Weakly supervised low-dose computed tomography denoising based on generative adversarial networks

Peixi Liao[1], Xucan Zhang[2], Yaoyao Wu[3], Hu Chen[4]^, Wenchao Du[4], Hong Liu[4], Hongyu Yang[4], Yi Zhang[5]

[1]Department of Stomatology, The Sixth People's Hospital of Chengdu, Chengdu, China; [2]The National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu, China; [3]The School of Computer Science, Sichuan University, Chengdu, China; [4]The College of Computer Science, Sichuan University, Chengdu, China; [5]The School of Cyber Science and Engineering, Sichuan University, Chengdu, China

Contributions: (I) Conception and design: P Liao, X Zhang, Y Wu; (II) Administrative support: H Chen; (III) Provision of study materials or patients: P Liao, H Chen; (IV) Collection and assembly of data: P Liao, H Liu, H Yang, Y Zhang; (V) Data analysis and interpretation: X Zhang, Y Wu, W Du; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Hu Chen, PhD. The College of Computer Science, Sichuan University, First Ring Road South, Wangjiang Road Street, Wuhou District, Chengdu 610065, China. Email: huchen@scu.edu.cn.

Background: Low-dose computed tomography (LDCT) is a diagnostic imaging technique designed to minimize radiation exposure to the patient. However, this reduction in radiation may compromise computed tomography (CT) image quality, adversely impacting clinical diagnoses. Various advanced LDCT methods have emerged to mitigate this challenge, relying on well-matched LDCT and normal-dose CT (NDCT) image pairs for training. Nevertheless, these methods often face difficulties in distinguishing image details from nonuniformly distributed noise, limiting their denoising efficacy. Additionally, acquiring suitably paired datasets in the medical domain poses challenges, further constraining their applicability. Hence, the objective of this study was to develop an innovative denoising framework for LDCT images employing unpaired data.

Methods: In this paper, we propose a LDCT denoising network (DNCNN) that alleviates the need for aligning LDCT and NDCT images. Our approach employs generative adversarial networks (GANs) to learn and model the noise present in LDCT images, establishing a mapping from the pseudo-LDCT to the actual NDCT domain without the need for paired CT images.

Results: Within the domain of weakly supervised methods, our proposed model exhibited superior objective metrics on the simulated dataset when compared to CycleGAN and selective kernel-based cycle-consistent GAN (SKFCycleGAN): the peak signal-to-noise ratio (PSNR) was 43.9441, the structural similarity index measure (SSIM) was 0.9660, and the visual information fidelity (VIF) was 0.7707. In the clinical dataset, we conducted a visual effect analysis by observing various tissues through different observation windows. Our proposed method achieved a no-reference structural sharpness (NRSS) value of 0.6171, which was closest to that of the NDCT images (NRSS =0.6049), demonstrating its superiority over other denoising techniques in preserving details, maintaining structural integrity, and enhancing edge contrast.

Conclusions: Through extensive experiments on both simulated and clinical datasets, we demonstrated the superior efficacy of our proposed method in terms of denoising quality and quantity. Our method exhibits superiority over both supervised techniques, including block-matching and 3D filtering (BM3D), residual encoder-decoder convolutional neural network (RED-CNN), and Wasserstein generative adversarial network-VGG (WGAN-VGG), and over weakly supervised approaches, including CycleGAN and SKFCycleGAN.

---

^ ORCID: 0000-0001-9300-6572.

Introduction

Low-dose computed tomography (LDCT) has clinical value due to its ability to mitigate risks related to radiation exposure during patient screening and diagnosis (1). Excessive ionizing radiation from X-rays may be potentially harmful, necessitating their strategic implementation. However, the reduction of X-ray radiation dose through adjustments in tube current (2) introduces unwanted noise and streaks, adversely impacting clinical diagnostics and compromising the quality of reconstructed CT images. Consequently, LDCT denoising has emerged as a pivotal research focus in medical imaging. However, constructing a statistical model for removing noise and artifact from LDCT images is challenging when their distribution patterns resemble image details of normal human tissues and low-density lesions. Hence, effective CT denoising methods are indispensable in clinical practice. Various studies have proposed strategies to enhance LDCT image quality, and these include sinogram filtering, iterative reconstruction, and image postprocessing algorithms (3-6). Despite their capacity to enhance CT image quality to a certain degree, these methods often exhibit slow computational convergence and yield oversmoothed images.

In recent years, supervised deep learning reconstruction have enabled low-dose imaging while preserving image quality (7). The residual encoder-decoder convolutional neural network (RED-CNN) (8), employing a CNN and residual learning, has demonstrated remarkable LDCT denoising outcomes. Similar approaches, such as the structurally sensitive multiscale deep neural network (SMGAN) (9) and the wavelet residual network (WavResNet) (10), also use neural networks for LDCT denoising. However, a limitation of these methods is their dependence on paired LDCT and normal-dose CT (NDCT) data for training, as these may be challenging to obtain in real medical settings. The difficulty arises from factors such as voluntary breathing and scan position variations during consecutive scans, hindering consistent pixel-to-pixel correspondence. The shortage of well-matched data poses a significant challenge in the LDCT denoising field. To address this issue, unsupervised GAN variations have been proposed. Examples of these include CycleGAN (11)—which employs cycle consistency for unsupervised learning—selective kernel-based cycle-consistent GAN (SKFCycleGAN) (12), unsupervised dual learning for image-to-image translation (DualGAN) (13), conditional GAN (CGAN) (14), and deep convolutional GAN (DCGAN) (15), which utilize unpaired data for model training. Although CycleGAN and SKFCycleGAN have been applied to LDCT denoising, challenges persist. The complex network model and numerous hyperparameters of CycleGAN pose training difficulties, while SKFCycleGAN, based on CycleGAN, may introduce blurring in the generated noise due to network structure defects, resulting in new artifacts. Moreover, the SKFCycleGAN model has issues with training stability. Despite addressing LDCT denoising to some extent, these unsupervised methods are not yet on par with existing supervised learning approaches.

Deep learning denoising methods effectively bypass the uncertainty of noise distribution, enabling the learning of high-level features and representations from local image patches. Consequently, numerous CNN-based LDCT denoising methods have been proposed, aiming to identify the pixel-level relationships between LDCT and corresponding NDCT images (16). Early adopters, such as Chen *et al.* (17), pioneered the integration of CNNS for LDCT denoising, presenting a model that not only reduced computational overhead but also surpassed previous methodologies. In a subsequent work, they introduced RED-CNN, a conventional LDCT image denoising network combining autoencoder and residual learning strategies (8), further enhancing the denoising performance. Seeking more effective image detail recovery, Wolterink *et al.* (18) incorporated the use of GANs. Their approach involved generating synthetic NDCT images through generators and training discriminators to distinguish between authentic and synthetic NDCT images. The Wasserstein GAN (WGAN) (19), proposed by Yang *et al.* (20), and the perceptual loss-based LDCT

denoising algorithm WGAN-VGG proposed by Kim *et al.* (21), employ the VGG19 (22) network to extract CT image features, enhancing perceptual loss through an improved generative adversarial loss function from WGAN (23). This approach significantly advanced the representation of details in denoised CT images. Singh *et al.* (24) proposed a CNN denoising method based on noise, employing a three-layer CNN network to handle residual components between noisy CT images and denoised images. Yang *et al.* (25) proposed an LDCT denoising network that employs hierarchical feature refinement and multiscale dynamic convolution to enhance denoising performance, with the purpose of fully exploiting hierarchical features. Yan *et al.* (26) introduced transfer learning densely connected convolutional dictionary learning (TLD-CDL), a convolutional denoising network that enhances feature extraction by integrating multiscale inception modules and dense connections.

As supervised learning algorithms require strictly aligned CT images which are difficult to obtain, unpaired image denoising methods (27-29) have emerged as a promising alternative. Weakly supervised learning is a comprehensive term encompassing various approaches to constructing predictive models under weak supervision. Three principal categories are present in weakly supervised learning: incomplete supervision, wherein merely a subset of training data is provided with labels; inexact supervision, where the training data are furnished solely with coarse-grained labels; and inaccurate supervision, wherein the provided labels may not consistently represent the ground truth (30). Kang *et al.* (31) proposed a weakly supervised low-dose CT image denoising model based on the CycleGAN framework. This method does not require one-to-one training data and relies on cycle loss for training on unpaired datasets, which enables the model to learn the mapping from LDCT to NDCT images. To progressively improve the denoising effect, the SKFCycleGAN (12) was proposed, which injects a two-sided network into selective kernel network (SK-NET) to adaptively select features and uses the patchGAN discriminator to generate CT images with more detail maintenance, which is aided by added perceptual loss. However, this approach struggles to preserve the original information of the image while separating noise, and it has a large network model, which makes training challenging. To address these issues, the unpaired image denoising network (UIDNet) (32) was proposed as an end-to-end unpaired image denoising framework that uses CGAN to learn the noise distribution and generate clean pseud noise pairs

for training. WGAN gradient penalty (WGAN-GP) (33) loss, a modified version of WGAN with a gradient penalty, was used to ensure training stability. Moreover, an image-sharpening technique was employed to better capture texture information. Liao *et al.* (34) proposed an unsupervised artifact separation network (ADN) that can separate artifacts from CT images in the potential space. ADN leverages generative models and decomposition networks to construct noisy images in the absence of paired CT data and achieves excellent results. The denoising algorithms based on the previously mentioned GANs all use noise simulators to generate pseudo-LDCT data. For example, the generative networks in CycleGAN and UIDnet can be used for both LDCT denoising and the denoising of natural images, showcasing relatively complex model architectures and superior performance in simulating noise in natural images (35). Zhao *et al.* (36) introduced a dual-scale similarity-guided cycle GAN (DSC-GAN) for unsupervised LDCT denoising, which leverages similarity-based pseudopairing to enhance denoising performance. However, simpler generator structures are adequate for simulating noise in CT images, which accelerates the model training process while achieving satisfactory noise simulation effects.

