Low-Dose CT Image Denoising Using a Generative Adversarial Network With a Hybrid Loss Function for Noise Learning

YINJIN MA [ID]1,2, BIAO WEI [ID]1, PENG FENG [ID]1, PENG HE [ID]1, XIAODONG GUO1, AND GE WANG [ID]3, (Fellow, IEEE)

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

Corresponding author: Biao Wei (weibiao@cqu.edu.cn)

ABSTRACT Potential risk of X-ray radiation from computed tomography (CT) has been a concern of the public. However, simply decreasing the dose will degrade quality of the CT images and compromise diagnostic performance. In this paper, we propose a noise learning generative adversarial network coupling with least squares, structural similarity and L1 losses for low-dose CT denoising. In our method, noise distributed in the input low-dose CT image is learned by the generator network and then subtracted from the input to generate the final denoised version. The denoised CT images are penalized by the least squares loss function, and they are pulled toward boundary of the decision even though they are classified as normal-dose CT. Least squares stabilize the training process without regularization. Structural similarity and L1 losses are utilized to keep textural details and sharpness of the denoised CT images respectively. Experiments and results show that our method can effectively suppress noise and remove artifacts compared with the state-of-the-art methods. The texture statistical properties, which include mean, standard deviation, uniformity, and entropy, further confirm that the generated noise-reduced CT image is as close as to that of the normal-dose counterpart.

INDEX TERMS Deep learning, generative adversarial network, least squares, low-dose CT, denoising.

I. INTRODUCTION

For recent decades, X-ray computed tomography (CT) is one of the most practical imaging modalities, which is extensively utilized in medical imaging, industrial evaluation, and other applications [1], [2]. Numerous patients receive huge benefits from CT scans. With the widespread use of medical CT, the potential risk of ionizing X-ray radiation to patients has been raised concerns of the public [3], [4]. Under the well-known guideline of ALARA (as low as reasonably achievable), the minimization of X-ray dose has become an important research topic [4]. The most popular way to cut

The associate editor coordinating the review of this manuscript and approving it for publication was Hengyong Yu [ID].

down radiation dose is to reduce the X-ray flux by shortening the exposure time and decreasing the operating current of an X-ray tube. Frequently, the lower the X-ray, the noisier a CT image, which results in the degradation of the signal-to-noise ratio and compromise of diagnostic performance [5]. To tackle this inherent problem, many algorithms have been developed to improve low-dose CT images [6]. Usually, these algorithms are classified into three categories: (a) sinogram filtration, (b) iterative reconstruction, and (c) post-processing.

Raw data or its log-transformation can be processed by sinogram filtering before CT image reconstruction, for example, filtered back-projection (FBP) [7]. It is convenient that the characteristics of noise in the sinogram domain are

well-known. This classic type of methods contains bilateral filtering [8], filtering for adaptive structure [9], and penalty of weighted least squares (PWLS) [10]. However, the methods based on sinogram filtering suffer from the drawbacks of spatial resolution loss or edge blurring.

The methods of iterative reconstruction iteratively optimize an objective function, enforce the statistical properties of projection data, the prior information of real images, and the parameters of the imaging system. The issues of low-dose, limited angle, few view, and interior CT can be addressed by compressive sensing [11]. Iterative reconstruction methods include dictionary learning [12], nonlocal means [13], [14], low rank [15], total variance [16] and variances [17], [18], and so on. This type of methods considers the physics in the iterative model, and some of these methods were implemented on modern CT scanners. These methods are vender-specific since the geometric parameters and correction steps of CT scanners are not publicly accessible. Moreover, the computational cost of iterative approaches is intensive. While the iterative methods are computationally demanding, the resultant image quality is clearly.

Post-processing methods directly perform on reconstructed low-dose CT image. These methods require no knowledge on raw data and are practically convenient. Hence, intensive efforts have made in the image domain for reducing low-dose noise and artifacts. For example, adaptive nonlocal mean (NLM) [19] was utilized for CT image noise reduction. Block matching (BM3D) [20], [21], was adopted for several CT tasks. With post-reconstruction processing, the image quality can be significantly improved, but over smoothing and residual problem remaining in the processed images, because of the non-uniform noise in CT images.

Recently, deep learning techniques have been actively developed, and applied to various applications [22], [23]. In particular, deep learning also provides new thinking and tremendous potential in the field of medical imaging [24]–[26]. The problem of noise reduction for low-dose CT can be handled by deep learning [27]. Simply, a three-layer of convolutional neural networks (CNN) presented not long ago for low-dose CT noise reduction with a promising performance [6], [25]. Specifically, a convolutional auto-encoder-decoder with residual learning (RED-CNN) [28] was successfully developed for estimating a normal-dose CT image from a low-dose version. A deep CNN with a directional wavelet transform combined shortcut connections for low-dose CT [29].

Although these novel network architectures achieve significant performance gains, the resultant image still suffer from blurry appearance and missing textural details. It is noted that the network structures described above only target minimizing the mean square error between the restored CT image and the corresponding normal-dose CT [30]. To deal with this limitation, the generative adversarial network (GAN) [31], [32] was introduced for low-dose CT. For handling the difficulty in training GAN, Wasserstein distance was introduced to measure the discrepancy between distributions of

generated and real images, yielding the Wasserstien distance based GAN (WGAN) [33]. Furthermore, gradient penalty was employed as a regularization mean to accelerate the training process (WGAN-GP) [34]. Importantly, WGAN-VGG [35] was a method with a perceptual loss for low-dose CT, which achieved promising denoised CT images [36], and the perceptual loss was implemented by VGG [37] that pre-trained on natural images. Also, SMGAN [38] based on WGAN-GP combined multiscale structural loss [39] and L1 loss into the objective function, and took advantage of the information between adjacent inter-slices, produced results compared favorably with that of WGAN-VGG [35].

Although WGAN-GP [34] overcame the mode collapse and training convergence issues, the use of gradient penalty weakened the representation capacity of GAN [40].

To deal with the problem and enlightened by the results [38], [40], here we propose a residual-CNN-block generator and discriminator for noise learning with the least squares [41]. The structural similarity (SSIM) [36] and L1 losses integrated into the overall objective function. The noise in the input low-dose CT image is learned by the residual-CNN-block generator, and then subtracted from the input to obtain a clean CT image. With this GAN, the problem of gradient vanishing is solved without gradient penalty regularization, and the resultant CT images are practically satisfactory, since those reconstructed low-dose CT images far from the distribution of real normal-dose CT would be effectively penalized.

Specifically, low-dose CT noise reduction is treated as a conversion of low-dose CT image into normal-dose CT image. Our dedicated GAN performs the estimation of the distance of distribution between low-dose CT and normal-dose CT. In this process, the SSIM loss keeps structural and textural details after noise suppression, and L1 loss preserves the sharpness of the denoised image. In the rest of this paper, the proposed method is described in Section II. The experiments and the results are presented in Section III. Finally, discussions and the conclusion are in Section IV.

