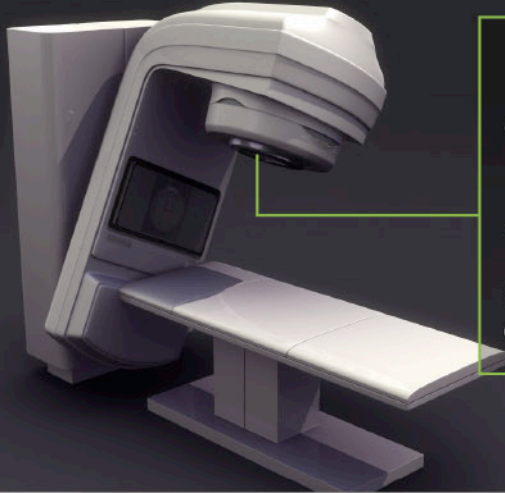
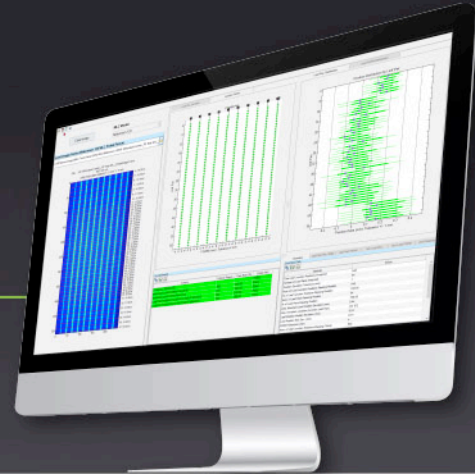


RIT[®]Gi42

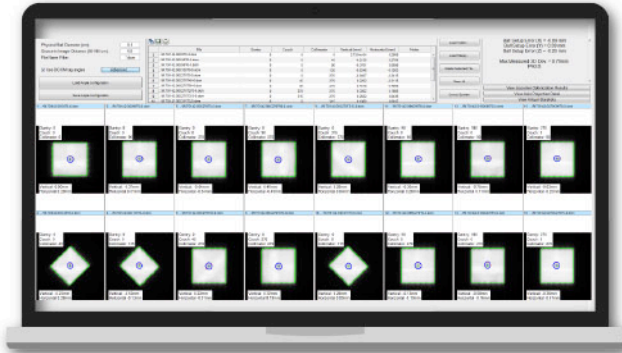
THE SINGLE-VENDOR SOFTWARE SOLUTION
FOR **EVERY QA TEST** RECOMMENDED IN TG-142



Perform comprehensive quality assurance of linear accelerators with confidence and ease, using an EPID and RIT software. Use either **RITGi42** or **RIT Complete** as your all-in-one software product for Machine QA, MLC QA, Imaging QA, and all the data tracking and trending you need to be in compliance with TG-142 and MPPG 8.b.



DAILY, MONTHLY, & ANNUAL LINAC MACHINE QA TESTS



- Enhanced 3D Winston-Lutz (Isocenter Optimization)
- Stereotactic Alignment (2D Winston-Lutz)
- Stereotactic Cone Profiles
- Field Alignment
- **Beam Measurements:**
Radiation/Light Field Coincidence, Fully-Automated Star Shot Analysis, Asymmetric Field/Matchline, Electron Energy (TG-25), Image Histogram, Isodose Contours, 3D Dose Profile, Quick Flatness & Symmetry, Water Tank Beam Measurement Analysis, and Depth Dose Profiles, Cross Profiles, & Orthogonal Profiles

FAST & EASY QUANTITATIVE MLC QA

- EPID Picket Fence Test for Varian & Elekta
- Leaf Speed Tests for Varian & Elekta
- Automated Varian RapidArc[®] Tests (0.1, 0.2, 1.1, 1.2, 2, and 3)
- Varian Halcyon[®] MLC Analysis
- Hancock Tests for Elekta
- Bayouth MLC Test
- TG-50 Picket Fence Test
- MSK Leaf Test
- Varian DMLC Test Patterns
- MLC Transmission Analysis

ONE CLICK, INSTANT IMAGING QA

- **EPID MV Imaging Module:** EPID QA, Las Vegas EPID, PTW EPID QC, and Standard Imaging QC-3 phantoms
- **CATPHAN[®]/OBI Module:** Catphan 500/600 Series, Elekta 503 Catphan (XVI), Varian 504 & 604 Catphan, and Siemens MV CT (OBI) phantoms
- **kV Imaging Module:** IBA Primus[®]-L, PTW NORMI[®]4, Leeds TOR 18FG, and Standard Imaging QC-kV1 phantoms
- **ISOCube Module:** Standard Imaging ISOCube Phantom
- **Penta-Guide Module:** Modus (IBA) Penta-Guide Phantom

RIT software simplifies the process of creating a routine QA program at your facility, streamlines the reporting process for accreditation, and provides you with a better understanding of measurement performance across multiple machines and facilities.

**SCHEDULE A PERSONAL DEMO OF
RIT'S TG-142 CAPABILITIES TODAY!**



RADIMAGE.COM[®]

+1.719.590.1077, Opt. 4 | sales@radimage.com | Connect with us on social media @RIT4QA



© 2023, Radiological Imaging Technology, Inc.

RapidArc[®] and Halcyon[®] are registered trademarks of Varian Medical Systems, Inc.
CATPHAN[®] is a registered trademark of The Phantom Laboratory.

Primus[®] is a registered trademark of IBA.
NORMI[®] is a registered trademark of PTW.

Learning low-dose CT degradation from unpaired data with flow-based model

Xuan Liu¹ | Xiaokun Liang² | Lei Deng² | Shan Tan¹ | Yaoqin Xie²

¹School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China

²Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

Correspondence

Shan Tan, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China.
Email: shantan@hust.edu.cn

Yaoqin Xie, Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China.
Email: yq.xie@siat.ac.cn

Funding information

Chinese Academy of Sciences Special Research Assistant Grant Program; Young S&T Talent Training Program of Guangdong Provincial Association for S&T, Grant/Award Number: SKXRC202224; National Natural Science Foundation of China, Grant/Award Numbers: U20A20373, 61871374, 62071197

Abstract

Background: There has been growing interest in low-dose computed tomography (LDCT) for reducing the X-ray radiation to patients. However, LDCT always suffers from complex noise in reconstructed images. Although deep learning-based methods have shown their strong performance in LDCT denoising, most of them require a large number of paired training data of normal-dose CT (NDCT) images and LDCT images, which are hard to acquire in the clinic. Lack of paired training data significantly undermines the practicability of supervised deep learning-based methods. To alleviate this problem, unsupervised or weakly supervised deep learning-based methods are required.

Purpose: We aimed to propose a method that achieves LDCT denoising without training pairs. Specifically, we first trained a neural network in a weakly supervised manner to simulate LDCT images from NDCT images. Then, simulated training pairs could be used for supervised deep denoising networks.

Methods: We proposed a weakly supervised method to learn the degradation of LDCT from unpaired LDCT and NDCT images. Concretely, LDCT and normal-dose images were fed into one shared flow-based model and projected to the latent space. Then, the degradation between low-dose and normal-dose images was modeled in the latent space. Finally, the model was trained by minimizing the negative log-likelihood loss with no requirement of paired training data. After training, an NDCT image can be input to the trained flow-based model to generate the corresponding LDCT image. The simulated image pairs of NDCT and LDCT can be further used to train supervised denoising neural networks for test.

Results: Our method achieved much better performance on LDCT image simulation compared with the most widely used image-to-image translation method, CycleGAN, according to the radial noise power spectrum. The simulated image pairs could be used for any supervised LDCT denoising neural networks. We validated the effectiveness of our generated image pairs on a classic convolutional neural network, REDCNN, and a novel transformer-based model, TransCT. Our method achieved mean peak signal-to-noise ratio (PSNR) of 24.43dB, mean structural similarity (SSIM) of 0.785 on an abdomen CT dataset, mean PSNR of 33.88dB, mean SSIM of 0.797 on a chest CT dataset, which outperformed several traditional CT denoising methods, the same network trained by CycleGAN-generated data, and a novel transfer learning method. Besides, our method was on par with the supervised networks in terms of visual effects.

Conclusion: We proposed a flow-based method to learn LDCT degradation from only unpaired training data. It achieved impressive performance on LDCT synthesis. Next, we could train neural networks with the generated paired data for LDCT denoising. The denoising results are better than traditional and weakly supervised methods, comparable to supervised deep learning methods.

KEYWORDS

computed tomography, flow-based model, noise synthesis, weakly supervised learning

1 | INTRODUCTION

Computed tomography (CT) has been widely used for noninvasive diagnoses in modern medicine.¹ At the same time, there have been growing concerns about the potential risks of excessive radiation to patients, and radiologists.^{2,3} Low-dose CT (LDCT) has attracted great interest for its significantly reduced imaging dose,^{4–9} which is commonly realized by cutting down the tube current and the exposure time of projection data acquisition. However, dose reduction results in electronic noise, quantum noise, and other degradations in projection data.^{10–12} After reconstruction, CT images always suffer from complex noise and artifacts. The low quality of LDCT images greatly limits its practical application.