To address the aforementioned challenges, we propose an innovative approach for denoising LDCT images using unpaired data. Our method is characterized by its swift training process and minimal number of parameters. It is important to note that we refer this form of unpaired learning as weakly supervised since it does not rely on precise label information. In unpaired scenarios, there is no direct correlation or matching between the input data and its corresponding output. In this paper, we introduce a novel denoising framework. Initially, a CT noise simulator GAN (NGAN) (*Figure 1A*), which is trained to learn the noise characteristics of LDCT, generating pseudo-LDCT images. Subsequently, the generated pseudo-LDCT images undergo denoising using a CT denoiser CNN (DCNN) (*Figure 1B*). A full-size discriminator is also introduced, which effectively suppresses artifacts and noise while recovering detailed information from LDCT images. This denoising method learns noise through a generative algorithm, leading to more effective LDCT image denoising. In essence, our model trains the CT denoiser by constructing training data pairs through the addition of noise to NDCT images rather than solely learning a direct mapping relationship from LDCT to NDCT. Detailed information about the model is provided in the Methods.
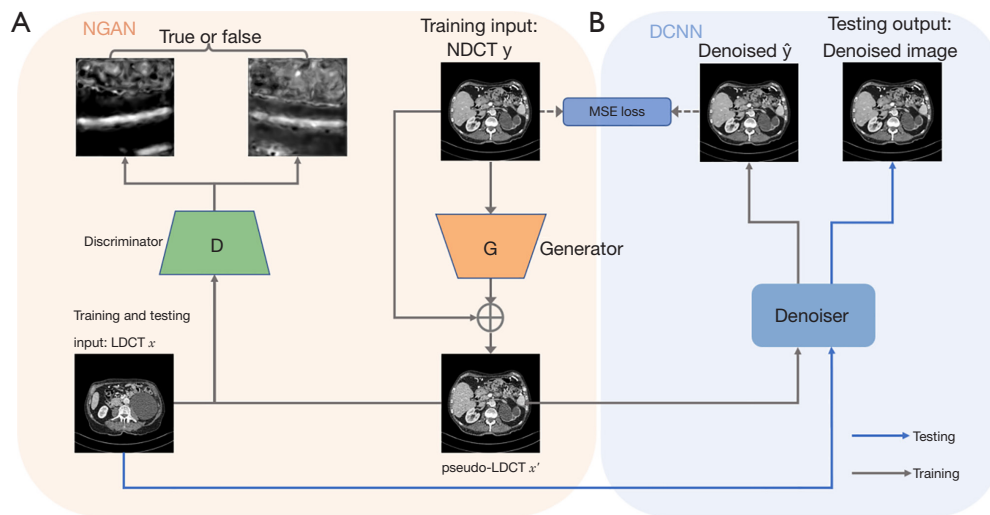
Figure 1 The overall structure of the proposed DNCNN framework is as follows. During the training process, (A) initially, unpaired LDCT input x and NDCT y are provided, where NGAN is employed to learn noise from unpaired data and establish paired data (x', y). The generator (G) is used to generate pseudo-LDCT image x', which is paired with the input NDCT image y. Subsequently, the discriminator (D) is employed to differentiate between the generated pseudo-LDCT image x' and the real LDCT image x, guiding the generator (G) to produce images x' that closely resemble the real LDCT image x. (B) Paired images (x', y) generated by NGAN are used for denoising training. During the testing process, the denoiser DCNN is applied to the input LDCT image for denoising, and thus the denoised image is obtained. $\oplus$, element-wise sum. NGAN, noise simulator GAN; GAN, generative adversarial network; LDCT, low-dose computed tomography; NDCT, normal-dose computed tomography; MSE, mean-squared error; DCNN, denoiser convolutional neural network.

In summary, our main contributions are as follows:

❖ We propose an autoencoder-based full-size discriminator to effectively guide the generator to produce more realistic images.
❖ We propose a DNCNN model that adopts a weakly supervised approach to training the model on unpaired data, thereby avoiding the challenge of directly learning the mapping from LDCT to NDCT.
❖ We demonstrate that our proposed DNCNN model achieves better results as compared to other weakly supervised methods while using fewer parameters.

Methods

Datasets

Our model undergoes training and testing on two datasets. The first dataset, known as the Mayo simulation dataset, comprises well-paired LDCT and NDCT images. During the experiment, the Mayo data are partitioned, with 90% allocated for training and 10% for testing. The second dataset, referred to as the CHCD clinic data from the Sixth People's Hospital of Chengdu, consisted of clinical data with a radiation dose of 30 mA, lacking paired LDCT and NDCT images; the same dataset was used for the partitioning approach. We conducted a comprehensive qualitative and quantitative evaluation of our weakly supervised denoising method on both simulated and clinical datasets. Despite the absence of strong supervision, our denoising method successfully produces high-quality CT images, comparable to those obtained using fully supervised learning methods. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The publicly available NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge dataset used in this study was originally collected from the Mayo Clinic with approval from their institutional review board. The CHCD Clinic low-dose CT data used in this study were collected from the physical examination population at the Sixth People's Hospital of Chengdu (2022–2023) and approved by the hospital (No. 2023-L-04). Individual consent for this retrospective analysis was waived.

Mayo clinic dataset

The performance evaluation of our method relied on the clinical dataset of low-dose CT officially endorsed by the Mayo Clinic (37). This dataset serves as a pivotal benchmark in contemporary research on low-dose CT image denoising algorithms. Comprising the standard-dose CT images from 10 patients and the corresponding simulated low-dose CT images, the dataset encompasses a total of 5,936 sets of meticulously aligned CT images. In the simulation of quarter-dose LDCT scans, Poisson noise was intentionally introduced, thereby generating a noise distribution akin to that observed in authentic LDCT images. The original dimensions of the images were 512×512, with a slice thickness of 1 mm.

CHCD clinic data

The clinical CT image dataset employed in our experiments originated from a real-world CT medical examination setting in China, consisting of three distinct low-dose CT datasets labeled as 10, 20, and 30 mA. All images within these datasets pertain to chest examinations and possess an original size of 512×512, with a slice thickness of 1 mm. The standard CT image dose was set at 160 mA. Unlike the Mayo Clinic dataset, these datasets lack an alignment relationship between LDCT and NDCT images, rendering supervised learning unfeasible.

This clinical CT image dataset, integral to our study, encompassed CT images from a wide array of patients. The 10-mA dataset included CT images from 6 patients, contributing to a total of 1,707 pairs of LDCT images and NDCT images. In the 20-mA dataset, CT images from 20 patients were included, totaling 5,545 LDCT images paired with an equivalent number of NDCT images. The 30-mA dataset consisted of CT images from 10 patients, resulting in a total of 2,834 pairs of LDCT and NDCT images.

Overall structure of the proposed method

In the proposed method, the projection data are conventionally acquired via tube currents that minimize noise while preserving clinical efficacy, and the data are subsequently processed logarithmically to approximate additive Gaussian noise (38). Nevertheless, distinct reconstruction algorithms may yield images with differing noise levels, potentially resulting in heightened noise in CT images. Although the consideration of reconstruction noise is beyond the scope of this paper, we posit that the noise model for CT images can be encapsulated as follows:

$$X = Y + N \qquad [1]$$

where $X \in R^{m \times n}$ denotes the image with noise (i.e., the low-dose image_, $Y \in R^{m \times n}$ is the original image to be recovered (i.e., the high-dose image), and N is the additive noise. It is worth noting that N does not necessarily follow a normal distribution even for regular dose CT images.

Building upon the previously outlined noise model, we added a GAN to acquire an understanding of the noise (N in Eq. [1]) present in LDCT images. Subsequently, this learned noise is introduced to the unpaired NDCT images, resulting in the creation of a set of paired images, comprising clean and noisy CT images, which serves as the training dataset for the denoiser. The denoising process involves using the DCNN to generate a clean image from its noisy counterpart. Under the assumption that $X \in R^{m \times n}$ is the LDCT images and $Y \in R^{m \times n}$ is the unpaired NDCT images, the problem can be formally expressed as follows:

$$X' = F_{NGAN}(X, Y) + Y \qquad [2]$$

$$\hat{Y} = F_{DCNN}(X') \qquad [3]$$

where $F_{NGAN}: R^{m \times n} \rightarrow R^{m \times n}$ denotes the real noise simulated by using unpaired CT images, X' denotes the generated pseudo-LDCT image, $F_{DCNN}: R^{m \times n} \rightarrow R^{m \times n}$ denotes the denoising process, and $\hat{Y}$ denotes the reconstructed LDCT image.

In our endeavor to effectively mitigate noise and artifacts in LDCT images, we propose an innovative framework centered on a GAN. The comprehensive architecture of our framework is depicted in *Figure 1* and comprises two pivotal modules. Aligned with previous studies on CT image denoising (8,15,39), our framework incorporates DCNN, a CNN module dedicated to the denoising process. This module undergoes training on pairs of LDCT and NDCT images as synthesized by the NGAN, with the aim of restoring the normal-dose images from their noisy counterparts. A detailed description of the network architecture is provided later in this paper.

Within our framework, an additional module employs a GAN, denoted as NGAN, for the generation of self-optimized noisy images through the collaborative learning of generators and discriminators. The primary objective of the NGAN module is to augment the network's responsiveness to diverse features present in LDCT images. Specifically, the generator within this module is tailored to amplify the network's capacity for discerning crucial features within LDCT images. In pursuit of heightened discriminative prowess and improved stability of the
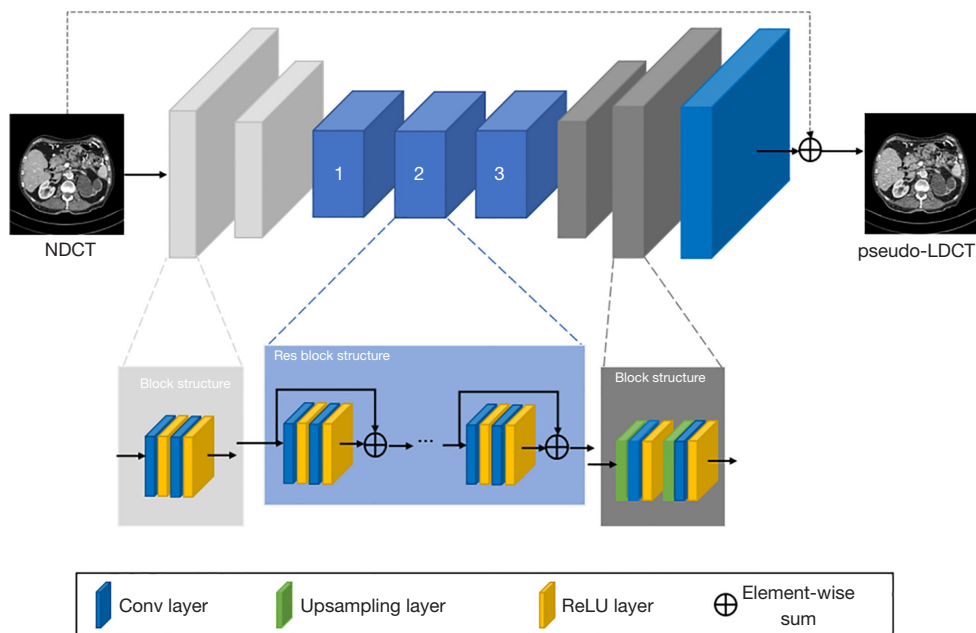
5576

Liao et al. Weakly supervised low-dose CT denoising based on GANs

Figure 2 Network structure of the generator of the NGAN. NDCT, normal-dose computed tomography; LDCT, low-dose computed tomography; ReLU, rectified linear unit; NGAN, noise simulator GAN; GAN, generative adversarial network.

