Tech Science Press

A C-GAN Denoising Algorithm in Projection Domain for Micro-CT

Lujie Chen[1], Liang Zheng[1], Maosen Lian[1] and Shouhua Luo[1,*]

[1]School of Biological Science and Medical Engineering, Southeast University, Nanjing, China.
*Corresponding Author: Shouhua Luo. Email: luoshouhua@seu.edu.cn.

Abstract: Micro-CT provides a high-resolution 3D imaging of micro-architecture in a non-invasive way, which becomes a significant tool in biomedical research and preclinical applications. Due to the limited power of micro-focus X-ray tube, photon starving occurs and noise is inevitable for the projection images, resulting in the degradation of spatial resolution, contrast and image details. In this paper, we propose a C-GAN (Conditional Generative Adversarial Nets) denoising algorithm in projection domain for Micro-CT imaging. The noise statistic property is utilized directly and a novel variance loss is developed to suppress the blurry effects during denoising procedure. Conditional Generative Adversarial Networks (C-GAN) is employed as a framework to implement the denoising task. To guarantee the pixel-wised accuracy, fully convolutional network is served as the generator structure. During the alternative training of the generator and the discriminator, the network is able to learn noise distribution automatically. Moreover, residual learning and skip connection architecture are applied for faster network training and further feature fusion. To evaluate the denoising performance, mouse lung, milkvetch root and bamboo stick are imaged by micro-CT in the experiments. Compared with BM3D, CNN-MSE and CNN-VGG, the proposed method can suppress noise effectively and recover image details without introducing any artifacts or blurry effect. The result proves that our method is feasible, efficient and practical.

Keywords: Micro-CT; projection domain denoising; C-GAN; variance loss

1 Introduction

In X-ray CT, image noise has a great impact on image quality. For traditional clinic CT, the noise mainly consists of two parts: Poisson noise brought by random X-ray photons and Gaussian noise generated during the detector acquisition. In terms of high-resolution micro-CT imaging, the photon noise is more severe as the power of micro-focus X-ray source is small and the X-ray flux is insufficient, making the noise distribution different from those of clinical CT images. Many researches have been carried out to reduce the micro-CT image noise and improve its quality. Some traditional state-of-art denoising methods like BM3D [1], K-SVD [2] have been applied to CT-image domain and achieved some effects. These approaches are in an iterative way that are quite time-consuming, thus cannot yet fulfill the needs for practical application. Recently, deep learning methods are emerging as a high-efficient alternative to solve such problems and obtain very impressive results [3-5]. However, the performance of above methods is strong dependent on the precise estimation of noise property. When transformed into image domain, the projection noise statistics is no longer distinct, resulting in the limited improvement and blurry effects on the reconstructed CT images. In views of the above problems, we use deep neural network to model the characteristics of the projection domain noise. With the merits of extracting image features and identifying the high-level semantics, the neural network can model the noise characteristics more precisely. By optimizing the loss function, neural networks can learn an end to end noise mapping automatically. Moreover, this method has the potential of real-time application. Once the network is trained successfully, the contaminated images can be denoised in a quite short time. The

specific procedures including methods, network architecture, experiment and results will be elucidated in the following chapters.

2 Method

In the previous researches, the deep learning based methods show advantages in the field of image processing [6,7]. Recently, conditional GAN(C-GAN) [8] is used in Pixel2Pixel network and achieves great performance in image translation. Inspired by the success of Pixel2Pixel network an improved C-GAN is applied as the framework for solving micro-CT denoising. C-GAN consists of two networks, the generator G and discriminator D. In the generator, the residual learning architecture [9] is used to learn a model mapping the noisy image x to the corresponding noise image n, where x is also the condition input. Consequently, the denoised image $\hat{y}$ can be expressed as $\hat{y} = x - n$; The role of the discriminator D is to distinguish the generated image pair $(x; \hat{y})$ from the real one $(x; y)$. Note that the input to D not only contains the denoised image $(\hat{y})$ and real high-quality image (y), but also contains the condition image (x). In this way, the output of D is constrained to specific condition to make accurate discrimination. The training procedure is illustrated in Fig. 1.
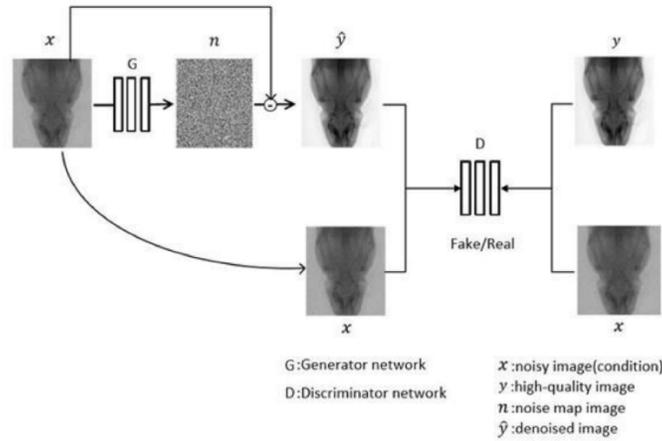
Figure 1: Overview architecture of proposed network

