LOW-DOSE CT IMAGE DENOISING USING DEEP LEARNING METHODS

by

ZEHENG LI

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE AND ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2021

Copyright © by ZEHENG LI 2021

All Rights Reserved

To my parents and friends, for their passionate support.

ACKNOWLEDGEMENTS

I wish to express my appreciation to people who helped and guided me during my master career.

I would like to thank my advisors, Dr. Junzhou Huang and Dr. Mingwu Jin, for their patient and professional guidance for my research works. It is of great pleasure to work with them.

In addition, during my research career in Scalable Modeling and Imaging and Learning Lab (SMILE), my labmates have been constantly discussing with me. Not only have they inspired me in research, but also in many other ways.

Last but not least, I owe a debt of gratitude to my family for their unconditional support in the last three years.

April 27, 2021

ABSTRACT

LOW-DOSE CT IMAGE DENOISING USING DEEP LEARNING METHODS

ZEHENG LI, M.S.

The University of Texas at Arlington, 2021

Supervising Professor: Dr. Junzhou Huang

Low-dose computed tomography (LDCT) has raised highly attention since the counterpart, full-dose computed tomography (FDCT), brings potential ionizing radiation influence to patients. However, LDCT still suffers from several issues such as relatively higher noise level, which limits its uses in practical applications. To improve LDCT image quality, conventional denoising methods, such as KSVD and BM3D, are first introduced to suppress noise in low-dose images. These methods, however, works under assumptions that are not robust to various data. In this paper, we conduct an extensive research on deep learning based denoising method in LDCT images. We mainly base on Generative-Adversarial Network (GAN) variants, such as CycleGAN, IdentityGAN and GAN-CIRCLE, and compare their performance in low-dose image denoising. Compared to supervised deep learning methods, these GAN based methods effectively learn image translation from the low-dose domain to the full-dose (FD) domain without the need of aligning FDCT and LDCT images, which is usually required in other methods for domain translation. Experiments on real and synthetic patient CT data show that these methods can achieve comparable peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) to, if not

better than, the other state-of-the-art denoising methods. Among CycleGAN, IdentityGAN, and GAN-CIRCLE, the later achieves the best denoising performance with the shortest wall clock time. In addition, we use GAN-CIRCLE to demonstrate that the increasing number of training patches and training patients can improve denoising performance. Finally, two non-overlapping experiments, i.e., no counterparts of FDCT and LDCT images in the training data, further demonstrate the effectiveness of unpaired learning methods. This work paves the way for applying unpaired GAN based methods to enhance LDCT images without requiring aligned FD and low-dose images from the same patient.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

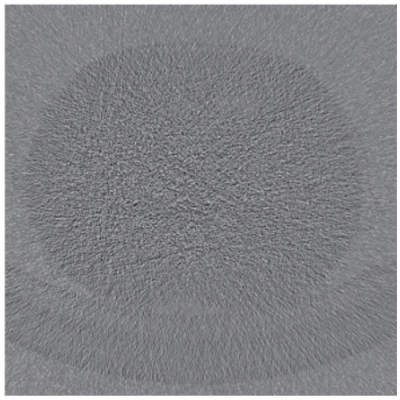
1.1 Low-dose CT denoising

In modern medicine, computed tomography (CT) is a common technique to produce cross-sectional images of the human body to find abnormalities through X-ray scanning around the patient. However, ionizing radiation in X-ray CT can potentially introduce adverse effects on patients [1], [2], especially for full-dose computer tomography (FDCT). To alleviate this problem, low-dose CT (LDCT) is pursued to scan the patient with the radiation dose lower than the regular dose without compromising the diagnostical accuracy of CT images [3], [4]. Compared to FDCT, the lower radiation dose of a CT scan features lowering X-ray tube cur- rent, lowering tube voltage, shortening exposure time, using sparse sampling, or combining the aforementioned techniques. In order to maintain the image quality, special image reconstruction and processing algorithms are needed to address the elevated noise and artifacts arisen in LDCT. In this work, we are particularly interested in addressing such high noise problem for LDCT using low tube current.

There are three major categories of LDCT denoising methods: 1) preprocessing of projection data [5]; 2) post- reconstruction image processing [6]–[12]; and 3) statistical iterative reconstruction [13]–[25]. All three categories of methods were under active research. For the preprocessing methods, the noise or artifacts in the projection data are either smoothed out or removed first. Then, the clean projection data can be used in a common reconstruction method such as filtered back-projection (FBP). For the statistical iterative reconstruction methods, the statistical nature of the projection

(a) full dose CT

(b) low dose CT



(c) noise

Figure 1.1: The examples of computed tomography

data and the desired image properties (using priors or regularizations) are modeled in the CT imaging process and solved iteratively. The methods in both categories have achieved much-improved image quality for LDCT. However, these methods require projection data, which is usually difficult to access due to the protection of the intellectual properties of the CT manufacturers. In addition, the iterative reconstruction methods usually suffer high computation complexity, which takes significant time if no special computational tools, such as high-end parallel hardware, are used. To address these issues, the post-reconstruction image processing methods directly work on the reconstructed images and provide an efficient tool for LDCT denoising. In this work, we extend the research along this direction.

1.2 Conventional Methods for Denoising

There are mainly two kinds of methods to suppress the noise in LDCT images: 1) conventional methods and 2) deep learning-based methods. The conventional methods usually first construct a simple relation between useful information and noise in images, then apply optimization algorithms to obtain denoised images. Toward this end, this type of method needs some prior knowledge or assumptions of noisy data. For example, a dictionary learning method, KSVD, was proposed to effectively reduce noise and artifacts in images using sparse representations of an overcomplete dictionary [26]. Both the atoms in the dictionary and the sparse representation of the dictionary are updated alternatively until the optimal denoising performance is achieved. Another popular and effective technique is the nonlocal means (NLM) filtering [9], [27]–[29], where the denoised pixel is computed as a weighted average of all the pixels in a search window, not like the conventional local filtering methods. NLM filtering can not only suppress noise in low-dose images but also preserve their details and texture. A patch-based NLM method (block-matching 3-D filter (BM3D) [30]) further improves denoising performance by block-matching and empirical Wiener filtering.

