



Sinogram denoising via attention residual dense convolutional neural network for low-dose computed tomography

Yin-Jin Ma^{1,2} · Yong Ren³ · Peng Feng¹ · Peng He¹ · Xiao-Dong Guo^{1,4} · Biao Wei¹

Received: 4 December 2020 / Revised: 21 January 2021 / Accepted: 22 January 2021 / Published online: 19 April 2021
© China Science Publishing & Media Ltd. (Science Press), Shanghai Institute of Applied Physics, the Chinese Academy of Sciences, Chinese Nuclear Society 2021

Abstract The widespread use of computed tomography (CT) in clinical practice has made the public focus on the cumulative radiation dose delivered to patients. Low-dose CT (LDCT) reduces the X-ray radiation dose, yet compromises quality and decreases diagnostic performance. Researchers have made great efforts to develop various algorithms for LDCT and introduced deep-learning techniques, which have achieved impressive results. However, most of these methods are directly performed on reconstructed LDCT images, in which some subtle structures and details are readily lost during the reconstruction procedure, and convolutional neural network (CNN)-based methods

for raw LDCT projection data are rarely reported. To address this problem, we adopted an attention residual dense CNN, referred to as AttRDN, for LDCT sinogram denoising. First, it was aided by the attention mechanism, in which the advantages of both feature fusion and global residual learning were used to extract noise from the contaminated LDCT sinograms. Then, the denoised sinogram was restored by subtracting the noise obtained from the input noisy sinogram. Finally, the CT image was reconstructed using filtered back-projection. The experimental results qualitatively and quantitatively demonstrate that the proposed AttRDN can achieve a better performance than state-of-the-art methods. Importantly, it can prevent the loss of detailed information and has the potential for clinical application.

This work was supported in part by the National Key R&D Program of China (Nos. 2016YFC0104609 and 2019YFC0605203). The Fundamental Research Funds for the Central Universities (Nos. 2019CDYGYB019 and 2020CDJ-LHZZ-075). And the Chongqing Basic Research and Frontier Exploration Project (No. cstc2020jcyj-msxmX0553).

Keywords Low-dose CT · Sinogram denoising · Deep learning · Attention mechanism

✉ Yong Ren
renyong@cqu.edu.cn

✉ Biao Wei
weibiao@cqu.edu.cn

Yin-Jin Ma
mayinjin@cqu.edu.cn

¹ Key Laboratory of Optoelectronic Technology and Systems, Ministry of Education, Chongqing University, Chongqing 400044, China

² School of Data Science, Tongren University, Tongren 554300, China

³ School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China

⁴ Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

1 Introduction

Computed tomography (CT) has recently become one of the most popular and indispensable medical imaging modalities [1], and it can be utilized for the visualization of anatomical structures of patients with high resolution without invading the human body [2]. However, the inherent X-ray radiation of CT induces potential cancer risks to patients once the cumulative exposure exceeds a certain value [3]. Therefore, the reduction of radiation dose in CT has been a hot research topic that requires imperative handling. Considering these radiation risks, researchers have made efforts to decrease the X-ray dose that a patient is exposed to during CT scanning [4]. In general, lowering

the radiation dose can be implemented by controlling the current of the X-ray tube or by reducing the exposure time to reduce the number of X-ray photons [5]. Although reducing the radiation dose significantly lowers the potential health hazards, such a technique compromises the quality of the reconstructed CT image owing to the low signal-to-noise ratio metrics, which induce severe noise and artifacts. Accordingly, the noise reduction technique determines the success of low-dose CT (LDCT) to a great degree.

To tackle the inherent problem of LDCT, researchers have made significant efforts and proposed various methods. These methods can be categorized into three types [6]: (a) projection data filtering before reconstruction, (b) iterative reconstruction, and (c) image domain filtering after reconstruction.

Projection domain filtering directly suppresses noise in raw projection data before inputting it into the analytic reconstruction. More than a decade ago, Balda et al. [7] and Manduca et al. [8] proposed structural adaptive (Adp-str) filtering and bilateral filtering, which are two efficient approaches. Li et al. investigated the model to determine the statistical property of projection data and presented a penalized likelihood method for quantum noise suppression for low-dose CT [9]. Wang et al. investigated the penalized weighted least-squares approach to address sinogram denoising and reconstruction for low-dose CT [10]. Although sinogram filtering is computationally effective and noise characteristics are modeled in the projection domain, the raw projection data of commercial CT scanners are often not available for research. In addition, projection data should be processed carefully since new artifacts may appear in the reconstructed CT images.

Iterative reconstruction approach estimates the reconstructed CT image using previous information in the image domain. Ordinarily, these methods optimize the objective function by incorporating the statistical properties of the system model, noise model, and previous image information. Recently, compressive sensing [11] has been adopted to address issues related to few-view, interior CT, and LDCT. Total variation (TV) minimization constraint is one of the most well-known methods used to concisely and robustly improve the reconstructed CT images [12]. Without considering the complex structures, the TV regularized method tends to cause blurred details and piecewise artifacts in the reconstructed images. Subsequently, researchers developed several methods that utilize a richer image of previous knowledge. These methods include dictionary learning [13], low rank [14], nonlocal mean [15], and TV variants [16]. Iterative reconstruction methods have been used to improve the denoising performance of LDCT images. Nevertheless, these iterative methods involve a high computational cost in the projection and

back-projection calculation steps; hence, the reconstruction procedure is time-consuming.

Image post-processing methods are an alternative to the two categories of denoising methods mentioned above. This technique directly manipulates the reconstructed LDCT image and is completely independent of projection data; it can be easily assembled into the workflow of the current CT scanner. Extensive efforts have been focused on exploring the image post-processing denoising techniques. Li et al. adapted the nonlocal means filtering (NLM) algorithm to reduce the noise for LDCT images [17]. The block-matching 3D method was used to restore CT images from a different type of noise on denoising tasks [18]. Chen et al. developed a fast dictionary learning [19] by adapting the K-SVD algorithm [20] for LDCT image denoising of the abdomen. However, the mottle noise and artifacts in the LDCT image are complicated and hardly modeled because they do not obey any specific distribution in the image domain. Hence, noise and artifacts in the LDCT image are too complex to be completely treated using conventional image post-processing methods.

In the past several years, there has been a rapid growth of machine learning, especially deep-learning techniques, in the fields of image processing and computer vision [21], which also brings up novel thinking and enormous potential for the medical imaging area [22]. Through a hierarchical multilayer architecture, deep neural networks can efficiently use high-level features at the pixel level [3]. Several deep network models have been presented for CT image restoration, resulting in expressive experimental results. For instance, Han et al. combined a U-Net with residual learning to estimate the artifacts in sparse-view reconstructed CT images [23]. Chen et al. were inspired by the idea of the auto-encoder and designed a convolutional neural network (CNN)-based residual encoder–decoder to address the problem of LDCT image denoising [24]. Because of the mean square error over-smoothing the denoised results, Ma et al. integrated the structural similarity and MSE losses into a deep CNN block model to prevent the over-smoothing issue [25]. A modularized deep CNN proposed by Shan et al. obtained a competitive performance for LDCT reconstruction compared with commercial algorithms [26].

With the popularity of the generated adversarial networks (GANs) [27], several GAN-based algorithms were also developed for LDCT image denoising and greatly enhanced the image quality and improved the diagnosis performance. Yang et al. [28] proposed Wasserstein GAN with perceptual loss for low-dose CT image denoising. Ma et al. [29] utilized a least-squares GAN with structural similarity and L1 losses for low-dose CT image denoising. However, these deep-learning methods are implemented in the image domain and directly operate low-dose CT

images, which easily lose partially detailed information during CT image reconstruction from raw low-dose sinograms [29]. In addition, a previous study [26] pointed out the importance of manipulating raw projection data. The deep-learning method for sinogram denoising can improve the signal-to-noise ratio of projection data in LDCT, which can recover more diagnostic details in reconstructed low-dose CT images.

Deep CNN-based methods for dealing with sinogram data are an emerging direction for CT denoising. Claus et al. restored contaminated projection data via a three-layer neural network and obtained the initial results [30]. By aiding data inconsistency, Park et al. presented a simplified U-Net (Sunet) to learn the correction of metal-induced beam hardening [31]. In contrast to the aforementioned deep-learning-based denoising methods for LDCT in the image domain, deep CNN-based denoising methods for low-dose projection data are scarcely reported owing to the limited availability of raw sinogram data. To improve the signal-to-noise ratio of LDCT projection data and preserve more diagnosis details during LDCT image reconstruction, we studied the deep-learning methods for projection data denoising.