Many algorithms have been proposed to improve the image quality of LDCT. These algorithms could be generally classified into three categories, which are sinogram processing in projection domain,^{13–16} iterative reconstruction,^{17–27} and postprocessing in image domain.^{4–7,28} Among the three kinds of methods, postprocessing, also known as LDCT denoising, is directly applied to the reconstructed CT images. Because it does not involve the reconstruction process, postprocessing methods can be easily integrated into different CT systems. Image prior modeling plays a key role in designing LDCT denoising algorithms. In early works, hand-designed priors were usually used in LDCT denoising methods.^{28–32} For example, nonlocal means (NLM)³³ utilized the prior of the similarity among nonlocal image patches and substantially improved the LDCT denoising performance.²⁸ K-SVD³⁴ was a dictionary learning-based algorithm using the sparsity of images as prior. It could effectively remove the noise of LDCT images.³² Similar to NLM, block-matching 3D (BM3D),³⁵ which utilized the prior of nonlocal block similarity, also achieved great performance in various CT image denoising tasks.^{29–31}

Although traditional methods for LDCT denoising have their advantages, their performance always depends on the accuracy of the hand-designed priors, which needs extensive expert knowledge and may not cover any cases or datasets. Therefore, the traditional algorithms will get unfavorable results when applied inappropriately. During the past 10 years, deep learning has achieved impressive success in all kinds of image processing tasks, including LDCT denoising.^{4–7,9} Deep learning-based methods generally train neural networks with the LDCT images as input and the normal dose CT (NDCT) images as targets. The image prior was learned from the paired training data. Compared

with traditional methods, deep learning-based methods always have better performance, and they can be used in more cases as long as there are paired training data.

The excellent performance of deep learning-based methods highly relies on the training data.³⁶ However, in LDCT denoising, it is not realistic to get paired training data of LDCT images and NDCT images because one patient has to accept two consecutive scans, which is not allowed for violation of the “as low as reasonably achievable” principle.¹ Therefore, even though the LDCT and NDCT of the same patient are acquired, the two reconstructed CT images are not aligned because of the movement or respiratory phase difference between the two scans. Existing methods are trained and tested on a popular benchmark generously provided by Mayo Clinic.^{5,6} In this dataset, LDCT images are simulated by inserting noise into normal-dose projection data and reconstructing. Although the simulated noise is based on a dedicated study and many effects are considered, it is only applicable to several CT systems, including GE Discovery CT750i, SOMATOM Definition AS+, and SOMATOM Definition Flash CT system. For those CT systems without dedicated noise study, paired data are hard to simulate.

To solve the above problem, we proposed a weakly supervised method to learn the degradation of LD CT from unpaired LDCT and NDCT images. In detail, the proposed method is based on an invertible network—SRFlow,³⁷ which can generate a high-resolution image given the low-resolution one and a latent code, also project a high-resolution image to the latent space given its low-resolution counterpart. Thus, given a NDCT or LDCT image, we can project the images to the latent space by SRFlow. Although the noise distribution of LDCT is quite complex and content-dependent,³⁸ it could be simplified in latent space. In latent space, we can assume that the noise between LDCT and NDCT follows any predefined distribution and make the neural network fit it adaptively. In our method, we assume its Gaussian noise.³⁹ Based on this assumption, the model can be trained with unpaired data. While the input is NDCT images, the latent space follows a predetermined distribution. While the input is LDCT images, the latent space follows the predetermined distribution plus a Gaussian distribution. Once we know the distribution of latent space, we can obtain the distribution of NDCT and LDCT images based on the invertibility of flow. Then, the NDCT and LDCT training images can be fed into the network separately to calculate the likelihood. Therefore, the whole model can be trained by maximizing the likelihood of NDCT and LDCT images

in an unpaired manner. After training, we can synthesize an LDCT image by first projecting an NDCT image into the latent space, then resampling a latent code according to the conditional probability distribution, and finally generating the LDCT image by the inverse network. The synthetic paired data can be used for any supervised LDCT denoising neural networks.

We validated the proposed method's effectiveness in two datasets that include different regions and different noise levels. The first one is abdomen CT images from the 2016 AAPM Grand Challenge, including normal-dose images and quarter-dose images.^{5,6} Second is chest CT images from LDCT and projection data (LDCT-PD),^{36,40} LD cases of which are provided at 10% of the routine dose. Compared with the most widely used image-to-image translation method, CycleGAN,⁴¹ our method achieved much better performance on LDCT image synthesis. Then, we validated the effectiveness of our generated image pairs on a classic convolutional neural network, REDCNN,⁵ and a novel transformer-based model, TransCT.⁹ The denoising network trained by our generated data outperformed several traditional CT denoising methods, the same network trained by CycleGAN-generated data, and a novel transfer learning method both quantitatively and visually. Besides, our method was on par with the supervised networks in terms of visual effects.

The paper is organized as follows. Section 2 introduces the related works, including flow-based models and other unsupervised or weakly supervised methods for LDCT denoising. Section 3 describes our method and the implementation details. Sections 4 and 5 introduce the datasets used for validation and show the experimental results. Conclusions are drawn in the last section.

2 | RELATED WORKS

2.1 | Supervised learning-based methods for LDCT denoising

Supervised deep learning-based methods generally train neural networks with LDCT images as input and NDCT images as targets. Kyong et al. proposed FBP-ConvNet to postprocess the CT images reconstructed by FBP algorithm.⁴ Chen et al. presented REDCNN to suppress noise on reconstructed CT images.⁵ Wu et al. proposed a Cascaded-CNN reconstruction algorithm using an improved DnCNN for noise suppression.⁴² Considering the visual effects of results, some works incorporated perceptual loss or GAN-based loss in the denoising networks.⁷ With the increasing popularity of transformer architecture in computer vision, Zhang et al. proposed a novel transformer-based network and achieved impressive performance.⁹

2.2 | Unsupervised learning-based methods for LDCT denoising

In recent years, more and more studies have noticed the training data collection burden of supervised deep learning and paid attention to the methods that need less labeled training data.^{43,44} In the field of image denoising, weakly supervised learning, self-supervised learning, and unsupervised learning have been increasingly popular.^{45–49} Noise2Noise⁴⁵ successfully trained denoising neural networks with pairs of noisy images instead of noisy/clean image pairs. Noise2Void⁴⁶ and Noise2Self⁴⁷ trained denoising networks with only noisy images. The main idea is to create input and target image pairs by a blind-spot strategy. Noisier2Noise⁴⁸ created training pairs by adding additional noise to the original noisy images and trained the denoising nets. Self2Self⁴⁹ achieved training a neural network from only one single noisy image based on a blind-spot strategy.

However, most of the above methods rely on the pixelwise independence of the noise, so that they can only be applied to some synthetic noise removal such as Gaussian noise. Noise in LDCT images is much more complex and spatially correlated.³⁸ Therefore, unsupervised or weakly supervised deep learning methods for LDCT denoising require dedication to design. One kind of method is based on Noise2Noise. For example, Yuan et al. proposed Half2Half that trained a neural denoiser with paired half-dose CT images.⁵⁰ Wu et al. proposed to train LDCT denoising nets with pairs of images reconstructed from odd and even projections.⁵¹ Niu et al. proposed a similarity-based unsupervised deep denoising approach to process the correlated noise in CT images.⁵² Another class of methods is based on the popular image-to-image translation model, CycleGAN. Kang et al. applied CycleGAN for the denoising of multiphase coronary CT angiography.⁵³ Gu et al. improved the double-generator structure of CycleGAN instead used an Adaln module to switch the direction of generators.⁵⁴ In addition, Kim et al. proposed a weakly supervised method plus a physics-based CT noise model to progressively denoise the LDCT images.⁵⁵ Shan et al. proposed a novel transfer learning framework for LDCT denoising.⁵⁶ Unlike most of the current unsupervised and weakly supervised methods, our proposed method concentrates on learning the degradation of LDCT images from unpaired data. With the learned degradation model, synthetic paired training data can train any supervised denoising networks.

2.3 | LDCT noise simulation methods

LDCT noise simulation plays an important role in LDCT studies, which can help determine the lowest scanning dose under the “as low as reasonably achievable”

principle and provide training pairs for supervised machine learning methods. Traditional LDCT noise simulation methods always insert noise in the projection domain.^{57–59} However, these methods require access to CT raw data and knowledge of the CT imaging system. Divel et al. proposed an accurate image domain noise insertion method and provide an alternative to projection domain methods.⁶⁰ In recent years, GAN-based methods also achieved impressive performance on LDCT noise simulation. Shan et al. introduced a method to simulate LDCT images at various noise levels with a GAN.⁶¹ Niu et al. also proposed a novel noise-entangled GAN that can create LDCT images quickly with no requirements of raw data or other proprietary CT scanner information.⁶²

2.4 | Flow-based model

A flow-based model is composed of a sequence of invertible modules which Jacobian determinant is easy to compute.^{63–65} Based on the change-of-variable law,⁶⁶ a flow-based network can transform a simple distribution into a complex one for generation, and project a complex distribution to a simple one for representation. Flow-based models are initially used for unconditional image generation. The earliest flow-based generative models are real-valued nonvolume preserving (RealNVP)⁶⁴ and nonlinear independent components estimation (NICE),⁶³ which utilize the coupling layers to calculate the inverse and the Jacobian determinant of the transformation function. OpenAI proposed Generative Flow (Glow) with invertible 1×1 convolutions and greatly improved the generated image quality of flow-based models.⁶⁵ Meanwhile, conditional flow-based models have also been investigated for different generative vision tasks such as point cloud generation, class-conditional image generation, and image superresolution.^{67–69} Among them, SRFlow⁶⁹ achieved diverse image superresolution by inputting a low-resolution image as the condition and a random latent code as the control of diversity. We noticed that the 4× low-resolution version of the LDCT image and NDCT image could be almost the same. Inspired by SRFlow, we considered generating the corresponding LDCT image from the low-resolution version of an NDCT image by adjusting the latent code. The reason for choosing the flow-based superresolution model is that we can precisely get the latent code of an NDCT image thanks to the invertibility of SRFlow.

2.5 | DeFlow

Wolf et al. proposed DeFlow to learn complex image degradations from unpaired data with conditional flows.³⁹ DeFlow models the realistic natural image noise,

generates synthetic training data for the downstream task of real-world superresolution, and achieves significant performance improvements. Our work mainly referenced DeFlow and generalized it to medical image denoising. Different from the realistic natural image noise that is quite slight, the degradation of LDCT images is much more complex and content-dependent. Besides, our follow-up task is not superresolution but denoising.

