Unsupervised Low-Dose CT Reconstruction With One-Way Conditional Normalizing Flows

Ran An, Ke Chen, and Hongwei Li

***Abstract*—Deep-learning techniques have demonstrated significant potential in low-dose computed tomography (LDCT) reconstruction. Nevertheless, supervised methods are limited by the scarcity of labeled data in clinical scenarios, while CNN-based unsupervised denoising methods often result in excessive smoothing of reconstructed images. Although normalizing flows (NFs) based methods have shown promise in generating detail-rich images and avoiding over-smoothing, they face two key challenges: (1) Existing two-way transformation strategies between noisy images and latent variables, despite leveraging the regularization and generation capabilities of NFs, can lead to detail loss and secondary artifacts; and (2) Training NFs on high-resolution CT images is computationally intensive. While conditional normalizing flows (CNFs) can mitigate computational costs by learning conditional probabilities, current methods rely on labeled data for conditionalization, leaving unsupervised CNF-based LDCT reconstruction an unresolved challenge. To address these issues, we propose a novel unsupervised LDCT iterative reconstruction algorithm based on CNFs. Our approach implements a strict one-way transformation during alternating optimization in the dual spaces, effectively preventing detail loss and secondary artifacts. Additionally, we propose an unsupervised conditionalization strategy, enabling efficient training of CNFs on high-resolution CT images and achieving fast, high-quality unsupervised reconstruction. Experimental results across multiple datasets demonstrate that the proposed method outperforms several state-of-the-art unsupervised methods and even rivals some supervised approaches.**

***Index Terms*—Low-dose CT, iterative reconstruction, unsupervised learning, conditional normalizing flows, generative models, regularized reconstruction.**

I. Introduction

COMPUTED tomography (CT) is a widely utilized medical imaging technique for visualizing the internal structures of objects. However, it is well known that excessive exposure to X-rays can pose significant health risks, including an increased likelihood of cancer and genetic damage [1]. In response, people have adopted the "as low as reasonably achievable" (ALARA) principle [2] to guide the minimization of X-ray dosage in diagnostic procedures. However, low-dose scanning will inevitably introduce noise into the projection data, thereby introducing severe noise and artifacts in the reconstructed images [3]. Therefore, LDCT reconstruction has always been a popular subject in the medical imaging community.

For LDCT, various denoising strategies have been proposed, including image post-processing techniques [4], [5], [6] and projection data pre-processing methods [7], [8], [9]. While these approaches are straightforward and convenient, they often neglect the consistency between the image and projection data during the reconstruction procedure, usually leading to blurring and secondary artifacts in the reconstructed images. To address these limitations, some other traditional methods have incorporated artificially designed priors, such as total variation (TV) minimization [10], into iterative reconstruction frameworks [11], [12], [13]. These methods achieve improved consistency and reconstruction quality but are often hindered by the need for extensive hyper-parameter tuning and a large number of iterations. Furthermore, the accuracy and universality of these artificial priors remain significant concerns that warrant further investigation.

In contrast to traditional methods, learning-based methods exhibit superior reconstruction performance, leveraging priors derived from "Big Data." Recently, deep learning (DL) techniques have shown remarkable capabilities in modeling and data fitting, leading to significant advancements in DL-based LDCT reconstruction. For image post-processing, various neural network (NN) architectures, such as FBPConvNet [14], RED-CNN [15], DIRE [16], and CTFormer [17], have achieved promising results. However, these purely post-processing strategies often overlook the consistency between image and projection data, easily resulting in over-smoothing and detail loss in the reconstructed images. To address this limitation, some dual-domain denoising networks have been proposed, which jointly process projection data and images within a unified framework. Notable examples include DRCNN [18], DDP-Net [19], DuDoUFNet [20], and DRONE [21]. Additionally, some methods unroll regularized iterative algorithms into NN to adaptively train hyperparameters, achieving outstanding performance. Representative approaches in this category include LEARN [22] and PD-Net [23]. Despite their exceptional performance, these methods rely on a supervised learning framework,

Received 28 October 2024; revised 29 January 2025; accepted 8 March 2025. Date of publication 19 March 2025; date of current version 4 April 2025. This work was supported in part by Beijing Natural Science Foundation under Grant Z210003, in part by the National Natural Science Foundation of China (NSFC) under Grant 61971292, in part by the Beijing Outstanding Young Scientist Program under Grant JWZQ20240101027, and in part by China Scholarship Council (CSC) under Grant 202307300001. The associate editor coordinating the review of this article and approving it for publication was Prof. Yang Chen. *(Corresponding author: Hongwei Li.)*

Ran An is with the School of Mathematical Sciences, Capital Normal University, Beijing 100048, China, and also with the Centre for Mathematical Imaging Techniques, University of Liverpool, L69 7ZL Liverpool, U.K.

Ke Chen is with the Department of Mathematics and Statistics, University of Strathclyde, G1 1XQ Glasgow, U.K., and also with the Centre for Mathematical Imaging Techniques, University of Liverpool, L69 7ZL Liverpool, U.K. (e-mail: k.chen@strath.ac.uk).

Hongwei Li is with the School of Mathematical Sciences, Capital Normal University, Beijing 100048, China (e-mail: hongwei.li91@cnu.edu.cn).

Digital Object Identifier 10.1109/TCI.2025.3553039

necessitating large amounts of labeled data for training. This requirement poses significant challenges in real-world CT due to ethical considerations and the practical difficulties of obtaining perfect consistent repeat scans.

To mitigate the reliance on labeled data, some unsupervised strategies have been proposed for LDCT reconstruction. One approach involves simulating paired data using generative adversarial networks (GAN) [24]. Representative methods in this category include GAN-CIRCLE [25], Cycle-Free CycleGAN [26], AdaIN-Based Tunable CycleGAN [27], and IdentityGAN [28]. However, these methods predominantly rely on cyclic consistency loss, which often results in unstable training. Furthermore, ensuring the authenticity and accuracy of the generated paired data remains a significant challenge. Alternatively, unsupervised LDCT image denoising methods inspired by the Noise2Noise (N2N) framework have been developed [29], [30], [31], [32]. Despite their potential, these methods still require pairs of noisy images of the same scene, which are difficult to obtain in practical applications. To address this limitation, self-supervised methods that leverage the noisy image itself have been proposed, such as Noise2Sim [33] and Noise2inverse [34]. Additionally, dual-domain denoising strategies have been introduced to ensure consistency between projection data and images, including ETSRP [35], SSDDNet [36], and SDBDNet [37]. These self-supervised methods demonstrate performance comparable to supervised approaches but often suffer from over-smoothing and loss of detail due to their pursuit of averaging. To enhance the sharpness and detail of reconstructed images, GAN have been employed to post-process the outputs of denoising networks. Notable examples include DD-UNET [38] and CLEAR [39]. However, this hybrid approach still faces challenges related to GAN instability and accuracy, limiting its performance in preserving structures and details.

Generative models, including normalizing flows (NFs) [40] and diffusion models [41], have been increasingly utilized to learn priors from clean images, serving as regularization terms in iterative reconstruction frameworks. For instance, Wei et al. introduced an alternating minimization algorithm featuring a two-way transformation strategy—mapping between noisy images and latent variables via NFs—to address imaging inverse problems [42]. Similarly, Fabian et al. proposed PatchNR [43], which trains NFs on patches of normal-dose images and incorporates NFs as the regularization term in iterative reconstruction. On the diffusion model front, He et al. developed EASEL [44] based on the score-based diffusion model [45], while Liu et al. and Xia et al. proposed Dn-Dp [46] and DPR-IR [47], respectively, leveraging the denoising diffusion probabilistic model (DDPM) [48] for iterative LDCT reconstruction. These methods adopt an unsupervised strategy that requires only normal-dose images for training, aligning well with practical applications. However, these methods still carry some challenges. Training NFs on high-resolution images (e.g., 512×512) is computationally intensive, limiting their applicability. Although dimension-reduction techniques [49] and conditional probability learning strategies [50] have been proposed to mitigate this issue, dimensionality reduction often leads to information loss, and existing conditionalization methods typically rely on labeled data, restricting their use in unsupervised LDCT reconstruction. Additionally, the two-way transformation between noisy images and latent variables in NFs can lead to distribution bias, resulting in detail loss and secondary artifacts in reconstructed images. Diffusion-model-based methods also face a well-recognized limitation: their reconstruction process is often time-consuming, requiring thousands of iterative sampling steps. While fast sampling techniques such as DDIM [51] and DPM-solver [52] have reduced the number of steps to as few as 50 or even 10-20, their efficacy in LDCT reconstruction has not yet been thoroughly validated. Consequently, achieving efficient and high-quality unsupervised LDCT reconstruction with generative models still poses challenges.