Inspired by the work [32], we utilized an attention deep residual dense CNN, referred to as AttRDN, for low-dose CT sinogram denoising. The AttRDN was aided by the attention mechanism and used the dense connection blocks (DCBs) and global residual dense network. The global residual learning was followed by the attention block, which is efficient for complex denoising tasks, to promote extraction of the noise feature hidden in the LDCT sinogram. The AttRDN first extracted the noise from contaminated LDCT projection data. Next, the attention mechanism is guided to extract the latent features from a complicated noisy sinogram. Then, the clean sinogram is recovered through the obtained noise and the given noisy sinogram. Finally, the CT image can be reconstructed from the denoised projection data.

The remainder of this article is organized as follows: The method of AttRDN is illustrated in Sect. 2. The experimental settings and results are presented in Sect. 3. The discussion is in Sect. 4, and conclusions are summarized in Sect. 5.

2 Methods

2.1 Model for noise reduction

Noise in the projection data originates from electronic and quantum noise. The quantum noise is approximated to a simple Poisson distribution in LDCT, and the electronic noise can be ignored owing to the improved performance

of the CT scanner [33]. The projection data are directly obtained by the CT scanner. If the signal-to-noise ratio of the projection data in LDCT is improved, we could recover more details in the reconstructed CT images, which is hard to restore by denoising in the image domain.

Assuming that P_{LDCT} is an LDCT projection measurement and P_{NDCT} is the corresponding normal-dose CT (NDCT) projection, their relationship can be described as.

$$P_{LDCT} = \sigma(P_{NDCT}), \quad (1)$$

where $\sigma: P \rightarrow \in P$ is a complicated degradation process involving photon starvation noise and other factors. Then, the problem can be converted to search for a function f

$$\operatorname{argmin}_f \|f(P_{LDCT}) - P_{NDCT}\|_2^2, \quad (2)$$

where f denotes the optimal inverse function of σ , which can be estimated using deep-learning techniques.

2.2 Attention residual dense convolutional neural network

2.2.1 Architecture of attention residual dense network.

As shown in Fig. 1, P_{LDCT} and $P_{Denoising}$ represent the low-dose CT and denoised projections, respectively, and serve as the input and output of the attention residual dense convolutional neural network (AttRDN). The first shallow feature, F_{-1} , was extracted from the input low-dose CT sinogram by the first convolutional layer.

$$F_{-1} = H_{SFE1}(P_{LDCT}), \quad (3)$$

where $H_{SFE1}(\cdot)$ denotes the first convolution operation. Then, F_{-1} was used as the input for global residual learning and the second convolutional layer. Hence, the second shallow feature is obtained as follows:

$$F_0 = H_{SFE2}(F_{-1}), \quad (4)$$

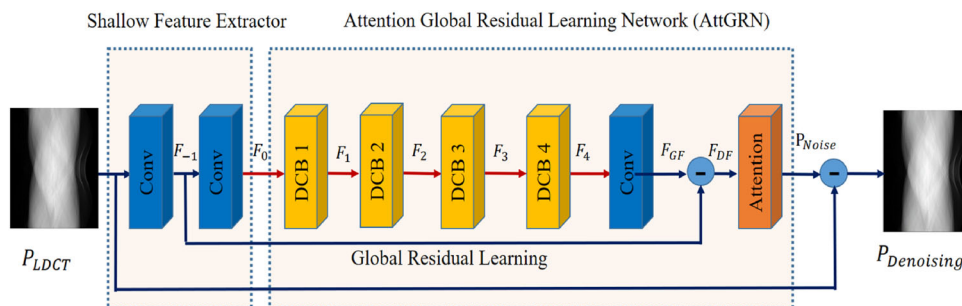
where $H_{SFE2}(\cdot)$ and F_0 represent the second convolution operator and its output, respectively.

We used the extracted shallow feature F_0 as the input to the DCBs. Assuming that our AttRDN contains N DCBs, the output of the n -th DCB is denoted by F_n . F_n can be determined as follows:

$$\begin{aligned} F_n &= H_{DCB,n}(F_{n-1}) \\ &= H_{DCB,n}(H_{DCB,n-1}(\dots(H_{DCB,1}(F_0))), \end{aligned} \quad (5)$$

where $H_{DCB,n}$ represents the operations of the n -th DCB. Each DCB included several layers for the convolution operation, leaky ReLU, and dense feature fusion. F_n takes advantage of each convolution layer contained in the block. As a result, F_n can be considered a local feature.

Fig. 1 (Color online) Architecture of the attention residual dense CNN (AttRDN). It contains shallow feature extractor, dense connection blocks (DCBs), global residual learning, and attention network



Noise-contaminated LDCT projection data can easily conceal the noise features, which prevents the extraction of key features when training deep neural networks. To overcome this difficulty, we introduced an attention mechanism to estimate the noise. The attention block takes the dense feature map F_{DF} as the input and outputs the predicted latent noise. The operation of the attention mechanism can be expressed as

$$P_{Noise} = H_{Att}(F_{DF}). \tag{6}$$

The residual mapping of LDCT projection noise is easier to learn than that of the original NDCT projections. As shown in Fig. 1, the proposed AttRDN utilizes a residual learning technique to reconstruct the predicted clean projection data. This process is formulated as follows:

$$P_{Denoising} = P_{LDCT} - P_{Noise} = P_{LDCT} - H_{AttGRN}(P_{LDCT}). \tag{7}$$

The AttRDN contains mainly three components: a shallow feature extractor, DCBs for local feature extraction and fusion, and attentional global residual learning for global feature fusion and attentional residual learning. Figure 1 shows the overall architecture of the AttRDN.

2.2.2 Dense connection block

Residual learning [34] addresses the performance degradation of training an extremely deep CNN, and a dense network connects each layer to other later layers. Each layer in the denseNet [35] benefits from both low-level and high-level features, alleviating gradient explosion and vanishment. The advantage of dense connection networks is its ability to fuse in dense connection blocks. A dense connection block has a dense connection, local feature fusion, and contiguous memory mechanism. Figure 2 presents the details of the dense connection block.

As shown in Fig. 2, the input signal from the previous DCB passes to every layer contained in the current DCB. Thus, a contiguous memory mechanism is implemented. Assume that F_{n-1} denotes the input and F_n symbols the output of the n -th dense connection block, respectively.

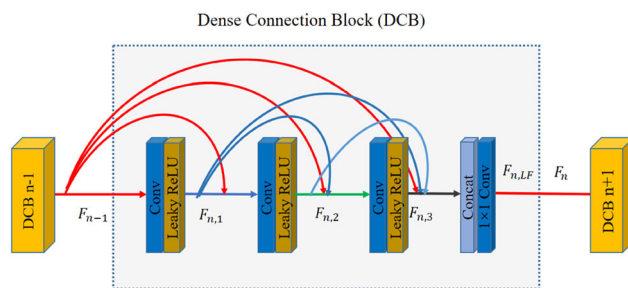


Fig. 2 (Color online) Architecture of a dense connection block (DCB). It integrates the advantages of both residual learning and denseNet

The output of the m -th convolution layer in the current (n -th) DCB can be described as follows:

$$F_{n,m} = \sigma(W_{n,m}[F_{n-1}, F_{n,1}, \dots, F_{n,m-1}]), \tag{8}$$

where σ denotes the activation function leaky ReLU. $W_{n,m}$ denotes the weights of the m -th convolution layer, in which the bias is ignored for brevity. $[F_{n-1}, F_{n,1}, \dots, F_{n,m-1}]$ represents the concatenation of the feature maps calculated based on the previous ($n-1$)-th DCB and convolution layers $1, \dots, m-1$ included in the current (n -th) dense connection block. The outputs of the previous DCB and each layer then directly concatenate to produce $F_{n,LF}$, which not only extracts the local dense feature, but also retains the characteristics of the feed-forward procedure.

For several convolution layers contained in one dense connection block and to further improve the signal processing flow, the feature output from each convolution layer is fused before the output is produced; this is referred to as local feature fusion. Finally, the output of the n -th DCB is obtained.

The local feature fusion and contiguous memory mechanisms can further enhance the representation ability of the neural network, leading to better performance.

