

Self-trained Deep Convolutional Neural Network for Noise Reduction in CT

Zhongxing Zhou, Akitoshi Inoue, Cynthia H. McCollough, Lifeng Yu

Department of Radiology, Mayo Clinic, Rochester, MN, 55905, USA

Abstract: Supervised deep convolutional neural network (CNN)-based methods have been actively used in clinical CT to reduce image noise. The networks of these methods are typically trained using paired high- and low-quality data from a large number of patients and/or phantom images. This training process is tedious, and the network trained under a given condition may not be generalizable to patient images acquired and reconstructed at different conditions. In this paper, we propose a self-trained deep CNN (ST_CNN) method which does not rely on pre-existing training datasets. The training is accomplished using extensive data augmentation through projection domain and the inference is applied to the data itself. Preliminary evaluation on patient images demonstrated that the proposed method could achieve similar image quality in comparison with conventional deep CNN denoising methods pre-trained on external datasets.

Keywords: Computed tomography (CT), low-dose CT, deep learning, convolutional neural network (CNN), supervised training

1. Introduction

In recent years, Deep convolutional neural network (CNN) has been one of the main driving forces for CT image denoising [1-7]. A majority of existing CNN-based denoising methods are supervised, which learn the mapping function between the low-quality image (e.g., low dose) and its high-quality (e.g., high dose) counterpart [1-5]. In order to have a CNN denoiser that generalizes well to new patient data, a large amount of low-/high-quality image pairs from a large number of patient and/or phantom images are needed to sufficiently cover the data distribution. However, this training process is costly, and the model trained from one dataset may not generalize well to another dataset acquired or reconstructed at a different condition. The inter-patient differences can also make it challenging to learn a model that generalizes well across patients.

To tackle this challenge, we propose a self-trained deep CNN (ST_CNN) method for noise reduction in CT which does not rely on pre-existing training datasets. This method trains the network directly using the data itself through extensive data augmentation (random rotation and noise addition) through projection domain and the inference is applied to the data itself. We demonstrated that this method could achieve similar performance as conventional deep CNN denoising methods trained on external datasets. There are three major potential benefits of this method. First, by removing the need of a large number pre-existing training dataset, it can be applied to any CT data, even if the data condition were not previously trained. Second, the self-training mechanism eliminates the generalizability issue that may occur for network models applied to datasets that are different from the training datasets. Third, the trained model can be applied to and finetuned for each individual patient if repeated CT exams are expected, which may maximize the benefit of image quality improvement and radiation dose reduction.

2. Methods

The proposed ST_CNN method belongs to the family of image-domain supervised deep learning techniques, but there is a distinct difference in the training scheme from the existing approaches, as described in Figs. 1 and 2. The availability of sufficient amount of patient cases for training is one key factor contributing to the performance of the conventional supervised deep learning methods.

The proposed ST_CNN method is trained based on the data acquired from one single patient by generating a large amount of paired low-quality and high-quality images from the same patient (Figure 2).

[*Yu.Lifeng@mayo.edu](mailto:Yu.Lifeng@mayo.edu)

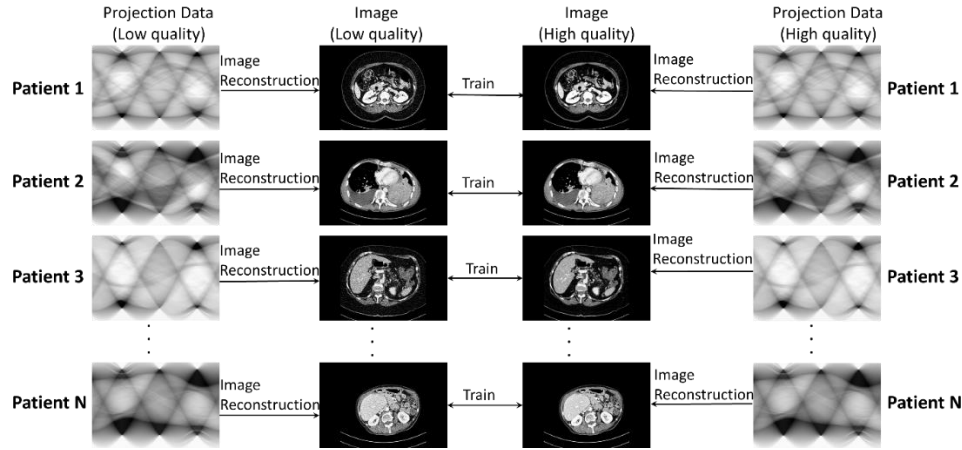


Fig. 1. Overview of the training scheme for conventional image-domain supervised deep learning techniques