This paper introduces a novel iterative reconstruction algorithm for LDCT by improving the current NFs-based methods. Our approach achieves high-quality and efficient reconstruction through an unsupervised framework that relies solely on normal-dose images for distribution learning. To fully exploit the regularization and generation capabilities of NFs, we implement a dual-space alternating iterative reconstruction in the data and latent space. Unlike existing methods that employ two-way transformation, our algorithm adopts strict one-way generation transformation, effectively preventing the introduction of secondary artifacts. Furthermore, to enable efficient training of NFs on high-resolution images for unsupervised LDCT reconstruction, we propose an innovative unsupervised conditionalization method, facilitating the training of CNFs. This allows our network to act on high-resolution images easily. By utilizing the linearization technique and the ordered-subset simultaneous algebraic reconstruction technique (OS-SART) [53] for incremental reconstruction, our method achieves computation-efficient reconstruction. As an unsupervised framework, our approach performs comparably to supervised methods and surpasses the speed of some popular generative-model-based iterative reconstruction methods. Experiments on two datasets demonstrate that our method effectively addresses the primary challenges associated with current NFs-based LDCT reconstruction. Compared to state-of-the-art unsupervised and even supervised LDCT reconstruction methods, our method shows promising performance. The key contributions of our work are summarized as follows:

- We propose a novel unsupervised iterative reconstruction algorithm for LDCT based on NFs, which integrates regularization in both the data and latent spaces of the NFs and utilizes a strict one-way transformation strategy. Our method effectively mitigates the issues of detail loss and secondary artifacts caused by the two-way transformation of noisy images.
- We propose an unsupervised conditionalization strategy for CNFs-based LDCT reconstruction without paired training data. Leveraging this strategy, we efficiently train unsupervised CNFs on high-resolution CT images and achieve fast iterative reconstruction. To the best of our knowledge, this is the first time CNFs have been incorporated into the LDCT reconstruction procedure.
- Experiments conducted on different datasets demonstrate that our method achieves relatively fast reconstruction

speed and superior performance compared to several state-of-the-art learning-based iterative reconstruction approaches.

II. Related Work

A. NFs-Based LDCT Reconstruction

The forward projection process of LDCT can be modeled as follows:

$$y = \mathbf{A}x + \eta, \tag{1}$$

where $y \in \mathbb{R}^m$ represents the vectorized low-dose projection data, $x \in \mathbb{R}^n$ denotes the vectorized ideal clean image, $\eta \in \mathbb{R}^m$ corresponds to the low-dose noise introduced during scanning, and $\mathbf{A} \in \mathbb{R}^{m\times n}$ is the known projection matrix.

Given the low-dose projection data y, the ideal reconstructed image $\hat{x}$ can be obtained by maximizing the following logarithmic probability:

$$\hat{x} = \arg\max_x \log p(y|x) + \log p(x). \tag{2}$$

where $p(y|x)$ is the posterior probability of y given x, and $p(x)$ denotes the prior probability of x. Assuming $p(y|x) \sim \mathcal{N}(\mathbf{A}x, \sigma_0^2)$, (2) transforms into a minimization problem:

$$\hat{x} \in \arg\min_x \frac{1}{2\sigma_0^2}\|y - \mathbf{A}x\|_2^2 - \log p(x). \tag{3}$$

The prior probability $p(x)$ plays a critical role in this minimization problem, and accurately modeling it is essential. A promising approach is to learn $p(x)$ from a large dataset of normal-dose images using generative models such as NFs [40]. In general, NFs learn a differentiable bijective mapping $\mathcal{F}_\theta = \mathcal{G}_\theta^{-1}$, parameterized by θ, between the data distribution $p(x)$ and a simple distribution $p(z)$, such as the standard Gaussian distribution $p(z) \sim \mathcal{N}(0, 1)$. Using a trained NFs $\mathcal{F}_\theta$, each image sample x corresponds to a unique latent variable z, and the two can be bidirectionally mapped through the invertible network:

$$\begin{cases} z = \mathcal{F}_\theta(x), \\ x = \mathcal{G}_\theta(z). \end{cases} \tag{4}$$

The prior probability $p(x)$ can then be expressed as:

$$\log p(x) = \log p(z) + \log\left|\det(\mathbf{D}\mathcal{F}_\theta(x))\right|, \tag{5}$$

where $\mathbf{D}\mathcal{F}_\theta(x)$ denotes the Jacobian of $\mathcal{F}_\theta$ with respect to x, and its determinant $\det(\mathbf{D}\mathcal{F}_\theta(x))$ accounts for the change in density due to the transformation $\mathcal{F}_\theta$. It is worth noting that the second term on the right side of (5) is a constant and $p(z) \sim \mathcal{N}(0, 1)$. Substituting (5) and $x = \mathcal{G}_\theta(z)$ into (3), we obtain a new minimization problem with respect to z:

$$\hat{z} \in \arg\min_z \|y - \mathbf{A}\mathcal{G}_\theta(z)\|_2^2 + \lambda\|z\|_2^2. \tag{6}$$

where λ is a parameter that controls the regularization strength.

To address the optimization problem in (6), several methods have been proposed. In [54] and [55], the authors suggest optimizing z directly in the latent space using gradient descent. While this approach leverages the generative capabilities of NFs and produces detail-rich images, it lacks constraints in the x-space and often converges to local optima, resulting in suboptimal reconstruction accuracy. In [43] and [49], the authors propose optimizing in the x-space by setting $\mathcal{G}_\theta(z) = x$ and $z = \mathcal{F}_\theta(x)$ in (6) and updating x via gradient descent. This strategy improves reconstruction accuracy but fails to fully utilize the generative capabilities of NFs, often leading to excessive smoothing. To better utilize the regularization and generation capabilities of NFs, Wei et al. [42] introduced an alternating optimization strategy. They alternately update x and z in the data and latent spaces, performing domain transfer through the two-way transformation of NFs (i.e., $x = \mathcal{G}_\theta(z)$ and $z = \mathcal{F}_\theta(x)$). This method achieves effective dual-space alternating optimization and demonstrates promising results on natural image tasks. However, it exhibits limitations in LDCT reconstruction, where high accuracy in the reconstructed image is critical. Firstly, it does not constrain the consistency between x and z across adjacent iterations, which can negatively impact data fidelity and reconstruction accuracy.

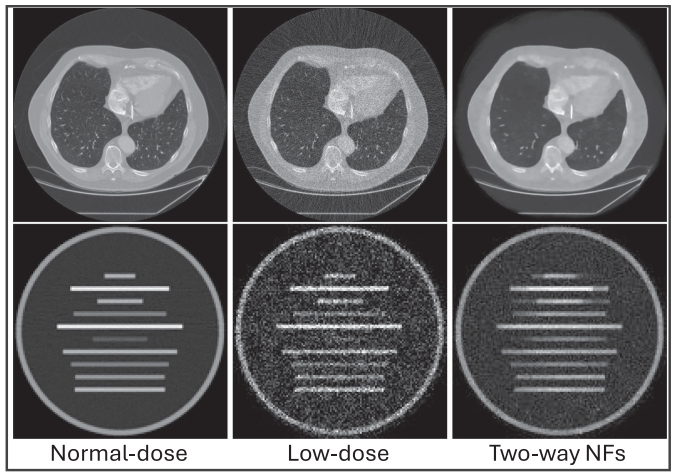


Fig. 1. Reconstructed images using the two-way transformation of NFs. Compared to the normal-dose images, structure distortion and noise residuals can be easily observed.

More importantly, the two-way transformation maps noisy images, rather than normal-dose images that follow the distribution of training samples, into the latent space using $z = \mathcal{F}_\theta(x)$. Such a mapping can cause a significant shift of the latent variable, adversely affecting subsequent iterations. In LDCT reconstruction, this issue would easily lead to detail loss and secondary artifacts, particularly under high noise levels. As shown in Fig. 1, compared to the normal-dose images, the reconstructed images using the two-way transformation exhibit severe artifacts and structural distortions. In [42], the authors propose to use the unfolding strategy to improve reconstruction performance. However, this brings the requirement for paired training data, which is challenging to obtain in clinical LDCT scenarios.

B. Conditional Normalizing Flows

Although NFs exhibit good mathematical properties, enabling efficient sampling and straightforward likelihood evaluation,

their complex network architectures with a large number of parameters make them computationally intensive. For generating low-resolution images such as those with a resolution of 64×64, NFs are more than capable. However, when applied to a high-resolution such as 512×512, the memory requirements and training time will be huge and unacceptable. For instance, training a Glow model [56] (a classic NFs model) on 5-bit 256×256 images can take approximately one week using 40 GPUs [49]. To address this challenge, several strategies have been proposed. Kothari et al. [49] introduced injective flows, termed Trumpets, which reduce the dimensionality of high-resolution images to a smaller scale, enabling the training of small-scale NFs. While this approach significantly reduces the computational costs and facilitates training NFs on high-resolution images, the dimensionality reduction will inevitably lead to information loss, resulting in detail loss in the generated images. Alternatively, conditional normalizing flows (CNFs) [57] have been proposed to address this issue by incorporating the already-known information as conditions. CNFs learn conditional probabilities rather than the full probability distribution. By introducing conditions into the training phase, CNFs can simplify the training process for high-resolution images. Additionally, by incorporating conditions during inference, CNFs can provide information for image generation, making it more efficient and adaptable to specific applications. Similar to (4), the bidirectional mapping of CNFs with a condition c can be expressed as:

$$\begin{cases} z = \mathcal{F}_\theta(x, c), \\ x = \mathcal{G}_\theta(z, c). \end{cases} \tag{7}$$

Although CNFs enable training on high-resolution images and enhance generation efficiency, their application in unsupervised LDCT reconstruction remains challenging. Current CNFs-based LDCT reconstruction methods often depend on paired data for conditionalization [50], [58], which is difficult to implement in an unsupervised framework. Additionally, these methods typically use the low-dose image itself as the condition, potentially introducing noise and artifacts into the inference process, thereby compromising reconstruction quality. While Wolf et al. [59] proposed a down-sampling conditionalization strategy for unpaired image denoising, this approach struggles to perform well when low-dose noise and artifacts are severe. Furthermore, existing methods primarily employ CNFs for pure image generation rather than integrating them into a reconstruction procedure [50], [58], [60]. This limitation often results in inconsistency between the projection data and the reconstructed image, leading to inaccuracy of the image structures.

