, a kind of multistage network is constructed which integrates spatial features and channel features of different dimensions, with setting the input of each stage to different sizes to extract multiscale features of LDCT images.

(2) An edge protection module is designed in this paper to protect the edge structures of different organs, nodes and lesions in LDCT images by extracting the edge features of LDCT images and generating multi-channel features. Feature fusion is performed from the extensive features extracted from multilevel network structure to realize edge recovery.

(3) An edge extractor containing Laplace operator is introduced as the edge loss, and it is combined with Charbonnier loss to construct a composite loss function as the final guide to train multistage networks.

The remainder of this paper is organized as follows. In Section 2, we present the proposed MEP-Net in detail. In Section 3, we show our results and analysis. Finally, we discuss some related issues and draw conclusions in Section 4.

2 METHODS

2.1 Denoising Model

The LDCT image has a lot of noise, which is not conducive to clinical diagnosis and quantitative analysis. In the image field, this paper removes the noise from LDCT images using a deep learning method to recover high-quality CT medical images, and the problem of LDCT image recovery problem of LDCT images can be solved by establishing the following mathematical model. Assume that $X \in R^{m \times n}$ is the low-dose CT image, $Y \in R^{m \times n}$ is the corresponding normal-dose CT image, and the relationship between them can be expressed as

$$X = \sigma(Y) \tag{1}$$

where σ denotes the denoising process. The recovery process from the LDCT image X to the NDCT image Y can be regarded as the inverse mapping of Eq. (1). The goal of the image reconstruction is to find the function F that maps the LDCT image X to the NDCT image Y, so it makes the cost function as shown below is minimized:

$$\arg\min_{F} \| F(X) - Y \|_2^2 \tag{2}$$

where F is the best approximation of σ^{-1} and is represented by a learnable convolutional neural network inside deep learning.

2.2 Network Architecture Design

Compared with the conventional single-stage network structure with single input and output, the network structure designed in this paper is based on a multi-stage network structure that can perform image recovery in a gradual manner. The designed network structure uses a codec-based network structure to extract shallow features in the first stage. In the second stage, deeper features are extracted and edge protection modules are introduced to protect edge information from loss. In the last stage of image restoration, a network structure that can restore full-size images is adopted to realize information interaction between different stages through cross stage feature fusion.

What's more, in order to enable the network at each stage to extract original features at different scales and achieve cross-stage feature fusion, inputs at different scales are used at each stage. The first stage divides the input image into four parts, of which 1/4-scale part is used as input, and 1/2-scale input is adopted in the second stage. Finally, the original image with the full size is used as input in the last stage.

As shown in Fig. 1, the image reconstruction framework proposed in this paper consists of three stages to gradually recover the image. The first two stages have similar structures, first generating multi-channel original image features by 3 × 3 convolution, then integrating the extracted multi-channel features through the channel attention mechanism (CAB), and generating a weighted feature map by maximum mean pooling to assign weights that mark different channel importance. The U-shaped encoder-decoder structure is used to extract the high-dimensional features of the image step by step through the encoder and then gradually recover to the input feature dimension through the decoder in reverse, so that the network can extract more contextual information. In the final stage, the original resolution block (ORB), which can directly process the original image, is used to generate features with high resolution and rich image information to

avoid the loss of image information. The implementation process is shown in Fig. 1b. This stage of the network mainly consists of edge protection module, Channel Attention Block (CAB) and convolution, and does not contain operations on image resolution, thus preserving more image details. At the same time, this paper also performs cross-stage feature fusion between each stage, as represented by the purple dashed arrows in Fig. 1, which are the features extracted by the supervised attention mechanism (SAM) in the previous stage. The features are concatenated with the multidimensional features obtained by CAB in the next stage, i.e., the "Cat" operation in Fig. 1. And the features from different stages are fused together to connect the information between the different stages.