II. METHODS
A. MODELING FOR NOISE REDUCTION

Noise reduction via post-processing is straightly implemented in the image domain. However, low-dose CT noise reduction differs from natural images, since the statistical property of low-dose CT images is difficult to be modeled in the image domain [41]. In general, noise reduction in natural image is noise-type-dependent, yet the data distribution of low-dose CT is considered as the mixture of quantum Poisson and electronic Gaussian noises which vary significantly in the whole CT image. This makes it suboptimal to use traditional denoising methods for low-dose CT. However, the uncertain noise model can be tackled using deep learning techniques, because of its strong ability of high-level features capturing.

Noise reduction for low-dose CT images can be modeled as follows. Assume that $x \in \mathbb{R}^{m \times n}$ denotes a low-dose CT image and $y \in \mathbb{R}^{m \times n}$ denotes the corresponding normal-dose

one, we can formulate their relationship as

$$x = \sigma(y) \tag{1}$$

where $\sigma : \mathbb{R}^{m \times n} \to \mathbb{R}^{m \times n}$ describes the contaminating process of normal-dose CT by quantum noise. m and n denote the width and height of a CT image respectively. The task for low-dose CT image denoising is to find a function f:

$$\arg\min_{f} ||f(x) - y||_2^2 \tag{2}$$

where f denotes an approximation function of σ^{-1}. In addition to the least square norm, the optimization problem can also be formulated under different objective functions. Then, we can invert the process using a deep learning method.

B. LEAST SQUARES GAN

In recent years, generative adversarial networks (GANs) have produced impressive results via unsupervised and semi-supervised learning. A GAN includes a pair of neural networks: a generator G and a discriminator D. The generator G takes a noisy z as the input and generates a fake sample, *i.e.*, $G(z)$. The discriminator D takes a real sample x or a fake sample $G(z)$ as the input, and makes the best effort to distinguish between real or fake samples. The generator and discriminator are in a game relationship. In the game, the generator is trained to generate the samples as realistic as possible, while the discriminator is trained to become as smart as possible for making a distinction between real or fake samples. The pair of networks are often trained alternatively, so that the competition encourages the generated samples to be hardly distinguished from real ones. Mathematically, G and D form a two-role minimax game with a value function V(G, D):

$$\min_{G} \max_{D} V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} \left[\log(D(x)) \right] \\ + \mathbb{E}_{z \sim p_z(z)}[(1 - log(D(G(z)))] \tag{3}$$

where $\mathbb{E}(\cdot)$ is the operator for expectation, p_{data} and p_z denote real and noise distributions respectively. Sampled noise is transformed by the generator G to mimic a real sample, its corresponding distribution is represented by p_G. For the parameters of G being fixed, the D can be trained to be an optimizer. In the regular GAN, z sampled from the predefined noise distribution is taken as the input of the G. For low-dose CT image noise reduction, G takes low-dose CT image as the input and output a denoised version. The minimization of the Jensen-Shannon (JS) divergence between p_{data} and p_G is equivalent to the minimization of G, which will easily give rise to a vanishing gradient with respect to the parameters of G, leading to mode collapse.

There are several problems in regular GAN, such as convergence, mode collapse, and poor of the generated image. Hence, we used the least squares GAN (LSGAN), which solves the vanishing gradient issue when fake samples are still far from real ones but on the correct side of the decision boundary. In LSGAN, D utilizes the *a-b* coding, where a, b

denote real and fake data respectively. The objective function for LSGAN can be described as follows:

$$\min_{G} \max_{D} V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} \left[(D(x) - b)^2 \right] \\ + \mathbb{E}_{z \sim p_z(z)}[(D(G(z)) - a)^2] \tag{4}$$

where a and b are included in the objective function to avoid gradient vanishing. A fake sample is far away from the real distribution but discriminated as real will be penalized. If the distribution of fake samples matches the real counterpart, the gradient becomes zero.

LSGAN offers two benefits. First, fake samples in the correct side of the decision boundary, but which a long way from will be penalized. The penalization will make generated samples move toward the correct decision boundary as closer as possible, reducing the distance between the generated and real data. Second, the mechanism of penalization results in increased gradients in the training process. This makes the LSGAN obtain a higher stability.

C. OBJECTIVE FUNCTION
1) ADVERSARIAL LOSS
Compared with the regular GAN, LSGAN adopts the least squares loss to minimize the divergence of Pearson χ^2 instead of the Jensen-Shannon (JS) to approximate the distribution between generated samples and real ones. In our task, the least square loss forces the distribution of denoised CT images to that of normal-dose CT images. To make the generator produce denoised CT images as realistic as possible, we use the objective functions of D and G respectively described as follows:

$$\min_{D} V_{LSGAN}(D) = \frac{1}{2} \mathbb{E}_{x \sim p_{data}(x)} \left[\left(D(x) - 1 \right)^2 \right] \\ + \frac{1}{2} \mathbb{E}_{z \sim p_z(z)}[(D(G(z)))^2] \tag{5}$$

$$\min_{G} V_{LSGAN}(D) = \frac{1}{2} \mathbb{E}_{z \sim p(z)} \left[\left(D(G(z)) - 1 \right)^2 \right] \tag{6}$$

where x and z denote a normal-dose CT image and its low-dose version respectively.

2) SIMILARITY LOSS FUNCTION
Feature correlations are strong in medical CT images with difference dose levels. The structural similarity index (SSIM) includes three human visual components, which are luminance, contrast, and structure. SSIM performs better than mean squared errors (MSE) and peak signal-to-noise ratio (PSNR) in visual tasks including diagnostic ones. For measuring the similarity between denoised CT image and normal-dose version, the classic SSIM can be used, which is defined as follows:

$$\text{SSIM}(x, y) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} * \frac{\sigma_{xz} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{7}$$

where μ_x, μ_y, σ_x, σ_y, and σ_{xy} denote the means, standard deviations and the cross-correlation of two images under

comparison respectively, and C_1 and C_2 are the constants to eliminate numerical singularity. When x and y are more similar, the value of SSIM is closer to one. Thus, the loss function for SSIM can be set to the following:

$$\text{Loss}_{\text{SSIM}} = 1 - \text{SSIM}(x, y) \qquad (8)$$

It is worth noting that the SSIM loss can be back-propagated to update the parameters of the neural networks, given its property of differentiability.

SSIM is employed to calculate the overall similarity between a normal-dose CT image and the denoised version.

3) L1 LOSS FUNCTION

The L_1 loss is also referred to as the least absolute error (LAE) [43]. Similar to the L_2 loss (mean-squared error, MSE) [44], the L_1 loss is a mean-based metric, but has a different effect on noise reduction. Larger errors between a denoised image and the ground truth are not over-penalized compared to the MSE loss. Some drawbacks, such as blurring and unnatural associated with the MSE loss can be alleviated by the L_1 loss.

In our low-dose CT image denoising task, the L_1 loss function is described as follows:

$$Loss_{L_1} = \frac{1}{mnb} |x - y| \qquad (9)$$

where x and y stand for a denoised CT image and the ground truth (normal-dose) respectively. m, n, b denote the width, height of the CT image, and the batch-size respectively.