GAN, the discriminator is formulated as a pixel-to-pixel discriminator. This strategic design choice contributes to the overall stability during network training.

In the training phase, NDCT images are inputted into the NGAN generator, and the resulting output is used to train the DCNN, facilitating the learning process for LDCT image denoising. The NGAN module's generator produces associated noisy images, which are subsequently assessed by the discriminator to produce more authentic synthetic images. Conversely, in the inference phase, only the DCNN is needed to execute image deblurring.

In summary, our network comprises two sequential steps: noise addition and denoising. In the noise addition step, the NDCT image undergoes processing within the NGAN module, producing a pseudo-LDCT image. This pseudo-LDCT image retains the fundamental structure and noise distribution characteristics observed in actual LDCT images. Subsequently, in the denoising step, the pseudo-LDCT image serves as the input to the DCNN module, resulting in the generation of an NDCT image closely resembling the authentic counterpart.

The NGAN module

The NGAN module is a critical component of our proposed

framework, as it discerns the noise patterns in LDCT and produces paired data conducive to the training of the DCNN module. Following the convention of other GAN-based models, the NGAN module comprises both a generator network and a discriminator network. In this section, a comprehensive examination of the architecture and loss functions employed by the NGAN module is provided.

Generator

The primary purpose of this network is to produce pseudo-LDCT images. As previously noted, LDCT images can be conceptualized as NDCT images affected by noise corruption. Consequently, the network is specifically configured to judiciously omit certain detailed information while preserving the structural edge characteristics of the initial NDCT image. This design ensures that the generated pseudo-LDCT images uphold structural information equivalent to that of the NDCT images.

We further introduced a nimble residual encoder-decoder network designed to overcome the challenges posed by NGAN tasks. The generator network employed in the NGAN module is illustrated in *Figure 2*. In order to guarantee that the noise distribution map acquired by the network aligns with the structural content of the initial NDCT image, the generator embraces residual learning
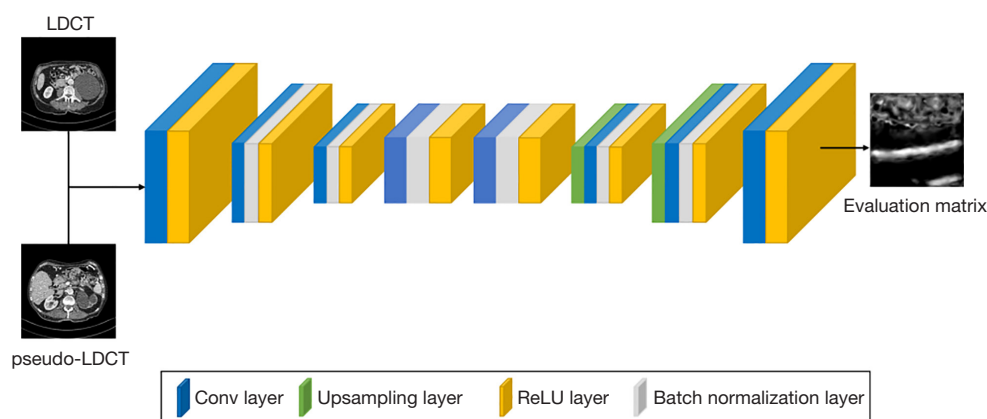
Figure 3 Network structure of the discriminator of NGAN. LDCT, low-dose computed tomography; ReLU, rectified linear unit; NGAN, noise simulator GAN; GAN, generative adversarial network.

comprehensively. This strategy ensures the inclusion of surplus information within the network, thereby facilitating the more effective learning of noise distribution.

The network architecture employs a 3×3 convolution with a stride of 2 for down-sampling and extracting image features, eschewing the use of pooling layers. Given the relatively straightforward composition of the CT images, the network restricts the downsampling to two stages, preserving more valuable information in the extracted feature map. Following these downsampling stages, three residual blocks are employed to modify the feature map, facilitating subsequent upsampling for partial information recovery. The generator uses residual learning to comprehend the LDCT noise distribution map, eschewing direct mapping of the entire network, and integrates the input NDCT image to generate the pseudo-LDCT image.

Discriminator

Drawing inspiration from the autoencoder architecture (38), we considered enhancing the discriminative capability of the discriminator network by incorporating a decoder network. The decoder is tasked with decoding and reconstructing the feature maps extracted by the encoder network. Specifically designed to take feature maps as input and regenerate the original image, the decoder's output plays a pivotal role in discerning the authenticity of the input image regardless of whether it is real or generated. The structure of the proposed discriminator network is depicted in *Figure 3*.

The discriminator network architecture is structured as a fully convolutional encoder-decoder, with the encoder handling downsampling and image feature extraction

and the decoder executing upsampling to recover feature information. To prevent excessive downsampling, which may result in the loss of high-frequency information and detrimentally impact network performance, the network restricts itself to two downsampling and two upsampling layers. With 3×3 convolution kernels, all kernels, except for the last convolutional layer, are set to 64. Incorporating zero-padding is essential to preserving the image size and preventing information loss. Batch normalization (BN) (40) is implemented postconvolution to expedite convergence and enhance network performance by aligning data within a specific distribution range. The chosen activation function is leaky rectified linear unit (LeakyReLU). The final layer of the network generates a full-size evaluation matrix, guiding the training of the generator.

Loss function

In the design of GANs, the discriminator assumes a pivotal role in guiding the generator toward producing high-quality images through an adversarial interplay. Consequently, the discriminators' design significantly influences the enhancement of denoised image quality. The discernment of image quality is chiefly executed by the discriminators exerting a profound impact on both the network's performance and stability.

In a standard GAN model, the architecture involves a minimization operation between a generator G, where its parameters map the sample z from the noise distribution $p_z(z)$, and a discriminator D, responsible for indicating the probability that the sample x belongs to the real data distribution $p_{data}(x)$. The optimization function of the initial GAN model is expressed as follows:
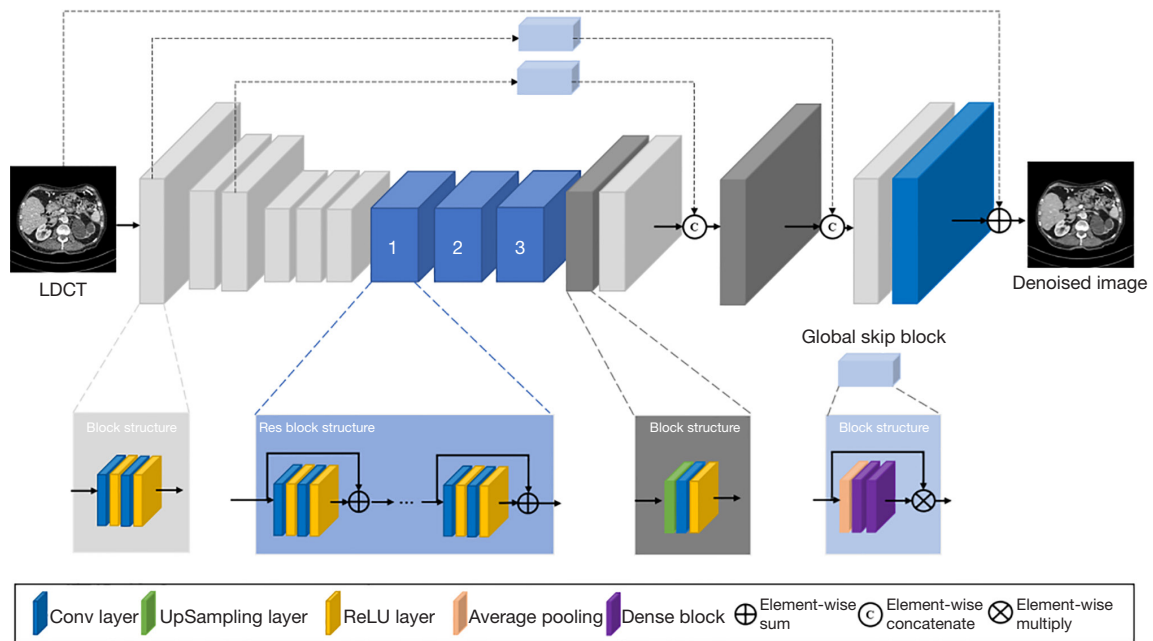
Figure 4 Network structure of DCNN. LDCT, low-dose computed tomography; ReLU, rectified linear unit; DCNN, denoiser convolutional neural network.

$$\min_{G} \max_{D} = \mathbb{E}_{x \sim p_{\text{data}}(x)}\left[\log D(x)\right] + \mathbb{E}_{z \sim p_z(z)}\left[\log\left(1 - D\left(G(z)\right)\right)\right] [4]$$

In this study, we adopted the least squares GAN (LSGAN) (41) as the primary loss function. Two key considerations underlie this choice. First, LSGAN employs the least squares loss for both discriminators and generators, compelling the generated pseudo-samples toward the decision boundary. In essence, the LSGAN loss effectively produces samples that closely approximate real data. Second, the use of LSGAN loss enhances training stability by penalizing samples based on their distance from the decision boundary, yielding a more informative gradient. The formulation of the LSGAN loss is expressed as follows:

$$\min_{D} L_{LSGAN}(D) = \frac{1}{2}\mathbb{E}_{x \sim p_{\text{data}}(x)}\left[\log(D(x)-b)^2\right] + \\ \frac{1}{2}\mathbb{E}_{z \sim p_z(z)}\left[\log(D(G(z))-a)^2\right] \quad [5]$$

$$\min_{G} L_{LSGAN}(G) = \frac{1}{2}\mathbb{E}_{z \sim p_z(z)}\left[\log(D(G(z))-c)^2\right] \quad [6]$$

where $p_{data}(x)$ signifies the authentic noise distribution, and $p_z(z)$ represents the distribution of random noise. As in other frameworks (42), the constants a, b, and c are assigned values of 1, 0, and 1, respectively. Notably, due to

our discriminator's output being a two-dimensional tensor, the variables a, b, and c also manifest two-dimensionally, necessitating padding with values 1, 0, and 1, respectively. Following the computation of the loss, the resultant output undergoes an averaging process to transition from a two-dimensional tensor to a scalar value.