2.2.3 Attention mechanism

As illustrated in Fig. 3, our attention block takes feature maps F_{DF} as the input, first performs the channel attention, and then follows the spatial attention. Both the channel and

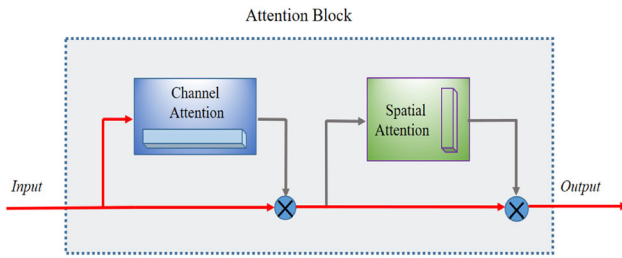


Fig. 3 (Color online) Diagram of attention mechanism utilized in the proposed AttRDN. Attention mechanism can guide our model for learning the noise information

spatial attentions separately learn “what” and “where” of the input feature maps to further push the model performance. We eventually exploited the generated attention feature maps to multiply the input feature maps F_{DF} to predict more significant features of LDCT projection noise, which can be transformed using formula (9):

$$P_{\text{Noise}} = Catt \times Satt \times F_{DF}, \quad (9)$$

where $Catt$ and $Satt$ denote the channel and spatial attention, respectively. By introducing the attention mechanism, we can improve the efficiency and complexity of our denoising model.

2.3 Loss functions

2.3.1 Multi-scale structural loss

In denoising tasks of LDCT projection data, a sinogram serves as an image to be processed, which contains strongly correlative features. The structural similarity measure (SSIM) is a perceptual metric that is more suitable for visual pattern recognition. To measure the structural similarity between two images, SSIM is defined as follows:

$$SSIM(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{\sigma_{xz} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (10)$$

where x, y denote the two compared images; μ_x and μ_y denote the means; σ_x , σ_y , and σ_{xy} represent the standard deviations and the cross-correlation; and C_1 and C_2 are constants for eliminating the singularity.

However, the SSIM is a single-scale metric; hence, we introduced multi-scale SSIM (MS_SSIM) for more flexibility in multi-scale variation. MS_SSIM is expressed using the following formula:

$$MS_SSIM(x, y) = \prod_{i=1}^M SSIM(x_i, y_i), \quad (11)$$

where M is the scale level number, and x_i, y_i are the contents of the i -th level local image. The MS_SSIM loss is usually denoted as follows:

$$MS_SSIM_{\text{Loss}} = 1 - MS_SSIM(x, y). \quad (12)$$

Note that the MS_SSIM loss can back-propagate to update the weights of the network model [36].

2.3.2 L1 loss

However, both L2 and L1 losses can effectively suppress the background noise and remove artifacts. L1 loss does not excessively penalize large errors and treats all errors linearly, which differs from L2 loss. Hence, in the image denoising tasks, L1 loss can alleviate blurring and unnaturalness, which cannot be performed well by L2 loss.

The L1 loss function is expressed using the following formula:

$$L1_{\text{Loss}} = \frac{1}{b \cdot m \cdot n} |x - y|, \quad (13)$$

where x and y denote the denoised CT and NDCT sinograms, respectively, and b , m , and n denote the batch size, height, and width of the sinogram, respectively.

In summary, we obtained the total joint objective function of AttRDN as follows:

$$\min_{\theta_{\text{AttRDN}}} \text{Loss} = \alpha L1_{\text{Loss}} + (1 - \alpha) MS_SSIM_{\text{Loss}}$$

where θ_{AttRDN} denotes the learnable parameters of the proposed AttRDN, and α is the weight balancing L1 loss and MS_SSIM loss. In the training process, the error between the denoised sinogram and the corresponding normal-dose version was calculated, and then, back-propagation was performed to optimize our AttRDN.

2.4 Metrics

For low-dose and denoised sinogram measurements, we used the root-mean-squared error (RMSE) and peak signal-to-noise ratio (PSNR) to quantitatively assess the quality of the projection data. For the reconstructed CT image, we exploited the RMSE, PSNR, and SSIM for the quantitative evaluation of the image quality.

3 Experiment designing and results

3.1 Data sources

To better understand the principle of low-dose CT and the procedure of low-dose simulation, we decided to simulate the low-dose data. With the assumption of a monochromatic X-ray source, the measured projection data can be approximated to a simple Poisson noise distribution, which is expressed as follows:

$$z_i \sim \text{Poisson}\{z_{0i}\exp(-s_i) + r_i\}, i = 1, \dots, I \tag{15}$$

where z_i is the measurement along the path of the i -th X-ray, z_{0i} denotes the photon intensity of the incident X-ray, s_i is the line integral of the attenuation coefficients, and r_i is the read-out noise. For low-dose simulation, we can utilize the parameter z_{0i} to control the noise level.

To evaluate the performance of the proposed AttRDN, a set of projection data were obtained using Radon transform from a realistic clinical CT dataset, which was created and provided by Mayo Clinics for “the 2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge” [37]. This CT dataset includes information on the cases of ten patients, 2,378 normal-dose CT images, and the corresponding simulated quarter-dose CT images with a resolution of 512×512 and slice thickness of 3 mm. In our study, normal-dose sinograms were obtained using Radon transform from 2,378 normal-dose CT images. Then, by setting $z_{0i} = 10^5$ in Eq. (15), we added Poisson noise to the normal-dose projection data to produce the corresponding low-dose versions. We randomly chose 1,943 sinogram pairs for training, 224 for validation, and 211 for testing. Sinogram patches 55×55 in size were used for the training.

3.2 Implementation and parameter setting.

We implemented the AttRDN in Python with the PyTorch 1.0 platform. All the experiments were performed on a personal computer with Intel CPU i7 9700 configuration and 32 G memory, and the training process was accelerated by an NVIDIA RTX 2080 TI graphics processing unit with 11 G video memory.

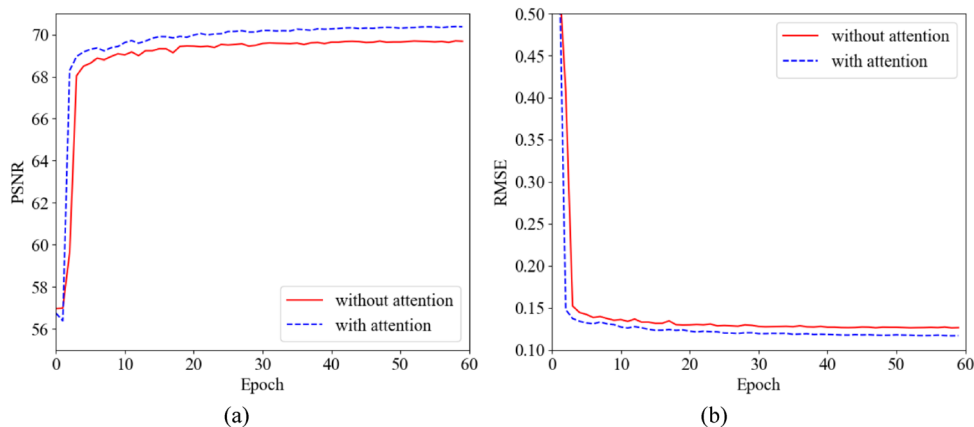
The AttRDN is an end-to-end deep-learning model optimized by minimizing the objective function (15). We adopted Adam to optimize the AttRDN. We set the base learning rate to 10^{-4} and then gradually decreased it to

10^{-5} . The mini-batch size was set to 75. All the convolutional kernel sizes were set to 3×3 and padded zeros to each side to maintain a fixed size, while the local and global feature fusions were set to 1×1 . The convolution layers in each DCB were set to four, and each convolution layer was followed by the activating function Leaky ReLU. The input channels in each DCB were set to 64, and the feature growth rates were 32. Because our task was sinogram denoising, the input and output channels of the entire AttRDN were set to one. To determine the parameter α in the loss function, the α was selected from the following numbers: 0, 0.01, 0.05, 0.1, 0.15, 0.2, 0.3, 0.5, 0.7, and 1.0. The results show that $\alpha = 0.15$ achieved the best performance. This is in agreement with the results of a previous study [36].

3.2.1 Convergence performance

In contrast to conventional convolution operations, the attention mechanism is utilized to excavate the noise components hidden in an intricate background, which helps handle complex denoising tasks, such as blind denoising and real scenario noisy images. Effectively extracting and selecting features are important for medical imaging applications. In this study, we introduced an attention mechanism to augment the representative capability of the denoising CNN model. We assessed the convergence of the network model with and without the attention mechanism, as shown in Fig. 4. The convergence of PSNR with the attention mechanism performed better than that without the attention mechanism. The RMSE with the attention mechanism was more stably convergent than that without the mechanism during the training stage. The attention mechanism improves the performance of the denoising neural network model.

Fig. 4 Convergences of PSNR and RMSE with and without attention mechanism on the testing dataset during training stage. **a** Convergence of PSNR; **b** convergence of RMSE



3.3 Experimental results

3.3.1 Performance improving of sinograms

Two representative results of the processed sinograms and the corresponding reconstructed CT images using the filtered back-projection (FBP) method are selected to demonstrate the denoising capability of the proposed AttRDN. The two examples are shown in Figs. 5 and 6, respectively.