4) OVERALL LOSS FUNCTION

Although minimizing the least squares loss can approximate the distribution of low-dose CT image to the normal-dose counterpart, yielding decent denoised image, it could not match well the corresponding normal-dose CT image in details. Often times, the L2 (MSE) loss is utilized in CNN-based methods, which attempts to minimize the error of per pixel between the predicted image and the target one. However, the L2 loss yields over-smoothened and blurry images, which results in distortion or loss of structural details. For restoration of high quality denoised CT images from low-dose CT, we combined different loss terms into a hybrid objective function for our task. The L1 loss helps further improve signal-noise-ratio (SNR).

For insuring the quality of restored CT image and keeping texture and structural details, our overall objective function incorporated the least squares loss, L1 loss, and structural loss. In summary, the overall objective function of our adapted LSGAN is formulated as follows:

$$Loss = \alpha Loss_{LSGAN} + \beta Loss_{SSIM} + \gamma Loss_{L_1} \qquad (10)$$

where α, β and γ are coefficients of the three loss terms. In the training stage, the total loss between denoised CT image and normal-dose one was calculated, which was then back-propagated for the neural network optimization. There are three hyper-parameters in the total loss function. To select the optimal parameters, first, we fixed β and γ, and select the
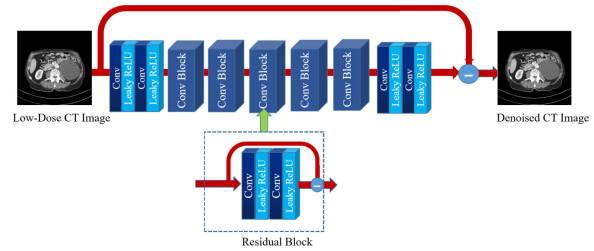

FIGURE 1. Architecture of the generator. It learns the noise, and then the original low-dose CT minus the noise is the denoised version.

optimal α. Then, we fixed α and γ, and determine the optimal β. Finally, we obtain the best value of γ based on the optimal α and β.

D. NETWORK ARCHITECTURE

1) GENERATOR NETWORK

The target of low-dose CT denoising is to generate a denoised CT image with improved quality, textural and structural details. The success of this denoising task depends on a proper network structure to produce the best denoised results.

Our deep generator G differs from traditional noise reduction models. The generator G is intended for residual learning [45] to obtain a mapping to a residual noise $N(x) = v$, where v stands for noise. Then, the denoised CT image is produced by $y' = x - v$. Our generator G was designed to learn the residual noise from the input low-dose CT image. A denoised CT image can be obtained by subtracting the learned noise from the input low-dose CT image. Fig. 1 illustrates the architecture of the generator G.

The generator G consists of two convolutional layers in the front and the rear respectively, and five residual convolutional blocks in the middle. The activation function for each convolutional layer is leaky rectified linear unit (Leaky ReLU) [46]. In the front, the first convolutional layer adopts 32 kernels of 7×7, and the second convolutional layer adopts 96 kernels of 5×5. In the rear, the first convolutional layer employs 96 kernels, and the second convolutional layer uses one kernel, all the kernels are of 3×3. In the middle, the first and second convolutional layers of each residual convolutional block all contains 96 kernels per layer, with the sizes of these kernels being set to 3×3 and 5×5 respectively.

2) DISCRIMINATOR NETWORK

The goal of the discriminator is to classify if an input is real or fake. Our discriminator D consists of six convolutional layers and three fully-connected layers, which use rectified linear unit (ReLU) [47]. In the front of the discriminator D, there are 64 kernels in two convolutional layers. The middle two convolutional layers have 128 kernels, and the rear two contains 256 kernels. The unit stride is used for all kernels being set to 3×3 in size. Three fully-connected layers are utilized after the six convolutional layers, of which the first layer has 1024 outputs, the second layer has 512 outputs, and the last layer holds a single output. Our implementation of the
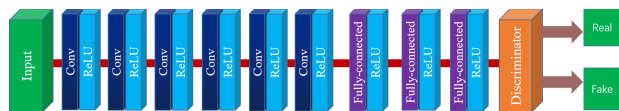
FIGURE 2. Architecture of the discriminator. It learns to classify if the input is a real normal-dose image or a generated one.

discriminator D differs from the original paper [41]. To avoid vanishing gradients and mode collapse, we utilize the least squares loss and ReLU instead of the Sigmoid function. The architecture of the discriminator is in Fig. 2.

Our adapted GAN with hybrid loss was trained with CT image patches, then applied on entire low-dose CT image. The details are described in Section III.

III. EXPERIMENTS AND RESULTS

A. DATASETS FOR EXPERIMENTS

To exhibit the capacity of our adapted GAN for low-dose CT image denoising, a real clinical CT image dataset was adopted in our study, which was authorized by Mayo Clinic for "2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge". The dataset contains 2, 378 normal CT images and the corresponding simulated low dose counterparts (quarter dose) from ten anonymous patients [48]. The matrix of each CT image is 512×512, with the thickness of 3.0 mm. In our experiments, we divided the dataset into two groups. One group contains 2, 167 image pairs from nine patients as the training set, and the other group contains 211 image pairs from the remaining one patient as the testing set.

In the training stage, image patches were extracted. The extracted patches of 55×55 were used. There were about 10^6 patch pairs extracted from 2, 167 image pairs as the training samples. Our patch based training captured local details, without a huge memory requirement. The deep learning approach needs a large number of training samples, but collecting medical images is constrained by complex formalities and extensive efforts.

B. PARAMETER SETTING

The generator and discriminator of our GAN both were optimized utilizing the adaptive momentum estimation (Adam) [49]. The size of mini-batch was 96. The original learning rate was set to 1×10^{-5}, then gradually decreased to 1×10^{-6} in the training process. On the basis of our experimental experience, the coefficients of overall objective function were set as $\alpha = 0.005$, $\beta = 0.995$, and $\gamma = 0.95$ respectively.

Our method were programed in Python with the platform of Pytorch. All experiments were implemented on a personal computer (Intel i7 9700 with 32 G random memory) and accelerated by a NVIDIA RTX 2080 TI GPU with 11 G memory.

Our low-dose CT denoising method was compared against five state-of-the-art different methods, including NLM, BM3D, RED-CNN, WGAN, and SMGAN. NLM and BM3D are two of the most popular traditional approaches which
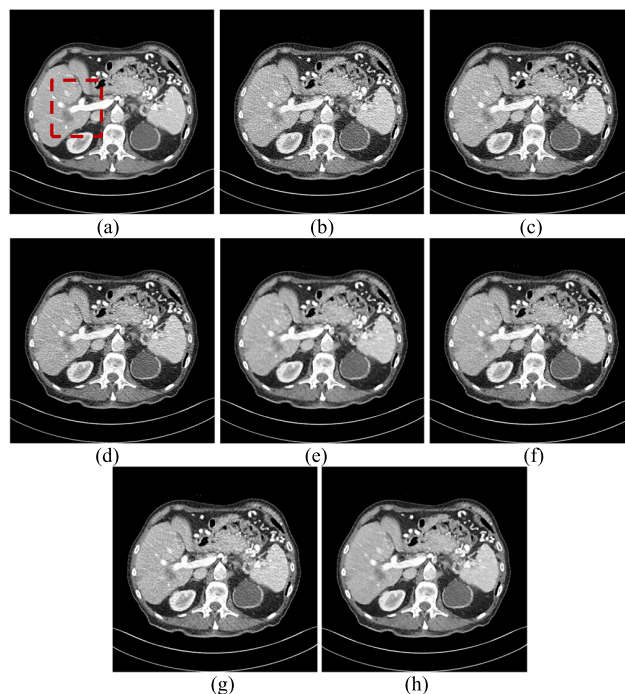

FIGURE 3. Results of an abdominal CT scan from the testing set using different methods. (a) Normal dose, (b) Low dose, (c) NLM, (d) BM3D, (e) RED-CNN, (f) WGAN-VGG, (g) SMGAN, (h) Proposed.

already utilized for low-dose CT image denoising. RED-CNN, WGAN-VGG and SM-GAN are three representative deep learning methods for low-dose CT. RED-CNN is an auto encoder-decoder convolutional neural network with mean squared error. WGAN-VGG and SM-GAN are working based on a GAN, and employing Wasserstein distance and sharing similar network architecture, but different loss function.