3 | METHOD

3.1 | Learning LDCT degradation from unpaired data

The degradation of LDCT images is

$$y = D(x), \quad (1)$$

where x is an NDCT image, y is the corresponding LDCT image of x , and $D(\cdot)$ denotes the degradation function. It is difficult to establish the physics model of LDCT degradation because the original corruption occurs in the projection domain. After reconstruction, the noise and artifacts in image domain become quite complicated and highly content-independent. We choose to model the degradation by statistical methods.

Assuming the NDCT images follow a distribution p_x , and the LDCT images follow another distribution p_y , for any image $x \sim p_x$, there is an unknown degradation function $D(\cdot)$ that makes $y = D(x) \sim p_y$. The degradation process can also be represented as the conditional distribution $p(y | x)$, where $x \sim p_x$ and $y \sim p_y$. What we want is to get $p(y | x)$ without paired training data, that is to say, given only the marginal distributions p_x and p_y . It is feasible to get a closed-form solution of $p(y | x)$ if p_x and p_y follow some distributions with explicit analytical expression such as Gaussian distribution.³⁹ However, the distribution of CT images is far more complicated than Gaussian distribution. It is impracticable to solve the problem directly in the image space. Thanks to the powerful representation capability and invertibility of flow-based generative networks, we can project the NDCT images and LDCT images to a much simpler latent space and then model the relation between x and y .

We use a flow-based generative network for the transformation between NDCT images $x \sim p_x$, LDCT images $y \sim p_y$ and the corresponding latent variables $z_x \sim p_{z_x}$, $z_y \sim p_{z_y}$, as

$$z_x = f_\theta(x), \quad x = f_\theta^{-1}(z_x), \quad (2a)$$

$$z_y = f_\theta(y), \quad y = f_\theta^{-1}(z_y), \quad (2b)$$

where $f_\theta(\cdot)$ denotes the flow-based generative network with parameters θ , and $f_\theta^{-1}(\cdot)$ denotes its inverse. According to DeFlow,³⁹ we let $z_x \sim \mathcal{N}(0, 1)$ and postulate the degradation process in the latent space is like

$$z_y = D_z(z_x) = z_x + u, \quad u \sim \mathcal{N}(\mu_u, \Sigma_u), \quad u \perp z_x, \quad (3)$$

where $D_z(\cdot)$ denotes the degradation of LDCT in latent space, which is set as an additive Gaussian noise u that is independent of z_x . According to Equations (2) and (3), the degradation function between x and y can be written as

$$y = f_\theta^{-1}(z_y) = f_\theta^{-1}(D_z(z_x)) = f_\theta^{-1}(z_x + u) = f_\theta^{-1}(f_\theta(x) + u),$$

$$z_x \sim p_{z_x} = \mathcal{N}(0, I), \quad u \sim p_u = \mathcal{N}(\mu_u, \Sigma_u), \quad z_x \perp u. \quad (4)$$

Now, we need to find the neural network $f_\theta(\cdot)$ that satisfies the assumptions above, that is, (1) $z_x \sim p_{z_x} = \mathcal{N}(0, 1)$ and (2) $u = (z_y - z_x) \sim p_u = \mathcal{N}(\mu_u, \Sigma_u)$. For the first assumption, we can derive the marginal distribution of the NDCT images, p_x from p_{z_x} , then fit the training data by maximizing a likelihood-based objective as Equation (5a). For the second assumption, it is hard to calculate p_u directly because the paired data (x, y) are not available and thus (z_x, z_y) is not available. Alternatively, we can deduce the marginal distribution p_y from p_u and p_x and then derive another maximum likelihood-based objective as Equation (5b). The probability density function of the two marginals p_x and p_y can be derived by the change of variables formula as

$$\begin{aligned} p_x(x) &= |\det Df_\theta(x)| \cdot p_{z_x}(f_\theta(x)) \\ &= |\det Df_\theta(x)| \cdot \mathcal{N}(f_\theta(x); 0, I), \end{aligned} \quad (5a)$$

$$\begin{aligned} p_y(y) &= |\det Df_\theta(y)| \cdot p_{z_y}(f_\theta(y)) \\ &= |\det Df_\theta(y)| \cdot \mathcal{N}(f_\theta(y); \mu_u, I + \Sigma_u), \end{aligned} \quad (5b)$$

where $p_x(\cdot)$ and $p_y(\cdot)$ denote the probability density function of distribution p_x and p_y , respectively, $p_{z_x}(\cdot)$ and $p_{z_y}(\cdot)$ denote the probability density function of distribution p_{z_x} and p_{z_y} . \det denotes the operation of determinant. $Df_\theta(x)$ denotes the Jacobian of the flow network. To fit the unpaired training data, including a set of NDCT images, $\{x_i\}_{i=1}^M$, where $x_i \sim p_x$, and a set of LDCT images, $\{y_i\}_{i=1}^N$, where $y_i \sim p_y$. We can train the flow network f_θ by applying a minimum negative log-likelihood objective as follows. At the same time, the mean and covariance matrix of p_u are learned.

$$L_x(\theta) = -\frac{1}{M} \sum_{i=1}^M \ln p_x(x_i), \quad (6a)$$

$$L_y(\theta; \mu_u; \Sigma_u) = -\frac{1}{N} \sum_{j=1}^N \ln p_y(y_j), \quad (6b)$$

$$L_{total}(\theta; \mu_u; \Sigma_u) = -\frac{1}{M} \sum_{i=1}^M \ln p_x(x_i) - \frac{1}{N} \sum_{j=1}^N \ln p_y(y_j). \quad (6c)$$

The learned flow-based generative network can be utilized for LDCT image synthesis by applying Equation (4).

3.2 | Flow with domain invariant condition

Although theoretically possible, it requires a large number of computational resources and time to directly model the distribution of an image domain with flow-based generative networks, which largely limits the implementation of the above method. To solve the problem, we use a conditional generative model similar to DeFlow.³⁹ Concretely, extract low-dimensional information from the images as a condition before transforming image space and latent space. The conditional version of Equations (2a) and (2b) is

$$z_x = f_\theta(x; h(x)), \quad x = f_\theta^{-1}(z_x; h(x)), \quad (7a)$$

$$z_y = f_\theta(y; h(y)), \quad y = f_\theta^{-1}(z_y; h(y)), \quad (7b)$$

where $h(\cdot)$ denotes the operation of condition extraction, such as downsampling and more. The conditional flow-based generative networks can be much easier to train. However, for Equation (4) to be still performable, $h(\cdot)$ must satisfy the following equation:

$$h(x) = h(y), \quad (8)$$

where x and y are one pair of NDCT image and LDCT image. Equation (8) means that the operation $h(\cdot)$ has to be invariant between NDCT image domain and LDCT image domain. On this premise, the degradation function in Equation (4) turns into

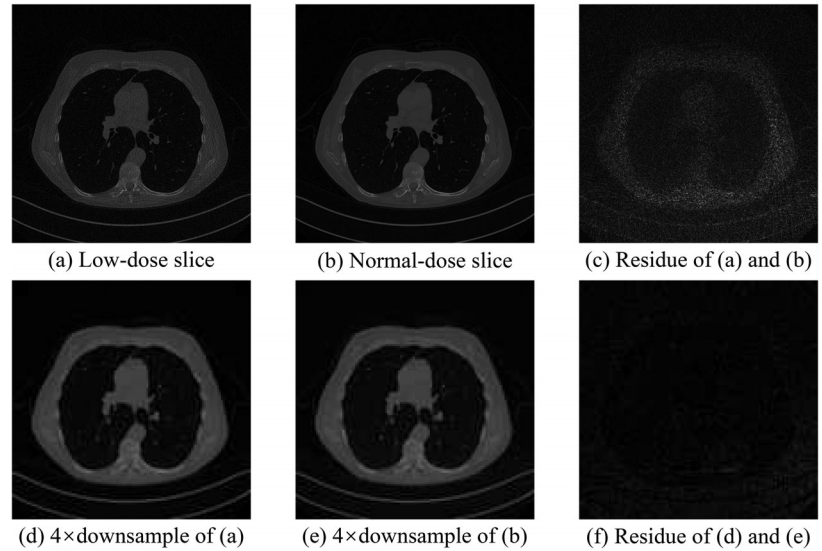
$$\begin{aligned} y &= f_\theta^{-1}(f_\theta(x; h(x)) + u; h(x)), \\ u &\sim p_u = \mathcal{N}(\mu_u, \Sigma_u). \end{aligned} \quad (9)$$

The marginal distributions of x and y can be rewritten as the conditional distributions given $h(x)$ and $h(y)$,

$$p_{x|h(x)}(x) = |\det Df_\theta(x; h(x))| \cdot \mathcal{N}(f_\theta(x; h(x)); 0, I), \quad (10a)$$

$$p_{y|h(y)}(y) = |\det Df_\theta(y; h(y))| \cdot \mathcal{N}(f_\theta(y; h(y)); \mu_u, I + \Sigma_u). \quad (10b)$$

FIGURE 1 (a–c) shows the difference of one pair of low-dose and normal-dose CT slices. (d–e) shows the difference of their 4× downsamples. Brighter pixels in the residue map denote larger gaps.



Similar to Equations (6a)–(6c), the conditional version of training objective is

$$L_x(\theta) = -\frac{1}{m} \sum_{i=1}^M \ln p_{x|h(x)}(x_i), \quad (11a)$$

$$L_y(\theta; \mu_u; \Sigma_u) = -\frac{1}{n} \sum_{j=1}^N \ln p_{y|h(y)}(y_j), \quad (11b)$$

$$L_{total}(\theta; \mu_u; \Sigma_u) = -\frac{1}{m} \sum_{i=1}^M \ln p_{x|h(x)}(x_i) - \frac{1}{n} \sum_{j=1}^N \ln p_{y|h(y)}(y_j). \quad (11c)$$

The choice of $h(\cdot)$ should follow two principles, that is, (1) satisfying Equation (8) and (2) making training of the flow as easy as possible. For NDCT and LDCT images, we choose the 4× downsampling as $h(\cdot)$ because paired NDCT images and LDCT images maintain similar structures and differ in high-frequency details. As shown in Figure 1, the 4× downsampling versions of an NDCT image and the corresponding LDCT image are almost the same. Using the low-resolution version of an image as a condition is a superresolution process when the flow-based network generates images from latent codes. Existing flow-based architectures for image superresolution can be applied, like SRFlow, which achieved diverse photorealistic image superresolution with paired training. In the training phase of our method, we at the same time feed the NDCT and LDCT images with the low-resolution condition to an SRFlow network and train the network with the negative log-likelihood loss in Equation (10).