The DCNN module

A pivotal element contributing to the efficacy of our framework is a denoising submodule. The NGAN module is designed with the objective of emulating authentic LDCT images and encompassing a broad spectrum of noise scenarios. Its overarching aim is to enhance the effectiveness of the DCNN module in the recovery of NDCT images from their low-dose counterparts. The architecture of the DCNN is illustrated in *Figure 4*.

The denoising network, in essence, mirrors the structure of the generator. Both entities employ a consistent design, featuring only two downsampling and two corresponding upsampling layers. A distinctive choice is the use of a 3×3 convolution with a stride of 2 to extract image features, eschewing the conventional pooling layer for downsampling. Within the denoiser structure, a notable component is the integration of a global residual module within its residual

skip connections. This module comprises a global average pooling layer and two dense layers, with the dense layers specifically operating across the channel dimension. The final step involves an element-wise multiplication between the module's input and output.

Loss function

In pursuit of enhanced denoising outcomes, we incorporated the mean-squared error (MSE) into the DCNN. MSE is a widely adopted metric used for assessing the accuracy in image processing via the measurement of the pixel-level discrepancy between LDCT and corresponding ground truth image pairs. It functions as a per-pixel loss metric. However, prior research indicates that sole reliance on MSE during training, as observed in models such as RED-CNN (8), may lead to the oversmoothing of tissues along the edges of CT images. Nevertheless, maintaining the use of MSE within a defined range yields improved metric results and can be mathematically expressed as follows:

$$L_{DCNN} = \mathbb{E}_x \left\| \hat{y} - y \right\|_2^2 \qquad [7]$$

The comprehensive loss function of the framework is formulated as the amalgamation of the denoiser loss and the LSGAN loss as follows:

$$L_{DNCNN} = L_{LSGAN}(D) + L_{LSGAN}(G) + L_{DCNN} \qquad [8]$$

Image quality assessment metrics

To evaluate image quality in the Mayo dataset, we used peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and visual information fidelity (VIF) as metrics. Due to the absence of paired LDCT and NDCT images in the clinical dataset, objective quantitative indicators are inadequate for assessing the experimental impact. Therefore, subjective visual perception becomes essential for analysis, entailing the observation of various tissues under different viewing conditions. The lack of reference images during testing necessitated the use of the no-reference structural sharpness (NRSS) metric (43) for image quality evaluation, which has proven to be highly effective in assisting radiologists in assessing medical images.

Experimental design

We commenced by conducting experiments to assess the efficacy of the NGAN network in the domain of CT image denoising and to clarify its influence on subsequent reconstruction procedures. Initially, we employed the NGAN network to generate synthetic LDCT images mimicking realistic noise distributions, which could serve as benchmarks to steer subsequent CT image reconstruction. To ascertain the network's proficiency in learning comparable noise distributions, we presented noise images acquired through the NGAN network in the experiments, exhibiting them with predefined window settings for validation.

Following this, our objective was to assess the efficacy of various discriminator types within the proposed model and ascertain their influence on network performance. We conducted comparisons of PSNR, SSIM, and VIF as evaluation metrics across various discriminators using the Mayo dataset. Furthermore, visual analysis was performed on regions of interest. We performed comparative experiments involving the conventional GAN discriminator, PatchGAN discriminator (44)—which includes the LSGAN discriminator—the WGAN-GP discriminator, and our full-size discriminator constructed on the UNet architecture, which were designated as *DNCNN-LS*, *DNCNN-W*, *DNCNN-Patch*, and *DNCNN-Unet*, respectively. The experimental outcomes were also examined under specific window settings.

To assess the efficacy of the proposed CT image denoising model, we conducted a comprehensive evaluation using the *Mayo* dataset. We used PSNR, SSIM, and VIF as evaluation metrics, comparing the model's performance with that of other state-of-the-art methods. The benchmark methods included block-matching and 3D filtering (BM3D) (45), RED-CNN, WGAN-VGG, CycleGAN, and SKFCycleGAN. BM3D represents a conventional approach for image noise processing, while RED-CNN utilizes MSE loss for image reconstruction. In contrast, the WGAN-VGG network integrates perceptual loss and adversarial loss. CycleGAN and SKFCycleGAN are exemplary types of weakly supervised deep learning methods. Supervised learning offers significant advantages over weakly supervised learning, primarily due to the data quality and label completeness. In supervised learning, models are trained using meticulously labeled data, facilitating a more in-depth exploration of feature-data relationships, thereby enhancing accuracy and reliability. Consequently, we anticipated that RED-CNN's metrics would surpass those of the weakly supervised methods in subsequent experimental results. For the CHCD Clinic dataset, to provide a rough evaluation of image quality, we employed NRSS, a gradient-based structural similarity metric, in which NRSS values

correspond to reduced noise and smoother images.

Finally, experiments were conducted on the Mayo dataset to evaluate the relationship between the denoising performance and the computational complexity of different algorithms. The evaluation metric was the time taken for training and testing.

Implementation details

To facilitate network model training, the original CT data were stored in npy format and underwent min–max normalization, followed by clipping to normalize Hounsfield Unit (HU) values within the range of 0 to 1. The minimum and maximum HU values in CT images were set to –1,000 and 2,000, respectively. Images were viewed using an observation window [e.g., (–150, 250)]. Additionally, due to device memory and CT data constraints, the original 512×512 CT images were cropped into multiple 128×128 images. LDCT and NDCT blocks were then constructed as training data using random pairing. To facilitate image quality assessment, the test set remained well-matched.

During model training on the Mayo Clinic and CHCD Clinic datasets, weights were initialized using a he-normal distribution with a standard deviation of 0.02. This he-*normal distribution*, introduced by He *et al.* in 2015 (46), refers to a method for initializing the weights of neural networks commonly employed in CNNs within the realm of deep learning. Weight updates occur after learning batches of size 48 in each iteration. The Adam optimizer was employed for training the entire network model during the network training phase, with hyperparameters set to $\beta\_1=0.5$ and $\beta\_2=0.9$. The learning rate of the proposed method undergoes annealing, starting at 1e–4 and decreasing to 1e–6 once the training loss has converged, with a total of 100 epochs for training iterations. The hyperparameters λ_a and λ_b in the total loss of the model were set to 1 based on previous research (32) and multiple experimental iterations. The proposed method operates on a patch-by-patch basis, promoting more efficient learning of the generated distribution. Specifically, random 128×128 patches are cropped from any position within the images, effectively increasing the number of training samples. Additionally, these images undergo random flips to further augment the sample pool. The parameters for these comparison methods were set based on recommendations from original papers.

Python 3.8 (Python Software Foundation, Wilmington, DE, USA) and Tensorflow 2.5 (Google, Mountain View,

CA, USA) were used for the experiments. The model was trained and tested on an Intel-Core i9 9960k processor (Intel, Santa Clara, CA, USA) and a GeForce 2080Ti graphics card (Nvidia Corp, Santa Clara, CA, USA) with 11 GB.

Results

Results of the NGAN

To assess the network's capability to learn a similar noise distribution, we analyze the noise images acquired by the NGAN network, displayed with a window setting of (–150, 250), as depicted in *Figure 5*.

Results of the comparison with different discriminators

We further aimed to assess the efficacy of the discriminator within our proposed model. The experimental outcomes are depicted in *Figure 6*, illustrating the observed effect within the window (–150, 250). Furthermore, *Figure 7* offers an enlarged view of the details within the red box in *Figure 6*.

Result of comparison with existing methods

Mayo clinic dataset

The quantitative evaluation of the experimental outcomes is presented in *Table 1*. Notably, the supervised learning approach, RED-CNN, demonstrated the most favorable evaluation metrics. Within the domain of weakly supervised methods, our proposed model exhibited superior objective metrics when compared to CycleGAN and SKFCycleGAN. Despite some variations in metrics between DNCNN and RED-CNN, the former showcased enhanced visual results. The experimental results are presented in *Figures 8,9*.

CHCD clinic dataset

The clinical dataset used in this experiment was nonaligned and lacked a genuine and reliable NDCT reference object for LDCT images. To provide a rough assessment of image quality, this study employed NRSS, with the experimental results being displayed in *Figures 10,11*.

Computational cost

To ascertain the connection between denoising performance and the computational complexity of diverse algorithms, we conducted a comparative analysis of the training and testing
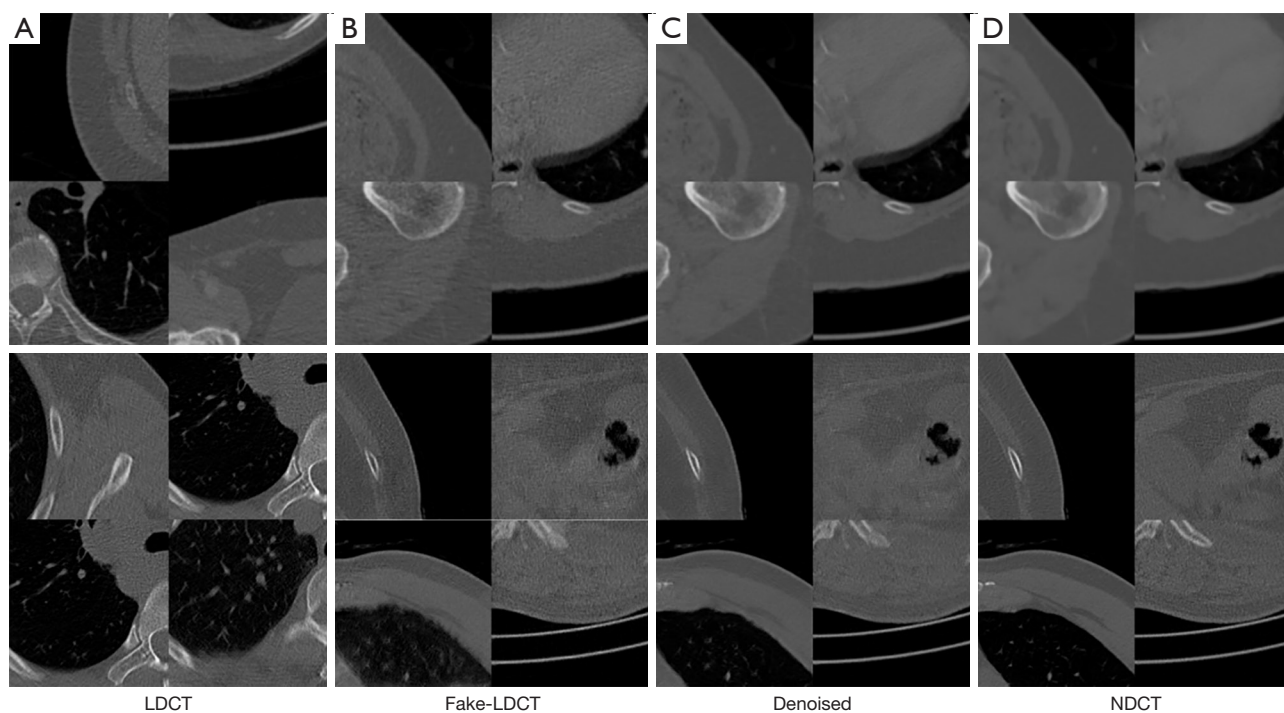
Figure 5 Partial results of the NGAN training process, with Mayo data in the first row and clinical data in the second row. (A,D) The unpaired LDCT and NDCT images that are used for NGAN to learn the noise. (B) The generated results of NGAN, which with (D), can form the paired LDCT and NDCT images for DCNN training. (C) The results of DCNN after the generated noisy images were denoised. LDCT, low-dose computed tomography; NDCT, normal-dose computed tomography; NGAN, noise simulator GAN; GAN, generative adversarial network; DCNN, denoiser convolutional neural network.

times on the Mayo dataset, as depicted in *Table 2*.

