MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI231065

A Multi-Resolution Denoising Method for Low-Dose CT Based on the Reconstruction of Wavelet High-Frequency Channel

Jinnan HU[a], Peijun HU[a], Yiwei GAO[a], Yanxia ZHAO[a] and Jingsong LI[a,b,1]
[a] *Research Center for Healthcare Data Science, Zhejiang Laboratory, Hangzhou, China*
[b] *Engineering Research Center of EMR and Intelligent Expert System, Ministry of Education, College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China*

Abstract. Computed tomography (CT) is widely applied in contemporary clinic. Due to the radiation risks carried by X-rays, the imaging and post-processing methods of low-dose CT (LDCT) become popular topics in academia and industrial community. Generally, LDCT presents strong noise and artifacts, while existing algorithms cannot completely overcome the blurring effects and meantime reduce the noise. The proposed method enables CT extend to independent frequency channels by wavelet transformation, then two separate networks are established for low-frequency denoising and high-frequency reconstruction. The clean signals from high-frequency channel are reconstructed through channel translation, which is essentially effective in preserving detailed structures. The public dataset from Mayo Clinic was used for model training and testing. The experiments showed that the proposed method achieves a better quantitative result (PSNR: 37.42dB, SSIM: 0.8990) and details recovery visually, which demonstrates our framework can better restore the detailed features while significantly suppressing the noise.

Keywords. Low-does CT, denoising, wavelet transformation, reconstruction

1. Introduction

Computed tomography (CT) as a non-invasive method, is significantly important in clinical examination [1]. However, due to the radiation risks, how to use the low-dose CT (LDCT) instead of the full-dose CT (FDCT), thereby reducing the amount of radiation absorbed by patients, has become a vital topic in academia and industrial community. Compared with the FDCT, it is more likely to observe serious noise and distortions in LDCT [2], which hinders the application of LDCT in clinic.

In the past few years, researchers have designed some post-processing algorithms aiming to improve the image quality of LDCT, which can be categorized into filtering-based [3,4] and deep learning-based methods [2,5,6]. The filtering algorithms usually gain limited improvement compared to the deep-learning based models, while the convolutional neural network will bring common problems like blurring effects [6].

[1] Corresponding Author: Jingsong LI, email: ljs@zju.edu.cn.

Therefore, denoising the LDCT while keeping the image details is a critical issue in the deep learning-based models.

Wavelet Transform (WT) can extract the image information in both spatial and frequency domains. The high-frequency wavelet coefficients contain the detailed information of the image but more likely to be polluted by noise. Therefore, some methods based on wavelet transform attempt to abandon, filtering by a threshold or directly denoising the high-frequency coefficients. We noticed that under the higher-level wavelet transform of the LDCT, the signals in some high-frequency channels are overwhelmed by noise, which makes it impractical to restore the original signal from the high-frequency channel. Hence, we propose to recover the high-frequency information from the reliable low-frequency channel, which is inferred as Channel Translation in this paper.

Recently, great progress has been made in style transfer [7-9], which enables images of one style convert to another. To achieve realistic texture pattern synthesis, based on the pix2pix network, the gauGAN proposes a spatially adaptive normalized SPADE model [9]. We notice that the detailed information in the form of oscillations contained in high-frequency coefficients, is similar to the textures in natural images. Therefore, we propose a novel model named Denoising and Channel Translation Neural Network (DCTNN) for the LDCT denoising problem. Specifically, we innovatively introduced SPADE layers to the pix2pix framework to realize the information recovery of high-frequency coefficients from the low-frequency channel, which can better retain details.

2. Methods

2.1. System overview

As shown in Figure 1, the proposed method includes wavelet grouping, low-frequency denoising network, high-frequency translation network and Inverse Wavelet Transformation (IWT). The LDCT and FDCT generate 16 wavelet coefficients respectively by second-level wavelet transform, which are divided into two groups according to their frequency channel. The low-frequency coefficients are denoised through the CNN-based network. The SPADE-based translation network is built to restore the information of high-frequency group from the low-frequency ones. Finally, the denoised whole CT images are obtained through IWT of two groups.

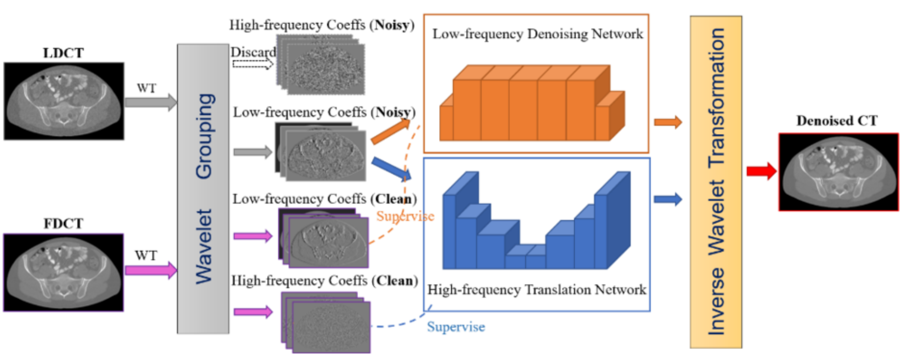

Figure 1. The system pipeline of the proposed method, where the coeffs is short for the coefficients.