The purpose of G is trying to confuse D, making the generated image $(\hat{y})$ close to real image (y), while D is aiming to figure out the difference between real and fake one. The opposite purpose of G and D forms the adversarial loss, which can be formulated as:

$$\mathcal{L}_{adv}(G, D) = \mathbb{E}_{x,y \sim p_{data}(x,y)}[\log D(x, y)] + \mathbb{E}_{x \sim p_{data}(x)}[\log(1 - D(x, x - G(x)))] \tag{1}$$

Previous researches have explored that it is beneficial to introduce a more traditional loss to GAN loss function [6]. L1 and L2 distance are the most commonly used loss function in regression problems. However, it is reported that L2 loss tends to result in blurring [6]. As a result, in this paper we select L1 distance as the traditional loss rather than L2 distance to encourage less blurring.

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,n}[\|n - G(x)\|_{L1}] \tag{2}$$

As we know, random noise can be modeled by mean and variance. Mean determines the overall property while variance controls the high-frequency information. To make most use of noise statistics, we firstly propose variance loss as follows:

$$\mathcal{L}_{var}(G) = \mathbb{E}_{x,n}[\|Var(n) - Var(G(x))\|_{L1}] \tag{3}$$

The motivation of introducing variance term is that L1 loss is calculated pixel by pixel while noise is in an overall statistical distribution. During the decreasing of L1 loss, the noise distribution may not be optimized at the same time. As a result, it is natural using a statistical indicator to describe the noise. The

purpose of adding variance term is to constrain the solution space and generate a more reasonable noise distribution. Combining these three losses together, the final loss function is as follows:

$$\mathcal{L}_{final}(G, D) = arg \min_G \max_D (\mathcal{L}_{adv}(G, D) + \lambda_1 \mathcal{L}_{L1}(G) + \lambda_2 \mathcal{L}_{var}(G)) \tag{4}$$

where λ_1 and λ_2 are the coefficient of L1-distance loss and variance loss respectively. Followed by the suggestion in Pixel2Pixel network [6], the λ_1 is set $1e^2$. In order to make the variance loss work, it is at least the same magnitude as L1 loss. Empirically, λ_2 is set $1e^5$ in our experiment.

3 Network Architecture

In GAN, traditional generator is derived from Encoder-Decoder structure, which is proved with outstanding performance in image-image translation. For our task, pixel-wised accuracy is strictly demanded since little mismatch may introduce severe artifacts in reconstructed CT image. To obtain the pixel-pixel precision, we use fully convolutional network [10] as the basis of generator. Dilation convolution [11] is introduced to enlarge the perception field and capture multi-scale features. For sake of promoting the integration of different feature levels, skip connection [12] is used to boost the feature expression. In addition, rather than transforming noisy to clean image, residual learning is applied to learn a mapping from noisy image to noise map which can significantly accelerate training. The generator structure is depicted as follow:
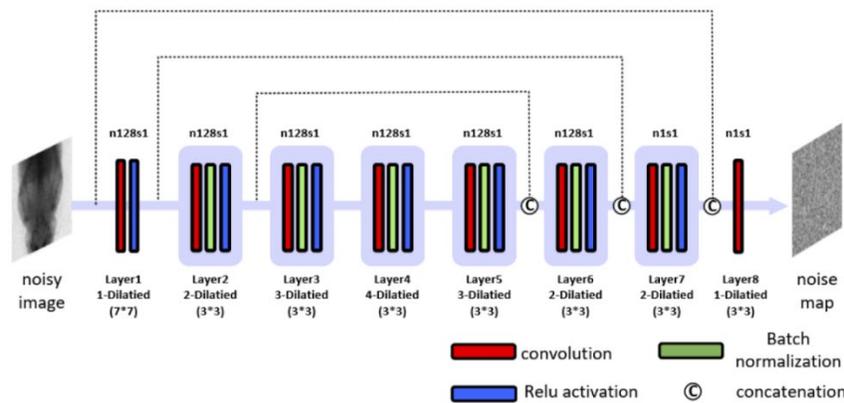
Figure 2: Proposed generator of the C-GAN