Discussion

Results of the NGAN

Figure 5 showed the effectiveness of NGAN. The LDCT images generated by NGAN indicate its ability to assimilate a comparable noise distribution while preserving the structural content of the NDCT image, thereby minimizing interference with the subsequent denoising process. The first row of the figure shows the results of the experiment with the Mayo dataset and suggests that even in the absence of strict alignment in the training data, the network can learn a noise distribution similarly to weak supervision, as evident in the learned noise distribution in LDCT images. The generated synthetic LDCT images closely emulate real LDCT images, effectively retaining the structural content of the NDCT images. The second row of the figure are the results from the experiment conducted on an authentic clinical dataset. A comparison between *Figure 5B* and *5C* reveals that despite

the clinical dataset's nonalignment, NGAN adeptly learns the noise distribution of actual LDCT images. Consequently, it produces synthetic LDCT images that closely resemble the visual characteristics of real LDCT images. Moreover, these generated images maintain the structural and edge information of the NDCT images. Thus, the noise image generated by NGAN proved effective in facilitating the training of subsequent denoising subnetworks.

Comparison with different discriminators

The experimental findings reveal a discernible impact of the NGAN discriminator on the denoising subnetwork's image generation quality. This was demonstrated by solely altering the NGAN discriminator while maintaining consistent conditions. The results underscore the enhanced network performance facilitated by the proposed discriminator. The four distinct discriminator networks effectively guide the generator to assimilate various degrees of a similar noise distribution, as depicted in *Figure 6*. Specifically, *Figure 6C*
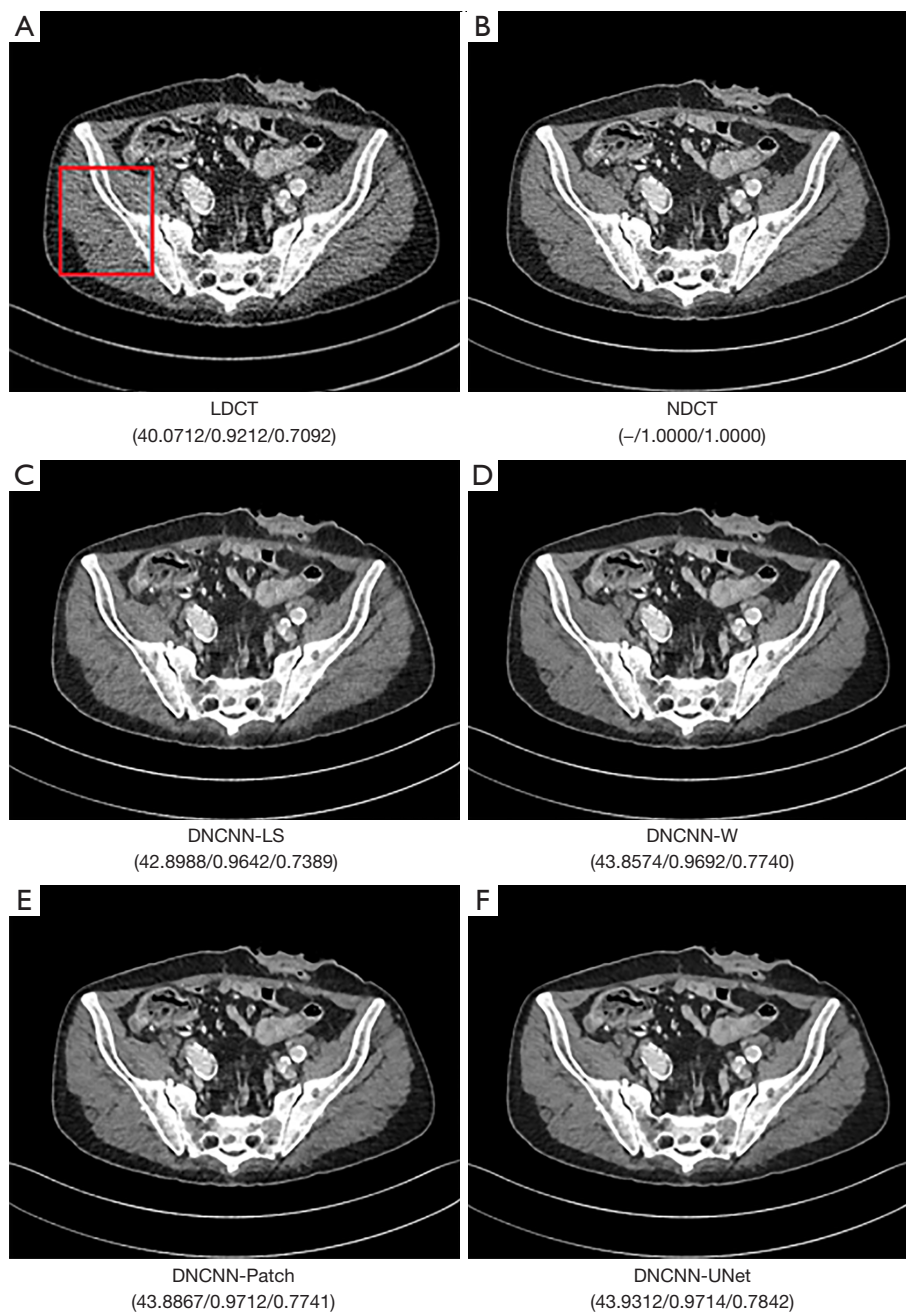
Figure 6 Results of comparison of different discriminators on the Mayo dataset (PSNR, SSIM, VIF). The red box indicates the region zoomed in from *Figure 7*. LDCT, low-dose computed tomography; NDCT, normal-dose computed tomography; LS, least squares generative adversarial networks discriminator; W, Wasserstein generative adversarial network with gradient penalty discriminator; Patch, patch generative adversarial network discriminator; UNet, U-shaped deep learning network architecture; PSNR, peak signal-to-noise ratio; SSIM, structural similarity index measure; VIF, visual information fidelity.
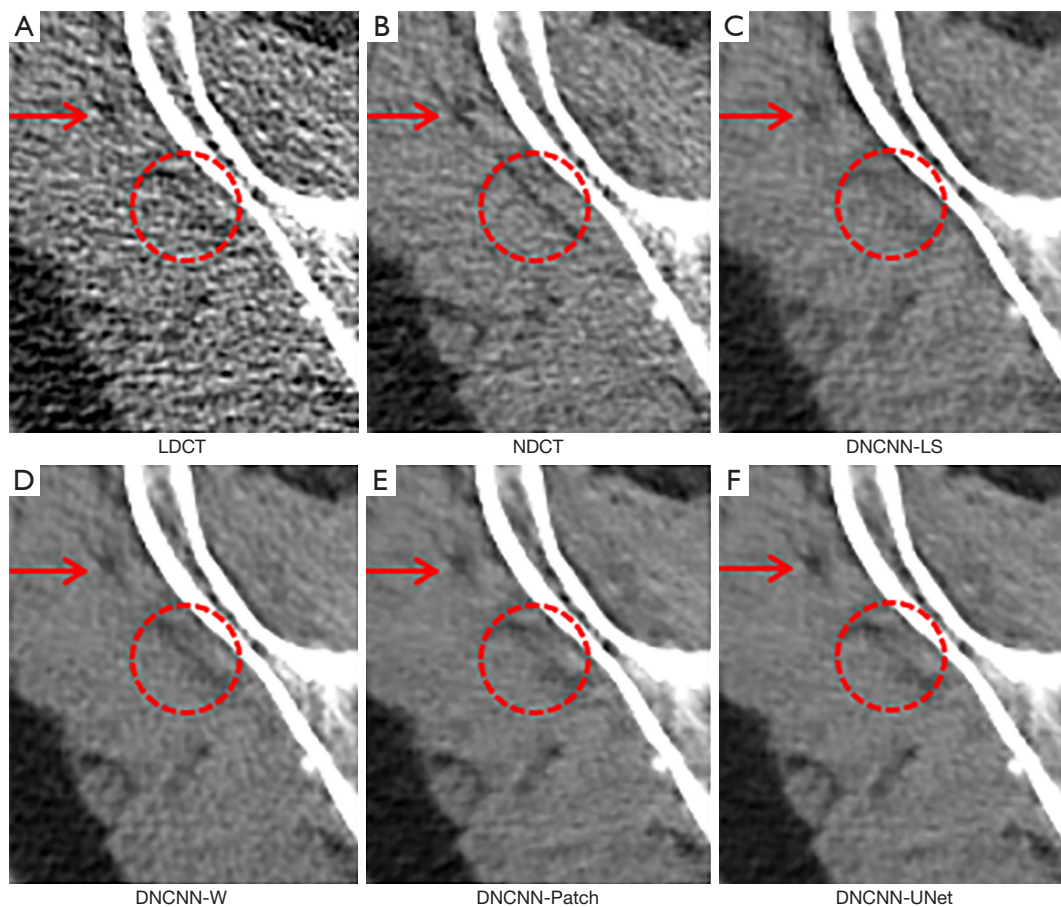
Figure 7 Zoomed in the region of interest results from *Figure 6*. The red arrows and circles indicate the two obvious differences. LDCT, low-dose computed tomography; NDCT, normal-dose computed tomography; LS, least squares generative adversarial network discriminator; W, Wasserstein generative adversarial network with gradient penalty discriminator; Patch, patch generative adversarial network discriminator; UNet, U-shaped deep learning network architecture.