1.3 Deep Learning-based Methods for Denoising

The conventional denoising methods though effectively suppress noise in LDCT images, some important and detailed local information is usually suppressed too due to the generic principle of removing high-frequency components as noise. Besides, these handcrafted algorithms are usually only effective to a certain type of noise, but may perform badly if the prior knowledge is not accurate or the assumptions do not hold. Considering the high variety of CT images, more robust methods are

needed to provide a general solution for low-dose image denoising. Recently, the deep-learning-based methods are emerging as a promising alternative for CT image denoising. Without strong assumptions and manual feature engineering, these methods can automatically extract useful features in image data, which makes it easier to evolve the trained model by accommodating new data. In [31] and [32], convolutional neural networks (CNNs) were applied to CT denoising, which shows CNNs are very effective on learning features in full-dose (FD) high-quality CT images in a supervised manner. A follow-up development, RED-CNN [33], adopts encoder and decoder strategies to extract detailed structures in residual images. A wavelet-based deep learning method also demonstrates superior denoising performance for low-dose CT [34]. In addition, CNNs are also used to denoise CT perfusion maps instead of raw data [35]. In addition to the CNN-based methods, generative adversarial networks (GANs) [36], [37] are also playing an increasingly significant role in CT denoising. In GANs, by conducting a min–max learning between the discriminator and the generator, the "fake" images generated by the generator can be indistinguishable from real images even for a well-trained discriminator. A another interpretation of GANs is that it learns data distribution of real images and manages to generate "fake" images with similar distribution. This characteristic enables GAN to effectively denoise CT image by learning data distribution of FD images, and generate indistinguishable FD images from low-dose images. Inspired by this powerful model, a GAN-based denoising method is proposed for cardiac CT images in [38]. Another GAN-based CT denoising method adapted the Wasserstein distance instead of the Jensen–Shannon divergence for the adversarial loss [39]. These methods deliver promising results on image denoising and post-reconstruction. However, they have a big limitation in practical use: paired images are required during supervised training. Not only does the

creation of paired images demand a huge amount of effort but also a large amount of unpaired data cannot be utilized for training.

1.4 Goal of Thesis

To overcome such limitation, we propose a novel LDCT denoising method by learning unpaired image-to-image translation based on cycle-consistent GAN (CycleGAN) [40], [41]. The essence of CycleGAN is to enforce a cycle-consistency loss to learn image translation between two different domains (e.g., LDCT images and FDCT images) without requiring paired images. In the meantime, we noticed that a similar algorithm, GAN-CIRCLE [42], has been developed for super resolution and denoising of CT images. Another variant, IdentityGAN, was also proposed for LDCT angiography [43]. The unique contributions of our work are as follows: 1) we propose a different network structure for LDCT denoising and compare it with two other variants, GAN-CIRCLE and IdentityGAN; 2) we compare these unpaired deep learning denoising methods with the conventional KSVD and BM3D methods and a paired deep learning method, RED-CNN; and 3) we thoroughly investigate the influence of the number of patches, the number of patients, and totally non-overlapping domain images in training data on denoising performance using GAN-CIRCLE because it provides the best performance and the shortest training time for the data used in this work. This work will provide important information on the further development of unpaired deep learning methods for image enhancement of LDCT images using unpaired data.

CHAPTER 2

RELATED WORK

In this chapter, we first provide necessary background about CT image and the noise model. Then, we briefly discuss related works on CT image denoising using both conventional methods and machine learning based methods.

2.1 CT Image Noise Model

For a regular dose of high-tube current with high minimum noise equivalent quanta, the projection data after logarithm operation can be well approximated by an additive Gaussian noise [18]. On the other hand, at a low dose of low-tube current, the compound Poisson model is the most accurate to date, which can be approximated by a shifted Poisson distribution [44]. In addition, the noise properties in the CT images also depend on the reconstruction algorithms. The noise model in the image domain can be generalized as Eq(2.1):

$$X = T(Y) + N, \tag{2.1}$$

where $X \in \mathbb{R}^{m \times n}$ denotes the distorted and noise corrupted image (i.e., the low-dose image), $Y \in \mathbb{R}^{m \times n}$ is the original image to be recovered (i.e., the high-dose image), T is a transform function that distorts the original image or creates content-related artifacts, and $N \in \mathbb{R}^{m \times n}$ is the additive noise. Note that N does not necessarily follow the normal distribution even for the regular dose CT images.

The basic filtering methods usually have a couple of parameters to enforce the smoothness in a trial-and-error fashion. With increased computing power, more

elaborate filters, such as NLM, can be optimized iteratively based on the certain optimal criterion. Generally, a denoising operator G can be obtained by minimizing the following loss function:

$$\hat{G} = \min_{G} Loss(G(X) - X, G(X)) \tag{2.2}$$

This loss function often balances the data fidelity, i.e., $G(X) - X$, and the desired properties in the denoised image $G(X)$. G usually takes a fixed formula and its parameters will be determined through the iterative process. Afterward, the denoised image is obtained by applying G on X. Since there is no prior knowledge of T and Y, the success of these methods heavily relies on the assumption of the randomness of the noise.

When the images in both domains are available and ample, the denoising problem can be further formulated as Eq(2.3)

$$\hat{G} = \min_{G} Loss(G(X) - Y) \tag{2.3}$$

It is easy to see that a good operator of G will first suppress the noise, then do an inverse transform of T to recover Y. The advance of deep learning provides a powerful tool to solve Eq(2.3) since any nonlinear operation can be achieved by a deep network with sufficient capacity. However, the paired data are needed in this case, which can be circumvented by a cycle-consistency strategy deliberated in Chapter 3.

2.2 Generative Adversarial Network

The GAN is originally proposed to generate fake data samples that resemble the real counterparts. A typical GAN comprises one generator G and one discriminator D. G is to generate near-real fake images, while D is designed to discriminate fake images from real ones. During training, the min–max learning scheme is adopted,

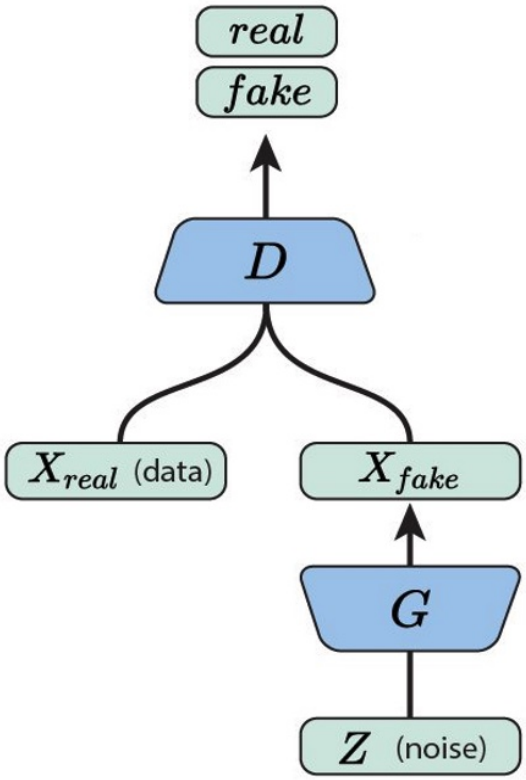


Figure 2.1: Basic GAN architecture

i.e., minimizing the classification error for G and maximizing it for D. The scheme will finally lead to a point where G generates images indistinguishable by D. Mathematically, the basic GAN objective function is defined as Eq(2.4)

$$\min_G \max_D \mathrm{L}_{GAN}(D, G) = \\ E_{y \sim p_{data}(y)}[\log D(y)] + E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \tag{2.4}$$

where y denotes real data samples, z is a random seed or data from another domain, E is the expectation operator. $p_{data}(y)$ captures the real data distribution, while $p_z(z)$ denotes the random distribution.