Although it is difficult to observe the differences in the sinograms from normal-dose CT data, low-dose CT data, and those processed by different noise reduction methods, one can easily differentiate the corresponding reconstructed CT images. From Figs. 5 and 6, we can see that various methods of projection domain suppress the noise to various extents. AttRDN and Sunet removed the most noise compared with other methods. The detailed textures indicated by the red arrows in the zoomed regions of interest (ROIs) shown in Figs. 5 and 6 demonstrate the advantages of AttRDN over other methods. The absolute difference maps of the proposed AttRDN are shown in Fig. 7. AttRDN yielded the smallest difference from the normal-dose sinogram data compared with the other methods in our study.

The quantitative results in Figs. 5 and 6 in the projection and image domains are listed in Tables 1 and 2, respectively. From Tables 1 and 2, we observe that the quantitative measurements for different sinogram denoising

methods followed similar trends with visual inspection, as shown in Figs. 5 and 6. Table 3 shows that, for our sinogram testing set, which contains 211 low-dose sinograms and the corresponding normal-dose targets, the average PSNR increased by 23.3892%, while the average RMSE decreased by 78.6915%. The AttRDN had the highest PSNR and the lowest RMSE, outperforming the other methods adopted in this study.

3.3.2 Visual evaluation on CT images

Regardless of both the projection and image domains, the goal of denoising is to restore high-quality CT images from LDCT data and meet wide clinical applications. We compared the AttRDN with state-of-the-art methods, which contain not only sinogram denoising approaches, but also image domain methods. Bilateral filter, Adp-str, and penalized weighted least-squares filtering (PWLS) are conventional methods for projection data. The Sunet represents the representativeness of the deep-learning method for sinograms. NLM is a popular conventional denoising method for image domains. RED-CNN is one of the most popular deep-learning methods that play an important role in the image-domain denoising.

To assess the power of denoising of the proposed AttRDN competing with the approaches mentioned above, we presented two representative results from the testing set and their corresponding zoomed ROIs shown in Figs. 8 to 11. Figure 8 shows a representative result from an

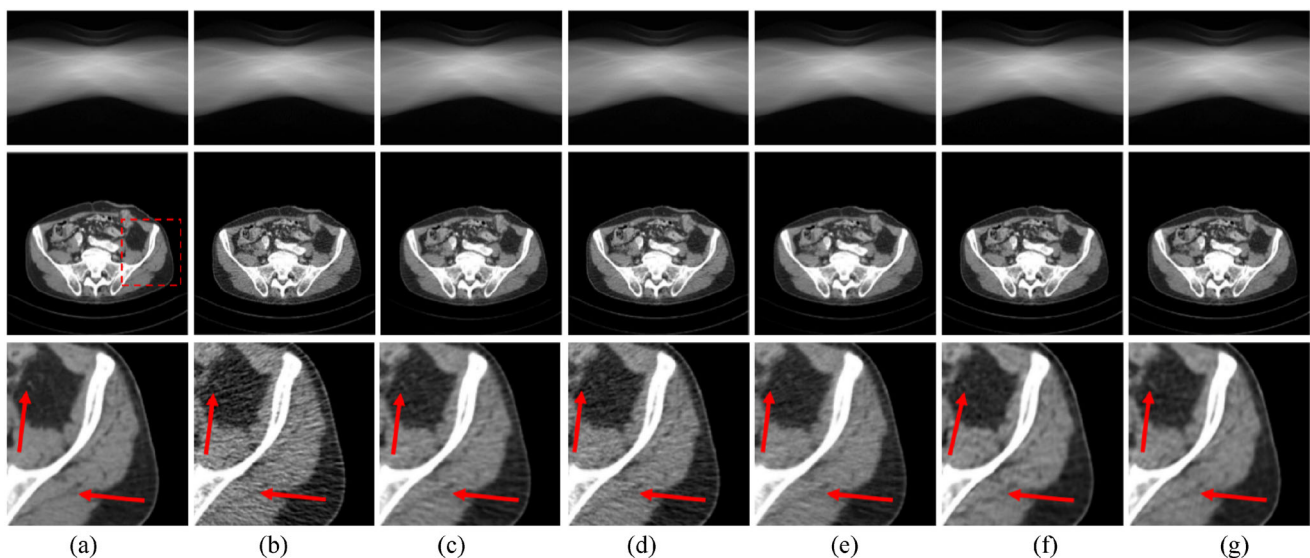


Fig. 5 Results of sinogram denoising of a pelvis slice using different methods from the testing set. The first row shows sinograms including normal dose, low dose, and those processed by different methods. The second row shows the corresponding reconstructed CT images via FBP. The third row shows the zoomed ROIs marked by a rectangle in the second row. **a** Normal dose, **b** low dose, **c** bilateral filter,

d adaptive structural filter, **e** penalized weighted least-squares filter, **f** Sunet, and **g** AttRDN. The display window ranges from -160 to 240 HU. Although hardly observing the differences of the sinograms from normal dose, low dose, and those processed by different noise reduction methods, one can easily differentiate the corresponding reconstructed CT images

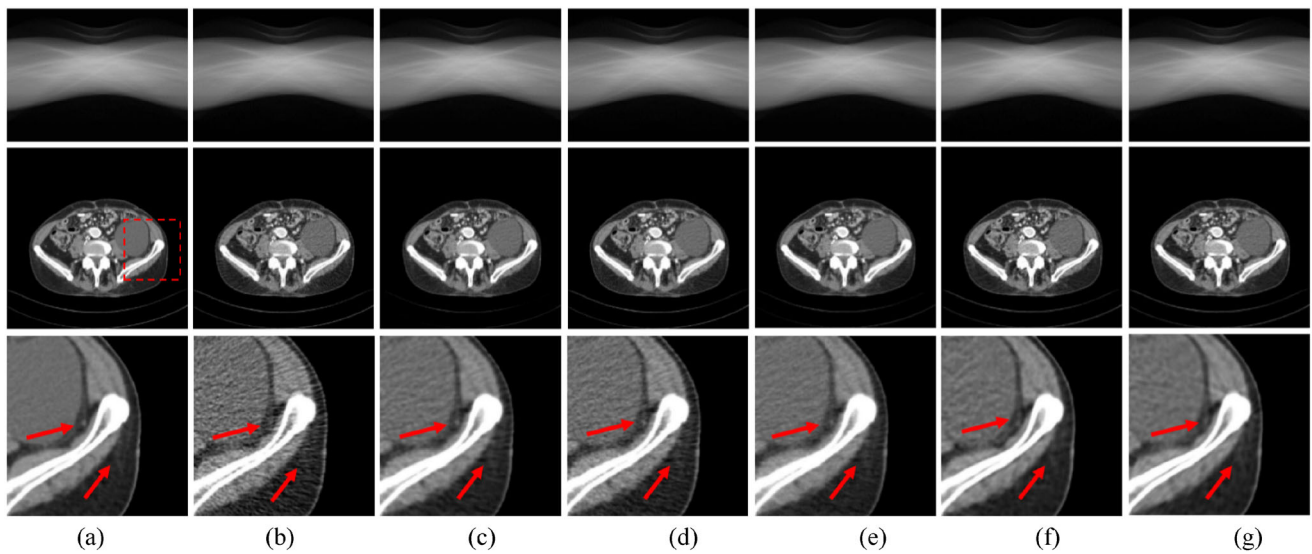


Fig. 6 Results of sinogram denoising of an abdomen slice using different methods from the testing set. The first row shows sinograms including normal dose, low dose, and those processed using different methods. The second row shows the corresponding reconstructed CT images via FBP. The third row shows the zoomed ROIs marked by a rectangle in the second row. **a** Normal dose, **b** low dose, **c** bilateral

filter, **d** adaptive structural filter, **e** penalized weighted least-squares filter, **f** Sunet, and **g** AttRDN. The display window ranges from -160 to 240 HU. Although hardly observing the differences of the sinograms from normal dose, low dose, and those processed by different noise reduction methods, one can easily differentiate the corresponding reconstructed CT images