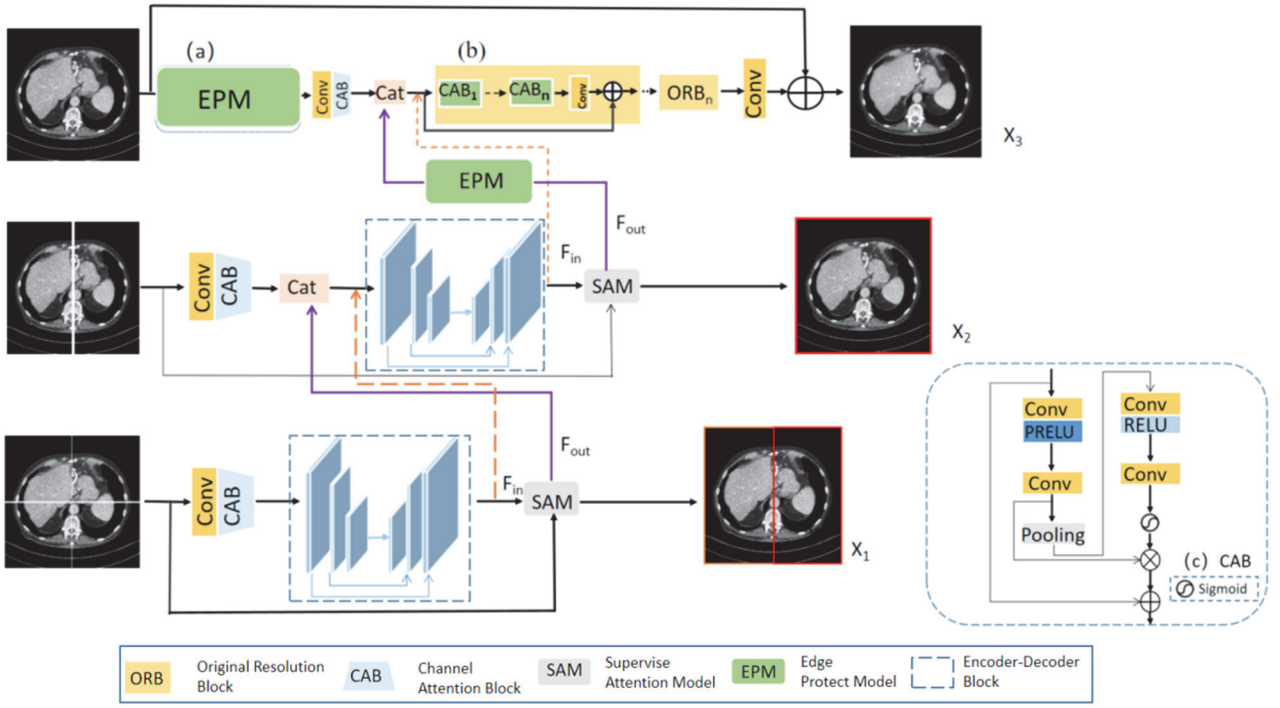


Figure 1 Multi-stage network architecture with emphasis on edge protection and it contains: a) Edge protection module; b) Original resolution module; c) Channel attention block

In order to provide a supervised forecast for the output of each stage while generating an attention map to improve network performance the attention mechanism supervision is introduced in the final output section of the first two stages, as shown in Fig. 2. Firstly, this module generates the loss of the recovered image with the normal dose image for the corresponding stage, providing supervision for the image recovery in each stage. Secondly, Sigmoid is used to generate attention maps to label the features with different levels of importance.

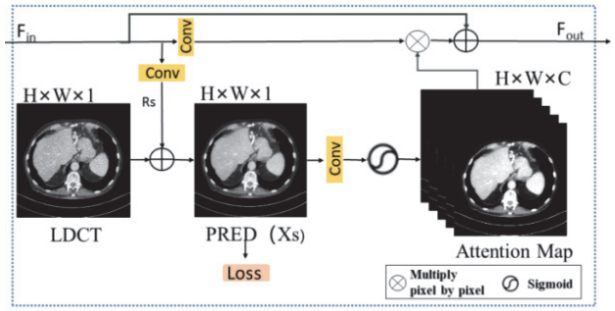


Figure 2 Supervise Attention Module (SAM)

The input F_{in} of SAM represents the features with $H \times W \times C$ dimensions output by the decoder, and in this paper, the *Rs* image (Shown in Fig. 2) is first generated by a 1×1 convolution. Then the low-dose image with dimension $H \times W \times 1$ is summed with the *Rs* image to generate the final prediction image at this stage. Then, the prediction image is generated through convolution and Sigmoid function to generate a pixel attention mask with C channel. These masks are multiplied by the input feature F_{in} refined by 1×1 convolution, and finally added to the input feature to generate an attention-guided feature. The output F_{out} of the supervised attention module is passed to the next stage for use.

2.3 Edge Protection Module

There are many noises in low-dose CT images, which seriously blur the edges of the internal tissues of LDCT images [26]. Although the traditional edge extraction process in the image processing field can effectively extract the edge details of images, the corresponding application scenarios are limited. In this paper, we propose the edge protection module (EPM), which integrates the traditional edge extraction process and deep learning into a unified module so that it can be universally embedded in the network structure and used to improve the quality of reconstructed low-dose CT images.

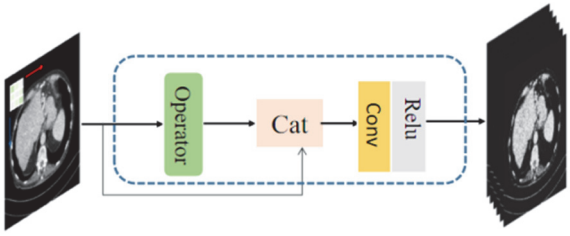


Figure 3 Edge Protection Module

The overall design structure of the module is shown in Fig. 3. Firstly, the gray scale weighting operation is performed on the input image in the pixel-centered 3 × 3 neighborhood, and the horizontal and vertical gradient approximations of the pixel point are calculated by the corresponding operator template to obtain the gradient value of the point, and the operation process is shown in Eq. (3):

$$G_i = \sqrt{G_x(i) + G_y(i)} \tag{3}$$

where G_i denotes the gradient value at the i-th pixel point, $G_x(i)$, $G_y(i)$ denote the gradient values in the x and y direction at that pixel point, respectively. Then, the gradient value of the point is mapped to the multidimensional channel feature by the multichannel convolution operation in deep learning; the process is shown in Eq. (4):

$$F_f = \mathrm{Conv}(G_i) \tag{4}$$

$$F_{\mathrm{eout}} = \mathrm{Re}\,LU(\mathrm{Conv}_{1\times1}(F_f + X(i))) \tag{5}$$

where $F_f \in H \times W \times C$. H, W are the height and width of the image respectively, and C is the number of channels. Finally, the extracted edge features are stitched with the input image $X(i)$ in the channel dimension, and the final edge feature output F_{eout} is obtained by 1 × 1 convolution layer and $ReLU$ activation layer. The process is shown in Eq. (5).

