Low-Dose CT Restoration with Deep Neural Network

Hu Chen, Yi Zhang*, Weihua Zhang, Peixi Liao, Ke Li, Jiliu Zhou and Ge Wang

***Abstract*—To reduce the potential radiation risk, low-dose CT has attracted much attention. However, simply lowering the radiation dose will lead to significant deterioration of the image quality. In this paper, we propose a noise reduction method for low-dose CT via deep neural network without accessing original projection data. A deep convolutional neural network is trained to transform low-dose CT images towards normal-dose CT images, patch by patch. Visual and quantitative evaluation demonstrates a competing performance of the proposed method.**

Keywords—low-dose CT; noise reduction; deep learning; convolutional neural network

I. INTRODUCTION

In the past decades, X-ray computed tomography has been widely used in diagnostic and interventional tasks. With the increasing number of CT scans, the potential radiation risk attracts an increasingly public concern [1]. Most commercial CT scanners utilize the filtered backprojection (FBP) method to analytically reconstruct images. One of the most used methods to reduce the radiation dose is to lower the operating current of the X-ray tube. However, directly lowering the current will significantly degrade the image quality due to the excessive quantum noise caused by an insufficient number of photons in the projection domain.

Many approaches were proposed to improve the quality of low-dose CT images. These approaches can be categorized into three classes: sinogram filtering, iterative reconstruction and image processing.

Sinogram filtering directly smoothens raw data before FBP is applied. Iterative reconstruction solves the problem iteratively, aided by some kinds of prior information on target images. Different priors were proposed, such as in terms of total variation (TV), nonlocal means (NLM) and dictionary learning [2-8]. Despite the successes achieved by these two approaches, they are often restricted in practice due to the difficulty of gaining well-formatted projection data since the vendors are not generally open in this aspect. Meanwhile, the iterative reconstruction methods suffer from heavy computational costs. In contrast to the first two categories, image processing does not rely on projection data, can be directly applied to low-dose CT images, and easily integrated into the current CT workflow. However, it is underlined that the noise in low-dose CT images does not obey a uniform distribution. As a result, it is not easy to remove image noise and artifacts effectively with traditional image denoising methods. Extensive efforts were made to suppress image noise via image processing for low-dose CT. Based on the popular idea of sparse representation, Chen et al adapted K-SVD to deal with low-dose CT images [9]. Also, a block-matching 3D (BM3D) algorithm has been proved powerful in image restoration for different noise types [10].

Recently, deep learning (DL) has generated an excitement in the field of machine learning and computer vision [11]. DL can efficiently learn high-level features from the pixel level through a hierarchical framework. In the field of medical image processing, there are already multiple papers on DL-based image analysis, such as image segmentation, nuclei detection, and organ classification. To our best knowledge, there are few studies proposed for imaging problems. In this regard, Wang et al introduced the DL-based data fidelity into the framework of iterative reconstruction for undersampled MRI reconstruction [12]. Zhang et al. proposed a limited-angle tomography method with deep convolutional neural network (CNN) [13]. Wang shared his opinions on deep learning for image reconstruction [14].

Inspired by the great potential of deep learning in image processing, here we propose a deep convolutional neural network to transform low-dose CT images towards corresponding normal-dose CT images. An offline training stage is needed using a reasonably sized training set. In the second section, the network details are described. In the third section, visual effects and quantitative results are presented. Finally, the conclusion is drawn.

II. MEHTHODS

A. Noise Reduction Model

Due to the encryption of the raw projection data, post-reconstruction restoration is a reasonable alternative for sinogram-based methods. Once the target image is reconstructed from a low-dose scan, the problem becomes image restoration or image denoising. A difference between low-dose CT image denoising and natural image restoration is that the statistical property of low-dose CT images cannot be easily determined in the image domain. This will significantly compromise the performance of noise-dependent methods, such as median filtering, Gaussian filtering, anisotropic diffusion, etc., which were respectively designed for specific noise types. However, learning-based methods are immune to this problem, because this kind of methods is strongly dependent on training samples, instead of noise type. We model the noise reduction problem for low-dose CT as follows.