The purpose of discriminator is to distinguish the fake from the real. PatchGAN [8] is adopted as the discriminator network structure. Instead of classifying the whole image as real or fake, it penalizes structure at the scale of patches and all the responses are averaged to generate an ultimate result of discrimination. Since it focuses attention to local high-frequency patches, it is reasonable in pixel-wised application. The architecture of PatchGAN is drawn in Fig. 3.
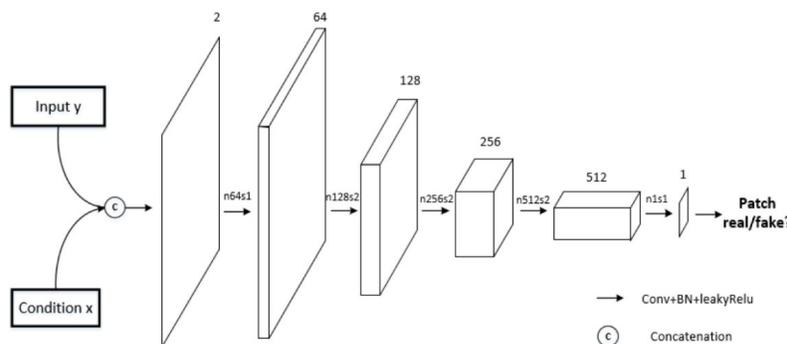
Figure 3: Structure of Patch-GAN discriminator

4 Results

To evaluate the effectiveness of our denoising method, experiments are carried out in both simulation and real data. All data used is acquired by Hiscan VM Micro CT system.