2.4 Cross-stage Feature Integration Fusion Mechanism

The multi-scale feature aggregation strategy can integrate features of different dimensions between different stages, and can effectively fuse the image features extracted by the encoder-decoder structure at different resolutions of the input and fuse them with the features tagged by the supervised attention mechanism.

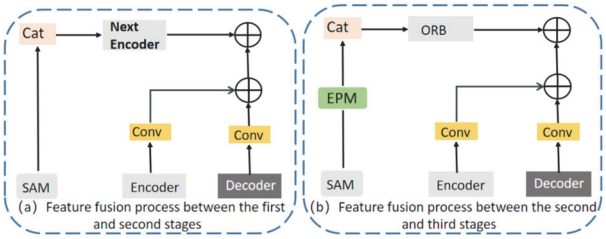


Figure 4 Cross-stage feature integration fusion mechanism a) Feature fusion process between the first and second stages; b) Feature fusion process between the second and third stages

Fig. 4a shows the feature fusion process between the first and second stages, which is divided into two parts, the first part is the multi-channel features provided by SAM, and the second part is the feature integration between the encoder and decoder of the upper and lower stages, and the features are refined by 1 × 1 convolution to make the network information not easily lost. Fig. 4b shows the feature fusion process between the second and third stages. Compared to the first and second stages, the EPM module is added to the output part of the second stage to extract the edge features of the second stage output and add them to the input part of the ORB.

The multiscale feature aggregation strategy can further enhance detail information to suppress noise, as the network extracts image information through the encoder-decoder structure. Meanwhile, the multiscale features extracted in the previous stage are added to the next stage segment, which helps to enrich the features in the next stage. By introducing the image edge features between stages 2 and 3 and extracting the edge features of the original image directly at the beginning position of stage 3, the edge features after noise removal in the first two stages and the original edge features are fused in the last stage to avoid extracting the edge features of a single noisy image and a smooth image after noise removal and to restore the real edge features to the maximum extent.

2.5 Loss Function

Since this paper uses a multi-stage network structure, the original input image of each stage is set as X_i, F is the mapping function in Eq. (6), and Y is the normal dose image. α is the trainable parameter introduced in the edge protection module. Charbonnier loss is introduced to optimize the end-to-end network structure proposed in this paper, and the Charbonnier loss formula is as follows:

$$L_{\mathrm{char}} = \sqrt{\left\| F\left(X_i, \alpha\right) - Y \right\| + \varepsilon^2} \tag{6}$$

Similarly, to the form of the MSE loss function, where ε denotes a very small term, usually set empirically and set to 10^{-3} in the experiments in this paper. Y denotes NDCT as the label to calculate the value of the loss function.

We also introduce the edge loss function L_{edge}, defined as:

$$L_{\mathrm{edge}} = \sqrt{\left\| \Delta\left(X_s\right) - \Delta\left(Y\right) \right\| + \varepsilon^2} \tag{7}$$

where Δ denotes the Laplace operator, and for all experiments, the constant ε is empirically set to 10^{-3}. At this point, we can obtain the composite loss function proposed in this paper.

For the multistage method proposed in this paper, the loss of each stage is considered comprehensively in this paper, and S is used to represent the different stages, and edge loss is introduced, and the final loss function can be expressed as:

$$L_{\mathrm{Compound}} = \frac{1}{N}\sum_{s=1}^{n}\left[L_{\mathrm{char}}(X_s, Y) + \lambda L_{\mathrm{edge}}(X_s, Y) \right] \tag{8}$$

3 EXPERIMENT AND RESULTS

3.1 Datasets

This section uses the clinical data set of the "2016 NIH-AAPM-Mayo Clinic Low-Dose CT Grand Challenge", licensed by Mayo Clinics, for training and testing the network. The dataset consists of projection data and slice thickness of 1 mm and 3 mm reconstruction data,

respectively. In this paper, we use 3 mm thick reconstructed images, which contain 2378 normal-dose CT images from 10 anonymous patients and their corresponding low-dose (quarter-dose) CT images, each of which is 512 × 512 pixels in size [27]. In the experiment, 267 pairs of images from 9 patients were selected as the training set, and 211 pairs of images from another patient were selected as the test set.

3.2 Experimental Setup

The software environment for implementation is Windows 10, and the hardware environment is an Intel Core i7 and Nvidia GeForce GTX 2080Ti GPU. The programming development environment is CUDA-Toolkit 10.0. The programming language is Python 3.6, and the deep learning framework is PyTorch.

The network in this paper is a full convolutional network with multi-stage struct, so it can simultaneously train image patches and test full-size images (512 × 512), which is helpful to reduce the amount of calculation, improve the training speed, increase the number of training samples and avoid over-fitting.