Table 1 Quantitative results from the different algorithms on the Mayo dataset

Category	Method	PSNR	SSIM	VIF
Supervised	LDCT	39.8547	0.9049	0.6969
	BM3D (45)	40.0178	0.9292	0.6960
	RED-CNN (8)	44.1726	0.9684	0.8030
	WGAN-VGG (20)	40.6627	0.9387	0.6971
Weakly supervised	CycleGAN (11)	41.5457	0.9467	0.6983
	SKFCycleGAN (12)	42.1581	0.9515	0.7069
	Proposed	43.9441	0.9660	0.7707

PSNR, peak signal-to-noise ratio; SSIM, structural similarity index measure; VIF, visual information fidelity; LDCT, low-dose computed tomography; BM3D, block-matching and 3D filtering; RED-CNN, residual encoder-decoder convolutional neural network; WGAN-VGG, Wasserstein generative adversarial network-VGG; SKFCycleGAN, selective kernel-based cycle-consistent GAN; GAN, generative adversarial network.
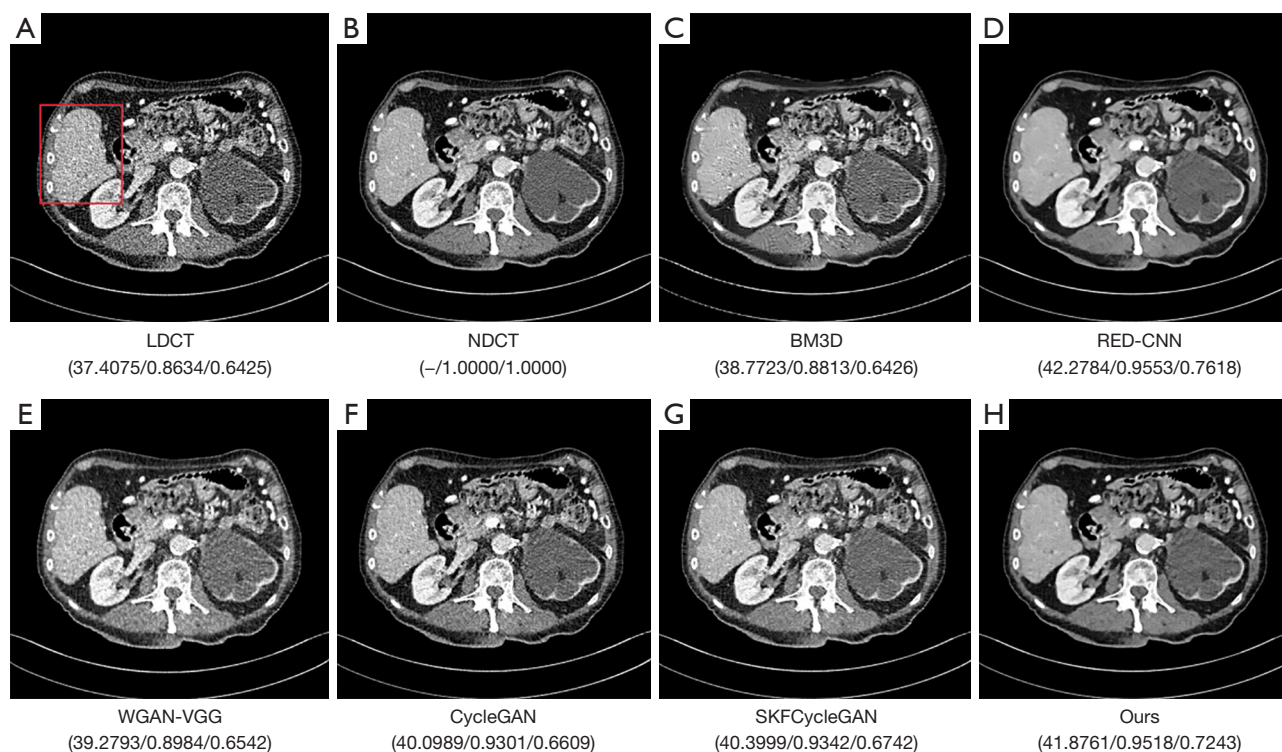
Figure 8 Results of comparison of different methods on the Mayo dataset (PSNR, SSIM, VIF). The red box indicates the region zoomed in from *Figure 9*. LDCT, low-dose computed tomography; NDCT, normal-dose computed tomography; BM3D, block-matching and 3D filtering; RED-CNN, residual encoder-decoder convolutional neural network; WGAN-VGG, Wasserstein generative adversarial network-VGG; SKFCycleGAN, selective kernel-based cycle-consistent GAN; GAN, generative adversarial network; PSNR, peak signal-to-noise ratio; SSIM, structural similarity index measure; VIF, visual information fidelity.

portrays the denoising outcome under the influence of DNCNN-LS. The denoised image, generated by DCNN, exhibits smooth and blurred edges, indicating suboptimal performance in recovering detailed information and texture nuances. Conversely, as illustrated in *Figure 6D*, DNCNN-W does not yield a substantial improvement in network denoising performance. The denoised image generated by the network under the guidance of BM3D exhibits fewer artifacts than does the noisy input; however, it still retains an excessive amount of noise artifacts. Similarly, the network guided by DNCNN-Patch significantly enhances the visual quality of the image, but the reconstructed CT image sacrifices more tissue, and the detail recovery is less effective compared to DNCNN. In contrast, the CT image reconstructed under the guidance of DNCNN closely resembles the standard dose CT image. It not only preserves relatively complete structural edge tissue information but also captures texture details more akin to the NDCT image, achieving a more comprehensive

detail information recovery. These outcomes underscore the positive impact of the proposed discriminator on enhancing the quality of reconstructed images, confirming the effectiveness of our approach.

Comparison with existing methods

Mayo clinic dataset

Figure 8 illustrates the impact within the observation window (−150, 250). A closer examination of the region enclosed in the red box in *Figure 8* is provided in *Figure 9*. The results presented in *Figure 8* demonstrate the superior performance of the method proposed in terms of reconstructed image quality, showcasing excellence in structural integrity, noise artifact suppression, and detail information retention. From *Figure 8C* and the corresponding details in *Figure 9C*, it is evident that BM3D struggles to fully preserve structure and suppress artifacts while effectively managing noise, often resulting in excessive blurring, and these results are
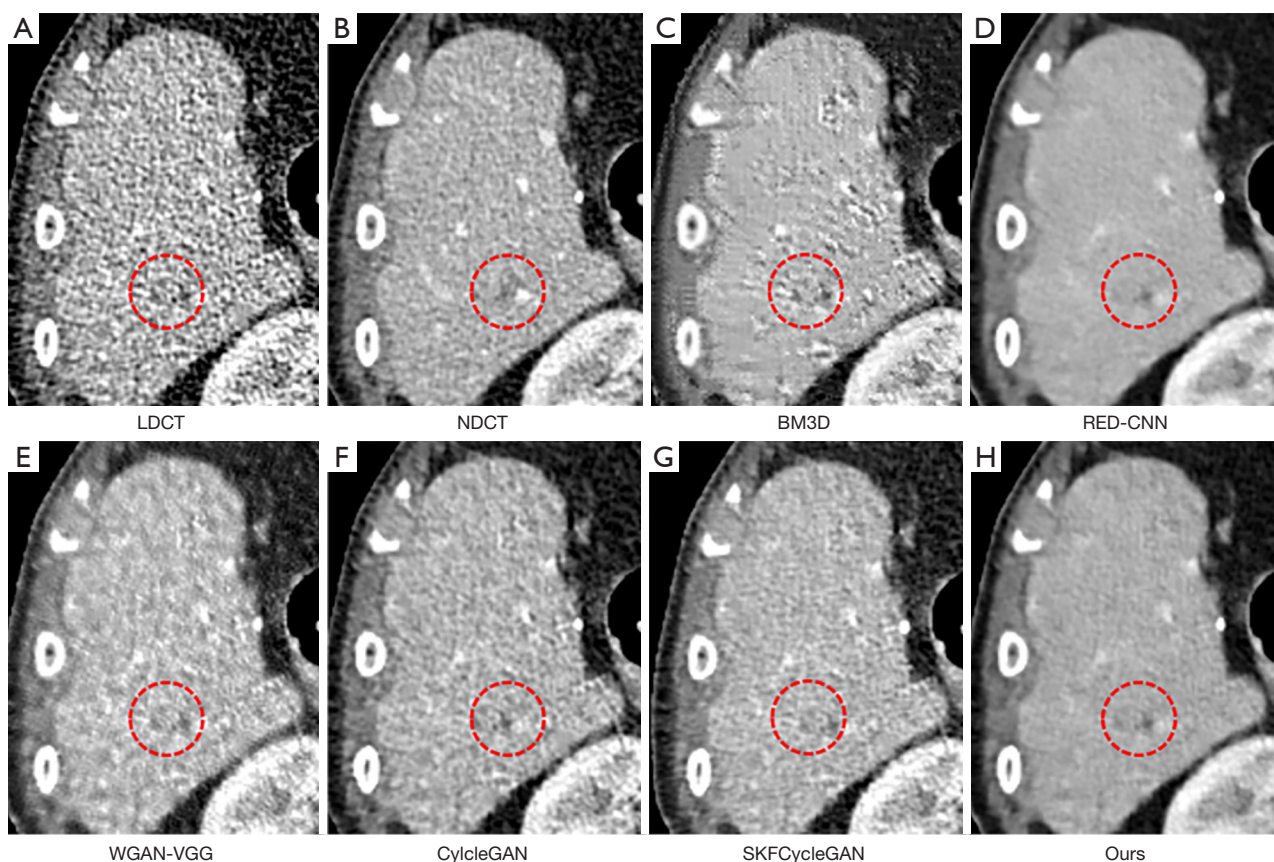
Figure 9 Zoomed in region of interest from the results in *Figure 8*. The red circle indicates one of the more obvious differences. LDCT, low-dose computed tomography; NDCT, normal-dose computed tomography; BM3D, block-matching and 3D filtering; RED-CNN, residual encoder-decoder convolutional neural network; WGAN-VGG, Wasserstein generative adversarial network-VGG; SKFCycleGAN, selective kernel-based cycle-consistent GAN; GAN, generative adversarial network.