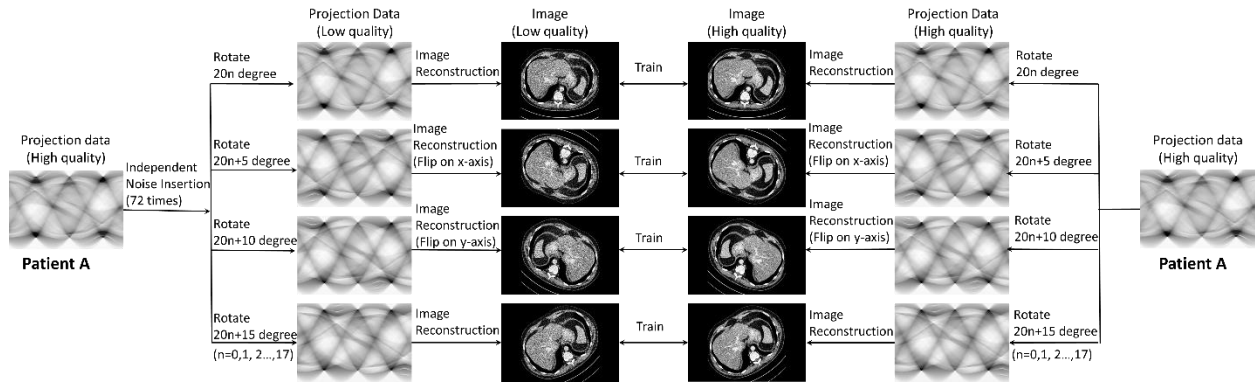


Fig. 2. Overview of the training scheme for the proposed ST_CNN method

The trained model is used to denoise the data acquired from the same patient. The training scheme is described as follows:

A. Low-quality image generation and augmentation:

1) Apply independent noise insertion multiple times (e.g., 72 times) on the original projection data of a specific patient to generate the corresponding low-quality projection data (e.g., 25% dose or 10% dose). The low dose levels can be randomized during the process of noise insertion.

2) Generate images with a large amount of different rotation angles. This cannot be accomplished by simply rotating the image itself as that will introduce errors through interpolation. Our approach is to apply the rotation angle directly on the projection data so that images are rotated at arbitrary angle without introducing additional errors. For example, we applied 72 different rotation angles (e.g., every 5 degrees on 360 degree) to the low-quality projection data. After reconstruction, 72 sets of images with different rotation angles are obtained.

3) Divide the images into 4 groups: group1 (rotate $20*n$ degree, $n=0, 1, \dots, 17$); group 2 (rotate $20*n+5$ degree); group 3 (rotate $20*n+10$ degree); group 4 (rotate $20*n+15$ degree). The images in group 2 are flipped on x-axis, and the images in group 3 are flipped on y-axis.

B. High-quality image generation and augmentation:

1) (Optional) Apply independent noise insertion multiple times (e.g., 72 times) on the original projection data. The amount of noise inserted is less than that to generate the low-quality images. One can also bypass this step so that the original dose level is directly used.

2) Apply rotation augmentation (e.g., every 5 degrees on 360 degree) on the original projection data of the same patient to generate multiple high-quality projections.

3) Follow step 3) in section A to generate 4 groups of high-quality images.

C. Low-/high-quality image pairs generation:

Generate matched high- and low-quality patches with multiple slices (e.g., $64 \times 64 \times 7$ voxels) from the reconstructed images in the first 3 groups for model training. The images in the 4th groups are used to generate the matched patches for model validation.

D. Model training:

The CNN denoising model can be based many of the popular network architecture. Here we employed a recently developed 2D residual-based CNN denoiser [3] for both ST_CNN and conventional deep CNN methods. The identical network architecture (Figure 3) was used for both methods so that any performance difference can be attributed to the different training methods. To optimize the performance of the CNN model, we used 7 adjacent CT slices as the channel input of the 2D residual CNN model [5]. The CNN inputs were first standardized (derived by subtracting the mean value and dividing by the standard deviation), and then subjected to initial layers that generated 128 feature maps using 2D convolutional layers. The feature maps were further processed by a series of 2D residual blocks, each of which consisted of repeated layers of 2D convolutional, batch normalization, and rectified linear unit activation. Then the output of residual blocks was projected back to a single-channel image by using a single convolutional layer with linear activation. This single-channel image was the estimated noise, which was further subtracted from the central input slice to get the final denoising result.