3.3 Ablation Experiments

To investigate the contribution of different components of MEPNet in improving recovered images, we selected different network architectures to conduct experiments on AAPM data sets. First, the effect of edge protection modules on network performance is verified and the effects of different operator templates are compared. Secondly, the performance of different loss combinations is verified and compared with the traditional loss function MSEloss and the single Charbonnier loss.

In order to objectively evaluate the performance of the method proposed in this paper, the peak signal-to-noise ratio (PSNR), spatial structure similarity (SSIM), and root mean square Error (RMSE) are used as evaluation indexes for the effect of low-dose CT image recovery for image quality assessment.

(1) The edge protection module

In order to verify the effectiveness of introducing the edge protection module, we compared the advantages and disadvantages of introducing the edge protection module and different operator templates. The first and last rows of Tab. 1 show that the introduction of the edge protection module improves approximately 0.4 dB in PSNR and approximately 2% in the SSIM. It can be seen that the edge protection module is effective.

Table 1 Introduction of Edge Protection Module and comparison of different operators

	PSNR	SSIM	RMSE
Without EPM (MPR)	26.2294 ± 1.2601	0.8120 ± 0.03904	0.0597 ± 0.0094
EPM with PREWITT	26.3598 ± 1.6864	0.8193 ± 0.03916	0.04925 ± 0.0106
EPM with Kirsch	26.4650 ± 1.7063	0.8202 ± 0.03871	0.04867 ± 0.0110
EPM with SOBEL (proposed)	**26.6437 ± 1.7924**	**0.8313 ± 0.3798**	**0.0468 ± 0.1110**

Meanwhile, in order to investigate the possible effects between different first-order differential operators, the commonly used first-order edge detection operator templates are selected for comparison in this paper. The Prewitt operator, Kirsch operator and Sobel operator, which are suitable for processing grayscale gradient images, are selected for comparison. And the comparison results are shown in Tab. 1, it can be seen that the Sobel operator achieves the best objective evaluation index. The results are shown in Tab. 1.

(2) Optimal loss function

In this part, we compare with the different selections of the loss function. The impact of different form on the final denoising results is considered. The results of different loss functions are shown in Tab. 2, where the loss function in the third row is a combination of the MSE loss function and the edge loss function (MSEloss + ELOSS), and the fourth row is a combination of the composite loss function proposed in this article (Closs + Eloss). It can be seen that introducing edge loss function can effectively improve the spatial structure similarity between the restored image and the normal dose image.

Table 2 Comparison of objective indicators of different loss functions

	PSNR	SSIM	RMSE
MSE Loss	25.7357 ± 2.7384	0.8025 ± 0.06872	0.0548 ± 0.0267
Closs	26.3412 ± 2.3657	0.8143 ± 0.07453	0.0473 ± 0.0194
MSEloss + ELOSS	26.1625 ± 2.6348	0.8130 ± 0.04527	0.0426 ± 0.0283
Closs+Eloss (proposed)	**26.6437 ± 1.7924**	**0.8313 ± 0.03798**	**0.0468 ± 0.0111**

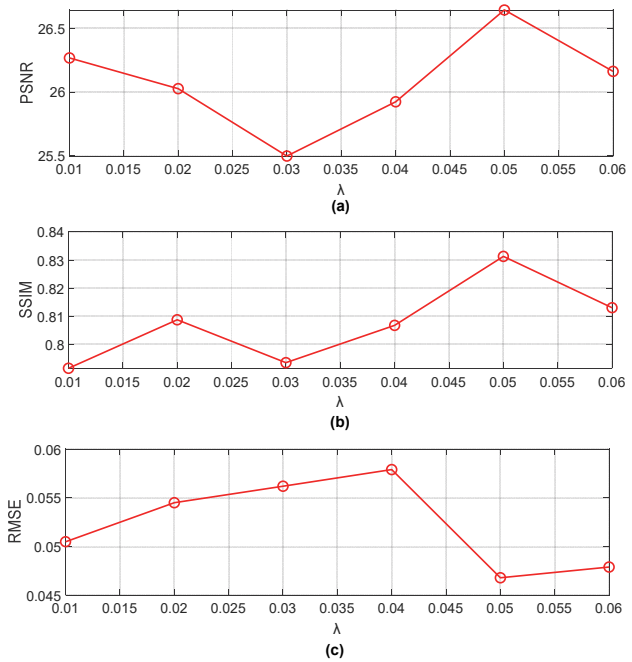


Figure 5 Comparison of objective indexes under different λ

We also conducted comparison experiments on the parameter λ in the compound loss function to determine the optimal value. As can be seen in Fig. 5, the fluctuation among the objective evaluation indicators is larger with the increasing λ. When λ is less than 0.03, the change trends of the three objective evaluation indicators are inconsistent, indicating that too small λ weights may not effectively balance the two different parts of the compound loss

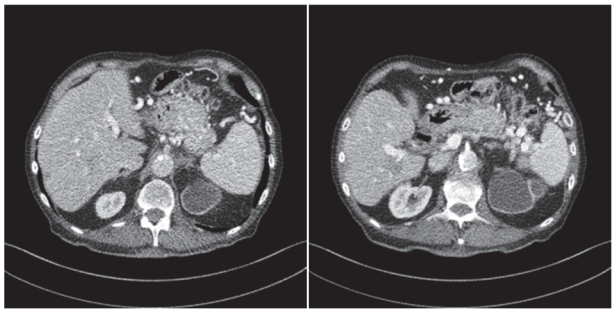
function, resulting in unbalanced indicators in the training results.