H. Chen, Y. Zhang, W. Zhang, K. Li and J. Zhou are with the School of Computer Science, Sichuan University, Chengdu 610065, China (e-mail: yzhang@scu.edu.cn).

P. Liao is with the Department of Scientific Research and Education, The Sixth People's Hospital of Chengdu, Chengdu 610065, China

G. Wang are with the Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA.

Asterisk indicates corresponding author.

DOI: 10.12059/Fully3D.2017-11-3202013

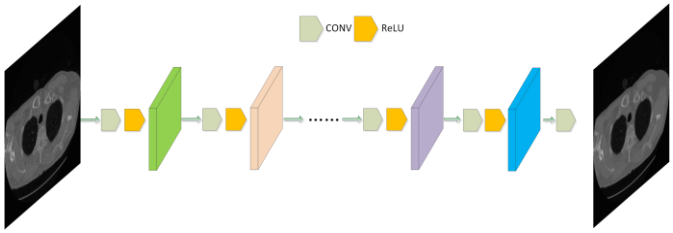

Fig.1. The architecture of the proposed network.

Let $\mathbf{X} \in \mathbb{R}^{m \times n}$ is a low-dose CT image and $\mathbf{Y} \in \mathbb{R}^{m \times n}$ is the corresponding normal-dose image, then the relationship can be formulated as:

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

where $\sigma : \mathbb{R}^{m \times n} \to \mathbb{R}^{m \times n}$ represents the corrupting process due to the quantum noise that contaminates the normal-dose CT image. Thus, the denoising problem can be converted to find a function f :

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

where f is treated as the best approximation of σ^{-1}.

B. Convolutional Neural Network

In this study, the low-dose CT denoising problem was solved in the three steps: patch coding, non-linear filtering, and reconstruction. The proposed architecture of network is illustrated in Fig. 1. Next, we introduce each step in details.

1) Patch encoding

Sparse representation (SR) is popular for image processing. The key idea of SR is to represent extracted patches of an image with a pre-trained dictionary. Such dictionaries can be categorized into two groups according to how dictionary atoms are constructed. The first group includes analytic dictionaries such as DCT, Wavelet, FFT, etc. The other one is learned dictionaries, which can preserve more application-specific details assuming proper training samples. This SR process can be often implemented as convolution operations with a series of filters, each of which is an atom. Our method is similar in the sense that SR is involved for patch encoding but with a neural network. First, we extract patches from training images with a fixed slide size. Second, the first layer to implement patch coding can be formulated as

$$C_1(\mathbf{y}) = \mathrm{ReLU}(\mathbf{W}_1 * \mathbf{y} + \mathbf{b}_1), \tag{3}$$

where $\mathbf{W}_1$ and $\mathbf{b}_1$ denote weights and biases respectively, $*$ represents the convolution operator, $\mathbf{y}$ is extracted patch from images, and $\mathrm{ReLU}(x) = \max(0, x)$ is the nonlinear activation function. In CNN, $\mathbf{W}_1$ can be seen as n_1 convolution kernels with a fixed size of $s_1 \times s_1$. After patch encoding, we embed the image patches into a feature space, and the output $C_1(\mathbf{y})$ is a feature vector, whose size is n_1.

2) Non-linear filtering

After processed by the first layer, a n_1 -dimensional feature vector is obtained from the extracted patch. In the second layer, we transform the n_1 -dimensional vector into n_2 -dimensional one. This operation is equivalent to a filtration on the feature map from the first layer. The second can be formulated as

$$C_m(\mathbf{y}) = \mathrm{ReLU}(\mathbf{W}_m * C_{m-1}(\mathbf{y}) + \mathbf{b}_m), m = 2, ..., 9 \tag{4}$$

where $\mathbf{W}_m$ is composed of n_m convolution kernels with a fixed size of $s_m \times s_m$. If the desired network only has two layers, the output of this layer is the corresponding cleaned patches for the final reconstruction. Generally, inserting more layers is a way to potentially boost the capacity of the network. However, a deeper CNN is at a higher cost of computation including longer training time. In this work, 8 layers constructed this part.