C. EXPERIMENTAL RESULTS

1) VISUAL EFFECT

For evaluating our method, two typical low-dose CT images from the testing set and their denoised results utilizing different methods are shown in Figs. 3 and 5 respectively. Their regions-of-interest (ROIs) are zoomed in Figs. 4 and 6.

Since insufficient photons of incident X-ray, the degradation is severe in the low-dose CT images, which are shown in Figs. 3(b) and 5(b). It is hard to see that the structure and details in the two low-dose CT images. From Figures 3 and 5, it can be seen that all methods suppressed noise to some degree. In Figs. 3(c) and 5(c), there were still some noise and artifacts in the whole denoised images. From Figs. 3(d) and 5(d), BM3D was a bit better than NLM method, but it suffered from significant blocky effects, some edges and small structures were blurred. It can be seen that deep learning methods effectively reduced noise and remarkably overmatched NLM and BM3D, they improved the effect of noise reduction and suppressed most artifacts. But in Figs. 3(e)−3(h) and 5(e)−5(h), there were some differences among these
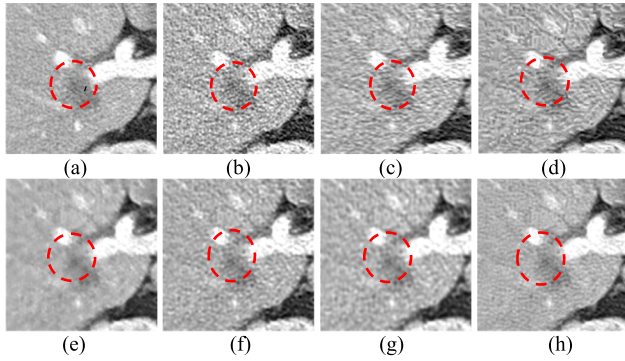
FIGURE 4. Zoomed red circle area as the region of interest (ROI) in Fig. 3. (a) Normal dose, (b) Low dose, (c) NLM, (d) BM3D, (e) RED-CNN, (f) WGAN-VGG, (g) SMGAN, (h) Proposed.
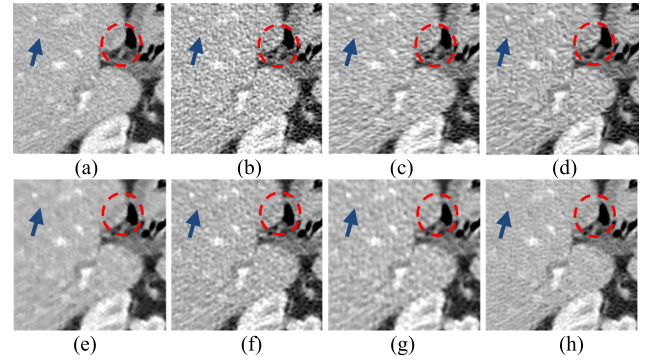

FIGURE 6. Zoomed red circle area as the region of interest (ROI) in Fig. 5. (a)-(h) correspond to those in Fig. 4.
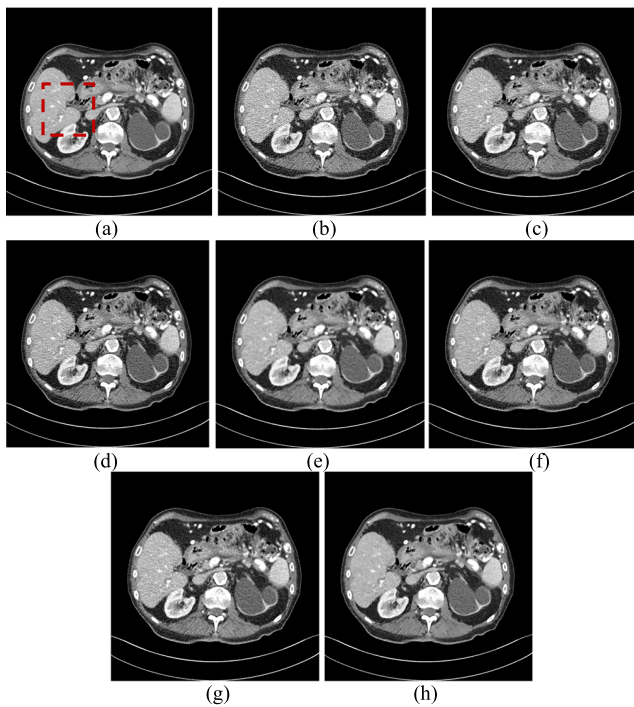

FIGURE 5. Results of another abdominal CT scan from the testing set using different methods. (a)-(h) correspond to the labels in Fig. 3.
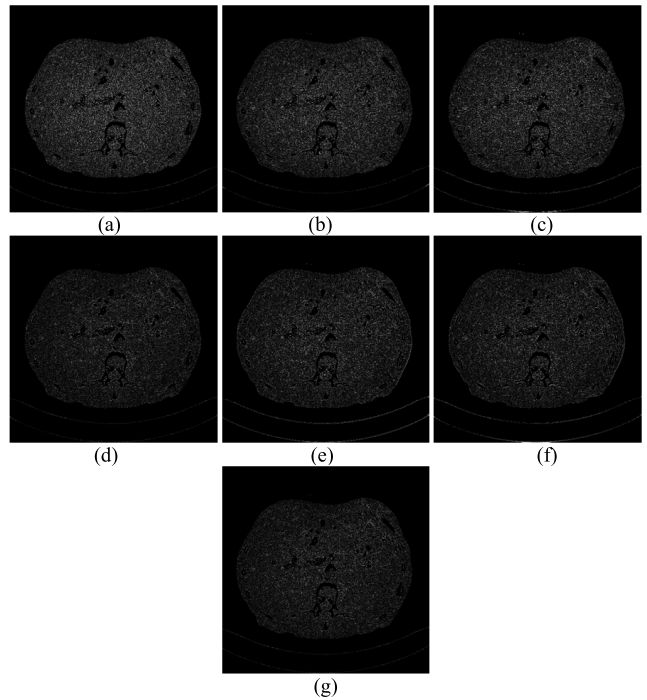

FIGURE 7. Absolute difference image relative to the corresponding normal-dose CT image in Fig. 3. (a) low-dose, (b) NML, (c) BM3D, (d) RED-CNN, (e) WGAN-VGG, (f) SMGAN, (g) Proposed.

networks for noise reduction. RED-CNN smoothened the result images excessively, some crucial structures, such as the region of porta were blurred. GAN based methods not only reduced most noise and artifacts, but also preserved structural details. From Figs. 3(h) and 5(h), it can be seen that our proposed method produced better content details and textural information thereby improving diagnostic performance compared to the other methods in our study.