III. Methods

Based on CNFs, we propose an end-to-end unsupervised iterative reconstruction algorithm for LDCT. Our method integrates three core components: the unsupervised conditionalization strategy, the one-way iterative LDCT reconstruction algorithm, and the CNFs network architecture. In this section, we provide a detailed description of each of these components.

A. Unsupervised Conditionalization Strategy

In the unsupervised training phase of CNFs, only normal-dose images are available. Consequently, the conditions must be derived from the normal-dose images themselves to maintain an unsupervised framework. On the other hand, during the inference stage, the only accessible information is the low-dose projection data, which implies that the condition must originate from the low-dose data. To ensure consistency between the training and inference phases, the conditions in both stages should follow the same distribution. Therefore, we propose the following two key criteria that the conditions should satisfy:

- The conditions should contain the majority of the structural and feature information in the ideal clean image, while avoiding the inclusion of redundant information such as excessive noise and artifacts.
- The two conditions corresponding to the normal-dose and low-dose data of the same object should exhibit a high degree of similarity and, ideally, be identical.

Based on the two key criteria outlined above, we propose a novel conditionalization method for CNFs-based unsupervised LDCT reconstruction. Specifically, for the low-dose projection data y, to get the intuitive image information, we first reconstruct it into an image x using a reconstruction operator $\mathcal{R}$:

$$x = \mathcal{R}(y). \tag{8}$$

Next, we remove the majority of noise and artifacts from the image using a plug-and-play denoiser $\mathcal{D}$ (e.g. BM3D [61], NLM [62] or DnCNN [63]) and a high-frequency-filter wavelet reconstruction operator $\mathcal{W}$:

$$c' = \mathcal{W}(\mathcal{D}(x)). \tag{9}$$

To enhance the robustness of the condition, we add a low-level Gaussian noise $n \sim \mathcal{N}(0, \sigma_1^2)$ to c':

$$c = c' + n, \tag{10}$$

where σ_1 is the standard deviation of the added noise. The entire process $\mathcal{C}$ for generating the condition c from the projected data y can be expressed as:

$$c = \mathcal{C}(y) = \mathcal{W}(\mathcal{D}(\mathcal{R}(y))) + n. \tag{11}$$

During the training stage, we replace $\mathcal{R}(y)$ in the above equation with the normal-dose image x and apply the same $\mathcal{W}$ and $\mathcal{D}$ to derive the conditions for the normal-dose images. It is important to note that although the denoiser is applied to the normal-dose image, it does not significantly compromise the image structures or details. Instead, it produces a condition that is consistent with the low-dose image, ensuring consistency between the training and inference phases.

Using our proposed conditionalization method, we can generate conditions for both normal-dose and low-dose data for unsupervised LDCT reconstruction. To assess the quality of the obtained conditions and fine-tune parameters—such as the noise level in BM3D or NLM—we require a pair of normal-dose and low-dose data. The evaluation criteria are based on the SSIM [64] between the paired conditions and the SSIM between these conditions and the clean normal-dose image. Higher SSIM

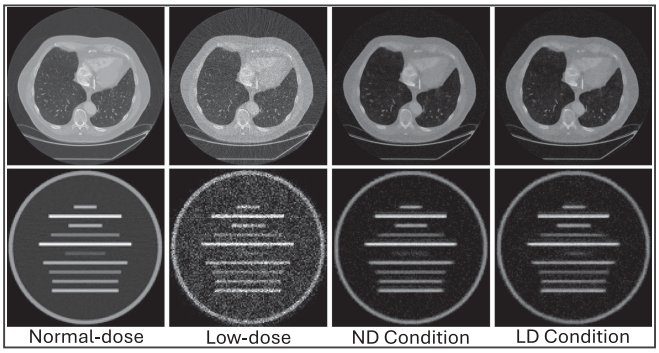


Fig. 2. Condition examples generated with the BM3D denoiser.

values generally indicate conditions with better quality. As an illustration, Fig. 2 showcases examples of conditions generated using the BM3D denoiser.

B. One-Way CNFs LDCT Reconstruction Algorithm

To address the limitations of current dual-space alternating optimization methods, we propose a novel iterative reconstruction algorithm with two key improvements: (1) To maintain consistency between x and z across iterations, we introduce constraint terms and update x and z separately; (2) Instead of the two-way transformation strategy, we adopt a strict one-way transformation between the image x and the latent variable z during alternating optimization, thereby effectively preventing secondary artifacts. Specifically, using the trained CNFs $\mathcal{G}_\theta$, we reformulated the problem in (6) into the following double-variable minimization problem with a constraint term:

$$(\hat{x},\hat{z}) = \arg\min_{x,z} \|y-\mathbf{A}x\|_2^2 + \lambda\|z\|_2^2 + \sigma\|x-\mathcal{G}_\theta(z,c)\|_2^2, \tag{12}$$

where λ and σ are model parameters that control the regularization strength and the proximity between x and $\mathcal{G}_\theta(z,c)$, respectively. Using the coordinate descent algorithms [65], this problem can be decomposed into the following two sub-problems:

$$x^{n+1} = \arg\min_x \|y-\mathbf{A}x\|_2^2 + \sigma\|x-\mathcal{G}_\theta(z^n,c)\|_2^2 + r_1\|x-x^n\|_2^2, \tag{13}$$

and

$$z^{n+1} = \arg\min_z \lambda\|z\|_2^2 + \sigma\left\|x^{n+1}-\mathcal{G}_\theta(z,c)\right\|_2^2 + r_2\|z-z^n\|_2^2, \tag{14}$$

where r_1 and r_2 are manual algorithm parameters that regulate the proximity of x and z between adjacent iterations. Notably, our method only incorporates the generation process $x=\mathcal{G}_\theta(z,c)$ of the CNFs, enforcing a strict one-way transformation from the latent variable z to the image x. This approach elegantly avoids the issues of detail loss and secondary artifacts associated with the two-way transformation.

The sub-problem in (13) admits a closed-form solution:

$$x^{n+1} = (\mathbf{A}^T\mathbf{A}+\sigma\mathbf{I}+r_1\mathbf{I})^{-1}(\mathbf{A}^Ty+\sigma\mathcal{G}_\theta(z^n,c)+r_1x^n). \tag{15}$$

To solve the sub-problem in (14), we need the derivative of $\mathcal{G}_\theta(z,c)$ with respect to z. However, despite the differentiability of CNFs, expressing the derivative of such a network explicitly is challenging. To address this, we linearize the second term in (14), which is difficult to differentiate directly. Let $s(z) = \|x^{n+1}-\mathcal{G}_\theta(z,c)\|_2^2$, based on the diffeomorphism property of $\mathcal{G}_\theta$, $s(z)$ can be approximated by its first-order Taylor expansion at z^n:

$$\begin{aligned} s(z) \approx\ & s(z^n) + < s'(z^n), (z-z^n) > = \left\|x^{n+1}-\mathcal{G}_\theta(z^n,c)\right\|_2^2 \\ & -2\nabla_z\mathcal{G}_\theta(z^n,c)(x^{n+1}-\mathcal{G}_\theta(z^n,c))(z-z^n). \end{aligned} \tag{16}$$

By substituting this approximation into the sub-problem in (14), the solution z^{n+1} is given by:

$$z^{n+1} = \frac{\sigma\nabla_z\mathcal{G}_\theta(z^n,c)(x^{n+1}-\mathcal{G}_\theta(z^n,c))+r_2z^n}{\lambda+r_2}, \tag{17}$$

where $\nabla_z\mathcal{G}_\theta(z^n,c)(x^{n+1}-\mathcal{G}_\theta(z^n,c))$ can be computed using the automatic derivation of the following loss function with respect to z^n using the **Pytorch** tools:

$$\mathcal{L} = -\tfrac{1}{2}\left\|x^{n+1}-\mathcal{G}_\theta(z^n,c)\right\|_2^2. \tag{18}$$

In summary, the iteration process of our unsupervised one-way CNFs reconstruction algorithm can be expressed as:

$$\begin{cases} x^{n+1} = (\mathbf{A}^T\mathbf{A}+\sigma\mathbf{I}+r_1\mathbf{I})^{-1}(\mathbf{A}^Ty+\sigma\mathcal{G}_\theta(z^n,c)+r_1x^n), \\ z^{n+1} = \dfrac{\sigma\nabla_z\mathcal{G}_\theta(z^n,c)(x^{n+1}-\mathcal{G}_\theta(z^n,c))+r_2z^n}{\lambda+r_2}, \\ \hat{x} = \mathcal{G}_\theta(\hat{z},c) = \mathcal{G}_\theta(z^{\mathrm{K}},c), \end{cases} \tag{19}$$

where $\hat{z}$ is the final output of z after the K_{th} iteration, and $\hat{x}$ is the final reconstructed image generated by the CNFs using $\hat{z}$ and the condition c.