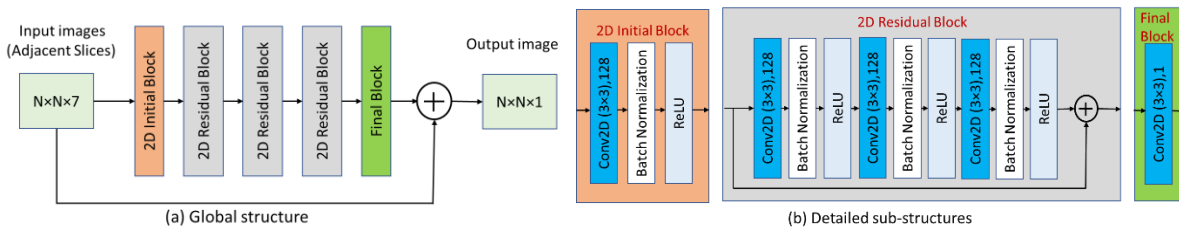


Figure 3. The architecture of residual-based 2D CNN denoiser. (a) Global structure of the network containing a 2D initial block, three 2D residual blocks, and a final block. (b) Details regarding the convolutional layers and transformations used within each block. Conv2D = two-dimensional convolutional layer, N = arbitrary image size, ReLU = rectified linear units

3. Results

Figure 4 compares full-dose (FD) images reconstructed and denoised using 4 different methods: (a) filtered-backprojection (FBP), (b) iterative reconstruction (IR), (c) conventional CNN, and (d) ST_CNN. The images were from a patient case in the Mayo/AAPM Low-dose CT Grand Challenge data library (Case number: L291). In Figure 4 and the following figures of this article, “CNN” refers to the conventional residual CNN method [3]. The FBP and IR reconstructions used matched kernels of B30 and I30 at a strength setting of 3. The conventional CNN was trained and validated using FBP FD and FBP QD image pairs from a subset of totally 30 patient data (17 patients for training and 5 patients for validation). The residual network architecture was identical to that used in the ST_CNN. The trained conventional CNN model was applied to denoise the FBP FD images of the rest of the patient data (e.g., L291). The ST_CNN was trained and validated using augmented FBP QD dose and FBP FD image pairs of a specific patient (e.g., L291), and then was applied to denoise the original FBP FD images of the same patient. The performance of two CNN models were assessed visually by an experienced radiologist. For the overall image quality evaluation, the radiologist ranked the ST_CNN method better than the conventional one because of more homogeneous liver parenchyma and better low-contrast lesion visibility (Arrows in the figure point to two subtle malignant liver tumors).