When λ is less than 0.03, the change trends of the three objective evaluation indicators are not consistent, indicating that too small λ weights may not be able to effectively balance two different parts of the composite loss function. With λ further increasing to 0.05, the objective evaluation metrics reach the optimum between PSNR, SSIM and RMSE. As the value of λ continues to increase, the increase of edge loss weights in the composite loss function leads to a decrease of PSNR. In summary, when λ is 0.05, the composite loss function can effectively guide the network training and obtain the best denoising effect.

3.4 Comparative Experiments and Analysis of Results

3.4.1 Comparison of Subjective Effects

In order to compare the performance of the proposed algorithm, two abdominal CT images with rich details were selected as samples for comparison which are shown in Fig. 6a and Fig. 6b.



a) Sample 1 b) Sample 2

Figure 6 Low dose CT image

The comparison algorithms selected in this paper include BM3D algorithm, RED-CNN algorithm, WGAN-VGG algorithm and MPRnet algorithm, all of which have excellent denoising effects. The experimental results are evaluated using both subjective effects and objective metrics.

The subjective effect is used to visually highlight the structural form of the image and then the objective metrics are used to evaluate the degree of superiority of the algorithms. Fig. 7 shows the general comparison of Fig. 6a, and Fig. 8 shows the zoomed-in comparison of the ROI region in Fig. 7. In the overall comparison of Fig. 7 and Fig. 8, in terms of visual effect, the selected comparison algorithms all possess a certain denoising effect. Fig. 7e is the result of the BM3D algorithm, which can remove part of the noise. Figs. 7c, 7d, and 7f show the results of the MPRNet algorithm, the W-GAN algorithm, and the REDCNN method, respectively. Compared to the LDCT image shown in Fig. 7b, it can be seen that both images have denoising effects, and the deep learning method has improved visual effects. The comparison shows the method recovery results of the proposed method in this paper, as shown in Fig. 7g, which can present more detailed information in the visual effect, and the image recovery results are clearer. Noise and artifacts are eliminated while edge details of the image are better preserved, and the image quality is closest to that of NDCT image.

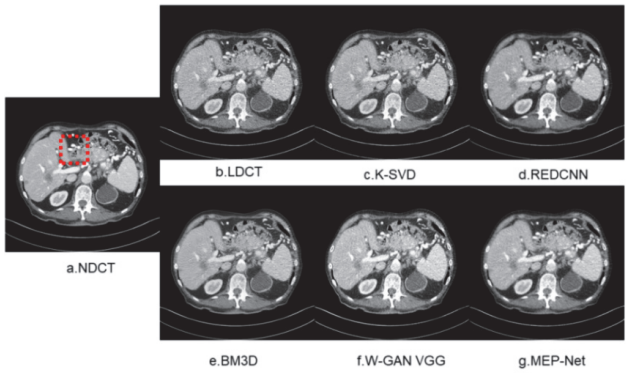


Figure 7 Overall comparison of sample 1 a) NDCT image; b) LDCT image; c) MPRNet result image; d) REDCNN result image; e) BM3D result image; f) W-GAN VGG result image; g) Results of the MEP-Net （Proposed）

The enlarged area of the combined ROI in Fig. 8 shows that the edges of the area circled by the red dotted line are too fuzzy and the clear details of the edge part are lost. Also, we compare the ROI enlargement in Fig. 8, the difference between the other four denoising algorithms and the NDCT image is also larger. In terms of the overall visual effect, the proposed method in this paper has better results in terms of intuitive noise comparison and is closer to the NDCT image. By comparing the ROI region enlargements of several denoising methods in Fig. 8, the protection of the edge of the image can be achieved.

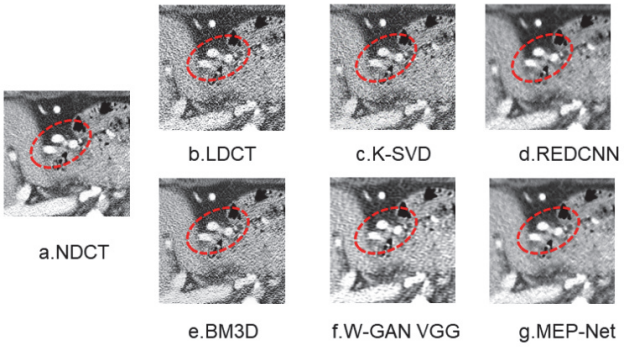


Figure 8 Enlarged comparison of ROI area of sample 1 a) Enlarged NDCT ROI; b) LDCT ROI enlarged; c) MPRNet ROI enlarged; d) REDCNN ROI enlarged; e) BM3D ROI enlarged; f) W-GAN VGG ROI enlarged; g) The ROI of the MEP-Net (Proposed)