In the iteration of x^{n+1}, computing the inverse of $(\mathbf{A}^T\mathbf{A}+\sigma\mathbf{I}+r_1\mathbf{I})$ is required. However, since the projection matrices $\mathbf{A}$ vary for each projection angle, using the matrix form of $\mathbf{A}$ in computation is impractical due to excessive storage and computational demands. On the other hand, if $\mathbf{A}$ is treated as an operator, solving the inverse of $(\mathbf{A}^T\mathbf{A}+\sigma\mathbf{I}+r_1\mathbf{I})$ is difficult. To address this, we employ a two-step incremental reconstruction strategy to approximate x^{n+1}. First, we reconstructed an intermediate image $x^{n+\frac{1}{2}}$ using the ordered-subset simultaneous algebraic reconstruction technique (OS-SART) [53], with x^n as the initial value:

$$x^{n+\frac{1}{2}} = \mathbf{OS}-\mathbf{SART}(x^n,y,\omega) \tag{20}$$

where ω is the relaxation factor. This step can be seen as minimizing the first term in (13). Notably, using OS-SART to solve the fidelity term is more appropriate than minimizing the l_2 norm, as modeling the low-dose noise in projection data with a fixed-variance Gaussian distribution is often inaccurate. In contrast, the reconstruction process of OS-SART better respects the noise characteristics of the projection data. In the second step, the solution x^{n+1} is computed without requiring matrix

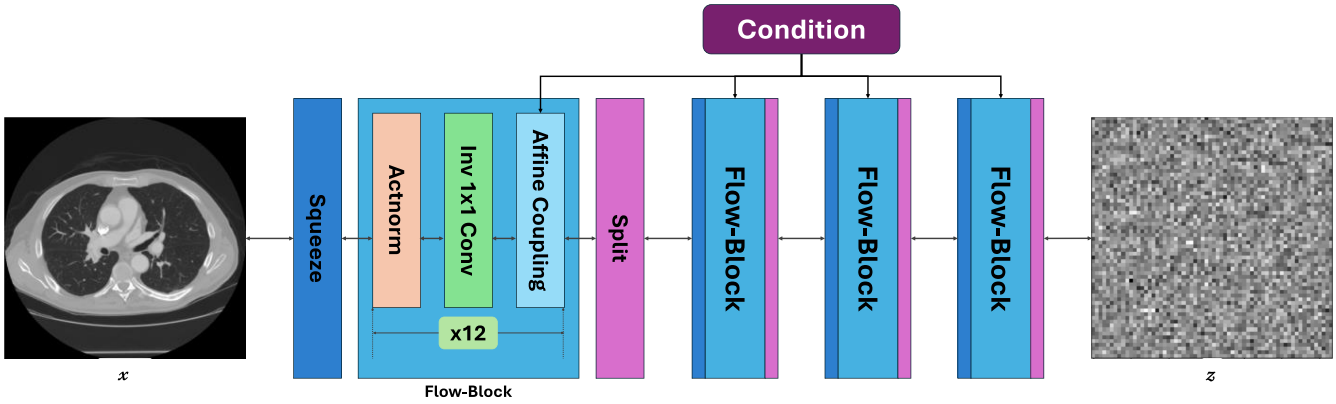


Fig. 3. The network structure of our conditional normalizing flows.

Algorithm 1: One-Way Conditional Normalizing Flows (OW-CNFs) Unsupervised LDCT Reconstruction Algorithm.

Input: The noisy projection data y, the hyperparameters λ, σ, r_1 and r_2, the relaxation parameter ω of OS-SART, the random Gaussian initialization z^0 and $x^0 = \mathcal{G}_\theta(z^0, c)$, and the number of iterations K.
Operators: The NFs $\mathcal{G}_\theta$ and **OS-SART**.
Output: The final reconstructed image $\hat{x}$.
for $n = 0, \ldots, \mathrm{K}-1$:
$x^{n+\frac{1}{2}} = \mathbf{OS-SART}(x^n, y, \omega)$
$x^{n+1} = \dfrac{x^{n+\frac{1}{2}} + \sigma\mathcal{G}_\theta(z^n, c) + r_1 x^n}{1 + \sigma + r_1}$
$z^{n+1} = \dfrac{\sigma\nabla_z \mathcal{G}_\theta(z^n, c)(x^{n+1} - \mathcal{G}_\theta(z^n, c)) + r_2 z^n}{\lambda + r_2}.$
end
$\hat{x} = \mathcal{G}_\theta(z^{\mathrm{K}}, c)$
return $\hat{x}$

inversion, as follows:

$$x^{n+1} = \frac{x^{n+\frac{1}{2}} + \sigma\mathcal{G}_\theta(z^n, c) + r_1 x^n}{1 + \sigma + r_1}. \tag{21}$$

This step can be seen as minimizing the remaining terms in (13). For OS-SART, we set the number of iterations to 1 and carefully tuned the relaxation factor ω to achieve a good balance between data fidelity and the CNFs priors. This ensures that structure distortion is avoided while achieving adequate denoising. The overall flow of our one-way CNFs LDCT reconstruction algorithm is summarized in Algorithm 1.

C. Network

The overall network of our CNFs consists of two main components: a backbone NFs network and a condition module. The backbone NFs network comprises 4 flow blocks, each containing 12 tandem combinations of the Actnorm layer, the invertible 1×1 convolutional layer, and the affine coupling layer, following a structure similar to the well-known Glow model [56]. Each convolutional layer in the network has 512 feature channels. In the condition module, the condition c is integrated into each affine coupling layer of the backbone NFs by concatenating it with the mainstream features. The overall network structure of our CNFs is illustrated in Fig. 3.

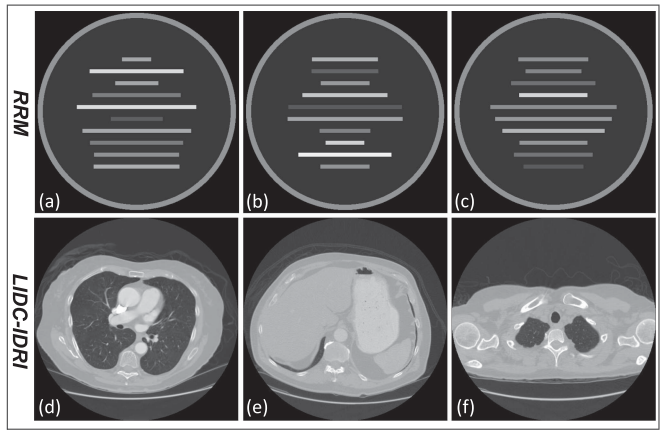


Fig. 4. Some image examples in the RRM dataset ((a)–(c)) and the LIDC-IDRI dataset ((d)–(f)).

IV. Experiments

A. Experimental Setup

To evaluate the performance of the proposed one-way conditional normalizing flows (**OW-CNFs**) unsupervised LDCT reconstruction algorithm, we conducted experiments on two datasets. We assessed the reconstructed images using both visual inspection and quantitative metrics, comparing the results with several state-of-the-art deep-learning methods. These include popular unsupervised generative-model-based methods such as PatchNR [43], EASEL [44], DPR-IR-I, and DPR-IR-II [47], as well as supervised LDCT denoising networks like EDCNN [66], DESDGAN [67], CoreDiff [68], and CTFormer [17].

The experiments were conducted on two datasets: (1) The random rectangle models (RRM) dataset and (2) The LIDC-IDRI dataset [69]. The RRM dataset is home-made, consisting of randomly generated images, each containing a large ring and ten parallel stripe rectangles with varying lengths and gray values, some examples are shown in Fig. 4(a)–(c). The reconstruction performance of different methods can be directly evaluated

based on the reconstruction of these rectangular stripes. This dataset includes 1024 images for training, 128 images for validation and 128 images for testing. Each image has a resolution of 128×128 and a gray value range of $[0, 1.0]$. The LIDC-IDRI dataset is a public CT image dataset containing 241689 normal-dose slices from 1012 patients. For our experiments, we selected 4409 slices from patients 001 to 025 for training, 272 slices from patient 052 for validation, and 133 slices from patient 051 for testing. Examples of these slices are shown in Fig. 4(d)–(f). All selected CT slices have a resolution of 512×512 and a Hounsfield units (HU) range of $[-1024, 2048]$.

The normal-dose and low-dose projection data were generated using a simulated projection algorithm with a fan-beam imaging system. For the RRM dataset, we employed 360 projection views uniformly distributed within the range $[0, 2\pi]$ and a linear detector comprising 256 cells. For the LIDC-IDRI dataset, both the projection views and the number of detector cells were set to 1000. We generated the clean projection data and added Poisson noise to the incident rays to emulate the low-dose projection data:

$$y_n = -\ln\left(\frac{I_d}{I_0}\right), I_d \sim \text{Poisson}\{I_0 \times e^{-y_c}\}, \qquad (22)$$

where y_c represents the clean projection data, while y_n the noisy low-dose one. I_0 and I_d signify the number of incident and collected photons, respectively. Typically, a smaller I_0 indicates a lower dose level and results in more noisy projection data. In accordance with this principle, we set $I_0 = 1 \times 10^6$ to simulate the normal-dose projection data, with the corresponding normal-dose images reconstructed using OS-SART for reference. For the generation of low-dose projection data, we assigned $I_0 = 1 \times 10^3$ for the RRM dataset, and $I_0 = 1 \times 10^4$ for the LIDC-IDRI dataset. The projection operator and OS-SART were coded using CUDA kernels wrapped by the **Cupy** library (https://github.com/cupy/cupy).