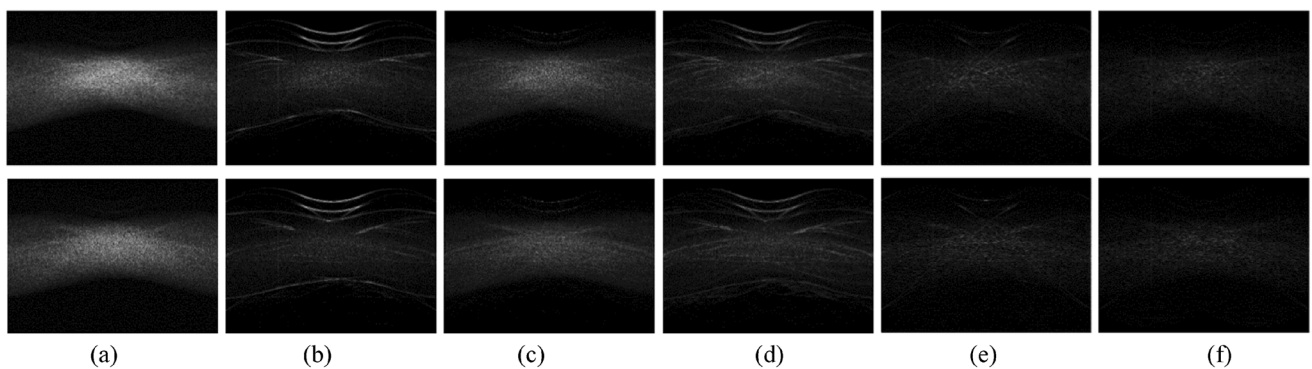


Fig. 7 Absolute difference maps related to the normal-dose CT projection data shown in Figs. 5 and 6. The first row corresponds to the projection data shown in Fig. 5, and the second row corresponds

to Fig. 6. **a** Low-dose CT data, **b** bilateral filter, **c** Adaptive structural filter, **d** Penalized weighted least-squares filter, **e** Sunet, and **f** AttRDN

Table 1 Quantitative results for Fig. 5 utilizing different methods

| | Sinogram | | Reconstructed CT image | | |
|-----------|----------|--------|------------------------|--------|---------|
| | PSNR | RMSE | PSNR | SSIM | RMSE |
| Low dose | 56.3112 | 0.5841 | 27.0602 | 0.8344 | 17.7439 |
| Bilateral | 65.4805 | 0.2033 | 30.3621 | 0.9151 | 12.1326 |
| Adp-str | 61.8737 | 0.3079 | 31.2678 | 0.8936 | 10.9312 |
| PWLS | 65.0111 | 0.2145 | 31.3632 | 0.9145 | 10.9151 |
| Sunet | 67.2282 | 0.1651 | 33.4923 | 0.9295 | 8.4614 |
| AttRDN | 70.1523 | 0.1187 | 33.9731 | 0.9347 | 8.0058 |

Table 2 Quantitative results for Fig. 6 utilizing different methods

| | Sinogram | | Reconstructed CT image | | |
|-----------|----------|--------|------------------------|--------|---------|
| | PSNR | RMSE | PSNR | SSIM | RMSE |
| Low dose | 55.9987 | 0.6055 | 26.0864 | 0.7846 | 19.8490 |
| Bilateral | 65.4381 | 0.2042 | 30.1876 | 0.8836 | 12.3788 |
| Adp-str | 61.5618 | 0.3191 | 30.3136 | 0.8532 | 12.2005 |
| PWLS | 65.0101 | 0.2146 | 31.0616 | 0.8834 | 11.1939 |
| Sunet | 66.9040 | 0.1725 | 32.4579 | 0.8918 | 9.5316 |
| AttRDN | 69.5402 | 0.1279 | 32.7767 | 0.8966 | 9.1881 |

Table 3 Average quantitative results for testing set utilizing different sinogram methods

| | PSNR | RMSE | | PSNR | RMSE |
|-----------|---------|--------|--------|---------|--------|
| Low dose | 56.9777 | 0.5533 | PWLS | 65.2113 | 0.2107 |
| Bilateral | 65.6640 | 0.1997 | Sunet | 67.3078 | 0.1652 |
| Adp-str | 62.3896 | 0.2948 | AttRDN | 70.3043 | 0.1179 |

abdominal CT image. In Fig. 8b, the noise is distributed in the whole abdomen, and streak artifacts appear near the tissues with high attenuation coefficient values, such as bone materials. All denoising methods of projection and image domains could remove noise and artifacts to some extent. Although PWLS effectively removed noise and outperformed NLM, adaptively structural, and bilateral filtering, it was obviously subjected to a blocky effect. As shown in Fig. 8f, PWLS filtering had a better effect than the other conventional methods, while still exhibiting small structural loss. Convolutional network-based deep-learning methods not only effectively eliminated most noise and artifacts, but also efficiently preserved the structural details better than traditional methods. However, the RED-CNN

blurred the denoised CT image, leading to over-smoothing of the subtle textures, because RED-CNN is based on the mean absolute error (MSE). For the capability of noise reduction and detail structure preservation, the proposed AttRDN seemingly exceeded the Sunet, which shows less noise in the top-left area in Fig. 8i than in Fig. 8h. Compared with the Sunet based on MSE, AttRDN adopted residual learning and dense connection, and aided by the attention mechanism, which is trained based on MS-SSIM and L1 losses. Hence, AttRDN performed better than the Sunet.

To further demonstrate the performance of the AttRDN, we provided zoomed images of the ROI labeled with a rectangular dashed line in Fig. 8, as shown in Fig. 9. Here, two white dots that are likely calcifications or calculi within the red circle were hardly observed with other methods, except for Sunet and AttRDN, and were also overly smoothed by RED-CNN. Regardless of the conventional projection domain or image domain approaches, the slim pathologic structures are easily lost. In our study, Sunet and the proposed AttRDN could partly recover them. AttRDN restored them better than Sunet. The fine anatomic textures indicated by the red arrow in Fig. 9 were also best

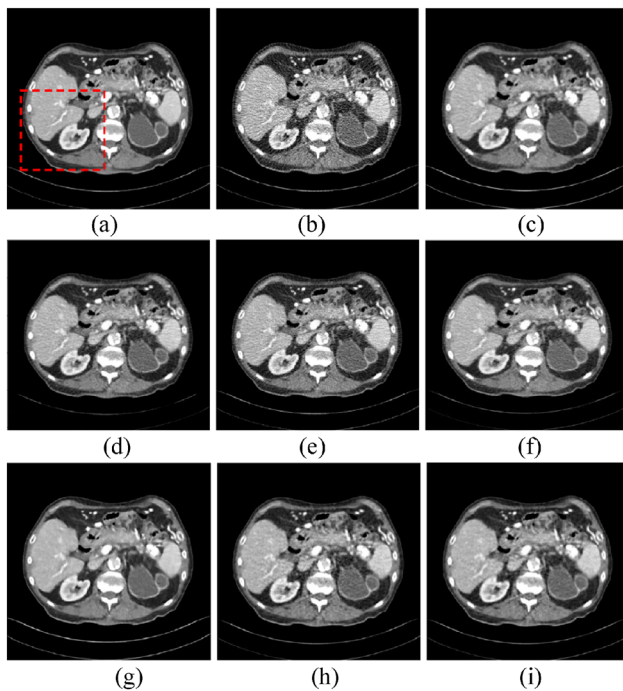


Fig. 8 Results of an abdominal slice from the testing set using different methods. **a** Normal dose, **b** low dose, **c** NLM, **d** bilateral filtering, **e** adaptively structural filtering, **f** penalized weighted least-squares filter, **g** RED-CNN, **h** simplified U-net, and **i** AttRDN. *Note* bilateral, Adp-str, PWLS, Sunet, and AttRDN are denoising methods in the projection domain. NLM and RED-CNN are denoising methods in the image domain. The display window ranges from -160 to 240 HU

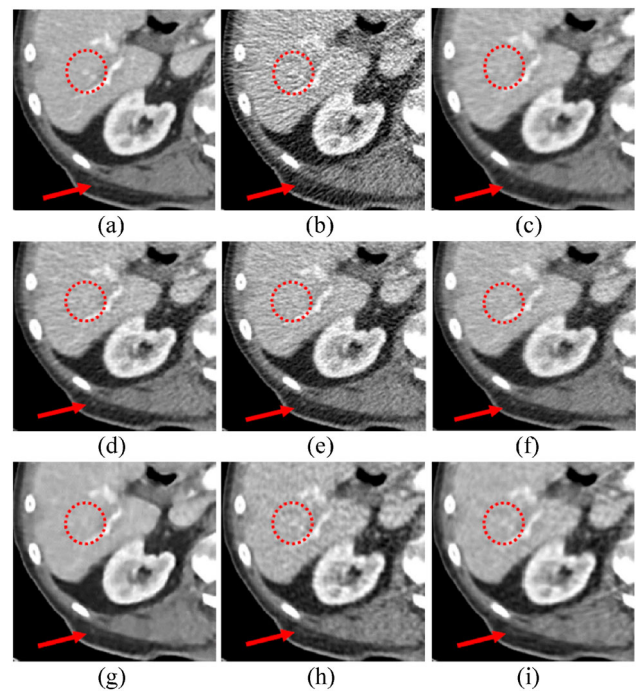


Fig. 9 Magnified region of interest (ROI) marked by a rectangular dashed line in Fig. 8. **a** Normal dose, **b** low dose, **c** NLM, **d** bilateral filtering, **e** adaptively structural filtering, **f** penalized weighted least-squares filter, **g** RED-CNN, **h** Simplified U-net, and **i** AttRDN. *Note* bilateral, Adp-str, PWLS, Sunet, and AttRDN are denoising methods in the projection domain; NLM and RED-CNN are denoising methods in the image domain. The display window ranges from -160 to 240 HU