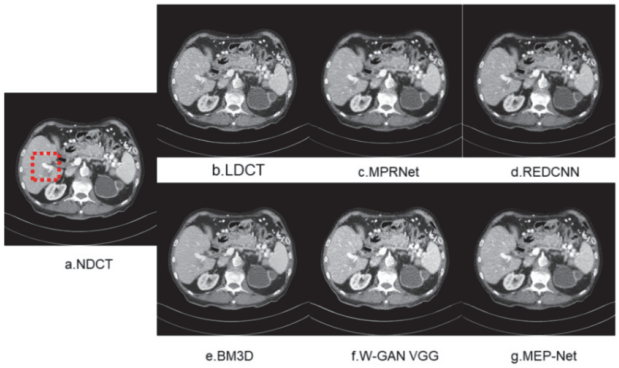


Figure 9 Overall comparison of sample 1 a) NDCT image; b) LDCT image; c) MPRNet result image; d) REDCNN result image; e) BM3D result image; f) W-GAN VGG result image; g) Results of theMEP-Net (Proposed)

Fig. 9 shows an overall comparison of the selected sample (b), and Fig. 10 shows a zoomed in comparison of the ROI region in Fig. 9.

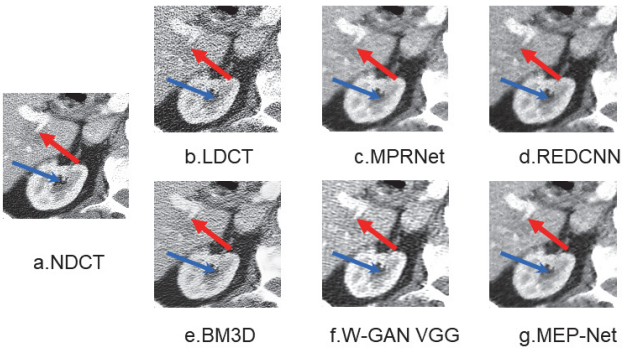


Figure 10 Enlarged comparison of ROI area of sample 1 a) Enlarged NDCT ROI; b) LDCT ROI enlarged; c) MPRNet ROI enlarged; d) REDCNN ROI enlarged view; e) BM3D ROI enlarged; f) W-GAN VGG ROI enlarged; g) The ROI of the MEP-Net (Proposed)

The comparison of the overall visual effect in Fig. 9 shows that the algorithm proposed in this paper also has a good denoising effect on CT images with different structures. Comparing the parts marked by the red and blue arrows, we have the proposed algorithm in Fig. 10g is clearer in edge details, and the contrast with the surrounding background is also more obvious. Comparing all the ROI zoomed region in Fig. 10, the proposed algorithm not only retains more edge details, but also comes closest to the NDCT image in terms of intuitive effect, and the overall image is cleaner. The comparison between Fig. 7 and Fig. 10 shows that the proposed algorithm achieves the best visual effect, retains detailed information similar to the normal dose CT images, and removes most of the noise. Therefore, the algorithm proposed in this paper has the best visual effect.

3.4.2 Comparison of Subjective Effects

Figs. 11 and 12 are histograms of objective evaluation indexes of test sample (a) and test sample (b). It can be seen that the method proposed in this paper achieves better performance metrics in both test plots.

Table 3 Comparison of objective indexes of different algorithms in the test set

	PSNR	SSIM	RMSE
LDCT	21.6048 ± 1.9739	0.7928 ± 0.04526	0.0858 ± 0.0219
BM3D	23.9134 ± 1.8372	0.8122 ± 0.3127	0.0664 ± 0.0210
RED-CNN	25.9159 ± 1.70845	0.8142 ± 0.04050	0.0487 ± 0.0110
WGAN-VGG	24.9213 ± 2.0376	0.8021 ± 0.05673	0.0744 ± 0.0109
MPRnet	26.2294 ± 1.2601	0.8120 ± 0.03904	0.0597 ± 0.0094
MEPNet (proposed)	**26.6437 ± 1.7924**	**0.8313 ± 0.03798**	**0.0468 ± 0.0111**

Tab. 3 shows the comparison of the average objective metrics of different algorithms in the test set. It can be seen that the method proposed in this paper also obtains the best results in terms of average evaluation metrics.