We selected an NFs-based unsupervised regularized iterative reconstruction method PatchNR [43], three diffusion-model-based unsupervised iterative reconstruction methods EASEL [44], DPR-IR-I and DPR-IR-II [47], and four recent supervised LDCT denoising methods EDCNN [66], DESDGAN [67], CoreDiff [68] and CTFormer [17], as the comparative methods, in which DPR-IR-II utilizes the DDIM [51] acceleration strategy for fast reconstruction. Additionally, we implemented the two-way iterative reconstruction algorithm [42] introduced in Section II, utilizing the same CNFs network as our proposed method, referred to as "TW-CNFs". For evaluation, we adopted the commonly used PSNR and SSIM [64] as image quality assessment metrics. Codes for all comparative methods are publicly available. During the training phase of each method, we used the loss function and optimizer specified in the original training methodology, selecting appropriate learning rates and mini-batch sizes. For the conditionalization part of our proposed OW-CNFs, we chose BM3D [61] as the denoiser $\mathcal{D}$. Given that generative-model-based iterative reconstruction methods, including EASEL, DPR-IR-I, DPR-IR-II, PatchNR, TW-CNFs, and our proposed OW-CNFs, exhibit randomness in their reconstructions, we performed 10 reconstructions simultaneously and averaged the results to obtain the final reconstruction. Regarding the hyper-parameters λ, σ, r_1 and r_2 in the proposed OW-CNFs, we empirically preset them and tuned them by the "trial and error" strategy, following the convergence state of the iteration. In experiments on the RRM dataset, these four parameters were set to 0.0005, 10.0, 0.001 and 0.01, respectively. For experiments on the LIDC-IDRI dataset, they were set to 0.0001, 1.5, 0.01 and 0.01, respectively. All experiments were conducted on a server running Ubuntu 20.04.5 with Python 3.11, PyTorch 1.12.1, and CUDA 11.3, equipped with an Nvidia Tesla V100 GPU card.

B. Results

On both the RRM and LIDC-IDRI datasets, the proposed OW-CNFs demonstrate outstanding performance. Fig. 5 illustrates one set of reconstruction results on the RRM dataset at dose level $I_0 = 1 \times 10^3$. The result of the proposed OW-CNFs exhibits comparable quality to the state-of-the-art generative-model-based methods and the supervised methods. As highlighted in the zoomed-in areas, the three diffusion-model-based methods—EASEL, DPR-IR-I, and DPR-IR-II—show effective reconstruction of the stripes but retain noticeable noise residues, with EASEL being particularly affected. The NFs-based method, PatchNR, performs well in denoising but introduces block artifacts in the reconstructed image. This issue likely arises because PatchNR learns priors from image patches rather than the entire image, leading to some loss of global information and a tendency to misinterpret large noise blocks as image structures. The TW-CNFs method exhibits noise retention and structural deformation in the reconstructed image. For instance, in the zoomed-in area, the length of the lower stripe is incorrect, and its right end is severely blurred or even missing. Given the simplicity and regularity of the image, the four supervised methods—EDCNN, DESDGAN, CoreDiff, and CTFormer—achieve satisfactory denoising results. Notably, the proposed unsupervised OW-CNFs perform comparably to these supervised methods and even surpass them in quantitative metrics. In addition to effective denoising, the proposed OW-CNFs accurately reconstruct every stripe in the image, demonstrating outstanding performance.

Figs. 6, 7, and 8 present the reconstruction results on the LIDC-IDRI dataset at dose level $I_0 = 1 \times 10^4$. The proposed OW-CNFs demonstrate superior performance. As shown in the zoomed-in areas, OW-CNFs exhibit high accuracy and sharpness in recovering image structures, while achieving high-quality denoising. In contrast, although the diffusion-model-based methods EASEL, DPR-IR-I and DPR-IR-II perform well in structure recovery, they are plagued by residual noise and artifacts that can obscure image details. The NFs-based method PatchNR still shows pseudo structures, while TW-CNFs suffer from severe structure deformation and loss of image detail. Among the four supervised methods, the GAN-based DES-DGAN exhibits severe artifacts, while the results of EDCNN and CTFormer show obvious structural distortions. Although the diffusion-model-based denoising method CoreDiff demonstrates better overall effect, it lacks precision in image details and contrast. The average metrics for the test sets are shown in Table I. Notably, as an unsupervised method, the proposed

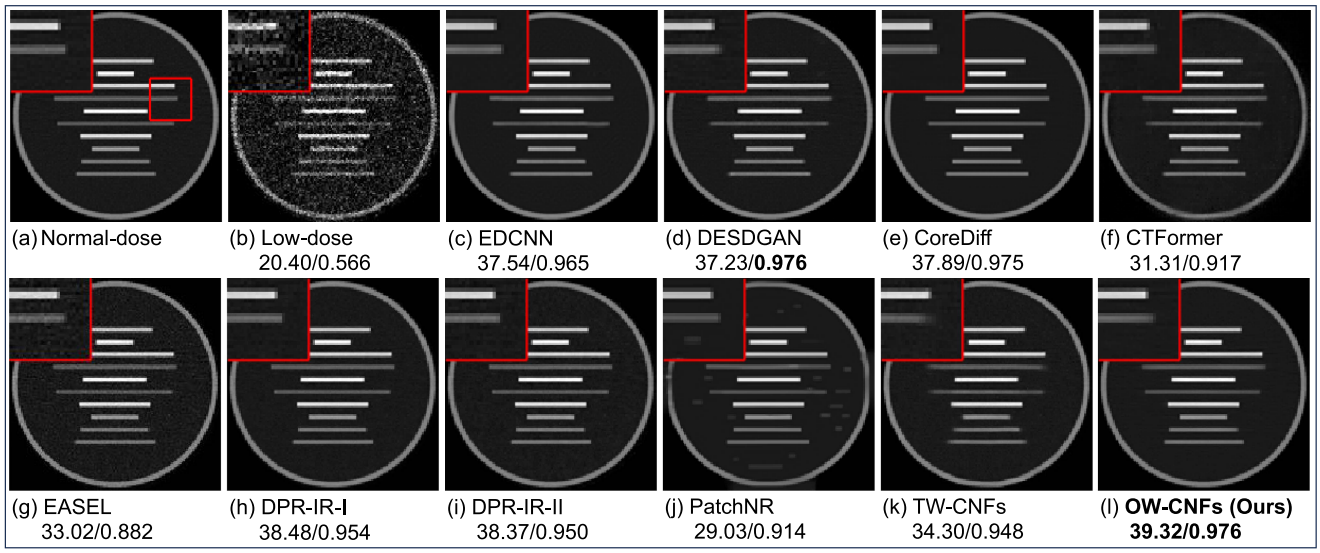


Fig. 5. Reconstruction results of each method on the RRM dataset at dose level $I_0 = 1 \times 10^3$. The display window of the gray value range is set to [0,1.0].

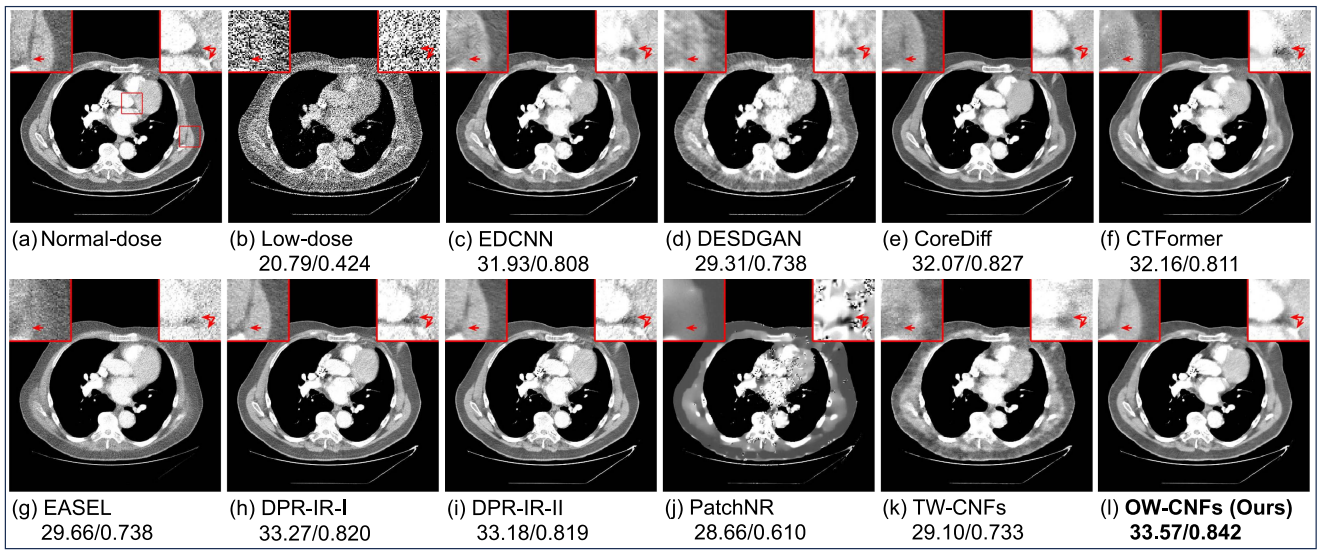


Fig. 6. Reconstructed image results of each method on the LIDC-IDRI dataset at dose level $I_0 = 1 \times 10^4$. The display window is $[-240, 160]$HU.