In simulation, mouse and milkvetch root are scanned for the high-quality projection images. The scanning parameters in our experiment are set as follows: tube voltage: 60 kV, tube current: 120 uA, exposure time: 100 ms. After acquiring the high-quality projection data, the Poisson noise $N \sim Poisson\{\alpha y_i\}$ is added to simulate photon starving situation, where y_i represents the projection value from i_{th} detector channel and α is a scaling factor. Then the projection images are cropped into overlapped 96 * 96 patches and a dataset containing over 500,000 image pairs is made for network training. After completing these processes, FDK algorithm is used to reconstruct CT images. We have compared our results with different approaches: BM3D, CNN-MSE [13] and CNN-VGG [14]. Figs. 4- 5 display the simulation results of CT images.
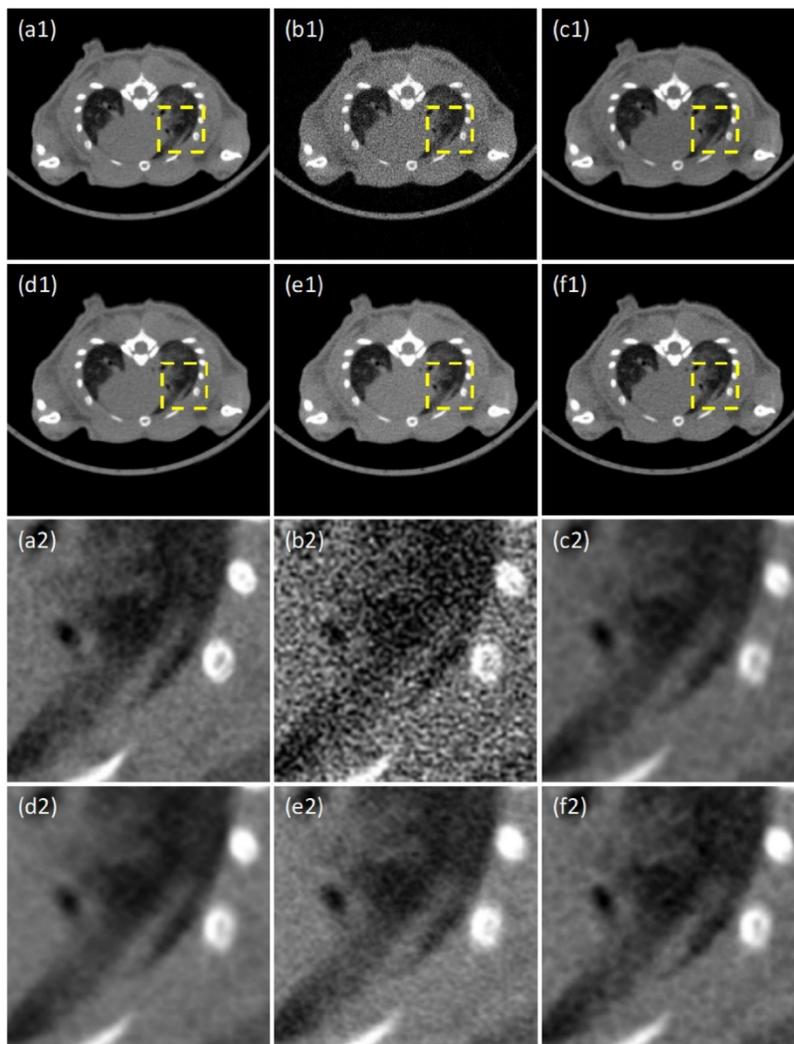
Figure 4: Mouse lung CT image: (a1) high-quality CT image; (b1) noisy CT image; (c1) BM3D; (d1) CNN-MSE; (e1) CNN-VGG; (f1) C-GAN. (a2)-(f2): Zoomed image inside the yellow box corresponding to (a1)-(f1)
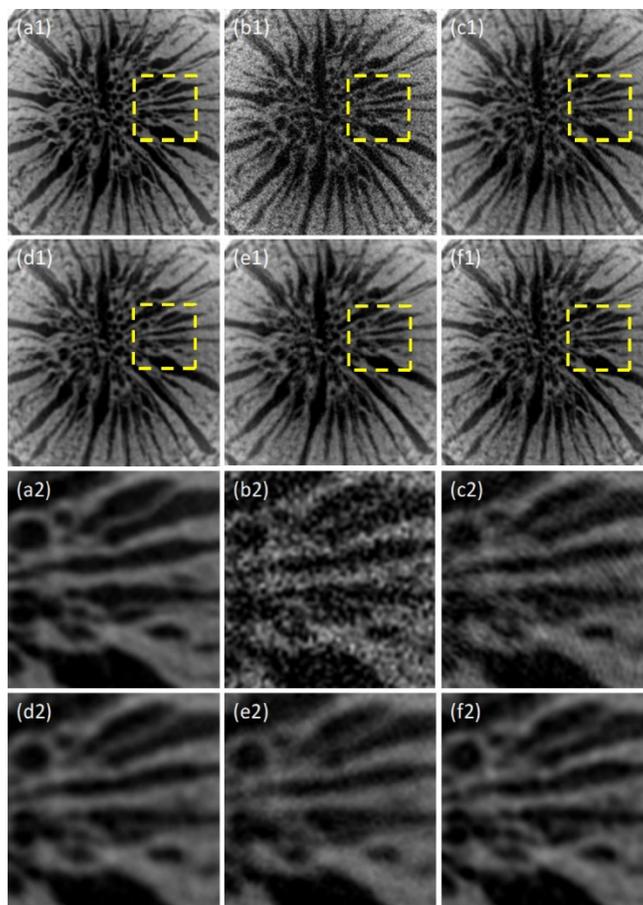
Figure 5: Milkvetch Root images: (a1) high-quality CT image; (b1) noisy CT image; (c1) BM3D; (d1) CNN-MSE; (e1) CNN-VGG; (f1) C-GAN. (a2)-(f2): Zoomed image inside the yellow box corresponding to (a1)-(f1)