During training, various optimizers can be used to optimize loss functions. In addition, other constraints such as L_2 regularization for generated data can be included. Essentially, GAN map the distribution of the source data (z in Eq(2.4)) to

that of the target data (y in Eq(2.4)). Relating to low-dose image denoising, given an LDCT image x as the source, GAN can be trained to obtain the corresponding FD image y.

Although GAN holds a great potential to LDCT denoising, its training requires sufficient number of data samples. Furthermore, in the original form of GAN, the paired images, namely LDCT and FDCT images denotes the same object, are needed to establish the relation between the low-dose domain and full-dose domain. The collection of a large number of paired images requires huge efforts, even in some case is an impossible job considering ethics and patient safety issues. To attack this issue, we introduce a CycleGAN based denoising method that uses *unpaired* images between LDCT and FDCT data domain. Specifically, LDCT and FDCT data used in CycleGAN represent different domains, but not necessarily denote the exactly same object.

CHAPTER 3

UNPAIRED DEEP LEARNING METHODS

3.1 Conditional GAN

Generative adversarial networks (GANs) use a generative model G to learn the distribution and generate images given random noise as input. On the other hand a discriminative model D in GANs is a classification network which is to evaluate model G.

Original GAN can only generate images similar with real counterpart, like handwritten digits in MNIST, but are unable to tell the label of which is being generated. By adding auxiliary information y to both generator and discriminator, latter conditional generative adversarial nets can generate exact digits given certain label. y could be class labels or other relevant information. The objective function of conditional GAN is as Eq(3.1).

$$\begin{aligned} \min_G \max_D V(D,G) = & E_{x\sim p_{data}(x)}[logD(x \mid y)] \\ & + E_{z\sim p_z(z)}[log(1 - D(G(z \mid y)))] \end{aligned} \tag{3.1}$$

where, x are real images, z are random noise.

The structure of conditional GAN is shown as in Figure 3.1. Input of generator includes more information in addition to the noise. In Figure 3.1, we use a cartoon building as an example, which guide the generator to generate a fake building almost the same as the real one.

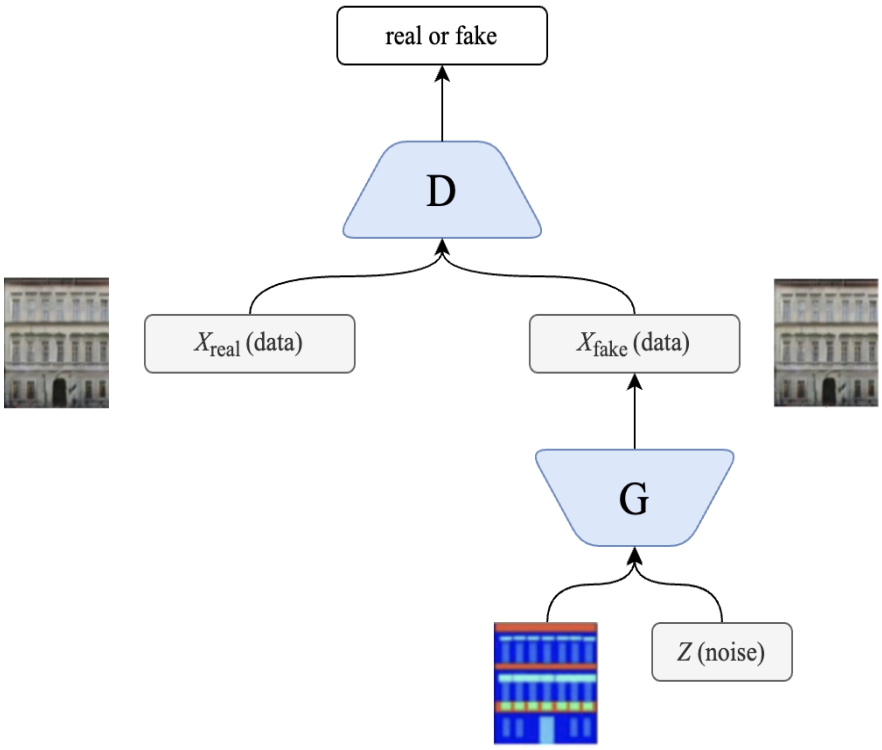


Figure 3.1: Conditional GAN architecture

3.2 CycleGAN

CycleGAN [41] is designed to effectively translate images from one domain X to another domain Y without requiring paired images. Figure3.2 shows an interesting application of CycleGAN, which transfer zebras to horses (and the other way around). In the following context, we denote X as the LDCT domain and Y as the full-dose CT (FDCT) domain. CycleGAN consists of two generators (G and F), as well as two discriminators (D_X and D_Y). The goal is to build two mappings between X and Y. Specifically, the generator G is trained to map images from the domain X into the domain $Y(G : X \rightarrow Y)$, and the generator F is trained to map images from the domain Y to the domain $X(F : Y \rightarrow X)$. Two discriminators are used to differentiate images from different domains: D_X aims to discriminate images in the

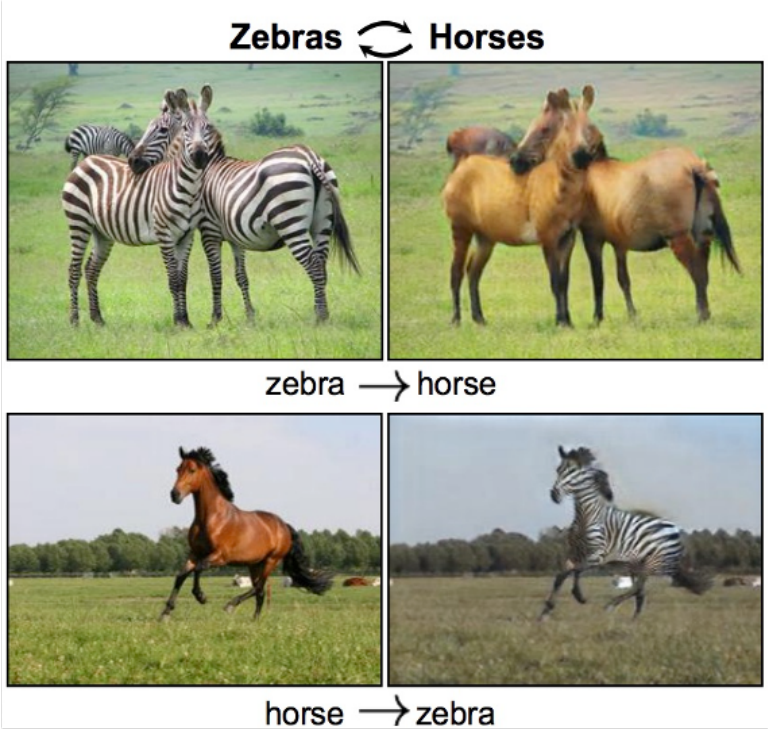


Figure 3.2: Application of CycleGAN

domain X from generated images from the domain Y, while D_Y distinguishes images in the domain Y from generated images from the domain X.