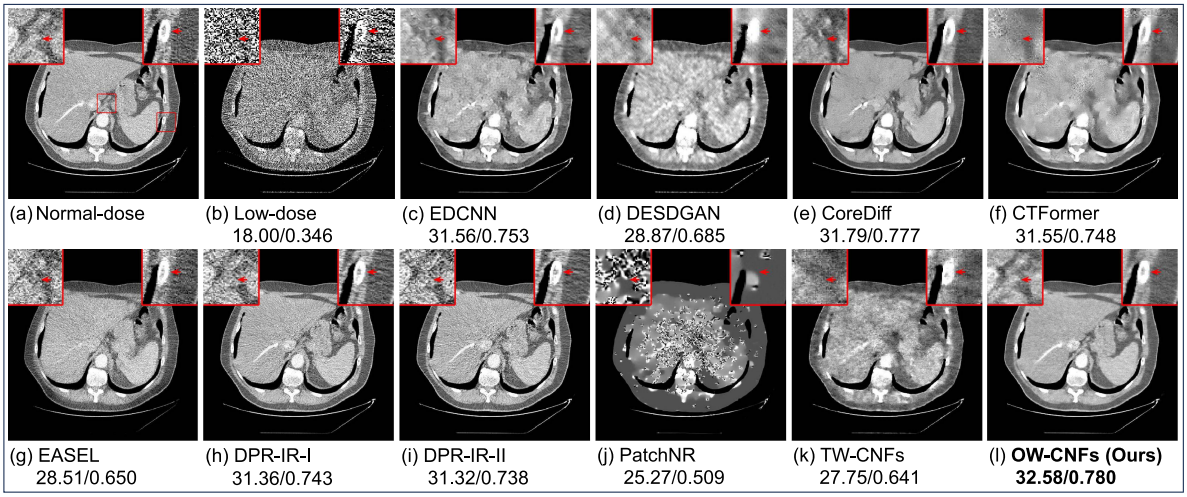


Fig. 7. Reconstructed image results of each method on the LIDC-IDRI dataset at dose level $I_0 = 1 \times 10^4$. The display window is $[-240, 160]$HU.

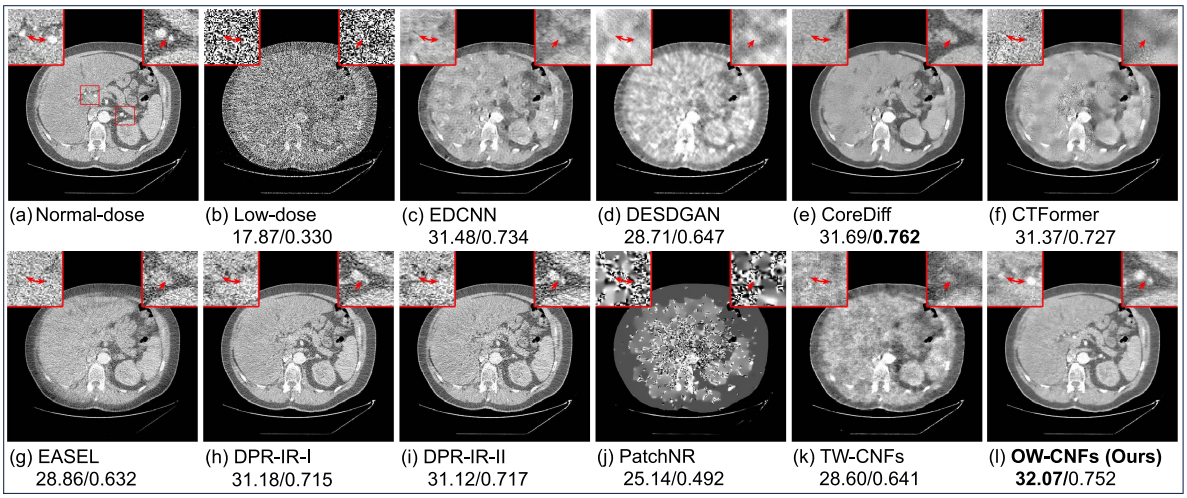


Fig. 8. Reconstructed image results of each method on the LIDC-IDRI dataset at dose level $I_0 = 1 \times 10^4$. The display window is $[-240, 160]$HU.

TABLE I
THE AVERAGE METRICS (PSNR(DB)/SSIM) OF EACH METHOD ON THE TEST SETS OF THE RRM AND THE LIDC-IDRI DATASETS

Dataset	RRM		LIDC-IDRI	
Method	**PSNR**	**SSIM**	**PSNR**	**SSIM**
SART	20.48	0.567	19.84	0.395
EASEL	33.06	0.884	29.33	0.703
DPR-IR-I	38.38	0.954	32.62	0.806
DPR-IR-II	38.31	0.951	32.40	0.801
Patch-NR	29.42	0.921	28.62	0.707
TW-CNFs	32.74	0.934	28.89	0.712
OW-CNFs (Ours)	**38.97**	**0.972**	**32.95**	**0.813**
EDCNN	37.31	0.964	31.79	0.786
DESDGAN	35.35	0.970	28.99	0.784
CoreDiff	38.77	0.970	31.91	0.807
CTFormer	31.32	0.916	31.77	0.782

TABLE II
THE AVERAGE METRICS (PSNR(DB)/SSIM), ITERATION NUMBERS AND INFERENCE TIME OF EACH GENERATIVE-MODEL-BASED METHOD BY PERFORMING SINGLE AND 10 SIMULTANEOUS RECONSTRUCTIONS, ON THE LIDC-IDRI DATASET

10 Reconstructions	**PSNR**	**SSIM**	**Iters**	**Time**
EASEL	29.33	0.703	1800	4650s
DPR-IR-I	32.62	0.806	1000	1814s
DPR-IR-II	32.40	0.801	200	361s
Patch-NR	28.62	0.707	409	509s
TW-CNFs	28.89	0.712	**33**	**131s**
OW-CNFs (Ours)	**32.95**	**0.813**	71	344s
1 Reconstruction	**PSNR**	**SSIM**	**Iters**	**Time**
EASEL	26.88	0.502	1800	620s
DPR-IR-I	32.23	0.800	1000	364s
DPR-IR-II	32.17	0.799	200	72s
Patch-NR	28.03	0.681	424	176s
TW-CNFs	28.15	0.688	**36**	**26s**
OW-CNFs (Ours)	**32.48**	**0.803**	76	67s

OW-CNFs achieve the highest average PSNR and SSIM values, surpassing all the four supervised methods.

To evaluate the reconstruction speed of the proposed OW-CNFs, we compared the average iteration numbers and reconstruction time of each generative-model-based iterative reconstruction method on 512×512 CT images from the LIDC-IDRI dataset. In previous experiments, we adopted a strategy of performing 10 reconstructions simultaneously and averaging the results to ensure robustness. However, since all these iterative methods involve time-consuming reconstruction procedure (e.g. OS-SART) in each iteration, the number of simultaneous reconstructions significantly impacts their overall reconstruction time. To provide a comprehensive comparison of performance and speed, we compared the results of each method under two scenarios: performing 1 reconstruction and performing 10 simultaneous reconstructions. As shown in Table II, the proposed OW-CNFs achieve a good balance between performance and reconstruction speed. The diffusion-model-based methods, EASEL and DPR-IR-I, require iterations at every predefined noise level, resulting in a large number of iterations and prolonged reconstruction times. Although DPR-IR-II, which employs the DDIM acceleration strategy for fast sampling, reduces the sampling steps to 200 without significant performance degradation, our method demonstrates comparable speed while delivering superior reconstruction performance. PatchNR, while faster than non-accelerated diffusion-model-based methods, exhibits relatively lower performance. TW-CNFs achieve the fastest reconstruction speed due to fewer iterations, but their performance is significantly inferior to the proposed OW-CNFs. In contrast, the proposed OW-CNFs combine high performance with relatively fast reconstruction speed, highlighting their advantages. Notably, performing 10 simultaneous reconstructions and averaging the results can enhance the performance of these generative-model-based methods, but this approach also increases the reconstruction time by multiples.

In the proposed OW-CNFs, the relaxation parameter ω of OS-SART plays a critical role in balancing data fidelity and CNFs priors. Variations in ω directly impact reconstruction quality. Fig. 9 illustrates the trend of average PSNR on the test sets for different values of ω. Overall, ω should remain small, but excessively small values can compromise data fidelity and reduce reconstruction accuracy. To address this, we utilized the automatic backpropagation of **Pytorch** to adaptively adjust ω. By initializing ω with a small value (e.g. $\omega = 0.1$) and using

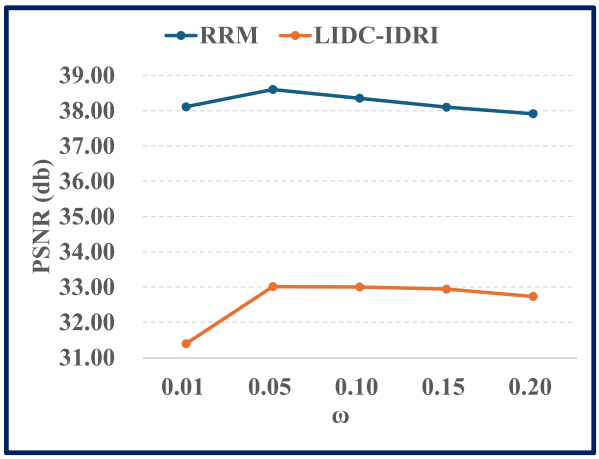


Fig. 9. The trend of average PSNR under different ω.

the minimization problem in (12) as the loss function, our reconstruction process can adaptively learn appropriate relaxation factors for different reconstructed images.