3) Reconstruction

In this step, the processed overlapping patches are merged into the final image. These overlapping patches are properly weighted before their summation. This operation can be considered as filtration by a pre-defined convolutional kernel formulated by

$$C(\mathbf{y}) = \mathbf{W}_{10} * C_9(\mathbf{y}) + \mathbf{b}_{10}, \tag{5}$$

where $\mathbf{W}_{10}$ is composed of only 1 convolution kernel with a size of $s_{10} \times s_{10}$, and $\mathbf{b}_3$ has the same size as that of $\mathbf{W}_{10} * C_9(\mathbf{y})$.

Eqs. (3)-(5) all use convolutional operations, although they have been designed for different purposes. That is the reason why the CNN architecture is in use for our low-dose CT image denoising.

4) Training

Once the network is configured, the parameter set, $\Theta = \{\mathbf{W}_1, ..., \mathbf{W}_{10}, \mathbf{b}_1, ..., \mathbf{b}_{10}\}$, of the network must be estimated to learn the function C . Given the training dataset $D = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \ldots, (\mathbf{x}_N, \mathbf{y}_N)\}$ with $\{\mathbf{x}_i\}$ and $\{\mathbf{y}_i\}$ denoting normal-dose image patches and its corresponding noisy versions respectively, and N being the total number of training samples, the estimation of the parameters can be achieved by minimizing the following loss function:

$$L(D; \Theta) = \frac{1}{N} \sum_{i=1}^{N} \| \mathbf{x}_i - C(\mathbf{y}_i) \|^2. \tag{6}$$

The loss function is optimized using the stochastic gradient descent method.

III. Experiments

To evaluate the performance of the proposed approach, the state-of-the-art methods, both iterative reconstruction and post-reconstruction processing, were selected for comparison, including ASD-POCS, KSVD and BM3D. The parameters for the completing methods were set according to the recommendation in the original references. The peak signal to noise ratio (PSNR), root mean square error (RMSE) and structure similarity index (SSIM) were used as the quantitative metrics. All the experiments were conducted using MATLAB 2015b on a PC (Intel i7 6700K CPU, 16 GB RAM and GTX 980 Ti graphics card).

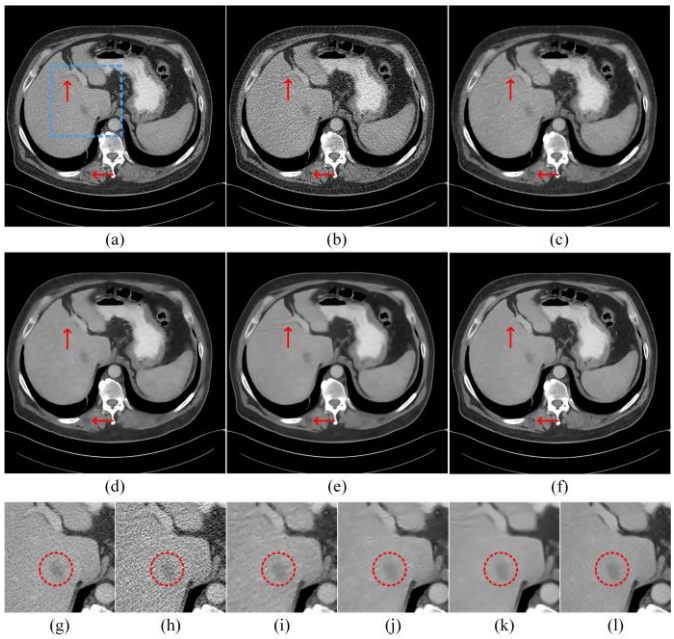

Fig. 2. Results of a slice of abdomen images. (a) Original normal-dose image; (b) the low-dose images; (c) the ASD-POCS reconstructed image; (d) the KSVD processed low-dose image; (e) the BM3D processed low-dose image; (f) the CNN processed low-dose image; and (g)-(l) the zoomed regions specified with the red boxes in (a)-(f).