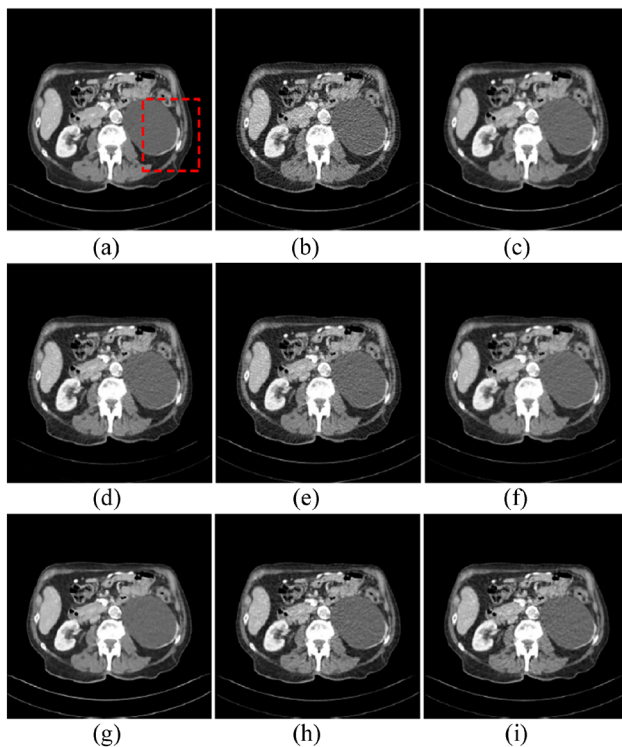


Fig. 10 Results of another abdominal slice from the testing set using different methods. **a** Normal dose, **b** low dose, **c** NLM, **d** bilateral filtering, **e** adaptively structural filtering, **f** penalized weighted least-squares filter, **g** RED-CNN, **h** Sunet, and **i** AttRDN. *Note* bilateral, Adp-str, PWLS, Sunet, and AttRDN are denoising methods in the projection domain; NLM and RED-CNN are denoising methods in the image domain. The display window ranges from -160 to 240 HU

preserved by the AttRDN with less remaining noise than that of other methods in this study.

Another result of the testing set is presented in Fig. 10, and its ROI is shown in Fig. 11. Because of the reduced radiation dose, the noise inundated most of the small pathological tissue structure, making it difficult to observe them clearly. Although the NLM suppressed most of the noise in the region indicated by the red arrow, the edges between different organs and the details appear blurry in Fig. 11. The subcutaneous fat structures or lipomata indicated by the red arrows were also properly restored by the proposed AttRDN. From Figs. 10 and 11, we observed that the images processed by the adaptively structural filter retained some noise in our study. The textures indicated by the two red arrows were over-smoothed by NLM and RED-CNN, which are consistent with the trends shown in Figs. 8 and 9. In summary, the Sunet and the proposed AttRDN, which is based on the projection domain and deep-learning techniques, directly processing the raw data, enable the effective reduction of noise and preserve more clinical information in contrast to the other methods, while the AttRDN performs slightly better than the Sunet.

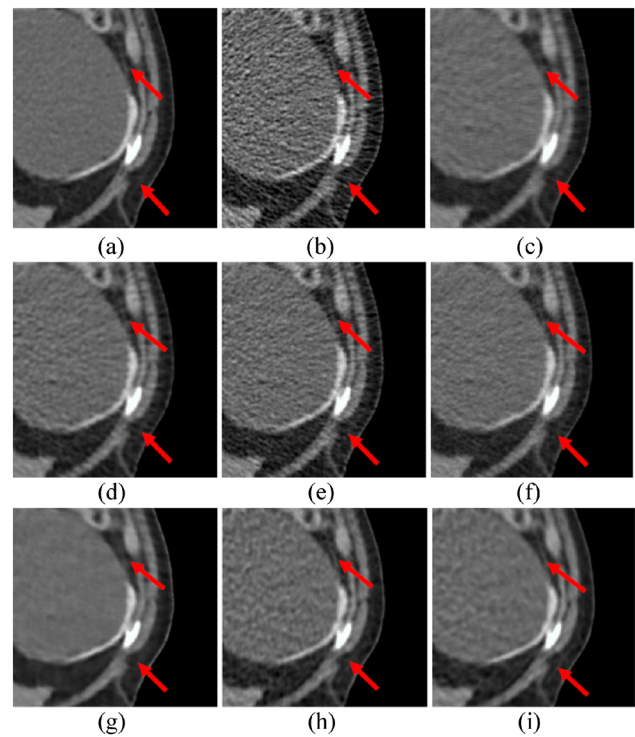


Fig. 11 Magnified region of interest (ROI) marked by a rectangular dashed line in Fig. 10. **a** Normal dose, **b** low dose, **c** NLM, **d** bilateral filtering, **e** adaptively structural filtering, **f** penalized weighted least-squares filter, **g** RED-CNN, **h** Sunet, and **i** AttRDN. *Note* bilateral, Adp-str, PWLS, Sunet, and AttRDN are denoising methods in the projection domain; NLM and RED-CNN are denoising methods in the image domain. The display window ranges from -160 to 240 HU

3.3.3 Quantitative evaluation on CT images

To quantitatively evaluate the different methods in this study, we calculated the PSNR, SSIM, and RMSE of the reconstructed CT images using different methods. For the sinogram denoising methods, we reconstructed the images from the predicted sinogram data via FBP. For image-post processing methods, we first reconstructed the images from the simulated low-dose projection data and then denoised the reconstructed LDCT images using image-domain denoising methods.

The quantitative measurements for the entire CT images shown in Figs. 8 and 10 are listed in Table 4. The adaptive structural filtering achieved low PSNR and SSIM and high RMSE. In Figs. 8 and 10, the CT images processed with adaptive structural filtering still had more remnant noise than those processed using other methods. The AttRDN obtained the best scores in both Figs. 8i and 10i. These results were consistent with those of the visual evaluation. The statistical average values for the metrics of the 211 samples included in the testing set are listed in Table 5. AttRDN obtained the best PSNR, SSIM, and RMSE.

Table 4 Quantitative measurements for whole CT images shown in Figs. 8 and 10 utilizing different methods

| | Figure 8 | | | Figure 10 | | |
|-----------|----------|--------|---------|-----------|--------|---------|
| | PSNR | SSIM | RMSE | PSNR | SSIM | RMSE |
| Low dose | 23.7817 | 0.7322 | 25.8806 | 28.7416 | 0.8530 | 14.6210 |
| NLM | 29.3363 | 0.8426 | 13.6535 | 31.8445 | 0.9111 | 10.2290 |
| Bilateral | 29.0859 | 0.8616 | 14.0529 | 30.8941 | 0.9221 | 11.4118 |
| Adp-str | 28.1601 | 0.8173 | 15.6334 | 32.7610 | 0.9057 | 9.2047 |
| PWLS | 29.8151 | 0.8619 | 12.8619 | 32.0379 | 0.9225 | 10.0038 |
| RED-CNN | 31.1732 | 0.8751 | 11.0510 | 34.4502 | 0.9353 | 7.5779 |
| Sunet | 30.9162 | 0.8770 | 11.3829 | 34.2774 | 0.9306 | 7.7301 |
| AttRDN | 31.2055 | 0.8819 | 11.0099 | 34.7131 | 0.9357 | 7.3520 |

Bilateral, Adp-str, PWLS, Sunet, and AttRDN are denoising methods in the projection domain; NLM and RED-CNN are denoising methods in the image domain

Table 5 Average quantitative results for testing set utilizing different methods

| | PSNR | SSIM | RMSE |
|-----------|---------|--------|---------|
| Low dose | 27.3008 | 0.8207 | 17.8004 |
| NLM | 31.7179 | 0.8918 | 10.9868 |
| Bilateral | 30.4159 | 0.9064 | 12.0971 |
| Adp-str | 31.4613 | 0.8821 | 10.9808 |
| PWLS | 31.3717 | 0.9050 | 10.8618 |
| RED-CNN | 33.7019 | 0.9174 | 8.3810 |
| Sunet | 33.4863 | 0.9171 | 8.6169 |
| AttRDN | 33.9796 | 0.9226 | 8.1682 |

Bilateral, Adp-str, PWLS, Sunet, and AttRDN are denoising methods in the projection domain; NLM and RED-CNN are denoising methods in the image domain

Moreover, the SSIM is a more suitable assessment of clinical information in medical images.