After training, we can simulate the corresponding LDCT image y given an NDCT image x following Equation (9) for test. First, an NDCT image is input to the trained flow network $f_\theta(\cdot)$ with its downsampled counterpart $h(x)$ as the condition to get its latent code z_x .

Then, we get z_y by adding Gaussian noise u with mean μ_u and covariance matrix Σ_u to z_x . At last, z_y is input to the inverse flow $f_\theta^{-1}(\cdot)$ also with the low-resolution NDCT image as the condition. The output is the simulated LDCT image.

The overview of our method is shown in Figure 2.

3.3 | Network architectures

The network architecture is based on SRFlow, which integrates a low-resolution encoder into an unconditional glow architecture.⁶⁹ The invertible modules include conditional affine coupling, affine injector, invertible 1×1 convolution, channel-wise normalization, split, and squeeze. The low-resolution images are encoded and plugged into the conditional affine coupling layers and affine injector layers. The whole architecture is shown in Figure 3. The backbone of a flow-based network consists of L submodules of different scales. In each submodule except the last one, there is a squeeze layer, a split layer, N unconditional flow steps, each consisting of an invertible 1×1 convolution layer and an act-norm layer, and K conditional flow steps, each consisting of a conditional affine coupling layer, an affine injector layer, an invertible 1×1 convolution layer, and an act-norm layer. In the last submodule, there is no split layer. As follows are the brief introductions of the modules.

Conditional affine coupling: The affine coupling layer is a simple and efficient architecture for flow-based models. It is convenient to merge conditional information in the affine coupling layer as follows:

$$z_A^{n+1} = z_A^n, \quad z_B^{n+1} = \exp(\tau_s^n(z_A^n; u)) \cdot z_B^n + \tau_b^n(z_A^n; u), \quad (12)$$

which represent the transformation from n th layer to $(n+1)$ th layer. (z_A, z_B) is a partition of the input feature

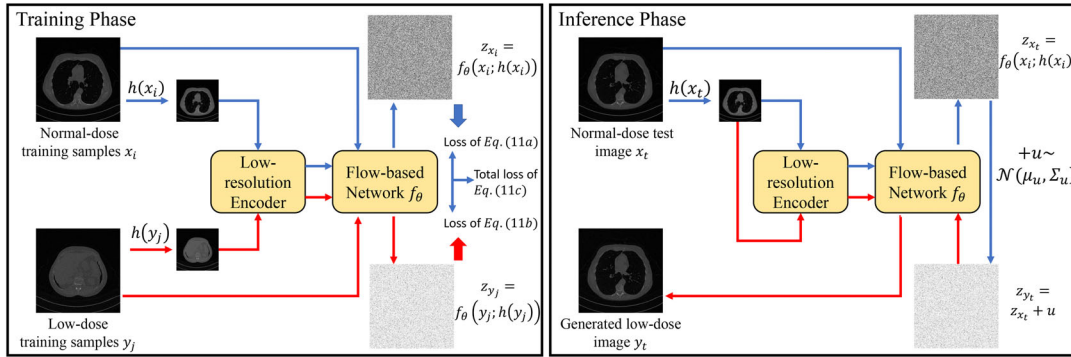


FIGURE 2 Overview of the proposed method. The blue pipelines denote NDCT processing and the red pipelines denote the LDCT processing. In the training phase, NDCT and LDCT images are simultaneously fed into the network. In the inference phase, the NDCT images are input to generate the LDCT images.

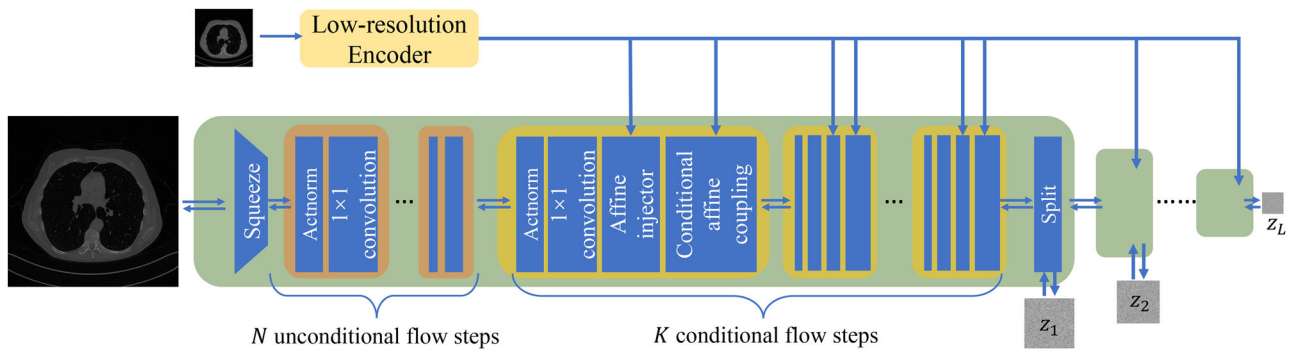


FIGURE 3 Architecture of the flow-based network, which consists of L different scales of subnetworks (green parts). In each subnetwork, there are N unconditional flow steps (orange parts) and K conditional flow steps (yellow parts). The details of modules in the flow steps are introduced in Section 3.3.

map z in the channel dimension. The first part z_A goes to the next layer after an identity mapping. The second part z_B is processed by an affine transformation, where τ_s and τ_b are networks to calculate the bias and scaling. $u = E(h(x))$ is the conditional variable encoded by the low-resolution images. The Jacobian of conditional affine coupling layer is triangular and efficient to compute.

Affine injector: Similar to the conditional affine coupling layer, the affine injector layer performs affine transformation to input z without partition as

$$z^{n+1} = \exp(\tau_s(u)) \cdot z^n + \tau_b(u). \quad (13)$$

Similarly, $\tau_s(\cdot)$ and $\tau_b(\cdot)$ can be arbitrary networks that transfer the conditional variable u . The inverse and Jacobian of the affine injector layer is also easy to compute.

Invertible 1×1 convolution: As demonstrated in Glow, a 1×1 convolution $z_{i,j}^{n+1} = W \cdot h_{i,j}^n$ that performs convolution at each spatial coordinate (i, j) can be used in flow-based neural networks. The Jacobian is a block-diagonal matrix.

Actnorm: The act norm layer performs a channel-wise batch normalization that is easily invertible.

Split and squeeze: The squeeze step halves the height and width of the feature maps by reshaping each 2 neighborhood into the channel dimension. First, the split layer splits the feature maps in channel dimension, then one half goes to the next flow step, and the other half becomes the latent variable.

Low-resolution encoder: The encoder of low-resolution images is based on residual-in-residual dense blocks (RRDB) proposed in ESRGAN.⁷⁰ The encoder captures multiscale features of the low-resolution images, which can be used in affine coupling layers and affine injectors of different scales.

3.4 | LDCT denoising

After synthesizing LDCT and NDCT image pairs, we can train any LDCT denoising networks in a supervised manner. In this paper, we chose two representative networks, REDCNN⁵ and TransCT.⁹ REDCNN, the residual encoder–decoder neural network, has achieved

impressive denoising performance with patch-based training. On the other hand, TransCT is a novel LDCT denoising method using a transformer-based neural network.

4 | EXPERIMENTAL SETTINGS

The experiments mainly include two parts. One is the synthesis of LDCT images by the learned flow-based generative model. The other is to train LDCT denoising networks with the synthetic paired data. As follows, we introduce the materials and settings of the experiments.

4.1 | Datasets

We validated our methods using two clinical datasets of different organs and different dose reductions. The first one is the widely used benchmark, that is, the abdomen CT images from the 2016 Low Dose CT Grand Challenge, sponsored by the AAPM, NIBIB, and Mayo Clinic.⁵ This dataset consists of 512×512 NDCT and quarter-dose CT image pairs of 10 patients. We chose NDCT images of five patients and LDCT images of the other five for training. The five patients' NDCT images were then used for paired data synthesis, and the other five patients' LDCT images were used for the test. The second one is the chest CT images from LDCT image and projection data (LDCT-and-Projection-data).^{36,40} This dataset provides paired 512×512 LD and ND chest CT images of 50 patients. Unlike the first one, the LD scans are provided at 10% of the routine dose. For this dataset, we chose 10 patients for training (half ND, half LD), five for paired data synthesis, and five for the test.

4.2 | Comparison methods

To validate the synthesis performance of our method, we compared it with the widely used weakly supervised image-to-image translation method, CycleGAN, which has been proven effective in diverse medical image synthesis tasks.^{71–73}

We compared the denoising performance of our method with four kinds of methods. The first kind is traditional algorithms that utilize hand-designed prior like nonlocal similarity to improve the image quality of LDCT. We chose BM3D and NLM as the comparison. The second is to train denoising networks with the paired training data generated by the above comparison methods. These comparisons can also validate the advantages of our method's synthesis performance. The third is a novel transfer learning method.⁵⁶ Concretely, we first trained an MAP-NN⁷⁴ with a roughly simulated pair dataset, that is, adding Gaussian–Poisson mixture noise in projection

domain, and then learned the weights of the denoised image sequence from MAP-NN with a few paired LDCT and NDCT images. We called this method MAP-NN-transfer. The fourth kind of comparison method is the supervised neural networks, REDCNN and TransCT, which relax the training requirements.