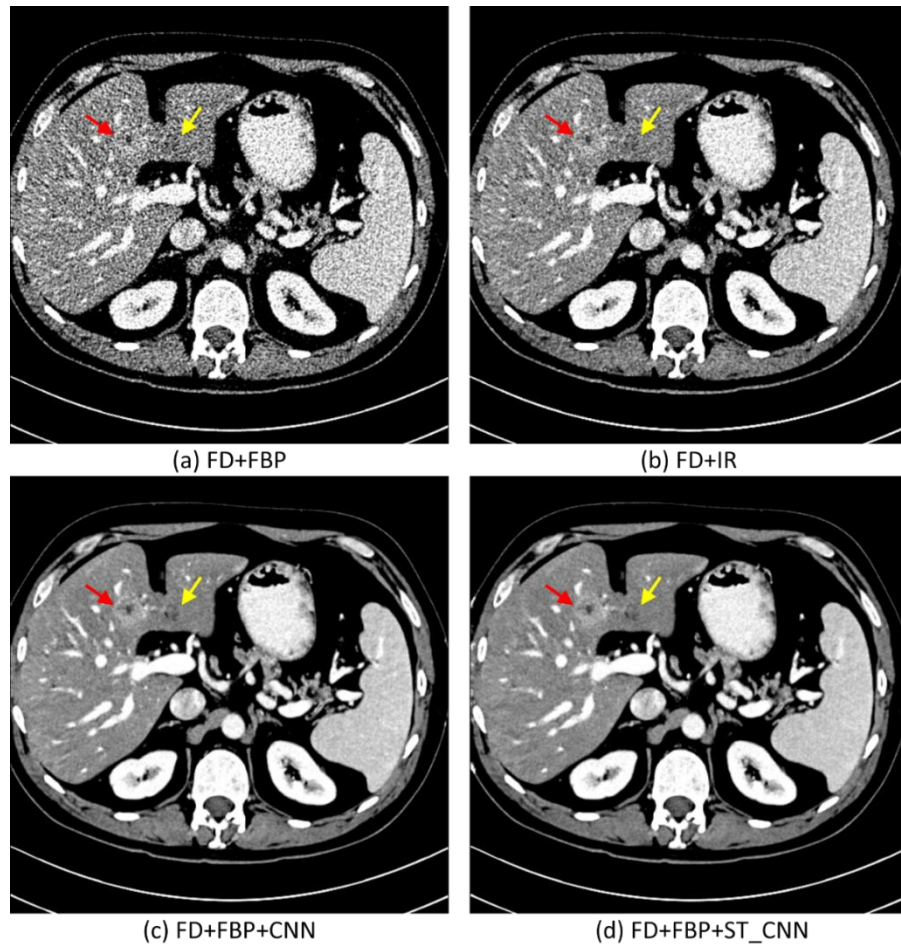


Figure 4. An example comparing the FBP and IR full dose (FD), conventional and self-trained CNN-denoised FBP FD images from a patient case in the Mayo/AAPM Low-dose CT Grand Challenge data library (Case number: L291). The arrows point to two subtle liver lesions. Slice thickness was 1 mm. To visualize the different appearance better, the display window was narrowed down to [60,200] HU.

To have a reference standard for quantitative evaluation of the performance, the self-trained CNN was trained and validated using augmented 10% dose (FBP) and QD (FBP) image pairs, and then applied to denoise the FBP QD images of the same patient. In this way, the original FD images can be used as the reference standard. The previously trained conventional CNN was used to denoise the same FBP QD images for CNN performance comparison. Figure 5 compares images reconstructed and denoised using 4 different conditions: (a) QD+FBP, (b) QD+IR, (c) QD+FBP+CNN, (d) QD+FBP+ST_CNN, and the two FD reconstructions were used as the reference standard: (e) FD+FBP and (f) FD+IR. Performance of the two CNN models were assessed visually by the same radiologist. In terms of low-contrast lesion visibility, conventional and self-trained CNN appeared to have a similar performance (Arrows in the figure point to two subtle malignant liver tumors). For the overall image quality evaluation, the radiologist ranked the self-trained CNN method better than the conventional one because of more homogeneous liver parenchyma and less false positive structures (A zoomed-in ROI in the liver parenchyma corresponding to the green box is shown in the bottom-right).