In Figs. 6(b) and 8(b), the metastases enclosed by the red dashed line circles were hard to be observed. The results restored by our adapted GAN was the most distinct compared with all the other methods shown in Figs. 6(h) and 8(h). The white spot which indicated by a blue arrow restored by our

method was not only clear but also lest surrounding disturbed compared with the other methods selected in our study.

To visualize the benefits of our proposed method, the absolute difference images of the results using different methods to the ground truth (normal-dose CT images) are provided in Figs. 7 and 8, and the absolute difference images of ROIs are shown in Figs. 9 and 10, respectively. From Figs. 7 to 10, it can be seen that our method could do better in suppressing artifacts and noise. The denoised image using our method was closer to the normal-dose image.

2) QUANTITATIVE MEASUREMENTS
The PSNR, SSIM, and RMSE were calculated for quantitative analysis. The quantitative results for the whole CT
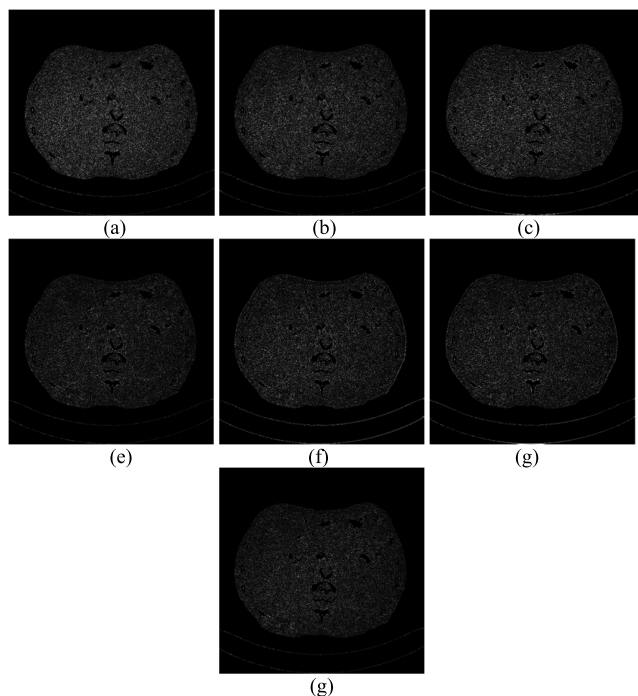
FIGURE 8. Absolute difference image relative to the corresponding normal-dose CT image for Fig. 5. (a)-(g) correspond to those in Fig. 7.
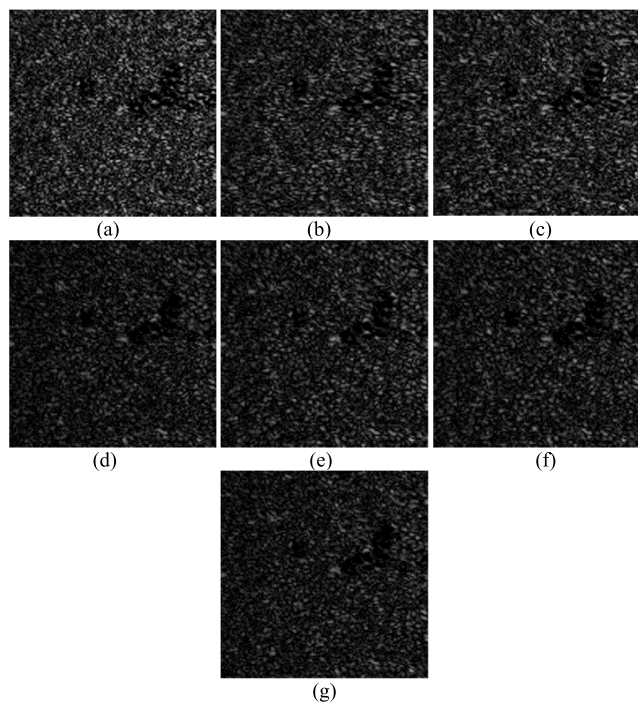

FIGURE 9. Absolute difference ROI image relative to the corresponding normal-dose ground truth in Fig. 4. (a) Low-dose, (b) NML, (c) BM3D, (d) RED-CNN, (e) WGAN-VGG, (f) SMGAN, (g) Proposed.

FIGURE 10. Absolute difference ROI image relative to the corresponding normal-dose ground truth in Fig. 6. (a)-(g) correspond to those in Fig. 9.

TABLE 1. Quantitative results for Fig. 3 utilizing different methods.

	PSNR	SSIM	RMSE
Low Dose	24.4411	0.7882	23.9885
NML	25.8902	0.8073	20.3024
BM3D	25.4434	0.7937	21.3743
RED-CNN	**28.7083**	0.8431	**14.6772**
WGAN-VGG	27.2029	0.8278	17.4548
SMGAN	27.5667	0.8333	16.7378
Proposed	28.4941	**0.8454**	15.0438

TABLE 2. Quantitative results for Fig. 5 utilizing different methods.

	PSNR	SSIM	RMSE
Low Dose	25.4773	0.8075	21.2910
NML	27.0287	0.8263	17.8085
BM3D	26.3803	0.8070	19.1886
RED-CNN	**29.6560**	0.8580	**13.1601**
WGAN-VGG	28.1414	0.8428	15.6671
SMGAN	28.4393	0.8494	15.1389
Proposed	29.4084	**0.8596**	13.5407