4.3 | Evaluation indexes

To evaluate the simulation performance, we compared the noise power spectrums(NPS) of the simulated LDCT images and the real LDCT images. The NPS is defined as

$$\text{NPS}(u, v) = \frac{d_i d_j}{N_i N_j} \cdot |\mathcal{F}[I(i, j) - P(i, j)]|^2, \quad (14)$$

where $I(i, j)$ is a 2D image or ROI, d_i and d_j denote pixel size (mm), and N_i and N_j are the number of pixels in the direction of i and j . u and v are the spatial frequency (mm^{-1}) in the two directions of i and j . $\mathcal{F}(\cdot)$ denotes the 2D Fourier transform, $P(i, j)$ denotes a second-order polynomial fit of $I(i, j)$. Further, the radial NPS, $\text{NPS}(r)$, is calculated by binning $\text{NPS}(u, v)$ according to radial distance $r = \sqrt{u^2 + v^2}$ and taking the average of each bin.

To validate the LDCT denoising results quantitatively, two widely used metrics, including peak signal-to-noise ratio (PSNR) structural similarity(SSIM), both pixel-wise and structural accuracy. The PSNR of a test image I_t and a reference image I_r is

$$\text{PSNR}(I_t, I_r) = 10 \cdot \log_{10} \frac{\mu_{\max}^2}{\text{MSE}(I_t, I_r)}, \quad (15)$$

where $\text{MSE}(I_t, I_r)$ is the mean-squared error between the test image I_t and the reference image I_r and μ_{\max} is the maximum possible value of the image. Higher PSNR means less error between the reconstructed image and the reference image. The SSIM is defined as

$$\text{SSIM}(a, b) = \frac{(2\mu_a \mu_b + C_1)(2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)}, \quad (16)$$

where a and b are two local windows of size 11 pixels in two images. The two windows have the same position, and μ_a and σ_a and μ_b and σ_b are their mean and standard deviation, respectively. σ_{ab} is the covariance between the two windows. C_1 and C_2 are two constants to avoid instability. In this study, C_1 and C_2 were chosen as $C_1 = (0.01\mu_{\max})^2$ and $C_2 = (0.03\mu_{\max})^2$. SSIM is used to measure the similarity in the structure between the two windows where a higher value means higher similarity. As the two windows move pixel-by-pixel over the test image and the reference image, we obtain an SSIM map. In practice, we use a single mean-SSIM

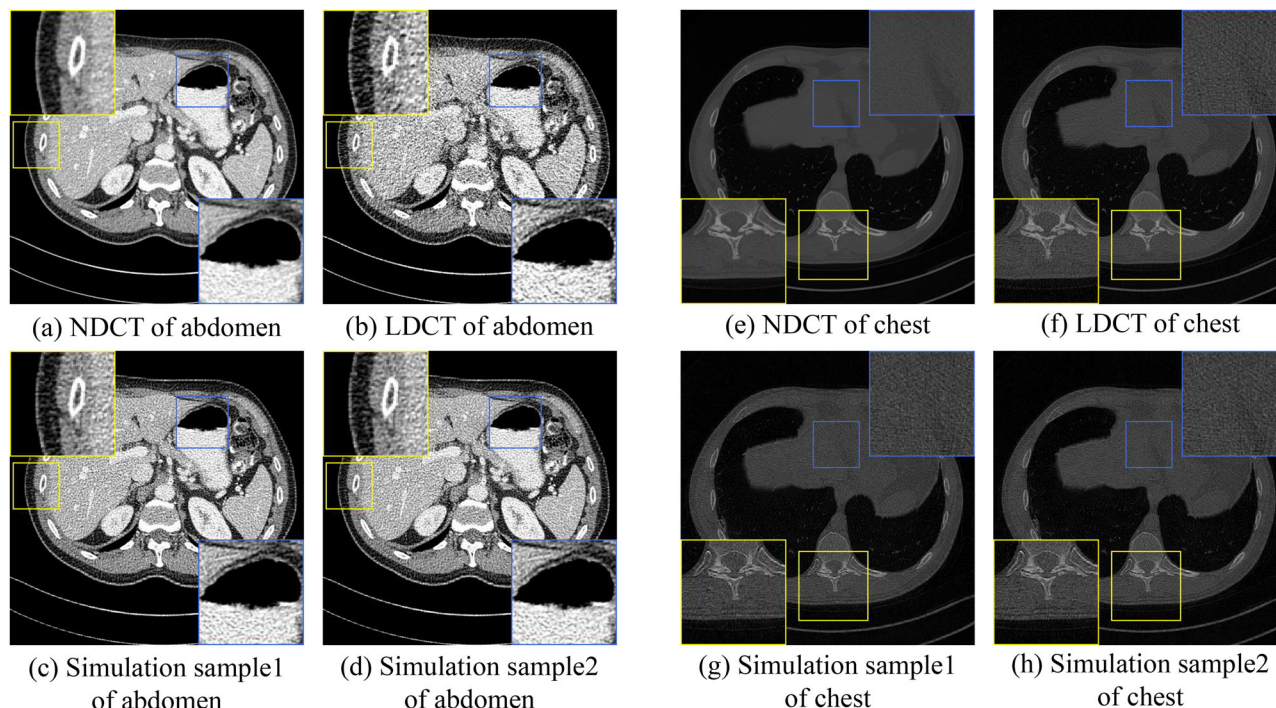


FIGURE 4 Multiple simulation samples of our method. The display window of abdomen CT images is $[-160, 240]$ HU, that of chest CT images is $[-1024, 3072]$ HU.

(MSSIM) value to evaluate the overall image quality by simply averaging the SSIM values.

5 | EXPERIMENTAL RESULTS

5.1 | LDCT simulation

This section shows our method's performance of simulating LDCT images from NDCT images. First, we show multiple simulation samples of our methods. According to Equation (9), we can generate multiple noise realizations given a single NDCT by sampling different u from the Gaussian distribution $\mathcal{N}(\mu_u, \Sigma_u)$. In Figure 4, we, respectively, show two different simulation samples of abdomen and chest CT slices.

We can also simulate LDCT images with different noise levels by rewriting Equation (9). We can multiply u by a coefficient δ as

$$y_\delta = f_\theta^{-1}(f_\theta(x; h(x)) + \delta \cdot u; h(x)), u \sim \mathcal{N}(\mu_u, \Sigma_u), \quad (17)$$

where y_δ denotes LDCT images with different noise levels. Figure 5 and 6 show the simulated LDCT images with different noise levels.

Furthermore, we compare the simulation performance of our method with CycleGAN. LD slices simulated by CycleGAN and our method are shown in Figure 7. In Figures 7(a)–(d), we can see that for LD abdomen CT

simulation, CycleGAN fails at learning the LDCT noise, whereas the noise simulated by our method is closer to real ones. In Figures 7(e)–(h) of chest CT slices, both of CycleGAN and our method capture the noise distribution according to the yellow zoom-in. However, there are artifacts at the contour of the CycleGAN-simulated chest slice, which destroys the accuracy of the simulation. The radial NPS results are consistent with the visual observations. As shown in Figure 8, the radial NPS of our method is much more similar to the real LDCT's than that of CycleGAN.

5.2 | LDCT denoising

To further demonstrate the advantage of the images generated by our method, we trained denoising neural networks with our simulated image pairs, named **REDCNN-Ours** and **TransCT-Ours**. The denoising performance is compared with two traditional denoising methods, **BM3D** and **NLM**, denoising networks trained with image pairs synthesized by CycleGAN, named **REDCNN-CycleGAN** and **TransCT-CycleGAN**, a novel transfer learning method, **MAP-NN-transfer**, and two supervised neural networks, **REDCNN** and **TransCT**.

In each dataset, that is, the quarter-dose abdomen CT dataset and 10%-dose chest CT dataset, we test the denoising performance on five patients, respectively. The quantitative results, that is, mean PSNR and SSIM

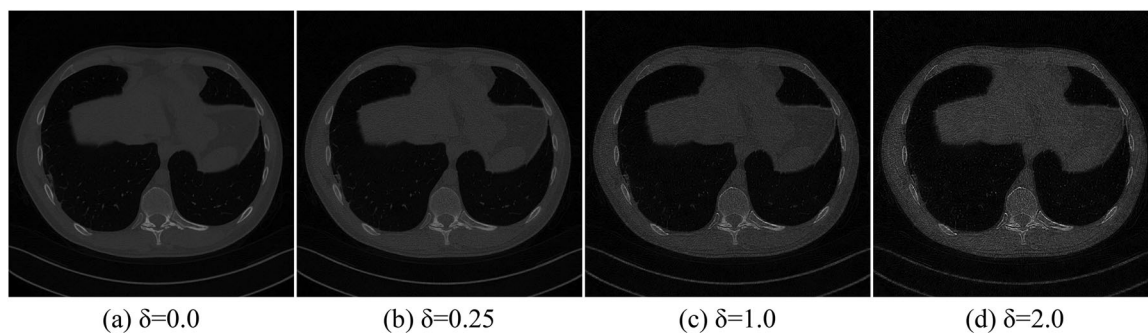


FIGURE 5 Simulated abdomen CT slices with different noise levels. The display window is $[-160, 240]$ HU.