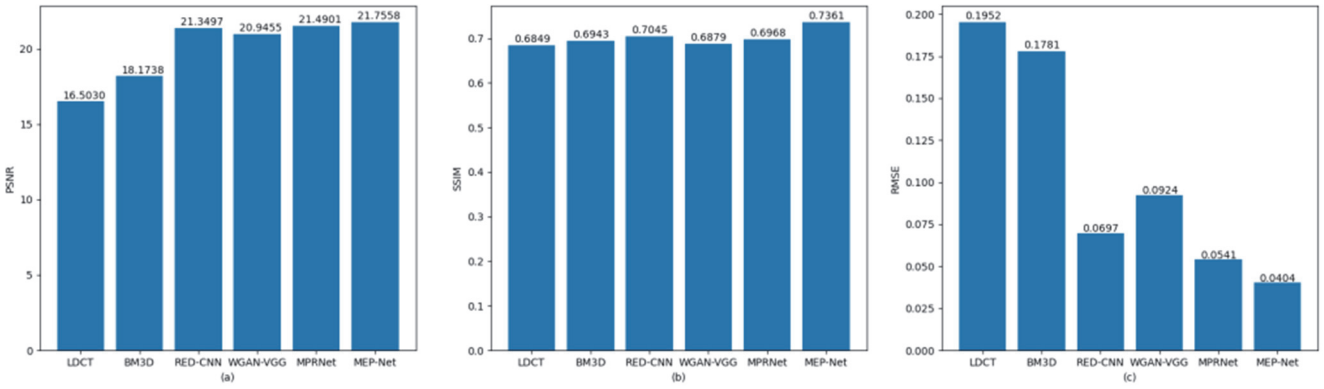


Figure 11 Sample 1 comparison of objective indexes

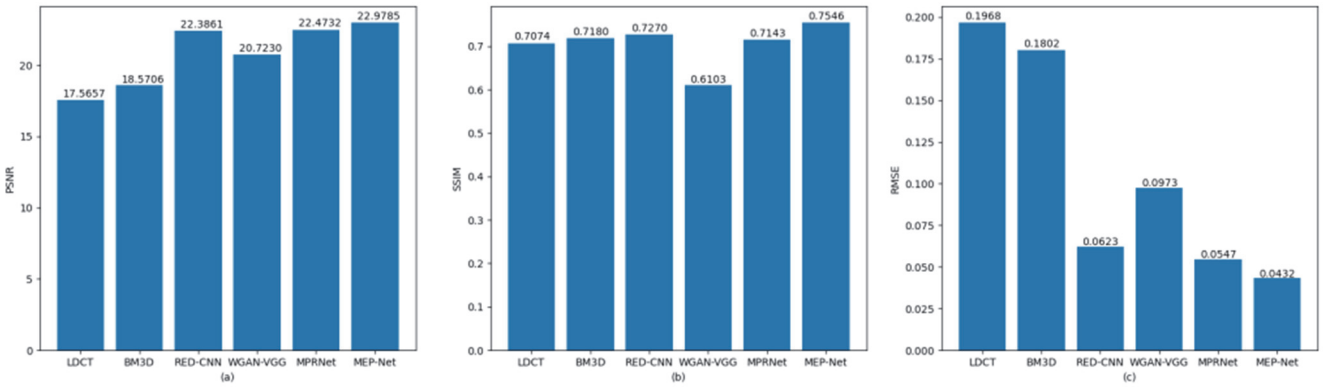


Figure 12 Sample 2 comparison of objective indexes

4 CONCLUSIONS

For the low image quality of low-dose CT images, the existing post-processing methods have the problems of too smooth images and insufficient edge details in the reconstruction process, and this paper proposes a low-dose CT image denoising method that combines the edge protect model and the deep learning. Compound loss is introduced

under the comprehensive consideration of the impact of each stage on the final recovered image. The proposed method makes full use of contextual links in the network space, effectively obtains the global information of spatial features, achieves a comprehensive grasp of the location of key features, and improves the richness of the information output by the network. In the comparison of subjective visual effects, the method can well reduce image noise and maintain the edge details of organs while maintaining image information. The proposed network achieves better visual results and the best objective evaluation metrics (PSNR, SSIM, RMSE) compared to other denoising algorithms on the AAPM dataset. In summary, our method is advanced in recovering normal-dose CT images from low-dose CT images. In the future, we plan to optimize the proposed EPM-Net to further investigate the edge features in LDCT images for better fusion with other basic image features to obtain better results.

5 REFERENCES

[1] Coxson, H. O., & Hogg, J. C. (2001). Erratum: a quantification of the lung surface area in emphysema using computed tomography. *American Journal of Respiratory and Critical Care Medicine*, *163*(6), 1500-1501. https://doi.org/10.1164/ajrccm.163.6.16362b

[2] Rodrigues, J. C. L., Pierre, A. F., Hanneman, K., Cabanero, M., & Nguyen, E. T. (2019). CT-guided microcoil pulmonary nodule localization prior to video-assisted thoracoscopic surgery: diagnostic utility and recurrence-free survival. *Radiology*, *291*(1), 181674. https://doi.org/10.1148/radiol.2019181674

[3] Huynh, E., Coroller, T. P., Narayan, V., Agrawal, V., Hou, Y., & Romano, J. et al. (2016). CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. Radiotherapy and Oncology. https://doi.org/10.1016/j.radonc.2016.05.024

[4] Shao, Y. H., Tsai, K., Kim, S., Wu, Y. J., & Demissie, K. (2020). Exposure to tomographic scans and cancer risks. *JNCI Cancer Spectrum*, *4*(1). https://doi.org/10.1093/jncics/pkz072

[5] Hart, D. & Wall, B. F. (2004). UK population dose from medical x-ray examinations. *European Journal of Radiology*, *50*(3), 285-291. https://doi.org/10.1016/S0720-048X(03)00178-5

[6] Kong, X. Q., Ping, H., & Feng, G. S. (2003). Low dose CT lung scanning technology is worth advocating and popularizing. *Journal of Clinical Radiology*, (07), 533.

[7] Zhang, Y., Yi, B., Wu, C., & Feng, Y. (2018). Low-dose CT image denoising method based on convolutional neural network. *Acta Optica Sinica*.