The bar graphs of the quantities of the two measured ROIs shown in Figs. 9 and 11 are shown in Fig. 12. The measured results of the two ROIs followed the same trends as those of the visual investigation mentioned above. The AttRDN had the highest PSNR and SSIM and the lowest RMSE for the two local ROIs.

4 Discussion

X-ray radiation may induce potential risks of cancer or genetic disease in patients, but a low radiation dose will decrease the signal-to-noise ratio of projection data in LDCT. This study aimed to investigate the projection data denoising in LDCT using a deep-learning method to obtain high-quality CT images reconstructed by the denoised projection data in LDCT. The LDCT denoising method restores the CT images from LDCT to approach the NDCT

images as much as possible. In this study, we investigated the integration of the advantages of residual learning, dense networks, and attention mechanisms and proposed the AttRDN for LDCT denoising in the projection domain. CNN-based methods have the potential to overcome the fundamental drawbacks of conventional methods. The essential challenge of introducing deep learning into the medical image field is to collect sufficient high-quality labeled training data. We mitigate this difficulty by adopting simulated sinogram training data from the Mayo Clinic Low Dose CT Dataset as the basis for sinogram denoising learning. To effectively suppress noise and remove artifacts while recovering more perfect projection data for subsequent CT image reconstruction, we utilized multi-scale structural loss and L1 loss as the objective function, which can achieve high performance and avoid over-smoothed denoised sinograms. Importantly, the introduced attention mechanism can guide the learning process by focusing on regional attention on feature extraction and augmenting the power of CNN for global information.

The experimental results demonstrated that the trained proposed AttRDN can effectively and efficiently restore the projection data from noise-contaminated LDCT raw data. Then, we can adopt a simple analytical reconstruction method, such as FBP, to reconstruct the sinogram into the image domain for diagnosis purposes. Although there are no significant discrepancies in processed sinograms by different methods, our results showed the best performance in terms of PSNR and RMSE compared with other projection methods used in this study. By comparing the CT images from the processed sinograms restored using different projection methods with those treated by the popular image post-processing approaches, the AttRDN obtained the best score on PSNR, SSIM, and RMSE compared with other methods in our study. Significantly, the AttRDN can restore slim structure and subtle detail from low-dose data.

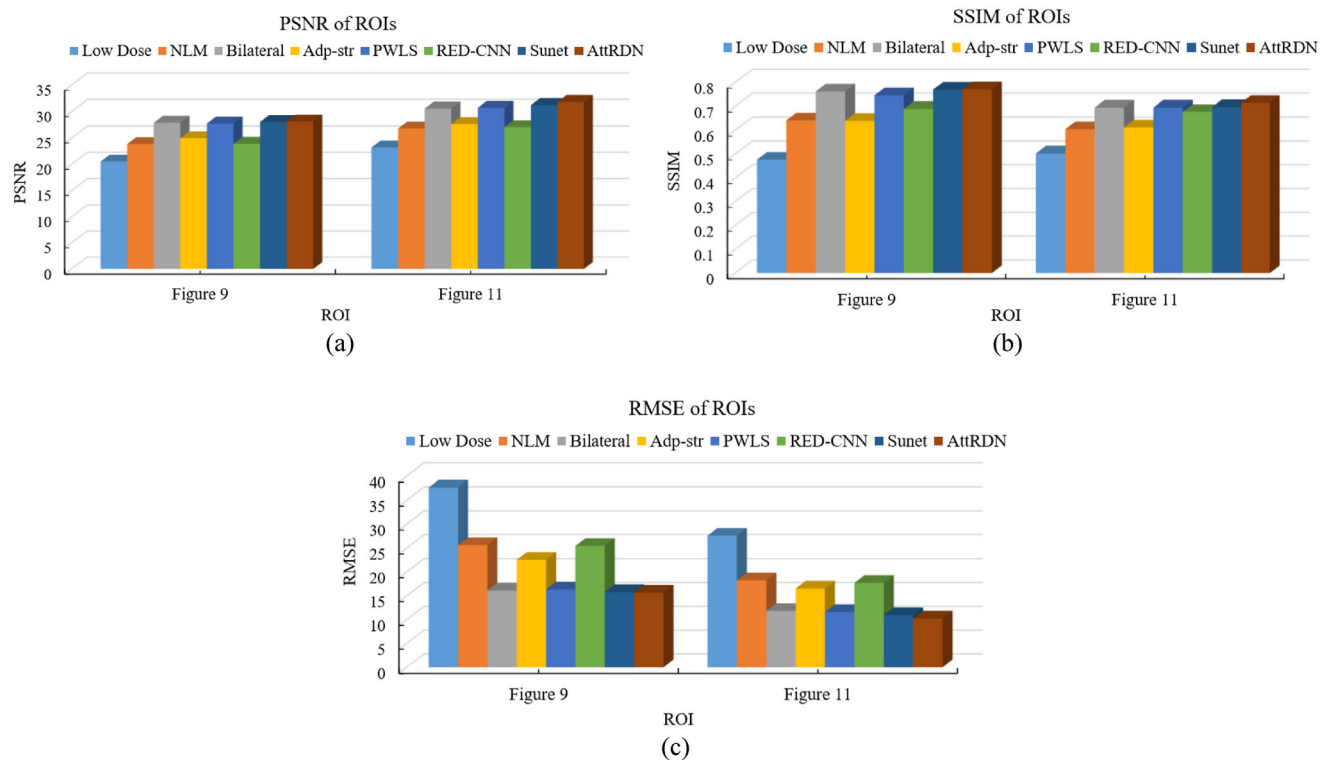


Fig. 12 (Color online) Performance comparison of low-dose CT and different methods over the ROI images in Figs. 9 and 11. **a** Comparison of PSNR, **b** comparison of SSIM, and **c** comparison of RMSE

(See Fig. 9; the two white dots indicating calcifications or calculi surrounded by a red dotted circle were restored by the AttRDN.) This weak but important texture information is more significant for clinical diagnosis, which can be easily lost by the image domain method, be it a traditional or deep-learning method. However, the AttRDN performed on the projection domain supplies only a gap for the shortage.

The training data contain only a single noise setting. Because the actual clinical situation is more complicated, network models should be retrained or re-adjusted for various samples to adapt to different noise levels. Meanwhile, the loss function of the AttRDN is a combination of multiple structural and L1 losses; thus, they should be carefully balanced. To some extent, although the proposed AttRDN can remedy the weakness of the image post-processing methods and generate a denoised sinogram approximating the standard normal-dose version, the reconstructed images do not completely match the corresponding normal-dose CT images. Owing to the LDCT projection data, there are many noises and artifacts in Figs. 9b and 11b. Although the AttRDN could suppress most of the noise, some noise or artifacts are still shown in Figs. 9i and 11i. Designing a network model that directly maps raw projection data into the final CT image is a better method, by which the fitting capacity of deep CNN and the

CT data completeness can be perfectly integrated, which should be our next target.

5 Conclusion

We have presented a CNN-based sinogram denoising method known as AttRDN for LDCT, which integrates the residual learning and dense network and is locally and globally guided by the attention mechanism. In place of concentrating on the intuitionistic image domain, great efforts were made on the projection data. Residual and dense networks leverage the advantage of feature fusion of the local and global feature information, augmenting the representative power. In addition, the attention mechanism is utilized to guide the filtering of the sinogram data. The experimental results demonstrated that the AttRDN outperformed the state-of-the-art methods in the projection domain or image domain and had the potential to improve the quality of low-dose CT images. To some degree, the AttRDN can cover a gap in the image post-processing methods. In the future, we plan to make further efforts to optimize the AttRDN, extend it to adversarial learning and reconstruction, and even adapt it to other medical imaging modalities.