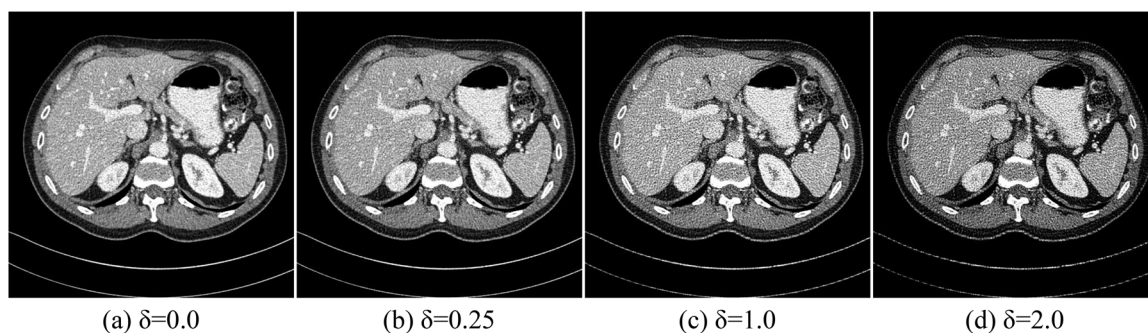


FIGURE 6 Simulated chest CT slices with different noise levels. The display window is $[-1024, 3072]$ HU.

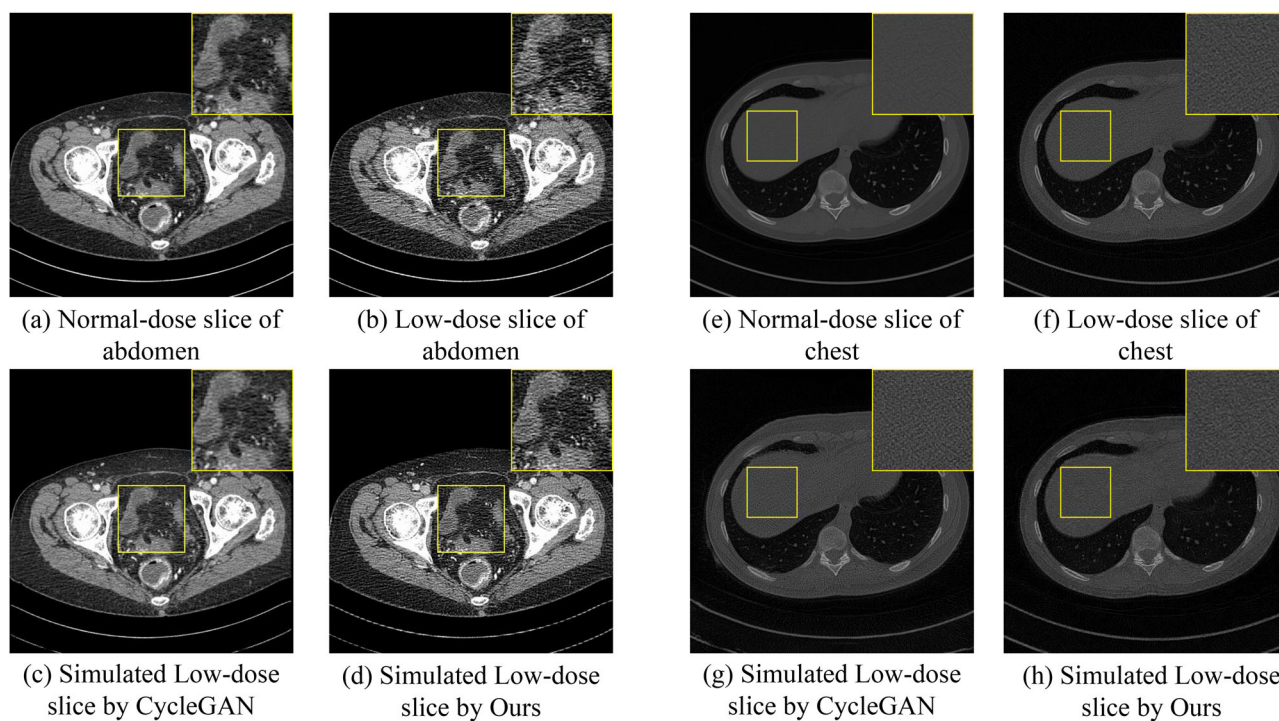


FIGURE 7 Simulation performance of our method and CycleGAN. The display window of abdomen CT images is $[-160, 240]$ HU, and that of chest CT images is $[-1024, 3072]$ HU.

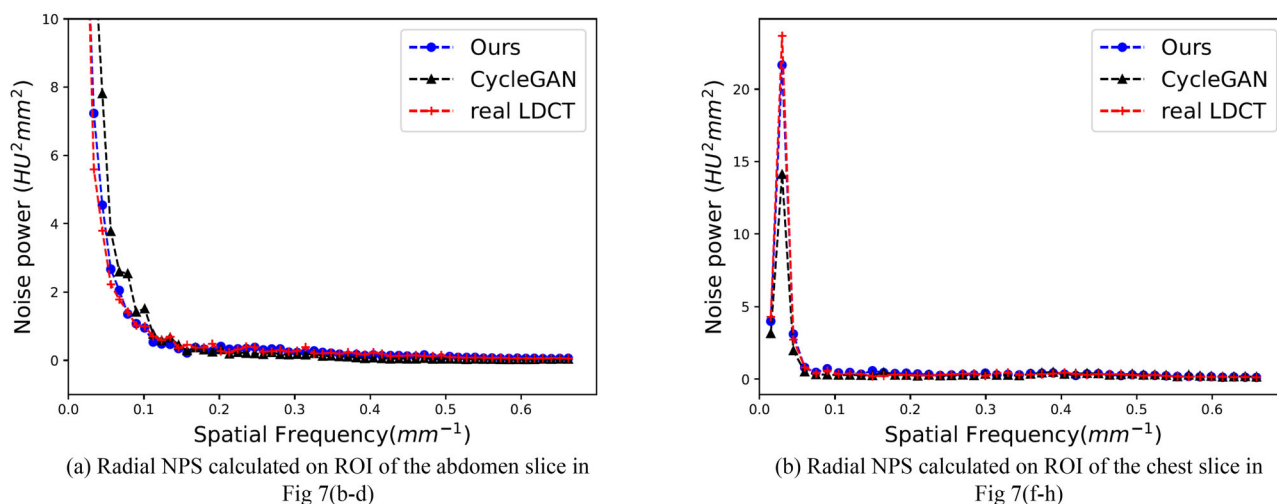


FIGURE 8 Radial noise power spectrum on ROIs in Figure 7

TABLE 1 Denoising results on quarter-dose abdomen dataset. Bolded indicates the best indices among unsupervised methods. Asterisk indicates the best indices among all of the methods

Patient Name	L109		L143		L291		L310		L333		ALL	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
LDCT	21.75	0.771	20.86	0.718	22.17	0.760	21.64	0.665	22.56	0.790	21.84	0.742
BM3D	24.13	0.803	23.21	0.755	24.61	0.794	24.38	0.714	25.00	0.818	24.32	0.778
NLM	23.81	0.800	22.74	0.750	24.51	0.793	24.45	0.720	24.87	0.821	24.14	0.778
REDCNN-CycleGAN	22.46	0.788	22.28	0.744	22.56	0.775	22.81	0.698	23.36	0.808	22.71	0.763
TransCT-CycleGAN	21.56	0.770	19.89	0.696	19.48	0.732	20.36	0.650	21.48	0.776	20.37	0.723
MAP-NN-transfer	23.84	0.807	23.57	0.764	24.24	0.798	24.16	0.718	24.60	0.827	24.12	0.783
REDCNN-Ours	24.24	0.808	23.68	0.768	24.61	0.799	24.37	0.721	25.04	0.827	24.43	0.785
TransCT-Ours	24.02	0.801	23.33	0.748	24.27	0.788	24.08	0.708	24.71	0.813	24.11	0.772
REDCNN	26.60*	0.821*	25.94*	0.774*	26.66*	0.806*	26.73*	0.744*	27.69*	0.836*	26.74*	0.796*
TransCT	25.36	0.805	24.70	0.748	25.34	0.787	25.49	0.720	26.21	0.815	25.42	0.775

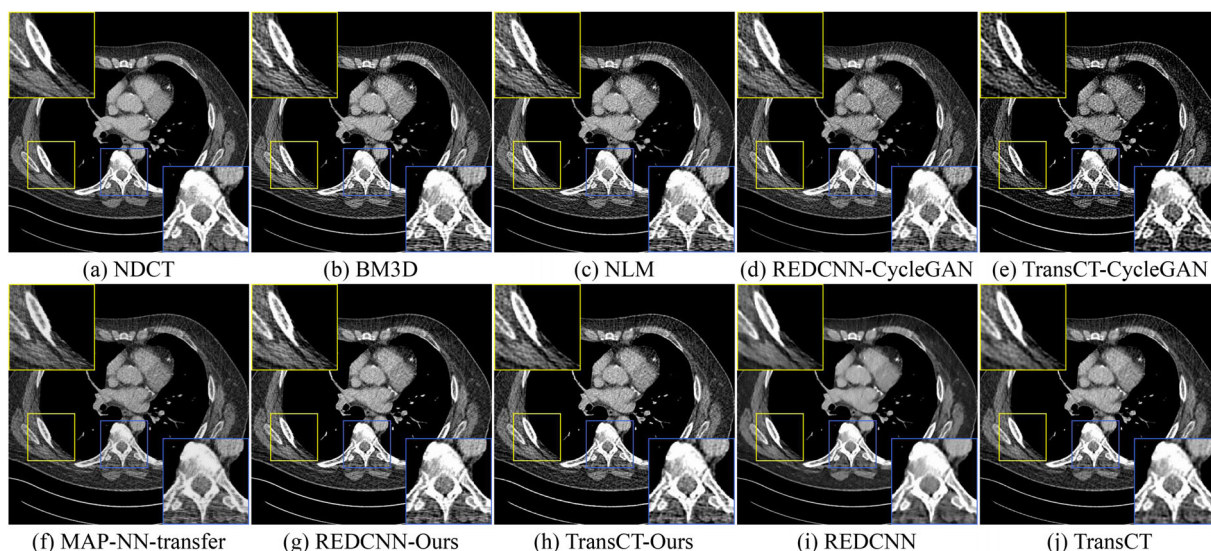
of each patient and all of the patients, are shown in Tables 1 and 2. The results of abdomen dataset are calculated with the HU window $[-160, 240]$ and the results of chest dataset are calculated with the HU window $[-1024, 3072]$. According to the indices, among the unsupervised and weakly supervised denoising methods, REDCNN-Ours and TransCT-Ours achieve the best PSNR and SSIM values. Although MAP-NN-transfer utilizes a few LDCT and NDCT image pairs for retraining, its results are not as good as our method, which does not use paired training data since the beginning. Notably, REDCNN-CycleGAN and TransCT-CycleGAN get extremely low indices, especially on the abdomen dataset, which contains relatively slight noise. This illustrates that it is not enough to train an effective denoising network, whereas the simulated LDCT images are only similar to the real ones in style. CycleGAN only learns the overall style of LDCT images but does not guar-

antee the accuracy of image details. Hence, it fails the training of denoising networks with a pixel-wise loss function. Our method does not have this problem, and it has both stylistic accuracy and pixel-wise precision. Therefore, the denoising networks trained with our simulated image pairs guarantee the accuracy of denoised images, which can be reflected in the indices of Tables 1 and 2.