[8] Zamyatin, A., Katsevich, G., Krylov, R., Shi, B., & Zhi, Y. (2014). Adaptive multi-scale total variation minimization filter for low dose CT imaging. Medical Imaging 2014: Image Processing. *International Society for Optics and Photonics*. https://doi.org/10.1117/12.2043841

[9] Ma, J., Huang, J., & Feng, Q. (2011). Low-dose computed tomography image restoration using previous normal-dose scan. *Medical Physics*, *38*(10). https://doi.org/10.1118/1.3638125

[10] Chen, L. L., Gou, S. P., Yao, Y., Jing, B., & Ke, S. (2016). Denoising of low dose CT image with context-based BM3D. *TENCON 2016 - 2016 IEEE Region 10 Conference*. https://doi.org/10.1109/TENCON.2016.7848089

[11] Mudan, L. V. (2017). *Research on low dose CT image restoration based on sparse representation and dictionary learning*. Master's thesis, Donghua University of Technology.

[12] Hu, C., Yi, Z., Weihua, Z. et al. (2017). Low-dose CT via convolutional neural network. *Biomedical Optics Express*. https://doi.org/10.1364/BOE.8.000679

[13] Chen, H., Zhang, Y., Kalra, M. K., Lin, F., Chen, Y., Liao, P. et al. (2017). Low-dose CT with a residual encoder-decoder convolutional neural network (red-cnn). *IEEE Transactions on Medical Imaging*, *36*(99), 2524-2535. https://doi.org/10.1109/TMI.2017.2715284

[14] Wu, D., Kim, K., Fakhri, G. E., & Li, Q. (2017). A cascaded convolutional neural network for x-ray low-dose CT image denoising.

[15] Wolterink, J. M., Leiner, T., Viergever, M. A., & Išgum, I. (2017). Generative adversarial networks for noise reduction in low-dose CT. *IEEE Transactions on Medical Imaging*, *36*(12), 2536-2545. https://doi.org/10.1109/TMI.2017.2708987

[16] Yang, Q., Yan, P., Zhang, Y., Yu, H., Shi, Y., Mou, X. et al. (2018). Low-dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss. *IEEE Transactions on Medical Imaging*, 1348-1357. https://doi.org/10.1109/TMI.2018.2827462

[17] Xin, Y. & Babyn, P. (2018). Sharpness-aware low-dose CT denoising using conditional generative adversarial network. *Journal of Digital Imaging*, *31*(5). https://doi.org/10.1007/s10278-018-0056-0

[18] Xiangrui, Z., Qianlong, Y., Yang, J. et al. Domain Progressive 3D Residual Convolution Network to Improve Low-Dose CT Imaging. *IEEE transactions on medical imaging*, *38*(12). https://doi.org/10.1109/TMI.2019.2917258

[19] Anwar, S. & Barnes, N. (2019). Real Image Denoising with Feature Attention. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. https://doi.org/10.1109/ICCV.2019.00325

[20] Zhang, Y., Li, K., Zhong, B., & Fu, Y. (2019). Residual non-local attention networks for image restoration.

[21] Ma, J. H., Huang, J., Feng, Q. et al. (2011). Low-dose computed tomography image restoration using a previous normal-dose scan. *Medical Physics*, 38(10), 5713-5731. https://doi.org/10.1118/1.3638125

[22] Brooks, T., Mildenhall, B., Xue, T., Chen, J., & Barron, J. T. (2019). Unprocessing Images for Learned Raw Denoising. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/CVPR.2019.01129

[23] Chen, Y., Wang, Z., Peng, Y., Zhang, Z., Yu, G., & Sun, J. (2018). Cascaded pyramid network for multi-person pose estimation. *CVPR*, *1*. https://doi.org/10.1109/CVPR.2018.00742

[24] Cheng, B., Chen, L. C., Wei, Y., Zhu, Y., Huang, Z., Xiong, J. et al. (2019). Spgnet: semantic prediction guidance for scene parsing. *IEEE*. https://doi.org/10.1109/ICCV.2019.00532

[25] Zamir, S., Arora, A., Khan, S., Hayat, M. Khan, F., Yang, M., & Shao, L. (2021). Multi-stage Progressive Image Restoration. https://doi.org/10.1109/CVPR46437.2021.01458

[26] Liang, T., Jin, Y., Li, Y., Wang, T., Feng, S., & Lang, C. (2020). Edcnn: edge enhancement-based densely connected network with compound loss for low-dose CT denoising. https://doi.org/10.1109/ICSP48669.2020.9320928

[27] Shan, H., Zhang, Y., Yang, Q., Kruger, U., Kalra, M. K., Sun, L. et al. (2018). 3d convolutional encoder-decoder network for low-dose CT via transfer learning from a 2d trained network. *IEEE Transactions on Medical Imaging*, *37*(6), 1522. https://doi.org/10.1109/TMI.2018.2832217

Contact information:

Zhitao GUO, Associate Professor
(Corresponding author)
Hebei University of Technology,
Xiping Road 5340#, Beichen District, TianJin
E-mail: mrnow@hebut.edu.cn

Feng ZHOU
Hebei University of Technology,
Xiping Road 5340#, Beichen District, TianJin
E-mail: 1226406144@qq.com

Yuqing CHEN
Hebei University of Technology,
Xiping Road 5340#, Beichen District, TianJin
E-mail: 970438119@qq.com

Jinli YUAN, Associate Professor
Hebei University of Technology,
Xiping Road 5340#, Beichen District, TianJin
E-mail: Jinli_yuan@hebut.edu.cn