Table 2 The training and testing time of the different algorithms from the Mayo dataset

Method	Training time/s	Testing time/s
BM3D (45)	–	133.08
RED-CNN (8)	2,248	0.10
WGAN-VGG (20)	78,349	0.22
CycleGAN (11)	55,635	0.25
SKFCycleGAN (12)	21,604	0.15
Proposed	17,872	0.05

BM3D, block-matching and 3D filtering; RED-CNN, residual encoder-decoder convolutional neural network; WGAN-VGG, Wasserstein generative adversarial network-VGG; SKFCycleGAN, selective kernel-based cycle-consistent GAN; GAN, generative adversarial network; 3D, three-dimensional.

consistent with the those of Tan *et al.* (12). Conversely, WGAN-VGG, as can be seen from *Figure 8E,9E*, adeptly suppresses noise and retains structural information but falters in preserving image details, occasionally leading to artifact production. Comparison with weakly supervised neural network algorithms, such as CycleGAN and SKFCycleGAN, revealed the superior performance of the proposed method in terms of structural integrity, detail information, and noise suppression, as evident in *Figures 8,9*. The proposed method effectively enhances the quality of reconstructed images. Notably, DNCNN exhibits significant improvement in reconstructed image quality concerning details, edges, and structure compared to the aforementioned algorithms.

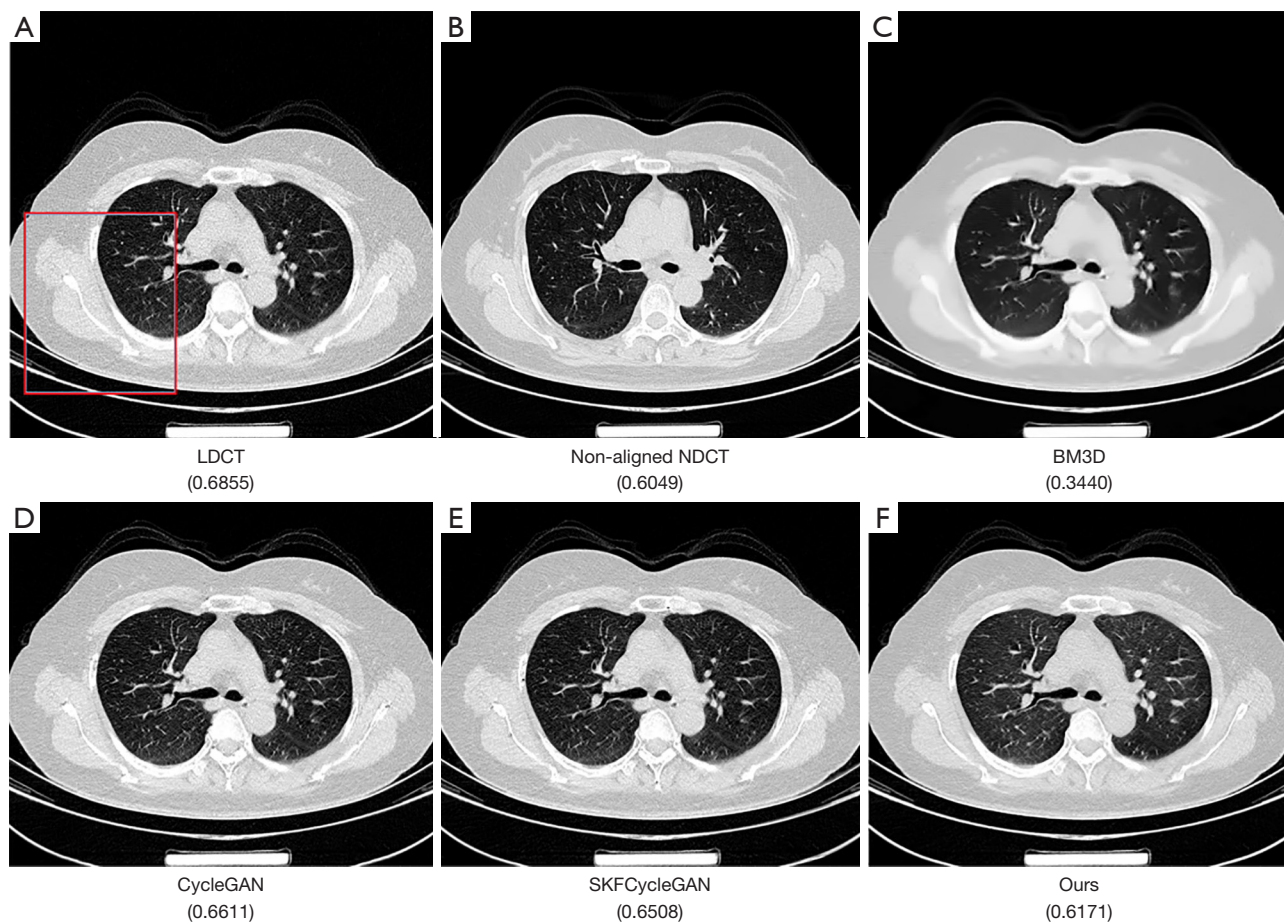In comparison to RED-CNN, the proposed method

Figure 10 Experimental results of the 30-mA clinical dataset with display window (–1,000, 200). The red box indicates the region zoomed in from *Figure 11*. The value in brackets represents NRSS, with smaller NRSS values corresponding to reduced noise and smoother images. LDCT, low-dose computed tomography; NDCT, normal-dose computed tomography; BM3D, block-matching and 3D filtering; NRSS, no-reference structural sharpness; SKFCycleGAN, selective kernel-based cycle-consistent GAN; GAN, generative adversarial network; 3D, three-dimensional.

represents a substantial advancement in detail information preservation and smoothing suppression in the reconstructed images, as demonstrated by the experimental results.

CHCD clinic dataset

Using a 30-Aa clinical dataset, we compared the proposed method with other methodologies. *Figures 10,11* show the experimental outcomes within the observation window (–1,000, 200), which allows for the clear observation of human tissue structures. The proposed method outperformed other denoising techniques in terms of detail, structural integrity, and edge contrast, as evident in *Figure 10* and the detailed view in *Figure 11*. Alternative denoising methods struggle to effectively suppress noise

artifacts, often exhibiting visible blurring. While the BM3D method can achieve partial noise suppression and retain relatively complete image edge and structural information, it sacrifices clarity, resulting in the poorest NRSS results. In contrast, the two weakly supervised algorithms, apart from the one proposed in this paper, enhance reconstructed image quality to a certain extent. However, CycleGAN exhibits suboptimal performance in noise artifact suppression and overall image quality, as depicted in *Figure 10D*, where images reconstructed by CycleGAN display noticeable artifacts, causing blurring and detail loss. Although SKFCycleGAN effectively suppresses noise artifacts, the denoised CT images exhibit blurred artifacts and indistinct edges.
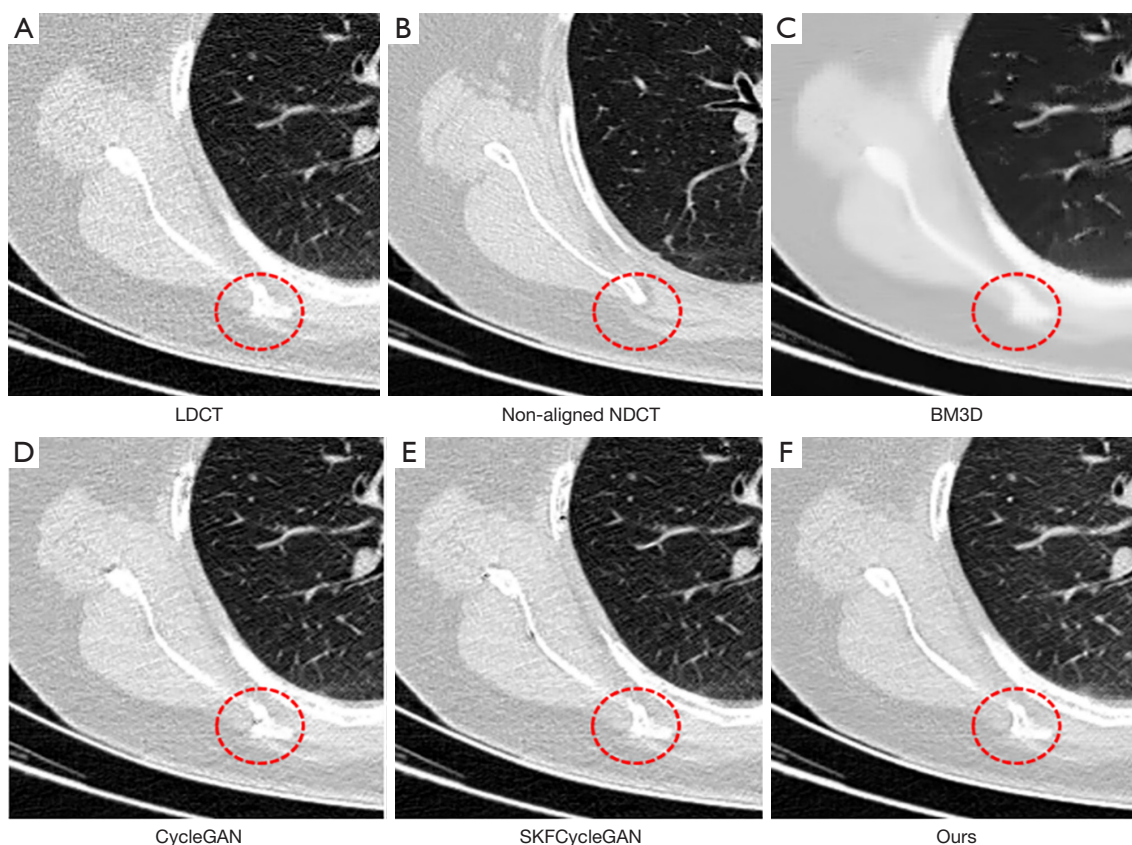
Figure 11 Zoomed in the region of interest for the results from *Figure 10*. The red circle indicates one of the more obvious differences. LDCT, low-dose computed tomography; NDCT, normal-dose computed tomography; BM3D, block-matching and 3D filtering; SKFCycleGAN, selective kernel-based cycle-consistent GAN; GAN, generative adversarial network; 3D, three-dimensional.

Smaller NRSS values correspond to reduced noise and smoother images. The method introduced in this paper exhibits NRSS values closest to those of NDCT images among the algorithms discussed. This confirm its outstanding performance in preserving structural integrity, enhancing edge contrast, and eliminating noise artifacts effectively.