The basic structure of CycleGAN is shown in Figure 3.3, where the upper GAN is to learn the forward mapping from X to Y (blue arrows) and the bottom GAN is to learn the backward mapping from Y to X (red arrows). If there are no paired images for training and two GANs are trained separately, both of them will not converge to a good mapping between two domains.

In order to learn two mappings ($G : X \rightarrow Y$ and $F : Y \rightarrow X$) without paired samples, the cycle consistency loss is added to the CycleGAN structure as shown on the red and blue dashed arrows in Figure 3.3. Specifically, a sample in the domain X (a real LDCT image) first goes through G to generate a sample $G(x)$ in the domain Y (a fake FDCT image). Then, $G(x)$ is fed into F (red dashed arrow in Figure 3.3) to

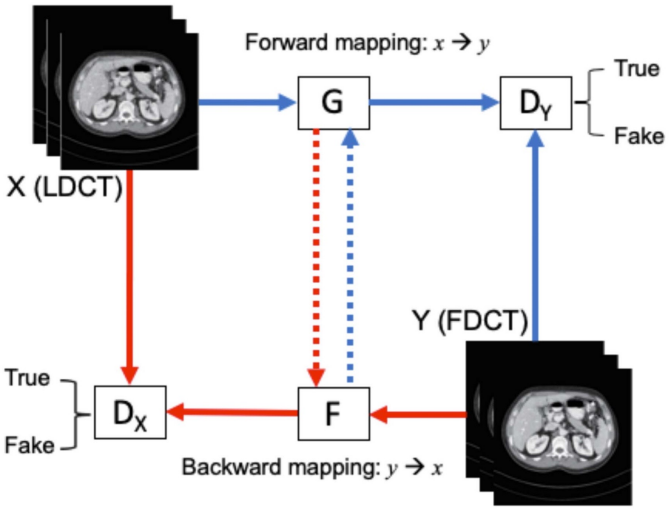


Figure 3.3: CycleGAN architecture

map it back to the domain X, i.e., $F(G(x))$. Finally, $F(G(x))$ will be discriminated with x in the domain X. A similar consistency loss is also applied to discriminate $G(F(y))$ with y in the domain Y (blue dashed arrow in Figure 3.3).

3.2.1 Loss Functions

In summary, the cycle consistency loss based on the forward consistency assumption, $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$, and the backward consistency assumption, $y \rightarrow F(y) \rightarrow G(F(y)) \approx$ y, can be defined as Eq(3.2).

$$\begin{aligned} L_{CYC}(G,F) = E_{x\sim P_{data}(x)}[\|F(G(x)) - x\|_1] \\ + E_{y\sim P_{data}(y)}[\|G(F(y)) - y\|_1] \end{aligned} \tag{3.2}$$

where $\|\cdot\|_1$ denotes the L1 norm, which can also be replaced by other metrics.

In addition, adding an identity loss was shown to significantly improve the quality of translated images. The function of the identity loss is to learn and keep

important structures and features in the image itself. The identity loss can be expressed as Eq(3.3)

$$\begin{aligned} L_{IDT}(G,F) = E_{y\sim P_{data}(y)}[\|G(y)-y\|_1] \\ +E_{x\sim P_{data}(x)}[\|F(x)-x\|_1] \end{aligned} \tag{3.3}$$

Finally, the objective function for CycleGAN can be written Eq(3.4).

$$\begin{aligned} L_{CycleGAN}(G,F,D_X,D_Y) = L_{GAN}(G,D_Y) + L_{GAN}(F,D_X) \\ +\lambda L_{CYC}(G,F) + \mu L_{IDT}(G,F) \end{aligned} \tag{3.4}$$

where the hyperparameters λ and μ control the relative importance of the original GAN losses, the cycle consistency loss, and the identity loss. In our work, we investigate the impact of λ on learning performance, while keeping μ as $0.5 \cdot \lambda$ (which can provide good results as long as it is not too small).

To avoid vanishing gradients when updating the generators G and F, we used the least-squares loss for L_{GAN} in Eq(3.4) [47] with the label "1" for the true data and the label "0" for the generated data, instead of the original negative log-likelihood loss function.

3.2.2 Network Structures

Different networks for the generators and discriminators in CycleGAN can lead to different performances. In this work, the generators (G and F) are built using residual blocks, while lightweight VGG-like networks are used for the discriminators (D_X and D_Y). Batch normalization is also used for stable learning for both generators and discriminators. To effectively discriminate the large size images, the discriminator takes the following process: first, the full image is broken down into 70×70 overlapping patches. Then, the patches are fed into the discriminator in a batch fashion. Instead of discriminating the full image directly, the discriminator first