images in Figs. 3 and 5 utilizing different methods are listed in Tables 1 and II respectively. The average values for the testing set (211 CT images) are listed in Table 3.
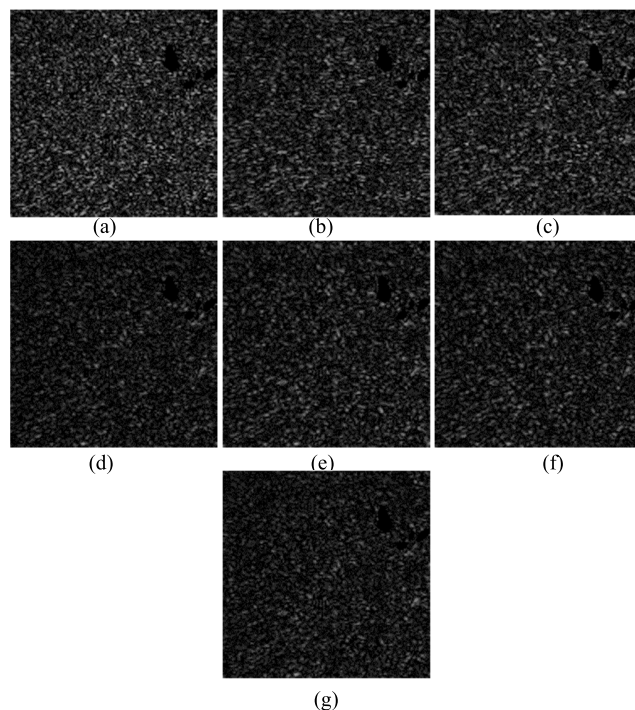
In Tables 1 and 2, our adapted GAN obtained the best scores in aspects of SIMM, but PSNR and RMSE were less than RED-CNN. It is not surprising that RED-CNN got the first place in terms of PSNR and RMSE. The reason is that PSNR and RMSE are equivalent to the loss of
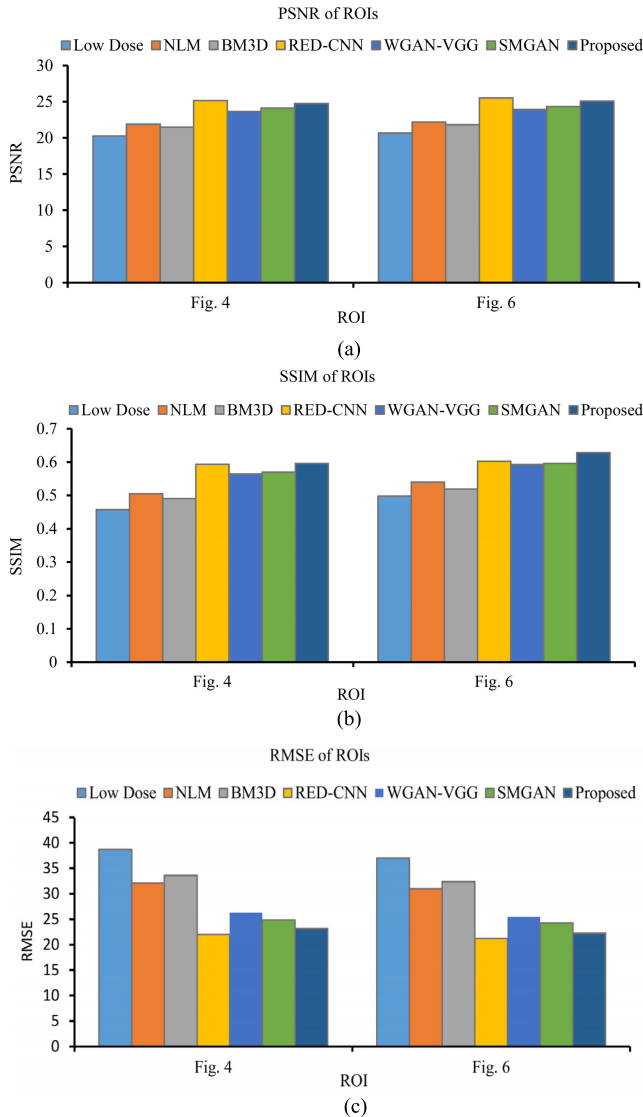
FIGURE 11. Performance comparison of low-dose CT and different methods over the ROIs images in Figs. 4 and 6. (a) The comparison of PSNR, (b) The comparison of SSIM, and (c) the comparison of RMSE.

TABLE 3. Average quantitative results for testing set utilizing different methods.

	PSNR	SSIM	RMSE
Low Dose	29.2489	0.8759	14.2416
NML	30.3964	0.8866	12.4114
BM3D	29.3693	0.8662	13.8754
RED-CNN	32.6261	0.9071	9.5349
WGAN-VGG	31.1007	0.8996	11.3596
SMGAN	31.1568	0.9027	11.2539
Proposed	**32.7052**	**0.9108**	**9.4868**

TABLE 4. The values textural statistics of Fig. 4 for different denoised methods.

	Mean (HU)	Std	Uniformity	Entropy
Full Dose	107.5181	67.9015	7.5429	0.0084
Low Dose	106.3259	74.5352	7.9390	0.0060
NML	106.6910	70.2706	3.9677	0.0839
BM3D	98.8041	70.8735	3.9965	0.0819
RED-CNN	108.1902	64.8363	7.4664	0.0092
WGAN-VGG	109.8331	**68.4109**	7.6990	0.0079
SMGAN	101.6183	65.3835	7.7630	0.0064
Proposed	**107.0674**	65.9830	**7.5856**	**0.0080**

measurements by calculating the statistical properties, containing mean CT number (Housnfield Unit), standard deviation (Std), uniformity, and entropy of two ROIs in Figs. 4 and 6. Ideally, denoising methods should generate latent clean CT image, in which the statistical properties are as close as to the corresponding normal-dose version. The textural statistics of Figs. 4 and 6 are shown in Tables 4 and 5 respectively.

From Table 4, it can be noted that our method obtained the best values except for standard deviation. While BM3D got highest standard deviation, which are consistent to visual appearance of Fig. 4. In Table 5, our adapted GAN obtained the best values except for the entropy. That RED-CNN generated over-smoothed images and BM3D had poor capacity of noise reduction for low-dose CT. It is important to point out that the statistical values of ROIs produced by our method were nearest to that of the normal-dose version and acquired the best matching textural statistics to normal-dose CT image in comparison of all other approaches selected in this study.

IV. DISCUSSIONS

By training a pair of competitive networks iteratively, GAN learns a deep representation, and achieves impressive successes in a variety of applications. GAN also provides an opportunity for medical imaging tasks, such as denoising, deblurring, reconstruction, detection, segmentation, and so on [50], [51].

The main goal of our study is to investigate low-dose CT deoising, approaching the gold standard normal-dose CT images as close as possible. We adapted GAN to tackle the low-dose CT problem, and compared it with the state-of-the-art traditional and deep learning methods. Because of Poisson

per-pixel. RED-CNN was trained to minimize the MSE loss, and excessed to WGAN-VGG, SMGAN and our adapted GAN, which trained to minimize other losses. The values of SSIM in Tables 1 and 2 are consistent to Figs. 3 and 5 respectively. The average measurements in Table 3 indicates that our adapted GAN attained the best quantitative results in the entire testing data sets.

The quantitative measurements of two ROIs in Figs. 4 and 6 are shown in Fig. 11. The comparative results of bar graphs followed similar trends toward the visual effect of Figs. 4 and 6, as well as Tables 1, 2, and 3. Although there were a bit of poor PSRN and RMSE than RED-CNN, our method gained the best scores in respects of SSIM for all the ROIs.

3) TEXTURAL MEASUREMENTS

For attaining further perspectives of our adapted GAN for noise reduction for low-dose CT, we conducted the textural

TABLE 5. The values textural statistics of Fig. 6 for different denoised methods.