V. Discussion and Conclusion

We propose a novel unsupervised LDCT iterative reconstruction algorithm, OW-CNFs, based on conditional normalizing flows. By introducing an iterative reconstruction algorithm that employs strict one-way transformation during alternating optimization in the data and latent spaces, we address the issues of detail loss and secondary artifacts inherent in the current two-way transformation strategy. To enable unsupervised LDCT reconstruction with CNFs, we propose a conditionalization method tailored for LDCT, facilitating efficient training on high-resolution CT images. Furthermore, we propose an efficient iterative reconstruction procedure by employing the coordinate descent algorithms and the linearization technique. By leveraging the incremental reconstruction of OS-SART, we achieve computationally efficient iterations and effectively balance data fidelity with CNFs priors. Experimental results demonstrate the outstanding performance of our method, which achieves competitive reconstruction speeds among generative-model-based iterative reconstruction methods and delivers performance comparable to supervised learning approaches.

As a generative-model-based method, the proposed OW-CNFs effectively overcome the primary limitations of CNN-based denoising networks, which often result in over-smoothing and detail loss in CT images. Instead, our method produces images that better align with diagnostic requirements. Compared to diffusion-model-based methods, which also stem from generative models, the key advantage of OW-CNFs lies in eliminating the need for a fixed, large number of sampling steps (often up to a thousand), thereby significantly reducing reconstruction time. This efficiency is largely attributed to our conditionalization method, which simplifies the iterative process by providing additional priors through the condition. While we are not the first to apply CNFs to LDCT problems, we are the first to effectively and deeply integrate CNFs into the unsupervised LDCT reconstruction procedure. Unlike other methods that use CNFs solely for image-domain post-processing and require extensive paired training data, our method integrates the priors learned by CNFs directly into the LDCT reconstruction procedure, avoiding the image-data inconsistency issues commonly associated with post-processing networks. Moreover, our method operates within an unsupervised framework, requiring only normal-dose images for training, making it highly applicable in practical CT scenarios where training data, especially paired data, is scarce.

Despite the promising performance, the proposed OW-CNFs could be further improved. First, the introduction of multiple hyperparameters necessitates extensive tuning, which can be burdensome. While we have addressed this partially by automating the adjustment of the relaxation parameter ω, how to process other hyperparameters needs further research. Second, the quality of the conditions significantly impacts the training and generation processes of CNFs. Developing methods to generate conditions with higher accuracy and stronger priors for LDCT reconstruction remains an open challenge. Although our proposed conditionalization method is simple and effective, it relies on basic quantitative indicators and manual adjustments. Incorporating more advanced information extraction techniques, such as pre-trained feature learning networks, could enhance the performance and speed of our method. Additionally, replacing OS-SART with faster reconstruction approaches, such as analytic reconstruction or sparse-angle reconstruction, could further accelerate the iteration and is a promising direction for future research.

Acknowledgment

The authors are also grateful to Beijing Higher Institution Engineering Research Center of Testing and Imaging for funding this research work.

References

[1] A. Lideståhl, G. Johansson, A. Siegbahn, and P. A. Lind, "Estimated risk of radiation-induced cancer after thymoma treatments with proton-or X-ray beams," *Cancers*, vol. 13, no. 20, 2021, Art. no. 5153.

[2] D. Z. Solomon, B. Ayalew, S. T. Dellie, and D. Admasie, "Justification and optimization principles of ALARA in pediatric CT at a teaching hospital in Ethiopia," *Ethiopian J. Health Sci.*, vol. 30, no. 5, pp. 761–766, 2020.

[3] L. W. Goldman, "Principles of CT: Radiation dose and image quality," *J. Nucl. Med. Technol.*, vol. 35, no. 4, pp. 213–225, 2007.

[4] Y. Chen et al., "Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior," *Comput. Med. Imag. Graph.*, vol. 33, no. 7, pp. 495–500, 2009.

[5] D. Kang et al., "Image Denoising of Low-Radiation Dose Coronary Ct Angiography by an Adaptive Block-Matching 3D Algorithm," in *Medical Imaging 2013: Image Processing*, vol. 8669. Bellingham, WA, USA, 2013, pp. 671–676.

[6] M. Green, E. M. Marom, N. Kiryati, E. Konen, and A. Mayer, "Efficient low-dose CT denoising by locally-consistent non-local means (LC-NLM)," in *Proc. Med. Image Comput. Comput.- Assist. Interv.: 19th Int. Conf.*, Athens, Greece, 2016, pp. 423–431.

[7] Y. Zhang, M. Salehjahromi, and H. Yu, "Tensor decomposition and non-local means based spectral CT image denoising," *J. X-ray Sci. Technol.*, vol. 27, no. 3, pp. 397–416, 2019.

[8] Q. Zhang, Z. Gui, Y. Chen, Y. Li, and L. Luo, "Bayesian sinogram smoothing with an anisotropic diffusion weighted prior for low-dose x-ray computed tomography," *Optik- Int. J. Light Electron Opt.*, vol. 124, no. 17, pp. 2811–2816, 2013.

[9] A. Manduca et al., "Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT," *Med. Phys.*, vol. 36, no. 11, pp. 4911–4919, 2009.

[10] L. I. Rudin and S. Osher, "Total variation based image restoration with free local constraints," in *Proc. IEEE 1st Int. Conf. Image Process.*, 1994, pp. 31–35.

[11] H. Yu and G. Wang, "Compressed sensing based interior tomography," *Phys. Med. Biol.*, vol. 54, no. 9, 2009, Art. no. 2791.

[12] 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, 2011, Art. no. 5949.

[13] X. Deng, W. Xu, and H. Li, "Fast ordered subsets chambolle-pock algorithm for CT reconstruction," in *Proc. 15th Int. Meeting Fully Three-Dimensional Image Reconstruction Radiol. Nucl. Med.*, 2019, pp. 454–458.

[14] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[15] H. Chen et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.

[16] J. Liu et al., "Deep iterative reconstruction estimation (dire): Approximate iterative reconstruction estimation for low dose CT imaging," *Phys. Med. Biol.*, vol. 64, no. 13, 2019, Art. no. 135007.

[17] D. Wang, F. Fan, Z. Wu, R. Liu, F. Wang, and H. Yu, "CTformer: Convolution-free token2token dilated vision transformer for low-dose CT denoising," *Phys. Med. Biol.*, vol. 68, no. 6, Art. no. 065012, 2023.

[18] Z. Feng, A. Cai, Y. Wang, L. Li, L. Tong, and B. Yan, "Dual residual convolutional neural network (DRCNN) for low-dose CT imaging," *J. X-Ray Sci. Technol.*, vol. 29, no. 1, pp. 91–109, 2021.

[19] R. Ge et al., "DDPNet: A novel dual-domain parallel network for low-dose CT reconstruction," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2022, pp. 748–757.

[20] B. Zhou, X. Chen, H. Xie, S. K. Zhou, J. S. Duncan, and C. Liu, "DuDoUFNet: Dual-domain under-to-fully-complete progressive restoration network for simultaneous metal artifact reduction and low-dose CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 41, no. 12, pp. 3587–3599, Dec. 2022.

[21] W. Wu, D. Hu, C. Niu, H. Yu, V. Vardhanabhuti, and G. Wang, "DRONE: Dual-domain residual-based optimization network for sparse-view CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3002–3014, Nov. 2021.

[22] H. Chen et al., "LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1333–1347, Jun. 2018.

[23] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1322–1332, Jun. 2018.

[24] 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.

[25] C. You et al., "CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE)," *IEEE Trans. Med. Imag.*, vol. 39, no. 1, pp. 188–203, Jan. 2020.

[26] T. Kwon and J. C. Ye, "Cycle-free cycleGAN using invertible generator for unsupervised low-dose CT denoising," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 1354–1368, 2021.

[27] J. Gu and J. C. Ye, "Adain-based tunable cycleGAN for efficient unsupervised low-dose CT denoising," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 73–85, 2021.

[28] Z. Li, S. Zhou, J. Huang, L. Yu, and M. Jin, "Investigation of low-dose CT image denoising using unpaired deep learning methods," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, no. 2, pp. 224–234, Mar. 2021.

[29] A. M. Hasan, M. R. Mohebbian, K. A. Wahid, and P. Babyn, "Hybrid-collaborative noise2noise denoiser for low-dose CT images," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, no. 2, pp. 235–244, Mar. 2021.

[30] W. Fang et al., "Iterative material decomposition for spectral CT using self-supervised noise2noise prior," *Phys. Med. Biol.*, vol. 66, no. 15, 2021, Art. no. 155013.

[31] N. Yuan, J. Zhou, and J. Qi, "Half2Half: Deep neural network based CT image denoising without independent reference data," *Phys. Med. Biol.*, vol. 65, no. 21, 2020, Art. no. 215020.

[32] C. Zhang, S. Chang, T. Bai, and X. Chen, "S2ms: Self-supervised learning driven multi-spectral CT image enhancement," in *Proc. 7th Int. Conf. Image Formation X-Ray Computed Tomogr.*, 2022, pp. 473–479.