Fig. 4 presents the CT images of mouse lung and Fig. 5 presents those of milkvetch root. In each figure, (a) is the high-quality CT image and (b) is the corresponding noisy CT image. (c)-(f) show the results by different denoising methods. The images in yellow box are zoomed view for more structure details, as shown in (a2)-(f2). Note that BM3D indeed reduces the noise level, though it introduces the streak artifacts around the bone areas. CNN-MSE tends to be over-smoothed, accompanied with blurry effects at the boundaries of the tissue. On the other hand, the perceptual loss is used in CNN-VGG to capture image features, which shows good performance in image-domain denoising. However, when applied to the projection domain, it cannot restore enough texture. In contrast, our proposed method generates clearer edges and reconstructs low-contrast structures, which is more likely to high-quality image. Due to the introducing of variance loss, less blurring occurs and more image details are recovered simultaneously.

Table 1: PSNR and SSIM of the denoised images

Object	indicator	Noisy	BM3D	CNN-MSE	CNN-VGG	C-GAN
Mouse lung	PSNR	27.4164	30.2533	37.6158	36.7754	37.8667
	SSIM	0.4918	0.9100	0.9301	0.8891	0.9153
Milkvetch Root	PSNR	17.6985	26.1765	30.1228	26.9995	28.2400
	SSIM	0.4614	0.7695	0.8967	0.8107	0.8635
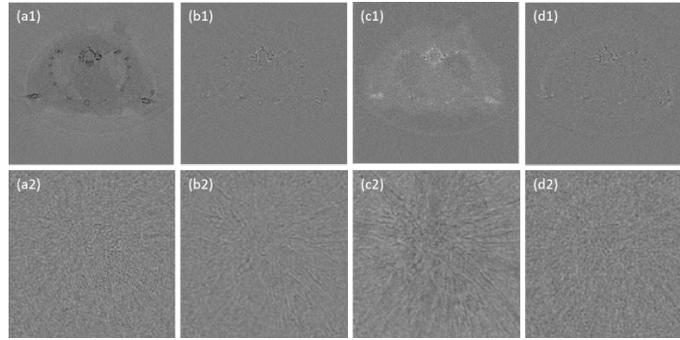
Figure 6: Residuals of the denoised CT images: (a) BM3D; (b) CNN-MSE; (c) CNN-VGG; (d) C-GAN; (a1)-(f1): Residuals of Mouse lung; (a2)-(f2): Residuals of Milkvetch Root

In views of the traditional PSNR and SSIM in Tab. 1, our method overrides other approaches except CNN-MSE. It is not surprising since the MSE loss function has the same pixel-wised formulation with PSNR. To evaluate the denoising performance intuitively, the differences between denoised images and the high-quality images are shown in Fig. 6. The upper row is the residual images of mouse lung and the second row is that of milkvetch root. It is apparent that the residuals of BM3D and CNN-VGG contain much image information, even some artifacts, which manifests these two methods are not suitable for projection domain CT image denoising. In terms of CNN-MSE and C-GAN, both methods show greater performance. While the structure information of C-GAN in residual image is less than CNN-MSE. The results prove that our proposed method can preserve more image details and display a more appealing performance.
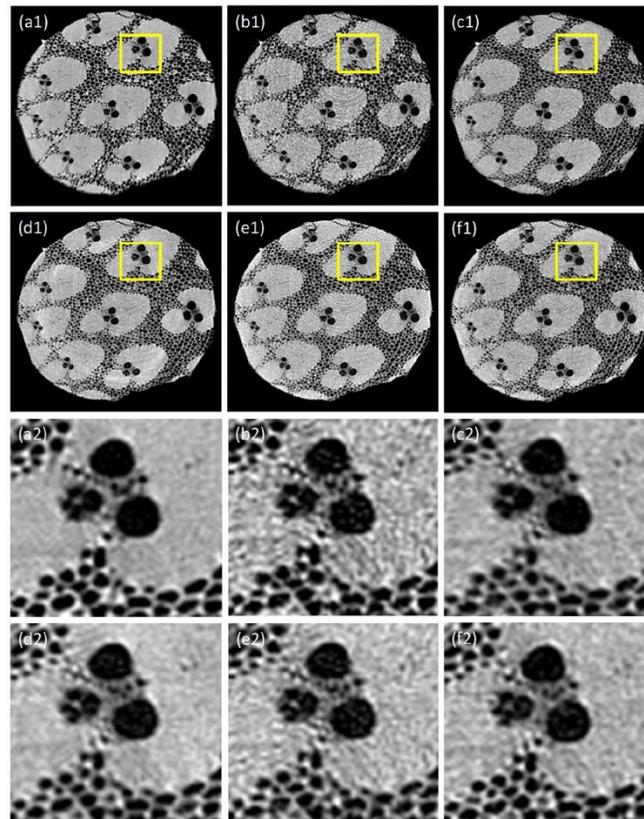
Figure 7: Bamboo stick CT images: (a) High-quality CT image; (b) Noisy CT image; (c) BM3D; (d) CNN-MSE; (e) CNN-VGG; (f) C-GAN. (a2)-(f2): Zoomed image inside the yellow box corresponding to (a1)-(f1)