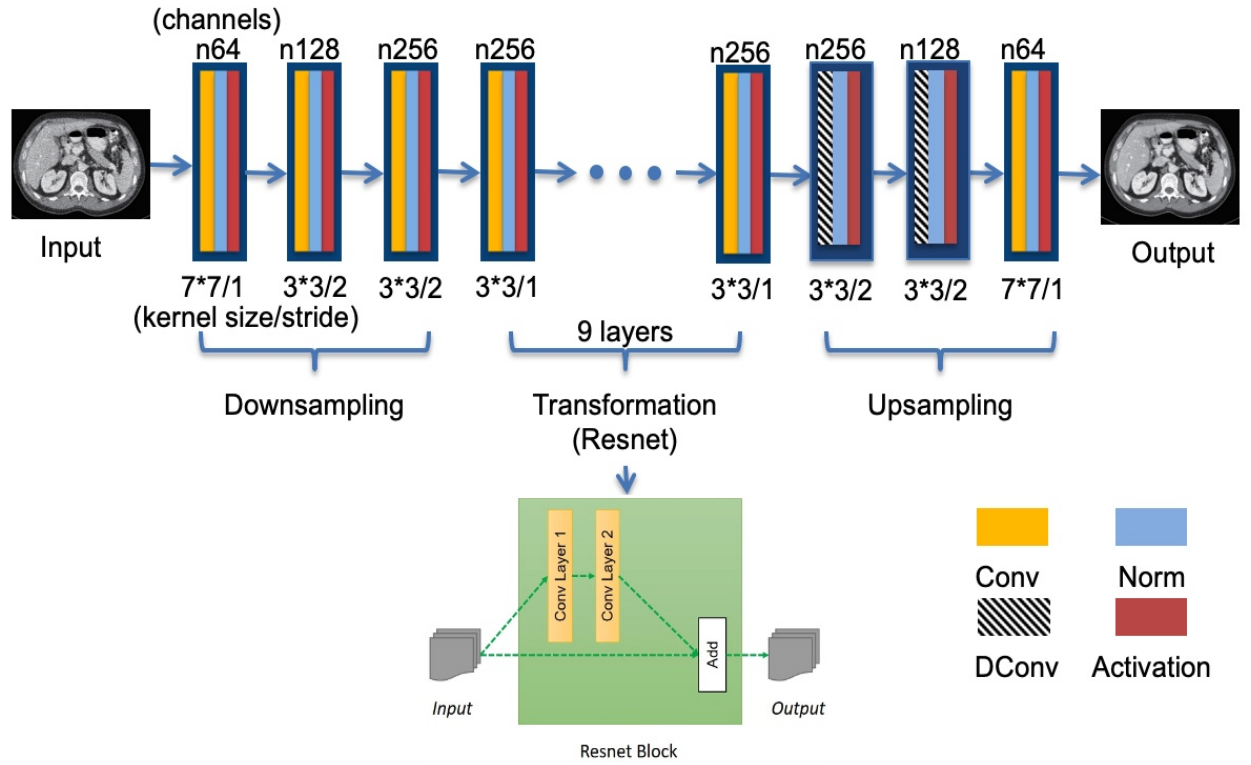


Figure 3.4: Generator structure of CycleGAN

makes a decision on every single patch, then synthesizes all patch results to get the final decision. The advantages of this process are that it greatly reduce the size of the discriminator, whose parameters are in a manageable order even for large images, and are adaptive to images with different sizes.

The detailed structures of the generator and the discriminator of CycleGAN used in this work are shown in Figure 3.4 and 3.5. For the generator in Figure 3.4, the yellow blocks represent convolutional layers, the blue ones are batch normalization, the red ones are rectified linear unit (ReLU) activation, and the shaded black-white ones are deconvolutional layers. The first three convolutional layers are used to down-sample original images, followed by a nine-layer residual network to further extract and synthesize features. The generators end with two deconvolutional layers for up-sampling and one convolutional layer for output. For the discriminator in Figure 3.5,

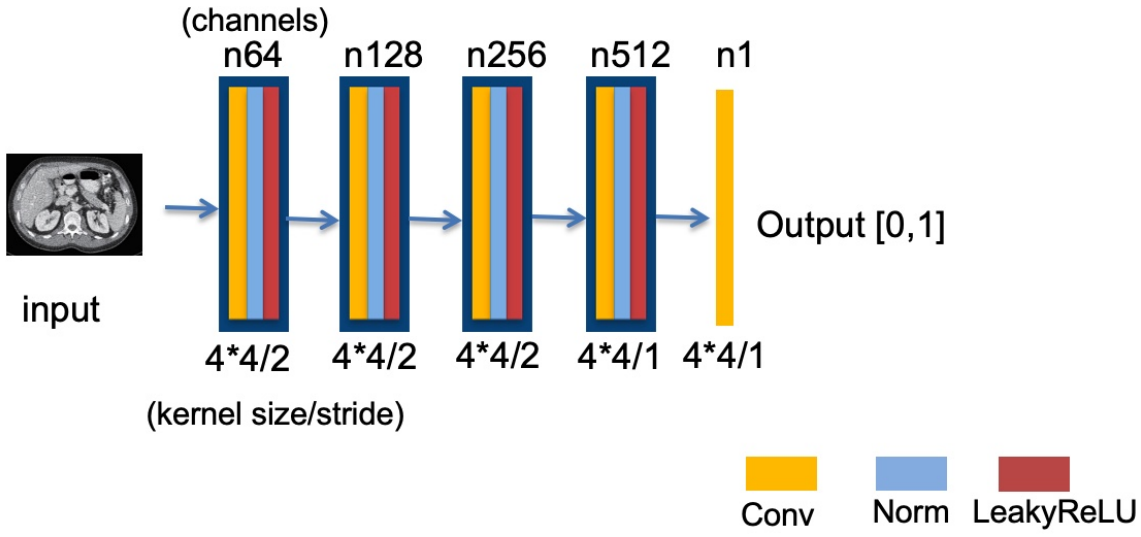


Figure 3.5: Discriminator structure of CycleGAN

the first four convolutional layers are used to extract features, then ending with a convolutional layer and a loss function 3.4 (greed block) for classification. Note that the red blocks represent leaky ReLU in the discriminator. The kernel sizes and stride sizes are noted in figures.

3.3 GAN-CIRCLE

In a recent work for CT super resolution and denoising, GAN-CIRCLE was proposed based on the CycleGAN principle. Here is a brief introduction and the details and the default hyperparameters can be found in [42].

3.3.1 Loss Functions

GAN-CIRCLE uses an additional loss function, joint sparsity transform (JST) loss, as shown in Eq(3.5).

$$L_{JST}(G) = \tau \left\| G(x) \right\|_{TV} + (1-\tau) \left\| y - G(x) \right\|_{TV} \tag{3.5}$$

where τ is a weighting factor and $\|\cdot\|_{TV}$ denotes the total variation (TV) norm. The two terms of this loss function serve two purposes: 1) the first term enforces sparsity on the generated images to remove artifacts and 2) the second term is to preserve anatomic details by minimizing the difference between y and $G(x)$. Thus, the overall objective function for GAN-CIRCLE becomes

$$\begin{aligned} L_{GAN-CIRCLE} = L_{WGAN}(D_Y, G) + L_{WGAN}(D_X, F) \\ +\lambda L_{CYC}(G, F) + \mu L_{IDT}(G, F) + \nu L_{JST}(G) \end{aligned} \tag{3.6}$$

where the GAN losses adopt the Wasserstein distance (for the forward mapping)

$$\begin{aligned} L_{WGAN}(D_Y, G) = E_x[D_y(G(x))] - E_y[D_y(y)] \\ +\alpha E_{\hat{y}}[(\|\nabla D_y(\hat{y})\|_2 - 1)^2] \end{aligned} \tag{3.7}$$

The first two terms in the Wasserstein loss function calculate the Wasserstein estimation, and the third term penalizes the deviation of the gradient norm of $D_y(\hat{y})$ from one, $\hat{y}$ is uniformly sampled along straight lines for pairs of G(x) and y, and α is a weighting parameter. A similar Wasserstein GAN loss can be defined for the backward mapping. Note that the JST loss was not used in this work since it did not improve the denoising performance significantly and increases the computation and the number of hyperparameters.