[33] C. Niu et al., "Noise suppression with similarity-based self-supervised deep learning," *IEEE Trans. Med. Imag.*, vol. 42, no. 6, pp. 1590–1602, Jun. 2023.

[34] A. A. Hendriksen, D. M. Pelt, and K. J. Batenburg, "Noise2inverse: Self-supervised deep convolutional denoising for tomography," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1320–1335, 2020.

[35] F. Wagner et al., "On the benefit of dual-domain denoising in a self-supervised low-dose CT setting," in *Proc. IEEE 20th Int. Symp. Biomed. Imag.*, 2023, pp. 1–5.

[36] C. Niu, M. Li, X. Guo, and G. Wang, "Self-supervised dual-domain network for low-dose CT denoising," in *Developments in X-Ray Tomography XIV*, vol. 12242. Bellingham, WA, USA, 2022.

[37] R. An, K. Chen, and H. Li, "Self-supervised dual-domain balanced dropblock-network for low-dose CT denoising," *Phys. Med. Biol.*, vol. 69, no. 7, Mar. 2024, doi: 10.1088/1361-6560/ad29ba. [Online]. Available: https://dx.doi.org/10.1088/1361-6560/ad29ba

[38] A. Zheng, H. Gao, L. Zhang, and Y. Xing, "A dual-domain deep learning-based reconstruction method for fully 3D sparse data helical CT," *Phys. Med. Biol.*, vol. 65, no. 24, 2020, Art. no. 245030.

[39] Y. Zhang et al., "CLEAR: Comprehensive learning enabled adversarial reconstruction for subtle structure enhanced low-dose CT imaging," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3089–3101, Nov. 2021.

[40] I. Kobyzev, S. J. Prince, and M. A. Brubaker, "Normalizing flows: An introduction and review of current methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 11, pp. 3964–3979, Nov. 2021.

[41] F.-A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 10850–10869, Sep. 2023.

[42] X. Wei, H. van Gorp, L. Gonzalez-Carabarin, D. Freedman, Y. C. Eldar, and R. J. van Sloun, "Deep unfolding with normalizing flow priors for inverse problems," *IEEE Trans. Signal Process.*, vol. 70, pp. 2962–2971, 2022.

[43] F. Altekrüger, A. Denker, P. Hagemann, J. Hertrich, P. Maass, and G. Steidl, "Patchnr: Learning from very few images by patch normalizing flow regularization," *Inverse Problems*, vol. 39, no. 6, 2023, Art. no. 064006.

[44] Z. He et al., "Iterative reconstruction for low-dose CT using deep gradient priors of generative model," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 6, no. 7, pp. 741–754, Sep. 2022.

[45] Y. Song and S. Ermon, "Improved techniques for training score-based generative models," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 12438–12448.

[46] X. Liu et al., "Diffusion probabilistic priors for zero-shot low-dose CT image denoising," *Med. Phys.*, vol. 52, no. 1, pp. 329–345, 2025.

[47] W. Xia, Y. Shi, C. Niu, W. Cong, and G. Wang, "Diffusion prior regularized iterative reconstruction for low-dose CT," 2023, *arXiv:2310.06949*.

[48] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 6840–6851.

[49] K. Kothari, A. Khorashadizadeh, M. de Hoop, and I. Dokmanić, "Trumpets: Injective flows for inference and inverse problems," in *Proc. Uncertainty Artif. Intell.*, 2021, pp. 1269–1278.

[50] A. Denker, M. Schmidt, J. Leuschner, P. Maass, and J. Behrmann, "Conditional normalizing flows for low-dose computed tomography image reconstruction," 2020, *arXiv:2006.06270*.

[51] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," in *Proc. Int. Conf. Learn. Representations*, 2021. [Online]. Available: https://openreview.net/forum?id=St1giarCHLP

[52] C. Lu, Y. Zhou, F. Bao, J. Chen, C. Li, and J. Zhu, "DPM-solver: A fast ode solver for diffusion probabilistic model sampling in around 10 steps," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 5775–5787.

[53] G. Wang and M. Jiang, "Ordered-subset simultaneous algebraic reconstruction techniques (OS-SART)," *J. X-ray Sci. Technol.*, vol. 12, no. 3, pp. 169–177, 2004.

[54] J. H. Park, J. Lee, and J. Hwang, "Solving inverse problems using normalizing flow prior: Application to optical spectra," *Phys. Rev. B*, vol. 109, no. 16, 2024, Art. no. 165130.

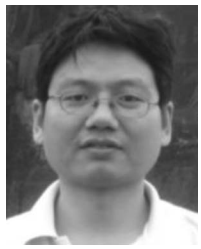
[55] L. Helminger, M. Bernasconi, A. Djelouah, M. Gross, and C. Schroers, "Generic image restoration with flow based priors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 334–343.

[56] D. P. Kingma and P. Dhariwal, "Glow: Generative flow with invertible 1x1 convolutions," in *Proc. 32nd Int. Conf. Neural Inf. Process. Syst.*, ser. NIPS'18. Red Hook, NY, USA: Curran Associates Inc. 2018, p. 10236–10245. [Online]. Available: https://dl.acm.org/doi/abs/10.5555/3327546.3327685

[57] C. Winkler, D. E. Worrall, E. Hoogeboom, and M. Welling, "Learning likelihoods with conditional normalizing flows," 2019, *arXiv:1912.00042*.
[58] L. Wei, A. Yadav, and W. Hsu, "CTflow: Mitigating effects of computed tomography acquisition and reconstruction with normalizing flows," in *Proc. Int. Conf. Med. Image Comput. Comput.- Assist. Interv.*, 2023, pp. 413–422.
[59] V. Wolf, A. Lugmayr, M. Danelljan, L. Van Gool, and R. Timofte, "DeFlow: Learning complex image degradations from unpaired data with conditional flows," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 94–103.
[60] X. Liu, X. Liang, L. Deng, S. Tan, and Y. Xie, "Learning low-dose CT degradation from unpaired data with flow-based model," *Med. Phys.*, vol. 49, no. 12, pp. 7516–7530, 2022.
[61] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[62] A. Buades, B. Coll, and J.-M. Morel, "Non-local means denoising," *Image Process. Line*, vol. 1, pp. 208–212, 2011.
[63] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
[64] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. 20th Int. Conf. Pattern Recognit.*, 2010, pp. 2366–2369.
[65] S. J. Wright, "Coordinate descent algorithms," *Math. Program.*, vol. 151, no. 1, pp. 3–34, 2015.
[66] T. Liang, Y. Jin, Y. Li, and T. Wang, "Edcnn: Edge enhancement-based densely connected network with compound loss for low-dose CT denoising," in *Proc.IEEE 15th Int. Conf. Signal Process.*, 2020, pp. 193–198.
[67] Z. Han, H. Shangguan, X. Zhang, P. Zhang, X. Cui, and H. Ren, "A dual-encoder-single-decoder based low-dose CT denoising network," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 7, pp. 3251–3260, Jul. 2022.
[68] Q. Gao, Z. Li, J. Zhang, Y. Zhang, and H. Shan, "CoreDiff: Contextual error-modulated generalized diffusion model for low-dose CT denoising and generalization," *IEEE Trans. Med. Imag.*, vol. 43, no. 2, pp. 745–759, Feb. 2024.
[69] S. G. Armato et al., "The lung image database consortium (lidc) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans," *Med. Phys.*, vol. 38, no. 2, pp. 915–931, 2011.

Ran An received the B.S. and M.S. degrees from the School of Mathematical Sciences, Capital Normal University, Beijing, China, in 2018 and 2021, respectively. He is currently working toward the Ph.D. degree through the collaborative program between Capital Normal University and the University of Liverpool, Liverpool, U.K. His research focuses on CT imaging and image processing, particularly in deep learning-based approaches for low-dose CT reconstruction, image denoising, and medical image segmentation.

Ke Chen received the B.Sc. degree from Dalian University of Technology, Dalian, China, the M.Sc. degree from the University of Manchester, Manchester, U.K., and the Ph.D. degree from the University of Plymouth, Plymouth, U.K., in applied and computational mathematics. He is a Computational Scientist specializing in the development of innovative algorithms for data and imaging challenges, with a particular focus on deep learning methods that leverage mathematical models. Driven by a passion for real-world applications and multidisciplinary research, he is the Director of two research centers—LCMH and CMIT—at Liverpool. He currently leads the Data Science and Learning Algorithms research group and is the Head of the Department of Mathematics and Statistics at the University of Strathclyde. He has authored more than 200 peer-reviewed articles, published seven books, and is a regular contributor to prestigious journals such as IEEE and SIAM.



Hongwei Li received the Ph.D. degree in the field of parallel computing from the institute of Software, Chinese Academy of Sciences, in 2002. From 2002 to 2005, he was with the Institute of Software as an Assistant Professor. In 2005, he joined the Center for Integrated Petroleum Research, University of Bergen, Norway, as a Postdoctoral Researcher for image processing and inverse problem computing. Since 2008, he has been with the Department of Mathematics, Capital Normal University, Beijing, as an Associate Professor and then a Full Professor. His research focuses on computed tomography reconstruction and related image processing problems since then, especially on limited-angle reconstruction, low-dose reconstruction, and CT image artifacts correction like ring artifacts and metal artifacts reduction.