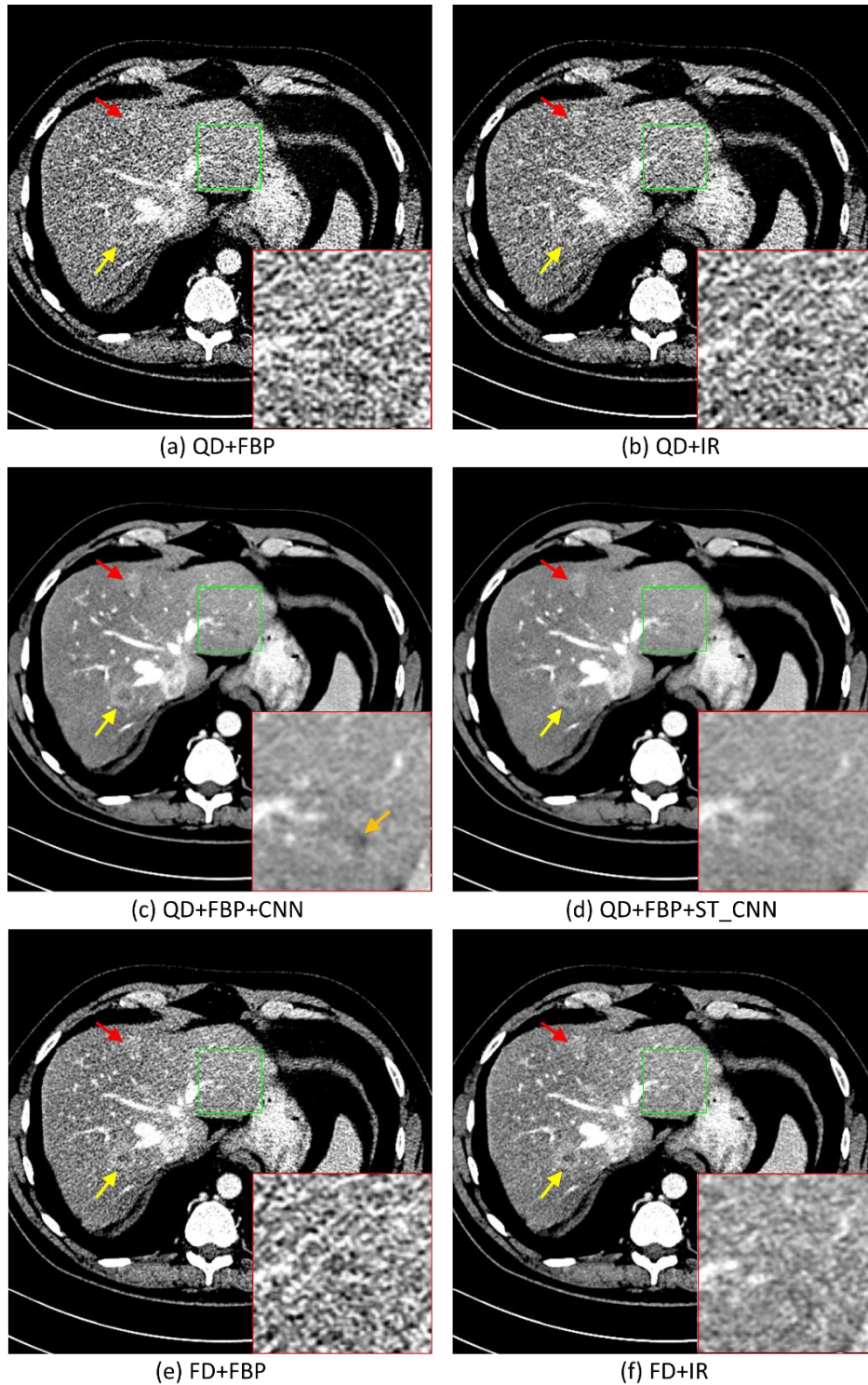


Figure 5. An example comparing the FBP and IR quarter dose (QD), FBP and IR full dose (FD), conventional and self-trained CNN-denoised QD images from a patient case in the Mayo/AAPM Low-dose CT Grand Challenge data library (Case number: L291). The red and yellow arrows point to two subtle liver lesions. A zoomed-in ROI within the green box is shown in the bottom-right corner. Note the orange arrow on the conventional CNN-denoised image points to a false positive lesion that does not exist on the self-trained CNN-denoised image. Slice thickness was 1 mm. To visualize the different appearance better, the display window was narrowed down to [60,200] HU.

Using FD+FBP as the reference, the root mean square error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM) were calculated for the conventional CNN and ST_CNN-denoised QD images (Table 1). The results provide clear evidence that the ST_CNN method has a performance similar to that of conventional deep CNN denoising methods without the need of a large number of training data.

TABLE I
Quantitative results (mean±SDs) associated with conventional and self-trained CNN methods for patient case L291

	PSNR	RMSE	SSIM
Conventional CNN	41.9±2.1	19.8±3.8	0.95±0.02
Self-trained CNN	41.8±2.1	20.1±3.7	0.95±0.02

4. Conclusions

We have designed a patient-specific self-trained CNN denoising method, aided by data augmentation through projection domain. Preliminary clinical evaluation demonstrated that the proposed method may achieve similar image quality in comparison with conventional deep CNN denoising methods pre-trained on a large number of patient cases. This new technique has the potential to overcome the generalizability issue of conventional training methods and to provide optimized noise reduction for each individual patient.

Acknowledgement

The authors acknowledge the computing facility (mForge) provided by Mayo Clinic for research computing. Dr. Zhou was supported by Mayo Radiology Research Fellowship.

References

1. H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang, "Low-dose CT via convolutional neural network," *Biomed. Opt. Express*, vol. 8, no. 2, pp. 679–694, 2017.
2. H. Chen, Y. Zhang, MK Kalra, F. Ling, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, 2017.
3. Nathan R. Huber, Andrew D. Missert, Lifeng Yu, Shuai Leng, Cynthia H. McCollough, "Evaluating a Convolutional Neural Network Noise Reduction Method When Applied to CT Images Reconstructed Differently Than Training Data," *Journal of Computer Assisted Tomography*, vol. 45, no. 4, pp. 544-551, 2021.
4. W. Yang, H. Zhang, J. Yang, J. Wu, X. Yin, Y. Chen, H. Shu, L. Luo, G. Coatrieux, and Z. Gui, "Improving low-dose CT image using residual convolutional network," *IEEE Access*, vol. 5, pp. 24698–24705, 2017.
5. Zhongxing Zhou, Nathan R. Huber, Akitoshi Inoue, Cynthia H. McCollough, Lifeng Yu, "Residual-based convolutional-neural-network (CNN) for low-dose CT denoising: impact of multi-slice input," in *SPIE Medical Imaging*. 2022. SPIE.
6. J. M. Wolterink, T. Leiner, M. A. Viergever and I. Isgum, "Generative adversarial networks for noise reduction in low-dose CT", *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2536-2545, 2017.
7. Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, and G. Wang, "Low dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1348 – 1357, 2018