Acknowledgements The authors would also like to thank Dr. Cynthia McCollough, the Mayo Clinic, the American Association of Physicists in Medicine, and grant EB017095 and EB017185 from the National Institute of Biomedical Imaging and Bioengineering for providing the Low-Dose CT Grand Challenge dataset.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Yin-Jin Ma, Peng Feng, Peng He, and Xiao-Dong Guo. The first draft of the manuscript was written by Yin-Jin Ma, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

- G. Wang, H.Y. Yu, B.D. Man, An outlook on x-ray CT research and development. *Med. Phys.* **35**, 1051–1064 (2008). <https://doi.org/10.1118/1.2836950>
- P. Feng, W.X. Cong, B. Wei et al., Analytic Comparison between X-ray Fluorescence CT and K-edge CT. *IEEE Trans. Biomed. Eng.* **61**(3), 975–985 (2014). <https://doi.org/10.1109/TBME.2013.2294677>
- H.K. Yang, K.C. Liang, K.J. Kang et al., Slice-wise reconstruction for low-dose cone-beam CT using a deep residual convolutional neural network. *Nucl. Sci. Tech.* **30**(59), 1–9 (2018). <https://doi.org/10.1007/s41365-019-0581-7>
- R.S.T. Kang, T. Wu, Z.H. Chen et al., 3D imaging of rat brain neural network using synchrotron radiation. *Nucl. Tech.* **43**(7), 070101 (2020). <https://doi.org/10.11889/j.0253-3219.2020.hjs.43.070101> (in Chinese)
- I. Barreto, N. Verma, N. Quails et al., Patient size matters: Effect of tube current modulation on size-specific dose estimates (SSDE) and image quality in low-dose lung cancer screening. *CT J. App. Clin. Med. Phys.* **21**, 87–94 (2020). <https://doi.org/10.1002/acm2.12857>
- H. Chen, Y. Zhang, W.H. Zhang et al., Low-dose CT via convolutional neural network. *Biomed. Opt. Exp.* **8**, 679–694 (2017). <https://doi.org/10.1364/boe.8.000679>
- M. Balda, J. Hornegger, B. Heismann, Ray contribution masks for structure adaptive sinogram filtering. *IEEE Trans Med Imag.* **31**, 1228–1239 (2012). <https://doi.org/10.1109/Tmi.2012.2187213>
- A. Manduca, L.F. Yu, J.D. Trzasko et al., Projection space denoising with bilateral filtering and CT noise modeling for dose reduction. *CT. Med. Phys.* **36**, 4911–4919 (2009). <https://doi.org/10.1118/1.3232004>
- T.F. Li, X. Li, J. Wang et al., Nonlinear sinogram smoothing for low-dose X-ray CT. *IEEE Trans. Nucl. Sci.* **51**, 2505–2513 (2004). <https://doi.org/10.1109/tns.2004.834824>
- J. Wang, T.F. Li, H.B. Lu et al., Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography. *IEEE Trans. Med. Imag.* **25**, 1272–1283 (2006). <https://doi.org/10.1109/tmi.2006.882141>
- M.F. Duarte, Y.C. Eldar, Structured compressed sensing: From theory to applications. *IEEE Trans. Sig. Process.* **59**, 4053–4085 (2011). <https://doi.org/10.1109/tsp.2011.2161982>
- H.W. Tseng, S. Vedantham, A. Karellas, Cone-beam breast computed tomography using ultra-fast image reconstruction with constrained, total-variation minimization for suppression of artifacts. *Phys. Med. Eur. J. Med. Phys.* **73**, 117–124 (2020). <https://doi.org/10.1016/j.ejmp.2020.04.020>
- Y. Chen, L.Y. Shi, Q.J. Feng et al., Artifact suppressed dictionary learning for low-dose CT image processing. *IEEE Trans. Med. Imag.* **33**, 2271–2292 (2014). <https://doi.org/10.1109/tmi.2014.2336860>
- J.F. Cai, X. Jia, H. Gao et al., Cine cone beam CT reconstruction using low-rank matrix factorization: algorithm and a proof-of-principle study. *IEEE Trans. Med. Imag.* **33**, 1581–1591 (2014). <https://doi.org/10.1109/Tmi.2014.2319055>
- Y. Zhang, Y. Xi, Q.S. Yang et al., Spectral CT reconstruction with image sparsity and spectral mean. *IEEE Trans. Comput. Imag.* **2**, 510–523 (2016). <https://doi.org/10.1109/tci.2016.2609414>
- L.Z. Deng, P. He, S.H. Jiang et al., Hybrid reconstruction algorithm for computed tomography based on diagonal total variation. *Nucl. Sci. Tech.* **29**(3), 45 (2018). <https://doi.org/10.1007/s41365-018-0376-2>
- Z.B. Li, L.F. Yu, J.D. Trzasko et al., Adaptive nonlocal means filtering based on local noise level for CT denoising. *Med. Phys.* **41**, 011908 (2014). <https://doi.org/10.1118/1.4851635>
- D. Kang, P. Slomka, R. Nakazato et al., Image denoising of low-radiation dose coronary CT angiography by an adaptive block-matching 3D algorithm. *Proc. SPIE.* **6869**, 1–6 (2013). <https://doi.org/10.1117/12.2006907>
- Y. Chen, X.D. Yin, L.Y. Shi et al., Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing. *Phys. Med. Biol.* **58**, 5803–5820 (2013). <https://doi.org/10.1088/0031-9155/58/16/5803>
- M. Aharon, M. Elad, A. Bruckstein, K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Sign. Process.* **54**, 4311–4322 (2006). <https://doi.org/10.1109/tsp.2006.881199>
- Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015). <https://doi.org/10.1038/nature14539>
- G. Zhao, H.L. Xiong, G.D. Wu et al., Evaluation study on comprehensive efficiency of physical protection system based on neural network. *Nucl. Tech.* **43**(2), 020602 (2020). <https://doi.org/10.11889/j.0253-3219.2020.hjs.43.020602> (in Chinese)
- Y. S. Han, J. Yoo, J. C. Ye, Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis. 2016.
- H. Chen, Y. Zhang, M.K. Kalra et al., Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans Med Imag.* **36**, 2524–2535 (2017). <https://doi.org/10.1109/tmi.2017.2715284>
- Y. Ma, P. Feng, P. He et al., Low-dose CT with a deep convolutional neural network blocks model using mean squared error loss and structural similar loss. *Proc. SPIE* **11209I**, 1–13 (2019). <https://doi.org/10.1117/12.2542662>
- H. Shan, A. Padole, F. Homayounie et al., Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction. *Nature Mach. Intel.* **1**, 269–276 (2019). <https://doi.org/10.1038/s42256-019-0057-9>
- I. J. Goodfellow, J. Pouget, M. Mirza et al, Generative adversarial nets. 2014.
- Q. Yang, P. Yan, Y.B. Yan et al., Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans. Med. Imag.* **37**, 1348–1357 (2018). <https://doi.org/10.1109/tmi.2018.2827462>
- Y. Ma, B. Wei, P. Feng et al, Low-dose CT image denoising using a generative adversarial network with a hybrid loss function for noise learning. *IEEE Access* **8**, 67519–67529 (2020). <https://doi.org/10.1109/access.2020.2986388>
- B.E.H. Claus, Y. Jin, L.A. Gjestebj et al., (2017) Metal-artifact reduction using deep-learning based sinogram completion: initial results. *Proc. Fully3D.* **45**: 631–635. <https://doi.org/https://doi.org/10.12059/Fully3D.2017-11-3110004>

31. H.S. Park, S.M. Lee, H.P. Kim et al., CT sinogram-consistency learning for metal-induced beam hardening correction. *Med. Phys.* **45**, 5376–5384 (2018). <https://doi.org/10.1002/mp.13199>
32. Y. Zhang, Y. Tian, Y. Kong et al., Residual dense network for image restoration. 2018.
33. I.A. Elbakri, J.A. Fessler, Statistical image reconstruction for polyenergetic X-ray computed tomography. *IEEE Trans. Med. Imag.* **21**, 89–99 (2002). <https://doi.org/10.1109/42.993128>
34. K. He, X. Zhang, S. Ren et al., 2016 Deep residual learning for image recognition. *Proc. IEEE Conf. Comput. vis. Pattern Recognit. (CVPR)* 770–778 (2016). <https://doi.org/https://doi.org/10.1109/Cvpr.2016.90>
35. G. Huang, Z. Liu, L. Maaten et al., Densely connected convolutional networks. 2017.
36. H. Zhao, O. Gallo, I. Frosio et al., Loss functions for image restoration with neural networks. *IEEE Trans. Comput. Imag.* **3**, 47–57 (2017). <https://doi.org/10.1109/Tci.2016.2644865>
37. C. McCollough, Overview of the low dose CT grand challenge. *Med. Phys.* **43**(6), 3759–3760 (2016). <https://doi.org/10.1118/1.4957556>