2.2. Discrete wavelet transformation and wavelet grouping

The N^{th}-level discrete WT can extend the two-dimensional CT image to 4^N frequency channels, in each a subgraph can be obtained, also known as the coefficient. In our work, 16 wavelet subgraphs of both LDCT and FDCT were generated through second-level WT, then divided into the low-frequency group and the high-frequency one according to the PSNR and SSIM. Further, the subgraphs in each group are concatenated by channels as the input to the subsequent networks.

2.3. Low-frequency denoising network and High-frequency translation network

The low-frequency denoising network is based on the deep-learning denoising model DnCNN [10], including ten convolutional layers, the batch normalization layer and the ReLU layer, also the residual learning is performed through the jump connection.

In Isola's study [7], it was illustrated that the UNet [11] structure has a strong superiority over the AutoEncoder and plain CNN architectures. Therefore, the high-frequency channel translation network proposed in this paper is based on the UNet, plus the SPADE [9] to help recover the detailed edges of the high-frequency subgraphs.

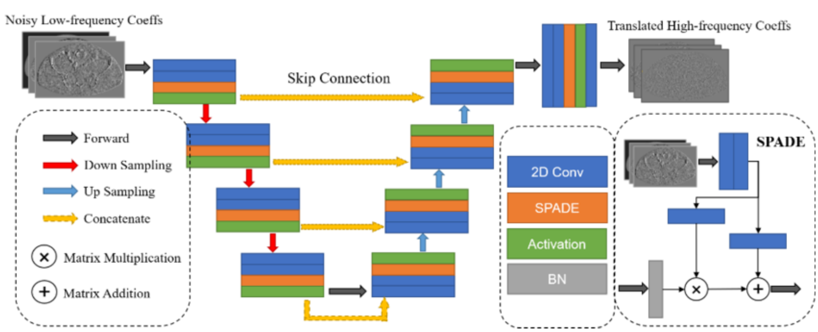

Figure 2. The structure of SPADE-based high-frequency translation network.

2D Spatial Adaptative Normalization (**SPADE**), Let $\boldsymbol{M} \in R^{H\times W\times N}$ be the union of all input low-frequency coefficients, where H, W, N are the height, width and channels of the input. Consider that h^i is the activation of the i^{th} layer of the network for a batch of n samples, and C^i is for the number of the input channels. Then the value after SPADE layer, can be written as:

$$\gamma^i{}_{c,x,y}(\boldsymbol{M})\frac{h^i{}_{n,c,x,y}-\mu^i{}_c}{\sigma^i{}_c}+\beta^i{}_{c,x,y}(\boldsymbol{M}) \quad , \quad (1)$$

where the subscripts in $h^i{}_{n,c,x,y}$ represent its dimensions, and $\mu^i{}_c$ and $\sigma^i{}_c$are the mean and standard deviation of the activation value for channel c. The parameters $\gamma^i{}_{c,x,y}(\boldsymbol{M})$ and $\beta^i{}_{c,x,y}(\boldsymbol{M})$ are obtained in SPADE.

$$\mu^i{}_c=\frac{1}{NH^iW^i}\sum_{n,x,y}h^i{}_{n,c,x,y} \quad (2)$$

$$\sigma^i{}_c=\sqrt{\frac{1}{NH^iW^i}\sum_{n,x,y}\left(\left(h^i{}_{n,c,x,y}\right)^2-\left(\mu^i{}_c\right)^2\right)} \quad (3)$$

3. Results

3.1. Data preparation

The dataset in this paper is from the Mayo Grand Challenge, including 10 patients with a z-axis spatial resolution of 3 mm. In the experiment, the Gaussian noise is applied to simulate LDCT images.

We apply the second-level Haar wavelet transform to obtain wavelet subgraphs of LDCT and FDCT in 16 frequency domains, and divide them into low-frequency group and high frequency group with 20dB of PSNR as the threshold. The low group contains 10 wavelet channels and the high group contains 6 channels.

3.2. Experiment setup

The experiments were conducted on a workstation with a Nvidia 3090 24G graphic card. For the denoising network training, the batch size is 64, the learning rate is $2\times 1e-4$, and the epochs are 200; For the translation network, the batch size is 32, the learning rate is $5\times 1e-5$, and the epochs are 500. Both networks utilized the mean squared error (MSE) as loss function, the Adam optimizer, the cosine annealing learning strategy, and the Gaussian kernel to initialize convolutional weights.

3.3. The denoising results of LDCT

Figure 3 and Table 1 show the visualized and quantitative results of DCTNN and compared methods. Filtering methods cannot clearly preserve the detailed edges, and the DnCNN also loses some texture. By translating the high-frequency subgraph, our method obtains an authentic result in terms of the SSIM.