3.3.2 Network Structures

We summarize the network structures of GAN-CIRCLE based on their shared code at GitHub repo in Figure 3.6 and 3.7. Note that they are slightly different from that described in [42]. GAN-CIRCLE simplifies the generator by shrinking the kernel and channel sizes and adding more bypass connections. It also omits the batch normalization following the convolutional layers in the generator. As shown in Figure 3.6, each output feature is garnered at the end of the feature extraction

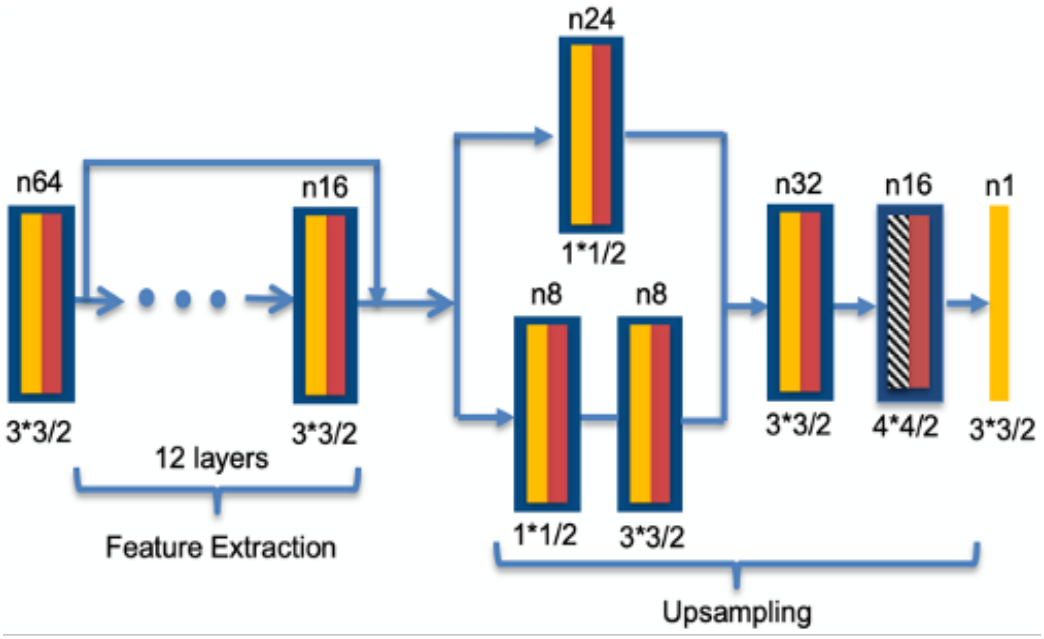


Figure 3.6: Generator structure of GAN-CIRCLE

stage, which can preserve information in previous convolutional layers. During upsampling, two-way convolutional layers are first used, then followed by a convolution layer and a deconvolution layer. The discriminator follows the same structure as that in CycleGAN, but with eight convolutional layers as shown in Figure 3.7. The block color and style represent the same functions as that in CycleGAN, except for the loss function [using Eq(3.6) instead of Eq(3.4)] (green block).

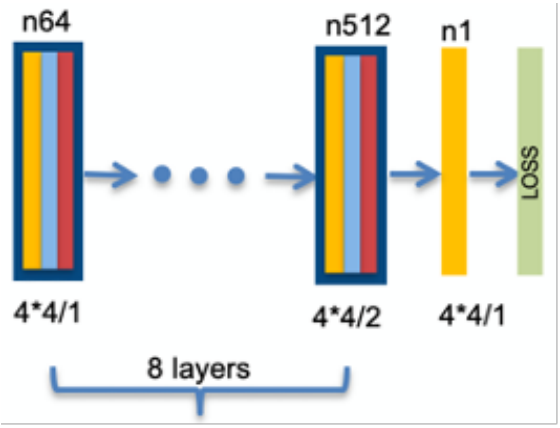


Figure 3.7: Discriminator structure of GAN-CIRCLE

CHAPTER 4

EXPERIMENTAL RESULTS

4.1 Experimental setup

To evaluate the denoising performance of CycleGAN, IdentityGAN, and GAN-CIRCLE, we performed the whole image-based learning, and then picked the best performer for patch-based training using a real clinical dataset from the 2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge (Link). Single-image iterative denoising methods, KSVD [26] and BM3D [30], and a paired deep learning denoising method, RED-CNN [33], were also implemented for comparison.

4.1.1 Training

The LDCT dataset contains ten anonymous patients' FD thoracic and abdominal CT data (close to 6000 2-D images for all patients) and corresponding low-dose (1/4 of the full-dose) CT data. The experiments are conducted in a system with NVIDIA Tesla V100 GPUs support. PyTorch ($\geq$1.4.0) and Tensorflow ($\geq$1.11) are used to train RED-CNN, CycleGAN, IdentityGAN, and GAN-CIRCLE.

1. **Whole Image Training**: The original images are resized from 512×512 to 256×256 to enable fast whole image-based training. We choose images of the 2nd to 10th patient as the training set (5376 images) and the 1st patient as the test data (560 images). Note that the training images are randomly scrambled, so there is no paired correspondence between the FD and low-dose images for each training batch.

2. **Patch-Based Training**: Since the original training dataset only has less than 6000 images, it is necessary to augment the training set to investigate the sample size influence on deep learning models for LDCT denoising. Specifically, if 64×64 overlapping patches were extracted from the original 512×512 images with a stride of 32, each image can produce 255 training patches. The whole process can lead to sufficient data (>1.5 million patches in total) to investigate the influence of the sample size on training performance. Since GAN-CIRCLE with simpler generator structures computes much faster (see Table 4.4) and has the best denoising performance among the three methods based on CycleGAN, we use GAN-CIRCLE for experiments in this part. Instead of fixing the training and test data sets, the tenfold cross-validation is used to get the estimate of expected performance on PSNR and SSIM on 512×512 images.

   First, we investigate the influence of the number of patches on the denoising performance of GAN-CIRCLE. One of ten patients is used as the test set and the other nine patients are used for training. The number of patches randomly selected from each image is 1, 4, 7, 13, 25, and 50, which leads to a total number of patches approximate 5000, 21000, 37000, 69000, 134000, and 268000 from images of nine training patients (where the exact number depends on the particular test patient selected). It repeats ten times until each of the ten patients was used as a test set. The quantitative evaluation metrics are averaged over ten trials for different numbers of training patches.