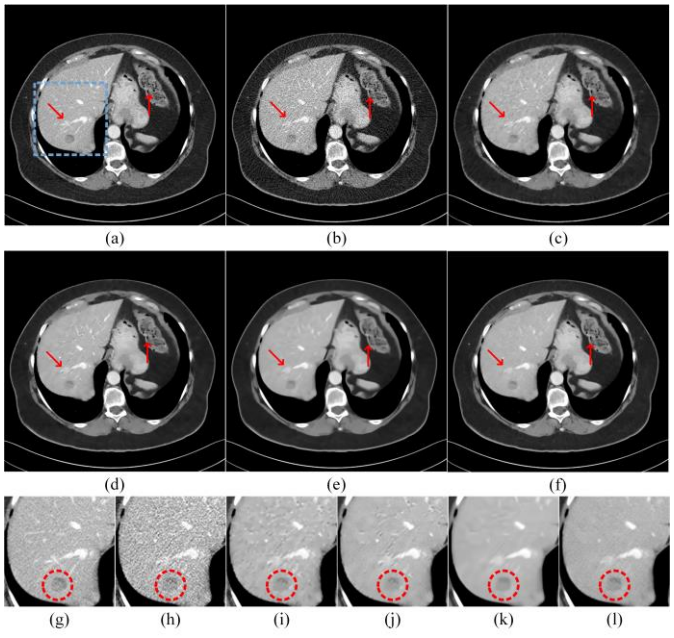

Fig. 3. Results of another slice of abdomen images. (a) Original normal-dose image; (b) the low-dose images; (c) the ASD-POCS reconstructed image; (d) the KSVD processed low-dose image; (e) the BM3D processed low-dose image; and (f) the CNN processed low-dose image; and (g)-(l) the zoomed regions specified with the red boxes in (a)-(f).

A. Dataset Preparation

To evaluate the clinical performance of RED-CNN, a real clinical dataset was used, which was authorized by Mayo Clinics for "The 2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge". The dataset contained 2378 3mm thickness full and quarter dose 512×512 CT images from 10 patients. The network was trained by part of 3mm thickness full dose and quarter dose image pairs. The rest of 3mm thickness quarter dose and full dose images were respectively regarded as the testing set and the standard of reference. For fairness, cross-validation was utilized in the testing phase. While testing on CT images from each patient, the images from the other 9 patients were involved in the training phase.

The input patches of the network were extracted from the original images with size $m = 100$. The slide step was 4. There are two reasons why patches were used, instead of whole images: one is that the images can be well represented by local structures; and the other is that deep learning requires a big training dataset and chopping the original images into patches can efficiently boost the number of samples.

B. Parameter Setting

In this paper, three layers were used in the proposed network. The filter number, we set $n_1 = n_2 = \ldots = n_9 = 96$, and the corresponding filter sizes, s_1, s_2,…, and s_{10} were set to 5. The initial weights of the filters in each layer were randomly set, which satisfies the Gaussian distribution with zero mean and standard deviation 0.001. The initial learning rate was 0.0001 and slowly decayed to 0.00001 during the training process.

C. Results

1) Visual inspection

For this purpose, we selected 2 representative slices from the abdomen. Figs. 2 and 3 gives the results from different methods. In both the figures, the noise and artifacts caused by the lack of incident photons severely degraded the image quality. Some details and structures cannot be discriminated. All of the methods can eliminate the noise and artifacts to different degrees. However, due to the piecewise constant assumption by the TV minimization, ASD-POCS caused blocky effects in the resultant images. KSVD and BM3D produced new artifacts as indicated by the left red arrow in Fig. 2. In Fig. 2(l), indicated by the red arrow in the bottom of the image, CNN preserved the details better than all other methods. Meanwhile, the zoomed parts (Fig. 2(g)-(l)) can show that CNN has better visual effect for the low contrast lesion marked by red dotted circle. In Fig. 3, a similar trend was observed. The red arrows point to small structures and vessels where only the proposed CNN method could recover the most details and in the zoomed parts (Fig. 3(g)-(l)), our

Table I. Quantitative measurements associated with different algorithms.