Table 1. The quantitative measurements of DCTNN and other approaches. **Bold** represents the greatest.

	PSNR (dB)	SSIM
LDCT	31.04	0.8308
Median Filtering	32.12	0.7718
Gaussian Filtering	32.98	0.8002
DnCNN [10]	35.43	0.8326
DCTNN(ours)	**37.42**	**0.8990**

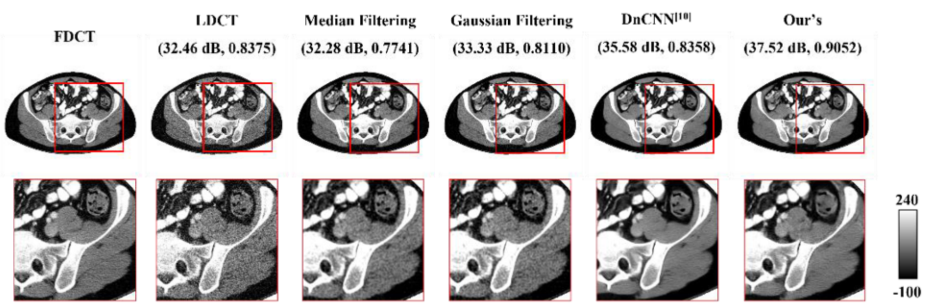

Figure 3. The visual comparison in abdominal window (-100, 240).

4. Discussion

The essential idea is to extend the images into independent frequency channels, then utilizing translation network to recover the clean signals in high-frequency channels, which helps to eliminate the blurring effects in denoising task.

5. Conclusions

In this article, we apply wavelet transformation to the LDCT and divide the coefficients into two groups, then establish denoising and translation network to recover the clean signals separately. The improvements over other methods, both in quantitative and visual results, demonstrate the effectiveness that the proposed DCTNN can better suppress the noise while keeping the detailed information.

Acknowledgements

The project is supported in part by the National Natural Science Foundation of China (No. 12101571) and Key Research Project of Zhejiang Lab (No. 2022ND0AC01).

References

[1] Diwakar M, Kumar M. A review on CT image noise and its denoising. Biomed Signal Process Control. 2018 Apr;42:73-88, doi: 10.1016/j.bspc.2018.01.010.

[2] You C, Yang Q, Shan H, Gjesteby L, Li G, Ju S, Zhang Z, Zhao Z, Zhang Y, Wenxiang C, Wang G. Structurally-sensitive multi-scale deep neural network for low-dose CT denoising. IEEE Access. 2018 Jul;6:41839-55, doi: 10.1109/ACCESS.2018.2858196.

[3] Zhao T, Hoffman J, McNitt-Gray M, Ruan D. Ultra-low-dose CT image denoising using modified BM3D scheme tailored to data statistics. Med Phys. 2019 Jan;46(1):190-8, doi: 10.1002/mp.13252.

[4] Li Z, Yu L, Trzasko JD, Lake DS, Blezek DJ, Fletcher JG, McCollough CH, Manduca A. Adaptive nonlocal means filtering based on local noise level for CT denoising. Med Phys. 2014 Jan;41(1):011908, doi: 10.1118/1.4851635.

[5] Chen H, Zhang Y, Zhang W, Liao P, Li K, Zhou J, Wang G. Low-dose CT via convolutional neural network. Biomed Opt Express. 2017 Jan;8(2):679-94, doi: 10.1364/BOE.8.000679.

[6] Yi X, Babyn P. Sharpness-aware low-dose CT denoising using conditional generative adversarial network. J Digit Imaging. 2018 Oct;31(5):655-69, doi: 10.1007/s10278-018-0056-0.

[7] Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: Proceedings of the 2017 IEEE conference on computer vision and pattern recognition (CVPR); 2017; Honolulu, HI, USA: p. 1125-34, doi: 10.1109/CVPR.2017.632.

[8] Gatys LA, Ecker AS, Bethge M. Image style transfer using convolutional neural networks. In: Proceedings of the 2016 IEEE conference on computer vision and pattern recognition (CVPR); 2016; Las Vegas, NV, USA: p. 2414-23, doi: 10.1109/CVPR.2016.265.

[9] Park T, Liu M-Y, Wang T-C, Zhu J-Y. Semantic image synthesis with spatially-adaptive normalization. In: Proceedings of the 2019 IEEE/CVF conference on computer vision and pattern recognition; 2019 Mar 18; arxiv; p. 2337-46, doi: 10.48550/arXiv.1903.07291.

[10] Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans Image Process. 2017 Feb;26(7):3142-55, doi: 10.1109/TIP.2017.2662206.

[11] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference; 2015 Oct 5-9; Munich, Germany; Springer International Publishing. p. 234-41, doi: 10.48550/arXiv.1505.04597.