   Second, we investigate the influence of the number of patients on the denoising performance of GAN-CIRCLE. We use one patient as the test data and randomly select 1, 2, 4, and 6 patients in the other patients as the training data. The training is repeated ten times for each of the ten patients as the test set. The last experiment is the ten-fold cross-validation with one patient as the test

data and the other nine patients as the training data. Twenty-five patches per image are used to form the training data.

Finally, we conduct two additional experiments to demonstrate the effectiveness of denoising methods using unpaired deep learning: (1) non-overlapping patient training: the training set comprises FDCT images of patients 2–6 and LDCT images of patients 7–10, and the test data are patients 1. (2) non-overlapping organ training: the training set is composed of FDCT images for thoracic slices and LDCT images for abdominal slices of patients 2–9 and the test data are patient 1. In both cases, there is no corresponding FDCT image for the LDCT image.

4.1.2 Evaluation Metrics

Two metrics are used to quantitatively evaluate different methods: 1) peak signal-to-noise ratio (PSNR) and 2) structural similarity index (SSIM). PSNR measures denoising performance by calculating the overall difference between the denoised LDCT image and the original FDCT image and is defined as

$$\text{PSNR} = 10 \cdot log_{10}(\frac{MAX_Y^2}{MSE}) \tag{4.1}$$

where MSE stands for mean-squared error and MAX_Y denotes the maximum intensity value, which was set as 4095 for 12-bit CT images in this work. The MSE is calculated as

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[x(i,j) - y(i,j)]^2, \tag{4.2}$$

where x denotes the LDCT image or the corresponding region of interest (ROI), y is the corresponding FDCT image or ROI, and m and n denote the size of the image or ROI.

In addition to PSNR, the SSIM is also used to measure the perceptual similarity between the denoised LDCT and FDCT images

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{4.3}$$

where μ_x and μ_y are the averages of x and y, σ_x^2, σ_y^2, and σ_{xy} are the variance and covariance, and c_1 and c_2 are two variables to stabilize the division operation. The higher PSNR and SSIM represent the better performance. Note that the results calculated on the down-sampled 256×256 images are usually better than those calculated on the original 512×512 images.

4.2 Results

4.2.1 Hyperparameter Tuning

Base on the metrics defined above, we explore influences of hyperparameters, including the learning rate (LR) and the regularization factor λ (for the cycle consistency loss), as well as the GAN loss functions (least-squares loss versus Wasserstein loss) by evaluating the test dataset (patient 1). We use fixed 150 epochs.

The LR is fixed for the first 100 epochs, and then linearly decays to a half for the last 50 epochs. Tables 4.1 and 4.2 show how the combinations of LRs and the regularization factor λ (for GAN-CIRCLE with the least-squares loss) change the PSNR and SSIM performance, respectively. The large LR causes the training divergent, while the small LR fails to efficiently update model parameters, thus leading to inferior PSNR and SSIM performance. We chose the LR of 1×10^{-4} and $\lambda = 100$ as the hyperparameters for GAN-CIRCLE for the following experiments unless otherwise stated.

Table 4.1: Influence of LR and λ on PSNR (dB)

LR/λ	0.1	1	10	100	500
1×10^{-5}	44.07	44.32	45.15	46.28	46.24
1×10^{-4}	43.30	47.80	47.68	47.90	47.80
1×10^{-3}	Divergent				

Table 4.2: Influence of LR and λ on SSIM

LR/λ	0.1	1	10	100	500
1×10^{-5}	0.9637	0.9610	0.9649	0.9686	0.9675
1×10^{-4}	0.9583	0.9696	0.9743	0.9753	0.9747
1×10^{-3}	Divergent				

For μ, we tested the values of $[0, 0.1\lambda, 0.5\lambda, 0.75\lambda, \lambda]$ and found that the good PSNR could be obtained as long as $\mu \neq 0$. Thus, we use $\mu = 0.5\lambda$ in the following experiments.

Compared Wasserstein loss to lease-square loss, we found that the least-squares loss works better than the Wasserstein loss based on PSNR. Therefore, the least-squares loss is used in the following experiments.

For CycleGAN and IdentityGAN, we conduct a similar hyper parameter exploration. The optimal setting is: for CycleGAN, LR $= 2 \times 10^{-4}$ and $\lambda = 10$; for IdentityGAN, LR $= 2 \times 10^{-4}$ and $\lambda = 10$. Similarly, we tuned the parameters for KSVD (noise $\sigma = 0.007$), BM3D (noise $\sigma = 0.009$) and RED-CNN (LR $= 1 \times 10^{-5}$) to get their best PSNR values averaged on all images of patient 1.

The detailed results (averaged on all image slices of patient 1) for different methods are listed in Table 4.3. All denoising methods lead to better PSNR and SSIM than the original LDCT images (44.89 dB and 0.9372). GAN-CIRCLE achieves the

Table 4.3: Quantitative Results Averaged over 560 Test Images of 1st Patient

Method	PSNR(dB)	SSIM
LDCT	44.89(1.01)	0.9372(0.0165)
KSVD	47.34(0.77)	0.9638(0.0085)
BM3D	47.76(0.73)	0.9662(0.0075)
RED-CNN	47.82(0.74)	0.9730(0.0070)
CycleGAN	47.09(0.98)	0.9717(0.0080)
IdentityGAN	47.59(0.72)	0.9674(0.0085)
GAN-CIRCLE	47.90(0.69)	0.9753(0.0074)

best performance in PSNR and SSIM, followed closely by RED-CNN. CycleGAN and IdentityGAN also yield decent results. KSVD and BM3D work well on PSNR, but suffer lower SSIM values. Note that the PSNR and SSIM values obtained from our implementation of (unsupervised) GAN-CIRCLE are much higher than that in [42], likely due to the resized images and different parameters used in the calculation of these metrics, such as the dynamic range of converted images and MAX_Y in PSNR. Thus, the head-to-head comparison is not meaningful.

4.2.2 Denoised Images

To visualize the denoising performance of different methods, we select one thoracic slice containing the lungs and one abdominal slice containing the liver from the test dataset (patient 1) as shown in Figure 4.1 and Figure 4.3, respectively. The zoomed-in ROIs for details in the lung and the liver are shown in Figure 4.2 [red rectangle in Figure 4.1(a)] and Figure 4.4 [red rectangle in Figure 4.3(a)], respectively. As shown in Figure 4.1, the LDCT image shows streak noise compared to the FDCT image. All denoising methods alleviate the streak artifacts. Among them, KSVD

and BM3D seem to suppress these artifacts more at the expense of oversmoothing. The deep learning methods not only suppress the noise effectively but also preserve structural details. The PSNR and SSIM for this slice are listed in the caption. RED-CNN achieves the best quantitative accuracy, closely followed by GAN-CIRCLE. This phenomenon is more apparent in zoomed-in views in Figure 4.2. KSVD and BM3D oversmooth the region (e.g., the red circle area) and lead to low PSNR and SSIM. The BM3D image shows smoother looking, but suffers a great loss of details. Thus, its PSNR and SSIM are worse than LDCT, because the parameters of BM3D were tuned for all images of a patient, instead of the current slice. The paired deep learning method, RED-CNN, achieves the best noise removal performance and three unpaired deep learning methods with CycleGAN perform similarly.