For further verification in actual practice, a piece of bamboo stick was taken as the sample and the noisy projections were obtained by reducing the exposure time. After processing by different pre-trained models, the denoised CT images are finally displayed in Fig. 7.

Table 2: PSNR and SSIM of the Bamboo CT images.

Object	Indicator	Noisy	BM3D	CNN-MSE	CNN-VGG	C-GAN
Bamboo	PSNR	20.8385	23.1004	24.4382	22.8693	23.7669
Stick	SSIM	0.6601	0.7235	0.8010	0.6693	0.7835

Figs. 7(a1)-7(f1) show the result of the denoised bamboo stick CT images and Figs. 7(a2)-7(f2) are the corresponding enlarged images of yellow box region. From the reconstructed images, it is apparent that the results of the real data are similar with those of simulation. PSNR and SSIM shown in Tab. 2 illustrate that though the real noise is somewhat different from simulated noise, the effect of variance loss still works. Compared with CNN-MSE and CNN-VGG, our method can reduce the noise level, as well as preserve sharper structure boundary. The results demonstrate that our approach has the same effect with the more complicated imaging of real Micro-CT.

5 Conclusion

In this paper, we propose a projection denoising method for micro-CT based on deep neural network. Conditional GAN is used as our framework. Residual learning and skip connection are applied in generator to accelerate training. To better exploit noise statistics, a novel variance loss is developed and fused into loss function which indeed upgrades the denoising performance. Through the comparison with other outstanding approaches, our method can effectively suppress noise level and improve CT image quality at the same time, especially outperforms in preserving detail structures. Although we have achieved some preliminary results, there is still room for improvement. In future work, we will further refine the proposed network for better adapting to practical application.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing* **2007**, 16(8): 2080-2095.

2. Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing* **2006**, 15(12): 3736-3745.

3. Chen H, Zhang Y, Kalra MK, Lin F, Chen Y et al. Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Transactions on Medical Imaging* **2017**, 36(12): 2524-2535.

4. Chen H, Zhang Y, Zhang W, Liao P, Li K et al. Low-dose CT via convolutional neural network. *Biomedical Optics Express* **2017**, 8(2): 679-694.

5. Yang Q, Yan P, Zhang Y, Yu H, Shi Y et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Transactions on Medical Imaging* **2018**, 37(6): 1348-1357.

6. Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **2017**: 1125-1134.

7. Zhang R, Yang J, Chen C. Tumor cell identification in Ki-67 images on deep learning. *Molecular & Cellular Biomechanics* **2018**, 15(3): 177-187.

8.   Chen C, Yang J. Image recognition of breast tumor proliferation level based on convolution neural network. *Molecular & Cellular Biomechanics* **2018**, 15(4).

9.   He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **2016**: 770-778.

10. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **2015**: 3431-3440.

11. Yu F, Koltun V, Funkhouser T. Dilated residual networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **2017**: 472-480.

12. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **2017**: 4700-4708.

13. Zhang K, Zuo W, Gu S, Zhang L. Learning deep CNN denoiser prior for image restoration. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **2017**: 3929-3938.

14. Yang Q, Yan P, Kalra MK, Wang G. CT image denoising with perceptive deep neural networks. arXiv:2017:1702.07019.