	Mean (HU)	Std	Uniformity	Entropy
Full Dose	96.4612	76.6386	7.7213	0.0070
Low Dose	95.7362	83.7266	8.0625	0.0048
NML	95.8261	79.8805	4.0788	0.0790
BM3D	87.6592	80.0537	4.0888	0.0782
RED-CNN	97.1172	75.0392	7.5335	0.0089
WGAN-VGG	97.7931	79.8322	7.7914	**0.0066**
SMGAN	90.9655	74.4002	7.8016	0.0063
Proposed	**96.0443**	75.9630	**7.6637**	0.0075

noise in low-dose CT, it is hard to be modeled precisely in the image domain, while NLM [19] and BM3D [20] are the typical ways appropriate for specific types of noise different from CT data noise. Hence, NLM [19] and BM3D [20] cannot achieve a satisfactory performance for low-dose CT denoising. In contrast to these classic methods, deep learning methods, especially convolutional neural networks are powerful for knowledge representation, and fit various complicated functions with an excellent generalizability. Once deep learning applies for image processing, high level abstract features can be captured from training data, enabling the state-of-the-art low-dose CT performance. Which is observed visually and quantitatively in our experiments. It is not surprising that RED-CNN [28], WGAN-VGG [35], SMGAN [38], and our adapted GAN outperformed NLM [19] and BM3D [20] in suppressing noise and removing artifacts significantly, and our GAN performed favorably relative to the other deep learning methods.

To a great extent, the loss function of deep learning influences the low-dose image restoration process. RED-CNN [28] with the MSE loss produced high values of PSNR and RMSE, but was featured by over-smoothened images. The background noise can be indeed suppressed by the MSE loss, but the denoised image is blurry and unnatural [36], [52]. Furthermore, the background noise is assumed as white Gaussian noise in general, independent of the local image features. Clearly, this is not suitable for low-dose CT image denoising, since the statistical fluctuations in a low-dose CT image reconstructed with FBP is correlated as determined by specific structures in the field of view.

WGAN-VGG [35] and SMGAN [38] both utilized the Wasserstein distance to learn the distribution of real data and the regularization term of gradient penalty to stabilize the training process and accelerate the convergence. While the WGAN-VGG [35] adopted the VGG loss for perceptual fidelity, mimicking the human visual system (HVS) at a cost of lower PSNR, SSIM, and RMSE than that of RED-CNN. The SMGAN [38] addressed low-dose CT denoising with the multi-scale structural similarity index (MS-SSIM), and generated CT images better than that of WGAN-VGG. Our adapted GAN utilized the least squares loss to decrease the distance between distributions of denoised CT image and standard normal-dose CT image, which relieves vanishing

gradients, penalizes data samples near the decision boundary, and facilitates contents learning with the structural similarity index (SSIM). Compared with WGAN-VGG [35], SMGAN [38], our adapted GAN was optimized for low-dose CT denoising, and delivered better results at a lower computational cost.

V. CONCLUSION

In conclusion, we have proposed a novel method for low-dose CT denoising, which adopted noise learning and enhanced a GAN with least squares loss, structural similarity and L1 losses. The generator learned noise distributed in the low-dose CT image and then subtracted it from the input contaminated low-dose CT image to obtain the clean one. After training offline with the pairs of low-dose and normal-dose CT images, our method could improve noise reduced CT images competitively relative to the state-of-the-art methods selected in this study. The results are promising in terms of both visual effects and quantitative measurements pertaining to suppressing noise and removing artifacts. The comparison of the textural statistics of the resultant images between our method and the other methods has further confirmed the effectiveness of our network design. In the future, deep learning methods for low-dose CT will be further advanced by taking additional priors and features into account in the low-dose sinogram domain.

REFERENCES

[1] M. M. Lell, J. E. Wildberger, H. Alkadhi, J. Damilakis, and M. Kachelriess, "Evolution in computed tomography: The battle for speed and dose," *Investigative Radiol.*, vol. 50, no. 9, pp. 629–644, Sep. 2015.

[2] G. Wang, H. Yu, and B. De Man, "An outlook on X-ray CT research and development," *Med. Phys.*, vol. 35, no. 3, pp. 1051–1064, Feb. 2008.

[3] C. M. Micallef, K. D. Shield, J. Vignat, E. Cléro, A. Kesminiene, C. Hill, A. Rogel, B. Vacquier, F. Bray, D. Laurier, and I. Soerjomataram, "The risk of cancer attributable to diagnostic medical radiation: Estimation for France in 2015," *Int. J. Cancer*, vol. 144, no. 12, pp. 2954–2963, Jun. 2019.

[4] D. J. Brenner and E. J. Hall, "Computed tomography—An increasing source of radiation exposure," *New England J. Med.*, vol. 357, no. 22, pp. 2277–2284, Nov. 2007.

[5] S. A. Kovalchik, M. Tammemagi, C. D. Berg, N. E. Caporaso, T. L. Riley, M. Korch, G. A. Silvestri, A. K. Chaturvedi, and H. A. Katki, "Targeting of low-dose CT screening according to the risk of lung-cancer death," *New England J. Med.*, vol. 369, no. 3, pp. 245–254, Jul. 2013.

[6] M. Diwakar and M. Kumar, "A review on CT image noise and its denoising," *Biomed. Signal Process. Control*, vol. 42, pp. 73–88, Apr. 2018.

[7] Y. Sagara, A. K. Hara, W. Pavlicek, A. C. Silva, R. G. Paden, and Q. Wu, "Abdominal CT: Comparison of low-dose CT with adaptive statistical iterative reconstruction and routine-dose CT with filtered back projection in 53 patients," *Amer. J. Roentgenol.*, vol. 195, no. 3, pp. 713–719, Sep. 2010.

[8] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," *Int. J. Comput. Vis.*, vol. 81, no. 1, pp. 24–52, Jan. 2009.

[9] M. Balda, J. Hornegger, and B. Heismann, "Ray contribution masks for structure adaptive sinogram filtering," *IEEE Trans. Med. Imag.*, vol. 31, no. 6, pp. 1228–1239, Jun. 2012.

[10] L. Ouyang, T. Solberg, and J. Wang, "Effects of the penalty on the penalized weighted least-squares image reconstruction for low-dose CBCT," *Phys. Med. Biol.*, vol. 56, no. 17, pp. 5535–5552, Sep. 2011.

[11] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[12] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1682–1697, Sep. 2012.

[13] Y. Chen, D. Gao, C. Nie, L. Luo, W. Chen, X. Yin, and Y. Lin, "Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior," *Computerized Med. Imag. Graph.*, vol. 33, no. 7, pp. 495–500, Oct. 2009.

[14] Y. Zhang, Y. Xi, Q. Yang, W. Cong, J. Zhou, and G. Wang, "Spectral CT reconstruction with image sparsity and spectral mean," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 510–523, Dec. 2016.

[15] J.-F. Cai, X. Jia, H. Gao, S. B. Jiang, Z. Shen, and H. Zhao, "Cine cone beam CT reconstruction using low-rank matrix factorization: Algorithm and a Proof-of-Principle study," *IEEE Trans. Med. Imag.*, vol. 33, no. 8, pp. 1581–1591, Aug. 2014.

[16] 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, Sep. 2008.

[17] Y. Liu, J. Ma, Y. Fan, and Z. Liang, "Adaptive-weighted total variation minimization for sparse data toward low-dose X-ray computed tomography image reconstruction," *Phys. Med. Biol.*, vol. 57, no. 23, pp. 7923–7956, Dec. 2012.

[18] Z. Tian, X. Jia, K. Yuan, T. Pan, and S. B. Jiang, "Low-dose CT reconstruction via edge-preserving total variation regularization," *Phys. Med. Biol.*, vol. 56, no. 18, pp. 5949–5967, Sep. 2011.