Although the supervised REDCNN and TransCT have an obvious advantage over unsupervised and weakly supervised methods in terms of metrics, the denoised images always suffer from the typical over smoothing effect. In Figures 9 and 10, we show the denoising results of different methods visually. As shown in the blue zoom-in of Figure 9, denoised images of REDCNN and TransCT are significantly over smoothing compared to the NDCT image. As one can see, the results of BM3D, NLM, and networks trained with

TABLE 2 Denoising results on 10%-dose chest dataset. Bolded indicates the best indices among unsupervised methods. Asterisk indicates the best indices among all of the methods

Patient Name	C218		C219		C224		C258		C268		ALL	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
LDCT	29.76	0.626	26.14	0.510	26.07	0.470	22.12	0.346	27.50	0.547	26.33	0.501
BM3D	35.70	0.852	32.91	0.782	33.80	0.787	29.31	0.677	34.28	0.817	33.22	0.784
NLM	34.85	0.833	32.14	0.756	33.04	0.767	27.67	0.641	33.69	0.802	32.32	0.761
REDCNN-CycleGAN	31.12	0.846	30.12	0.785	30.47	0.774	28.32	0.678	30.42	0.807	30.10	0.777
TransCT-CycleGAN	32.41	0.854	31.71	0.789	31.40	0.793	29.66	0.704	31.07	0.819	31.25	0.792
MAP-NN-transfer	28.99	0.735	26.57	0.612	26.39	0.588	23.79	0.463	27.33	0.663	26.63	0.613
REDCNN-Ours	35.76	0.852	33.28	0.777	33.43	0.764	29.91	0.662	34.60	0.824	33.42	0.777
TransCT-Ours	36.04	0.867	33.85	0.802	33.88	0.790	30.81	0.697	34.72	0.828	33.88	0.797
REDCNN	38.45*	0.910*	36.45*	0.870*	36.32*	0.861*	33.26*	0.788*	37.21*	0.887*	36.36*	0.864*
TransCT	36.90	0.890	35.32	0.855	35.29	0.842	32.61	0.755	35.89	0.867	35.22	0.843

**FIGURE 9** Visual denoising results of an abdomen CT slice. The display window is $[-160, 240]$ HU.

CycleGAN-generated images do not suppress the noise successfully. In Figures 9(b)–(e), the ROIs are still noisy. According to Figure 9, Our methods simultaneously suppress the noise and avoid over smoothing or blurring and achieve the best performance on the abdomen dataset. In the yellow zoom-in of Figure 10, the results of BM3D and NLM are locally blurred. The red circles in the blue zoom-in of Figure 10 mark out a subsolid, noncalcified (ground-glass or semisolid) lesion according to the clinical report.^{36,40} In the denoised images of our methods, this lesion is as clearly visible as the NDCT image. As a comparison, the lesions are invisible in the results of BM3D and NLM, and networks with CycleGAN do not recover the lesion as clearly as NDCT. This illustrates that clinical diagnosis can benefit from our denoising methods.

6 | CONCLUSIONS AND DISCUSSION

Dose reduction has been a long-standing concern in computed tomography. LDCT denoising helps patients acquire high-quality CT images with much less radiation. Although deep learning-based methods have achieved great success in LDCT denoising, the lack of large amounts of paired training data undermines their practicability. We proposed a weakly supervised flow-based learning method to learn the degradation of LDCT, which can generate LDCT and NDCT image pairs. The invertibility of flow-based networks guarantees both the stylistic accuracy and pixel-wise precision, whereas the latter is always ignored by GAN-based image translation networks such as CycleGAN. The synthetic image pairs can be further used to train supervised LDCT

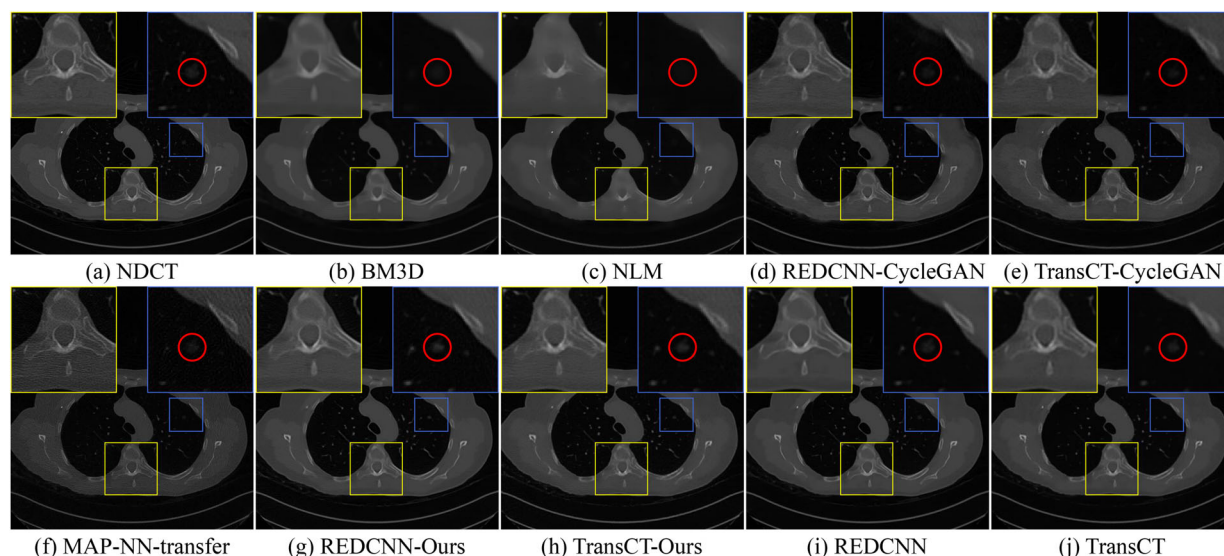


FIGURE 10 Visual denoising results of a chest CT slice. The display window is $[-1024, 3072]$ HU.

denoising neural networks. The denoising results of our methods are the best among several unsupervised and weakly supervised methods both quantitatively and visually and on par with the supervised networks in terms of visual effects.

In the future, we would like to generalize the weakly supervised learning-based methods to the projection domain and the reconstruction pipeline of CT to take full advantage of the projection information and acquire higher quality CT images in a LD manner. We also want to generalize our method to clinical data from different vendors.

ACKNOWLEDGMENTS

This work is supported in part by grants from the National Natural Science Foundation of China (U20A20373, 61871374, 62071197), the Chinese Academy of Sciences Special Research Assistant Grant Program, and the Young S&T Talent Training Program of Guangdong Provincial Association for S&T (SKXRC202224).

CONFLICT OF INTEREST

The authors have no relevant conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

The abdomen CT data are from “2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge.” The chest CT data are from Low Dose CT Image and Projection Data (LDCT-and-Projection-data),^{36,40} which is in public.

REFERENCES

- Brenner DJ, Hall EJ. Computed tomography—an increasing source of radiation exposure. *N Engl J Med*. 2007;357(22):2277-2284.
- De González AB, Mahesh M, Kim KP, et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. *Arch Internal Med*. 2009;169(22):2071-2077.
- Pearce MS, Salotti JA, Little MP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet*. 2012;380(9840):499-505.
- Jin KH, McCann MT, Froustey E, Unser M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Process*. 2017;26(9):4509-4522.
- Chen H, Zhang Y, Kalra MK, et al. Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans Med Imaging*. 2017;36(12):2524-2535.
- Chen H, Zhang Y, Zhang W, et al. Low-dose CT via convolutional neural network. *Biomed Opt Express*. 2017;8(2):679-694.
- Yang Q, Yan P, Zhang Y, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging*. 2018;37(6):1348-1357.
- Chun IY, Zheng X, Long Y, Fessler JA. BCD-Net for low-dose CT reconstruction: Acceleration, convergence, and generalization. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer; 2019:31-40.
- Zhang Z, Yu L, Liang X, Zhao W, Xing L. TransCT: dual-path transformer for low dose computed tomography. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer; 2021.
- La Rivière PJ, Bian J, Vargas PA. Penalized-likelihood sinogram restoration for computed tomography. *IEEE Trans Med Imaging*. 2006;25(8):1022-1036.
- Xu J, Tsui BM. Electronic noise modeling in statistical iterative reconstruction. *IEEE Trans Image Process*. 2009;18(6):1228-1238.
- Ma J, Liang Z, Fan Y, et al. Variance analysis of x-ray CT sinograms in the presence of electronic noise background. *Med Phys*. 2012;39(7Part1):4051-4065.
- Balda M, Hornegger J, Heismann B. Ray contribution masks for structure adaptive sinogram filtering. *IEEE Trans Med Imaging*. 2012;31(6):1228-1239.
- Manduca A, Yu L, Trzasko JD, et al. Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT. *Med Phys*. 2009;36(11):4911-4919.