		TV	KSVD	BM3D	CNN
Fig. 2	PSNR	44.623	46.4276	45.6449	**46.4848**
	RMSE	0.0059	0.0048	0.0052	**0.0047**
	SSIM	0.9793	0.9819	0.9805	**0.9849**
Fig. 3	PSNR	40.3023	40.6949	40.1139	**41.4684**
	RMSE	0.0097	0.0092	0.0103	**0.0084**
	SSIM	0.9303	0.9330	0.9197	**0.9465**
All	PSNR	42.3257	42.5578	42.5499	**43.0178**
	RMSE	0.0077	0.0072	0.0073	**0.0068**
	SSIM	0.9514	0.9588	0.9590	**0.9677**

method achieved best ability of lesion detection. However,

2) Quantitative measurement

To quantitatively evaluate the proposed CNN algorithm, PSNR, RMSE and SSIM were measured for all the images in the testing set. The results for the restored images in Figs. 2 and 3 are in Table I. It can be seen in Table I that for both Fig. 2 and 3, our proposed method obtained the best results, which are consistent to the visual inspection. The term "All" in Table I means the average values of the measurements for all the images in the testing set. It can be observed that all the metrics demonstrate the proposed CNN based method had the best performance.

IV. Conclusion

In this pilot study, we have evaluated the performance of a deep convolutional neural network for noise reduction in low-dose CT. The results demonstrate the potential of CNN based method for medical imaging. In the future, the proposed network structure will be optimized and applied to other CT topics, such as few-view CT reconstruction, metal artifact reduction, and interior CT. Another possible direction is to investigate other network architectures to deal with dynamic and spectral CT problems.

Acknowledgment

The authors would like to thanks Dr. Cynthia McCollough, the Mayo Clinic, and the American Association of Physicists in Medicine (AAPM) for providing the Low-Dose CT Grand Challenge data set.

This work is supported by National Natural Science Foundation of China (NSFC) (61202160, 61302028, 61671312); National Institute of Biomedical Imaging and Bioengineering (NIBIB)/National Institutes of Health (NIH) (R01 EB016977, U01 EB017140).

References

[1] D. J. Brenner and E. J. Hall, "Computed tomography–An increasing source of radiation exposure," New Eng. J. Med., vol. 357, pp. 2277–2284, 2007.

[2] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," Phys. Med. Biol., vol. 53, no. 17, pp. 4777-4807, 2008.

[3] Y. Zhang, Y, Wang, W. Zhang, et al., "Statistical iterative reconstruction using adaptive fractional order regularization," Biomed. Opt. Express, vol. 7, no. 3, pp. 1015-1029, 2016.

[4] Y. Zhang, W. Zhang, Y. Lei, and J. Zhou, "Few-view image reconstruction with fractional-order total variation," J. Opt. Soc. Am. A, vol. 31, no. 5, pp. 981-995, 2014.

[5] Y. Zhang, W.-H. Zhang, H. Chen, M.-L. Yang, T.-Y. Li, and J.-L. Zhou, "Few-view image reconstruction combining total variation and a high-order norm," Int. J. Imaging Syst. Technol., vol. 23, no. 3, pp. 249-255, 2013.

[6] J. Ma, H. Zhang, Y. Gao et al., "Iterative image reconstruction for cerebral perfusion CT using a pre-contrast scan induced edge-preserving prior," Phys. Med. Biol., vol. 57, no. 22, pp. 7519-7542, 2012.

[7] 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.

[8] Q. Xu, H. Yu, X. Mou, et al., "Low-dose x-ray CT reconstruction via dictionary learning," IEEE Trans. Med. Imaging, vol. 31, no. 9, pp. 1682-1697, 2012.

[9] Y. Chen, X. Yin, L. Shi, et al., "Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing," Phys. Med. Biol., vol. 58, no. 16, pp. 5803-5820, 2013.

[10] K. Sheng, S. Gou, J. Wu, et al., "Denoised and texture enhanced MVCT to improve soft tissue conspicuity," Med. Phys., vol. 41, no. 10, 101916, 2014.

[11] Y. LeCun, Y, Bengio and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015.

[12] S. Wang, Z. Su, L. Ying, et al., "Accelerating magnetic resonance imaging via deep learning," 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 514-517, 2016.

[13] H, Zhang, L. Li, K. Qiao, et al., "Image predication for limited-angle tomography via deep learning with convolutional neural network," arXiv:1607.08707, 2016.

[14] G. Wang, "A perspective on deep imaging," arXiv:1609.04375, 2016.