[19] Z. Li, L. Yu, J. D. Trzasko, D. S. Lake, D. J. Blezek, J. G. Fletcher, C. H. McCollough, and A. Manduca, "Adaptive nonlocal means filtering based on local noise level for CT denoising," *Med. Phys.*, vol. 41, no. 1, Dec. 2013, Art. no. 011908.

[20] P. F. Feruglio, C. Vinegoni, J. Gros, A. Sbarbati, and R. Weissleder, "Block matching 3D random noise filtering for absorption optical projection tomography," *Phys. Med. Biol.*, vol. 55, no. 18, pp. 5401–5415, Sep. 2010.

[21] D. Kang, P. Slomka, R. Nakazato, J. Woo, D. S. Berman, C.-C.-J. Kuo, and D. Dey, "Image denoising of low-radiation dose coronary CT angiography by an adaptive block-matching 3D algorithm," in *Proc. Med. Imag., Image Process.*, Mar. 2013, Art. no. 86692.

[22] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[23] T. Ching, "Opportunities and obstacles for deep learning in biology and medicine," *J. Roy. Soc. Interface*, vol. 15, no. 141, p. 47, Apr. 2018.

[24] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[25] G. Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016.

[26] H. Shan, A. Padole, F. Homayounieh, U. Kruger, R. D. Khera, C. Nitiwarangkul, M. K. Kalra, and G. Wang, "Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction," *Nature Mach. Intell.*, vol. 1, no. 6, pp. 269–276, Jun. 2019.

[27] T. Würfl, F. C. Ghesu, V. Christlein, and A. Maier, *Deep Learning Computed Tomography*. Cham, Switzerland: Springer, Oct. 2016, pp. 432–440.

[28] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, 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, Dec. 2017.

[29] E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *Med. Phys.*, vol. 44, no. 10, pp. e360–e375, Oct. 2017.

[30] 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, Dec. 2017.

[31] I. Goodfellow, "NIPS 2016 tutorial: Generative adversarial networks," 2017, *arXiv:1701.00160*. [Online]. Available: http://arxiv.org/abs/1701.00160

[32] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*. [Online]. Available: http://arxiv.org/abs/1511.06434

[33] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, *arXiv:1701.07875*. [Online]. Available: http://arxiv.org/abs/1701.07875

[34] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of wasserstein GANs," 2017, *arXiv:1704.00028*. [Online]. Available: http://arxiv.org/abs/1704.00028

[35] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, 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, Jun. 2018.

[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556

[38] C. You, W. Cong, G. Wang, Q. Yang, H. Shan, L. Gjesteby, G. Li, S. Ju, Z. Zhang, Z. Zhao, and Y. Zhang, "Structurally-sensitive multi-scale deep neural network for low-dose CT denoising," *IEEE Access*, vol. 6, pp. 41839–41855, Jul. 2018.

[39] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2003, pp. 1398–1402.

[40] D. Bau, J.-Y. Zhu, J. Wulff, W. Peebles, B. Zhou, H. Strobelt, and A. Torralba, "Seeing what a GAN cannot generate," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 4502–4511.

[41] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks," 2016, *arXiv:1611.04076*. [Online]. Available: http://arxiv.org/abs/1611.04076

[42] D. Zeng, J. Huang, Z. Bian, S. Niu, H. Zhang, Q. Feng, Z. Liang, and J. Ma, "A simple low-dose X-ray CT simulation from high-dose scan," *IEEE Trans. Nucl. Sci.*, vol. 62, no. 5, pp. 2226–2233, Oct. 2015.

[43] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 47–57, Mar. 2017.

[44] C. Zhang, H. Zhang, and X. Wei, "Deep bimodal regression for apparent personality analysis," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 311–324.

[45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[46] A. Maas, A. Hannun, and A. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. 30th. Int. Conf. Mach. Lear.*, vol. 30, no. 1, 2013, pp. 1–6.

[47] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th. Int. Conf. Mach. Lear.*, 2010, pp. 807–814.

[48] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang, "Low-dose CT via convolutional neural network," *Biomed. Opt. Expr.*, vol. 8, no. 2, pp. 679–694, Feb. 2017.

[49] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980

[50] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 53–65, Jan. 2018.

[51] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of wasserstein GANs," 2017, *arXiv:1704.00028*. [Online]. Available: http://arxiv.org/abs/1704.00028

[52] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

YINJIN MA received the M.E. degree in instrumentation engineering from Chongqing University, Chongqing, China, in 2015, where he is currently pursuing the Ph.D. degree in optical engineering. His research interests include the applications of deep learning in medical CT images and image processing.

BIAO WEI received the B.S. and Ph.D. degrees from the College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, Chengdu, China, in 1988 and 1996, respectively. He held a Postdoctoral position at Chongqing University, Chongqing, China, from October 1996 to May 1999. He was a Lecturer with Chongqing University, from June 1999 to November 2004, where he is currently a Professor and the Vice Dean of the College of Optoelectronics Engineering. He has published more than 60 articles in his research areas. His research interests include optoelectronic imaging and light energy detection of X-rays, neutrons and visible light, high-resolution optoelectronic imaging detection (sensing) technology with a scientific-grade CCD, EMCCD, and ICCD digital cameras, and optoelectronics image compression coding technology.

PENG FENG received the B.S. degree in mechanical and electronics engineering, and the Ph.D. degree in optics engineering from Chongqing University, Chongqing, China, in 2002 and 2007, respectively. From June 2008 to January 2012, he was a Postdoctoral Fellow of the School of Biomedical Engineering, Chongqing University, where he is currently an Associate Professor with the Department of Optoelectronics Engineering. He has published more than 30 peer-reviewed journal articles as a corresponding author. His interests include computed tomography, compressive sensing, image representation, and biomedical image processing.

PENG HE received the Ph.D. degree in optical engineering from Chongqing University, Chongqing, China, in 2013. He was a Lecturer of instrument science and technology with Chongqing University, from July 2015 to December 2016, where he is currently an Associate Professor with the Department of Optoelectronics Engineering. He has published more than 30 articles in his research areas. His research interests include X-ray spectral CT imaging, digital image processing, and big data artificial intelligence.

XIAODONG GUO received the B.S. degree from the University of Electronic Science and Technology of China, in 2014, and the M.S. degree from Southwest Jiaotong University, in 2018. He is currently pursuing the Ph.D. degree in optical engineering with Chongqing University, under the supervision of Prof. X. Lei. His research interests include the applications of deep learning in medical CT images and image processing.

GE WANG (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from The State University of New York, University at Buffalo, Buffalo, NY, USA. He is the Clark and Crossan Chair Professor and also the Director of the Biomedical Imaging Center, Rensselaer Polytechnic Institute, USA. He has authored the first article on the spiral/helical cone-beam/multi-slice CT algorithm, and over 430 peer-reviewed journal publications. Currently, there are over 100 million medical CT scans yearly with a majority in the spiral cone-beam mode. He has pioneered bioluminescence tomography. His group has published the first article on interior tomography and omni-tomography (all-in-one) to acquire diverse data sets simultaneously (all-at-once). His results were featured in *Nature*, *Science*, and PNAS, and are recognized with awards. He is a Fellow of the SPIE, OSA, AIMBE, AAPM, and AAAS.

• • •