15. 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(10):1272-1283.
16. Kim K, Soltanayev S, Chun SY. Unsupervised training of denoisers for low-dose CT reconstruction without full-dose ground truth. *IEEE J Sel Top Signal Process*. 2020;14(6):1112-1125.
17. Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys Med Biol*. 2008;53(17):4777-4808.
18. Chen Y, Gao D, Nie C, et al. Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior. *Comput Med Imaging Graphics*. 2009;33(7):495-500.
19. Ma J, Zhang H, Gao Y, et al. Iterative image reconstruction for cerebral perfusion CT using a pcontrast scan induced edge-preserving prior. *Phys Med Biol*. 2012;57(22):7519-7542.
20. Sun T, Sun N, Wang J, Tan S. Iterative CBCT reconstruction using Hessian penalty. *Phys Med Biol*. 2015;60(5):1965-1987.
21. Zhang Y, Zhang W, Lei Y, Zhou J. Few-view image reconstruction with fractional-order total variation. *JOSA A*. 2014;31(5):981-995.
22. Zhang Y, Wang Y, Zhang W, Lin F, Pu Y, Zhou J. Statistical iterative reconstruction using adaptive fractional order regularization. *Biomed Opt Express*. 2016;7(3):1015-1029.
23. Zhang Y, Xi Y, Yang Q, Cong W, Zhou J, Wang G. Spectral CT reconstruction with image sparsity and spectral mean. *IEEE Trans Comput Imaging*. 2016;2(4):510-523.
24. Xu Q, Yu H, Mou X, Zhang L, Hsieh J, Wang G. Low-dose X-ray CT reconstruction via dictionary learning. *IEEE Trans Med Imaging*. 2012;31(9):1682-1697.
25. Chen B, Xiang K, Gong Z, Wang J, Tan S. Statistical iterative CBCT reconstruction based on neural network. *IEEE Trans Med Imaging*. 2018;37(6):1511-1521.
26. Katsura M, Matsuda I, Akahane M, et al. Model-based iterative reconstruction technique for radiation dose reduction in chest CT: comparison with the adaptive statistical iterative reconstruction technique. *Eur Radiol*. 2012;22(8):1613-1623.
27. Liang X, Jiang Y, Zhao W, et al. Scatter correction for a clinical cone-beam CT system using an optimized stationary beam blocker in a single scan. *Med Phys*. 2019;46(7):3165-3179.
28. Li Z, Yu L, Trzasko JD, et al. Adaptive nonlocal means filtering based on local noise level for CT denoising. *Med Phys*. 2014;41(1):011908.
29. Feruglio PF, Vinegoni C, Gros J, Sbarbati A, Weissleder R. Block matching 3D random noise filtering for absorption optical projection tomography. *Phys Med Biol*. 2010;55(18):5401-5415.
30. Sheng K, Gou S, Wu J, Qi SX. Denoised and texture enhanced MVCT to improve soft tissue conspicuity. *Med Phys*. 2014;41(10):101916.
31. Kang D, Slomka P, Nakazato R, 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, pp. 86692G, International Society for Optics and Photonics, 2013.
32. Chen Y, Yin X, Shi L, et al. Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing. *Phys Med Biol*. 2013;58(16):5803-5820.
33. Buades A, Coll B, Morel JM. A non-local algorithm for image denoising. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 2. IEEE; 2005:60-65.
34. Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process*. 2006;54(11):4311-4322.
35. Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans Image Process*. 2007;16(8):2080-2095.
36. McCollough C, Chen B, Holmes D, et al. Low dose CT image and projection data [data set]. The Cancer Imaging Archive. 2020.
37. Lugmayr A, Danelljan M, Van Gool L, Timofte R. Srflow: learning the super-resolution space with normalizing flow. In: *European Conference on Computer Vision*. Springer; 2020:715-732.
38. Lu H, Hsiao T, Li X, Liang Z. Noise properties of low-dose CT projections and noise treatment by scale transformations. In: *2001 IEEE Nuclear Science Symposium Conference Record (Cat. No. 01CH37310)*, Vol. 3. IEEE, 2001:1662-1666.
39. Wolf V, Lugmayr A, Danelljan M, Van Gool L, Timofte R. Deflow: learning complex image degradations from unpaired data with conditional flows. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021:94-103.
40. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digital Imaging*. 2013;26(6):1045-1057.
41. Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision*, 2017:2223-2232.
42. Wu D, Kim K, Fakhri GE, Li Q. A cascaded convolutional neural network for x-ray low-dose CT image denoising. *arXiv preprint arXiv:1705.04267* 2017.
43. Jing L, Tian Y. Self-supervised visual feature learning with deep neural networks: a survey. *IEEE Trans Pattern Anal Mach Intell*. 2020;PP(99):1.
44. Taleb A, Loetzsch W, Danz N, et al. 3d self-supervised methods for medical imaging. *Adv Neural Inf Process Syst*. 2020;33:18158-18172.
45. Lehtinen J, Munkberg J, Hasselgren J, et al. Noise2Noise: learning image restoration without clean data. *arXiv preprint arXiv:1803.04189*. 2018.
46. Krull A, Buchholz TO, Jug F. Noise2void-learning denoising from single noisy images. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019:2129-2137.
47. Batson J, Royer L. Noise2self: blind denoising by self-supervision. In: *International Conference on Machine Learning*, PMLR, 2019:524-533.
48. Moran N, Schmidt D, Zhong Y, Coady P. Noisier2noise: learning to denoise from unpaired noisy data. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020:12064-12072.
49. Quan Y, Chen M, Pang T, Ji H. Self2self with dropout: Learning self-supervised denoising from single image. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020:1890-1898.
50. Yuan N, Zhou J, Qi J. Half2Half: deep neural network based CT image denoising without independent reference data. *Phys Med Biol*. 2020;65(21):215020.
51. Wu D, Gong K, Kim K, Li X, Li Q. Consensus neural network for medical imaging denoising with only noisy training samples. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2019:741-749.
52. Niu C, Li M, Fan F, et al. Suppression of correlated noise with similarity-based unsupervised deep learning. *arXiv preprint arXiv:2011.03384* 2020.
53. Kang E, Koo HJ, Yang DH, Seo JB, Ye JC. Cycle-consistent adversarial denoising network for multiphase coronary CT angiography. *Med Phys*. 2019;46(2):550-562.
54. Gu J, Ye JC. AdaIN-based tunable CycleGAN for efficient unsupervised low-dose CT denoising. *IEEE Trans Comput Imaging*. 2021;7:73-85.
55. Kim B, Shim H, Baek J. Weakly-supervised progressive denoising with unpaired CT images. *Med Image Anal*. 2021;71:102065.
56. Shan H, Kruger U, Wang G. A novel transfer learning framework for low-dose CT. In: *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and*

- Nuclear Medicine*, Vol. 11072, International Society for Optics and Photonics, 2019:110722Y.
57. Yu L, Shiung M, Jondal D, McCollough CH. Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols. *J Comput Assisted Tomogr*. 2012;36(4):477-487.
 58. Zeng D, Huang J, Bian Z, et al. A simple low-dose x-ray CT simulation from high-dose scan. *IEEE Trans Nucl Sci*. 2015;62(5):2226-2233.
 59. Žabić S, Wang Q, Morton T, Brown KM. A low dose simulation tool for CT systems with energy integrating detectors. *Med Phys*. 2013;40(3):031102.
 60. Divel SE, Pelc NJ. Accurate image domain noise insertion in CT images. *IEEE Trans Med Imaging*. 2019;39(6):1906-1916.
 61. Shan H, Jia X, Mueller K, Kruger U, Wang G. Low-dose CT simulation with a generative adversarial network. In: *Developments in X-Ray Tomography XII*, Vol. 11113, International Society for Optics and Photonics; 2019:111131F.
 62. Niu C, Wang G, Yan P, et al. Noise entangled GAN for low-dose CT simulation. *arXiv preprint arXiv:2102.09615*. 2021.
 63. Dinh L, Krueger D, Bengio Y. Nice: non-linear independent components estimation. *arXiv preprint arXiv:1410.8516* 2014.
 64. Dinh L, Sohl-Dickstein J, Bengio S. Density estimation using real nvp. *arXiv preprint arXiv:1605.08803* 2016.
 65. Kingma DP, Dhariwal P. Glow: generative flow with invertible 1x1 convolutions. *Adv Neural Inf Process Syst*. 2018;31.
 66. Lax PD. Change of variables in multiple integrals. *Am Math Mon*. 1999;106(6):497-501.
 67. Pumarola A, Popov S, Moreno-Noguer F, Ferrari V. C-flow: conditional generative flow models for images and 3d point clouds. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020:7949-7958.
 68. Liu R, Liu Y, Gong X, Wang X, Li H. Conditional adversarial generative flow for controllable image synthesis. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019:7992-8001.
 69. Lugmayr A, Danelljan M, Gool LV, Timofte R. SrfLOW: learning the super-resolution space with normalizing flow. In: *European Conference on Computer Vision*. Springer; 2020:715-732.
 70. Wang X, Yu K, Wu S, et al. Esrgan: enhanced super-resolution generative adversarial networks. In: *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018.
 71. Liang X, Chen L, Nguyen D, et al. Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. *Phys Med Biol*. 2019;64(12):125002.
 72. Dong G, Zhang C, Liang X, et al. A deep unsupervised learning model for artifact correction of pelvis cone-beam CT. *Front Oncol*. 2021;11.
 73. Dong G, Zhang C, Deng L, et al. A deep unsupervised learning framework for the 4D CBCT artifact correction. *Phys Med Biol*. 2022;67.
 74. Shan H, Padole A, Homayounieh F, et al. Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction. *Nat Mach Intell*. 2019;1(6):269-276.

How to cite this article: Liu X, Liang X, Deng L, Tan S, Xie Y. Learning low-dose CT degradation from unpaired data with flow-based model. *Med Phys*. 2022;49:7516–7530.
<https://doi.org/10.1002/mp.15886>