Computational cost

The results revealed that RED-CNN requires the shortest training time among the five deep-learning algorithms. Notably, WGAN-VGG, CycleGAN, SKFCycleGAN, and the DNCNN proposed in this paper, all being grounded in GANs, necessitate more time for training in contrast to RED-CNN. Compared with RED-CNN and DNCNN, DNCNN has a shorter inference time. This is because in GAN-based denoising methods (11,12,20), the discriminator does not participate in the image denoising

process after training is completed; instead, the process relies solely on the generator (16). CycleGAN is particularly challenging to train owing to the intricacies of its network architecture (11). In summary, the proposed model in this paper demonstrates efficiency in both training and inference durations, which can be attributed to its advantageous parameters and structure.

Conclusions

This study introduced an unpaired deep learning approach for LDCT image denoising. The method leverages a combination of GAN- and CNN-based denoising networks and maps NDCT images to LDCT images. In comparison to existing techniques, our model operates with minimal assumptions on noise distribution and data type, eliminating the need for additional prior knowledge. A significant advantage lies in its rapid convergence and efficacy, achieved

5588

Liao et al. Weakly supervised low-dose CT denoising based on GANs

through the use of a lightweight generator and discriminator structure, coupled with the adoption of the LSGAN loss function. Vital to the model is the generation of realistic LDCT images. To facilitate this, we propose an encoder-decoder network as a discriminator for NGAN, enhancing the overall model performance. The discriminator employs a full-size output matrix, allowing for a focused analysis of noise details in LDCT images for realistic noise simulation. The learned noise is incorporated into NDCT images, forming NDCT and pseudo-LDCT image pairs that are used to train a denoising network in similar fashion to that of previously developed methods based on paired images. All components are seamlessly integrated to enable end-to-end training. Thorough evaluations conducted on synthetic and CHCD Clinic datasets demonstrated that our model surpasses previous unpaired data-based methods.

However, there remains a discernible disparity between existing weakly supervised methods and their supervised counterparts. Optimal visual outcomes have yet to be attained. For instance, as depicted in *Figure 9*, this method moderately diminishes fine details. Hence, future research could follow two avenues: a transition from two-dimensional to three-dimensional CT images, wherein adjacent contextual slices could offer additional structural and contour data to preserve details during the reconstruction process; and the development of more efficient generative models, such as diffusion models, which replicate the noise diffusion process to progressively diminish noise and accomplish denoising. Moreover, the amalgamation of diverse generative models may further amplify the denoising efficacy.

Acknowledgments

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-24-68/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The publicly available NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge dataset used in this study was originally collected from the Mayo Clinic with approval from their institutional review board. The CHCD Clinic low-dose CT data used in this study were collected from the physical examination population at the Sixth People's Hospital of Chengdu (2022–2023) and approved by the hospital (No. 2023-L-04). Individual consent for this retrospective analysis was waived.

References

1. Brenner DJ, Hall EJ. Computed tomography--an increasing source of radiation exposure. N Engl J Med 2007;357:2277-84.
2. Naidich DP, Marshall CH, Gribbin C, Arams RS, McCauley DI. Low-dose CT of the lungs: preliminary observations. Radiology 1990;175:729-31.
3. Manduca A, Yu L, Trzasko JD, Khaylova N, Kofler JM, McCollough CM, Fletcher JG. Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT. Med Phys 2009;36:4911-9.
4. Balda M, Hornegger J, Heismann B. Ray contribution masks for structure adaptive sinogram filtering. IEEE Trans Med Imaging 2012;31:1228-39.
5. Wang J, Li T, Lu H, Liang Z. Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography. IEEE Trans Med Imaging 2006;25:1272-83.
6. Zhu Y, Zhao M, Zhao Y, Li H, Zhang P. Noise reduction with low dose CT data based on a modified ROF model. Opt Express 2012;20:17987-8004.
7. Tamura A, Mukaida E, Ota Y, Nakamura I, Arakita K,

Yoshioka K. Deep learning reconstruction allows low-dose imaging while maintaining image quality: comparison of deep learning reconstruction and hybrid iterative reconstruction in contrast-enhanced abdominal CT. Quant Imaging Med Surg 2022;12:2977-84.

8. Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, Zhou J, Wang G. Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network. IEEE Trans Med Imaging 2017;36:2524-35.

9. You C, Yang Q, Shan H, Gjesteby L, Li G, Ju S, Zhang Z, Zhao Z, Zhang Y, Wenxiang C, Wang G. Structurally-sensitive Multi-scale Deep Neural Network for Low-Dose CT Denoising. IEEE Access 2018;6:41839-55.

10. Kang E, Chang W, Yoo J, Ye JC. Deep Convolutional Framelet Denosing for Low-Dose CT via Wavelet Residual Network. IEEE Trans Med Imaging 2018;37:1358-69.

11. Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017:2223-32.

12. Tan C, Yang M, You Z, Chen H, Zhang Y. A selective kernel-based cycle-consistent generative adversarial network for unpaired low-dose CT denoising. Precis Clin Med 2022;5:pbac011.

13. Yi Z, Zhang H, Tan P, Gong M. DualGAN: Unsupervised Dual Learning for Image-to-Image Translation. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017:2868-76.

14. Mirza M, Osindero S. Conditional generative adversarial nets. arXiv: 1411.1784, 2014.

15. Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv: 1511.06434, 2015.

16. Zhang J, Gong W, Ye L, Wang F, Shangguan Z, Cheng Y. A Review of deep learning methods for denoising of medical low-dose CT images. Comput Biol Med 2024;171:108112.

17. Chen H, Zhang Y, Zhang W, Liao P, Li K, Zhou J, Wang G. Low-dose CT via convolutional neural network. Biomed Opt Express 2017;8:679-94.

18. Wolterink JM, Leiner T, Viergever MA, Isgum I. Generative Adversarial Networks for Noise Reduction in Low-Dose CT. IEEE Trans Med Imaging 2017;36:2536-45.

19. Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. Proceedings of the 34th International Conference on Machine Learning, PMLR 2017;70:214-23.

20. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, Kalra MK, Zhang Y, Sun L, Wang G. Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss. IEEE Trans Med Imaging 2018;37:1348-57.

21. Kim B, Han M, Shim H, Baek J. A performance comparison of convolutional neural network-based image denoising methods: The effect of loss functions on low-dose CT images. Med Phys 2019;46:3906-23.

22. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2014.

23. Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. In: Leibe B, Matas J, Sebe N, Welling M. editors. Computer Vision – ECCV 2016. Lecture Notes in Computer Science, Springer, 2016;9906:694-711.

24. Singh P, Diwakar M, Gupta R, Kumar S, Chakraborty A, Bajal E, Jindal M, Shetty DK, Sharma J, Dayal H, Naik N, Paul R. A method noise-based convolutional neural network technique for CT image Denoising. Electronics 2022;11:3535.

25. Yang S, Pu Q, Lei C, Zhang Q, Jeon S, Yang X. Low-dose CT denoising with a high-level feature refinement and dynamic convolution network. Med Phys 2023;50:3597-611.

26. Yan R, Liu Y, Liu Y, Wang L, Zhao R, Bai Y, Gui Z. Image denoising for low-dose CT via convolutional dictionary learning and neural network. IEEE Transactions on Computational Imaging 2023;9:83-93.

27. Wu X, Liu M, Cao Y, Ren D, Zuo W. Unpaired learning of deep image denoising. In: Vedaldi A, Bischof H, Brox T, Frahm JM. editors. Computer Vision – ECCV 2020. Lecture Notes in Computer Science, Springer, 2020;12349:352-68.

28. Huang Y, Xia W, Lu Z, Liu Y, Chen H, Zhou J, Fang L, Zhang Y. Noise-Powered Disentangled Representation for Unsupervised Speckle Reduction of Optical Coherence Tomography Images. IEEE Trans Med Imaging 2021;40:2600-14.

29. Yin Z, Xia K, He Z, Zhang J, Wang S, Zu B. Unpaired image denoising via Wasserstein GAN in low-dose CT image with multi-perceptual loss and fidelity loss. Symmetry 2021;13:126.

30. Zhou ZH. A brief introduction to weakly supervised learning. Natl Sci Rev 2018;5:44-53.

31. Kang E, Koo HJ, Yang DH, Seo JB, Ye JC. Cycle-consistent adversarial denoising network for multiphase coronary CT angiography. Med Phys 2019;46:550-62.

32. Hong Z, Fan X, Jiang T, Feng J. End-to-End Unpaired Image Denoising with Conditional Adversarial Networks. Proceedings of the AAAI Conference on Artificial Intelligence 2020;34:4140-9.

33. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC. Improved Training of Wasserstein GANs. Part of Advances in Neural Information Processing Systems 30 (NIPS 2017) 2017;30:5769-79.

34. Liao H, Lin WA, Zhou SK, Luo J. ADN: Artifact Disentanglement Network for Unsupervised Metal Artifact Reduction. IEEE Trans Med Imaging 2020;39:634-43.

35. Hossain S, Lee B. NG-GAN: A Robust Noise-Generation Generative Adversarial Network for Generating Old-Image Noise. Sensors (Basel) 2022.

36. Zhao F, Liu M, Gao Z, Jiang X, Wang R, Zhang L. Dual-scale similarity-guided cycle generative adversarial network for unsupervised low-dose CT denoising. Comput Biol Med 2023;161:107029.

37. Moen TR, Chen B, Holmes DR 3rd, Duan X, Yu Z, Yu L, Leng S, Fletcher JG, McCollough CH. Low-dose CT image and projection dataset. Med Phys 2021;48:902-911.

38. Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol PA, Bottou L. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. J Mach Learn Res 2010;11:3371-408.

39. Liang T, Jin Y, Li Y, Wang T. EDCNN: Edge enhancement-based Densely Connected Network with Compound Loss for Low-Dose CT Denoising. 2020 15th

40. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on Machine Learning, PMLR 2015;37:448-56.

41. Mao X, Li Q, Xie H, Lau R, Wang Z, Paul Smolley S. Least squares generative adversarial networks. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017:2794-802.

42. Li S, Kawale J, Fu Y. Deep collaborative filtering via marginalized denoising auto-encoder. CIKM '15: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management 2015;811-20.

43. Xie X, Zhou J, Wu Q. No-reference quality index for image blur. J Comput Appl 2010;30:921-4.

44. Kang E, Min J, Ye JC. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. Med Phys 2017;44:e360-75.

45. Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans Image Process 2007;16:2080-95.

46. He K, Zhang X, Ren S, Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015:1026-34.

IEEE International Conference on Signal Processing (ICSP), Beijing, China, 2020:193-8.