The abdominal slice in Figure 4.3 shows more organs and structural details. Again, all denoising methods effectively remove the grainy and streak noise and lead to improved PSNR and SSIM. Among all methods, BM3D provides the best PSNR and GAN-CIRCLE yields the best SSIM. BM3D gains an advantage in more uniform regions in this slice although the texture of its denoised image is different from the FDCT image. In the zoomed-in view of Figure 4.4 for the liver region, BM3D shows the smoothest look, which strongly boosts its PSNR performance. However, the texture inside the liver is largely lost. The deep learning-based denoising methods achieves good performance on both suppressing noise and preserving texture details. The smoothness of this ROI is in the order of RED-CNN, GAN-CIRCLE, CycleGAN, and IdentityGAN. From these images, we can see that: 1) the deep learning-based denoising can adapt to different targets in the images with different noise properties and structural details compared to the conventional methods, which are less adaptive and 2) the unpaired deep learning methods based on CycleGAN are comparable to, if not better than, the paired deep learning method on LDCT denoising.

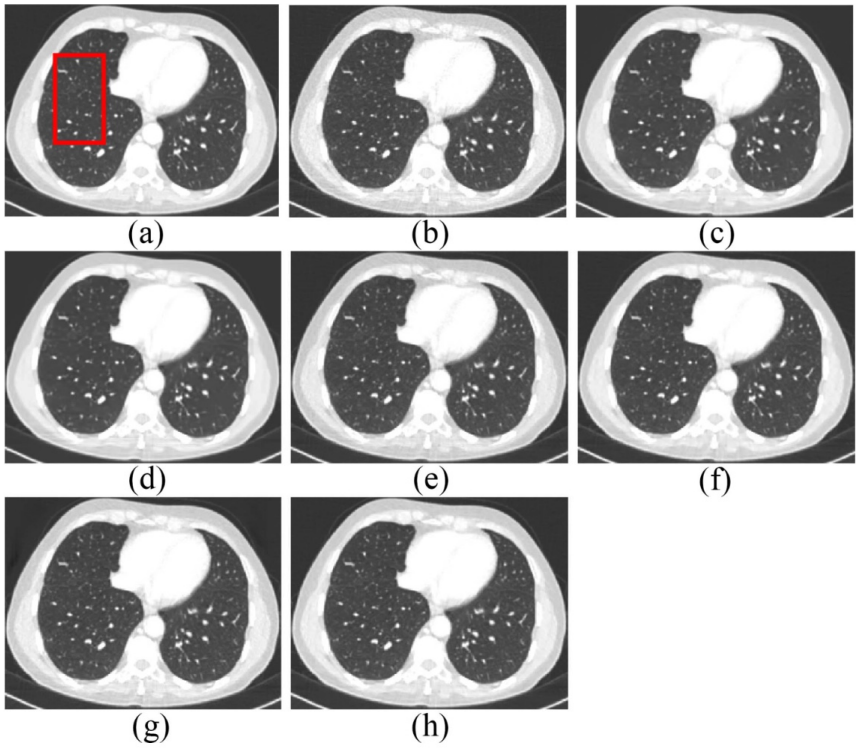


Figure 4.1: Thoracic transverse slice for different methods. (a) FDCT. (b) LDCT (PSNR: 44.64 dB and SSIM: 0.9565). (c) KSVD (46.25 dB and 0.9726). (d) BM3D (46.40 dB and 0.9719). (e) RED-CNN (46.57 dB and 0.9776). (f) CycleGAN (44.72 dB and 0.9756). (g) IdentityGAN (45.71 dB and 0.9619). (h) GAN-CIRCLE (46.31 dB and 0.9772). (Display window [-1350 150] HU).

4.2.3 Influence of the Size of Training Samples

GAN-CIRCLE has lightweight generators (0.168 million parameters for each generator), which is much more computationally efficient than CycleGAN and IdentityGAN (see Table 4.4). Furthermore, the hyperparameters are thoroughly investigated and optimized in [36] for patch-based learning. Therefore, we use GAN-CIRCLE to evaluate the influence of the size of training samples using image patches. Note that the PSNR and SSIM results in the following sections are calculated based on 512×512 images, which are different from previous sections using down-sampled

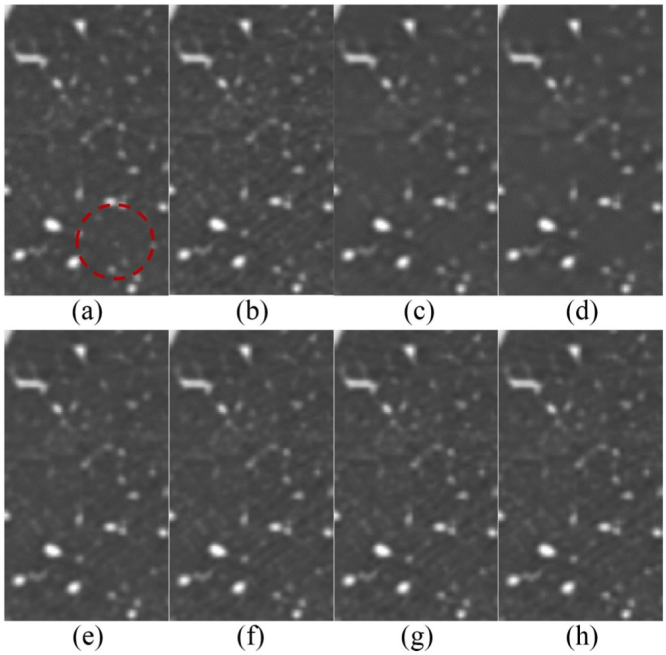


Figure 4.2: Zoomed-in lung region for different methods (the red rectangle in Fig. 4.1(a)). (a) FDCT. (b) LDCT (PSNR: 46.78 dB and SSIM: 0.9370). (c) KSVD (47.17 dB and 0.9409). (d) BM3D (46.59 dB and 0.9229). (e) RED-CNN (48.01 dB and 0.9515). (f) CycleGAN (47.55 dB and 0.9479). (g) IdentityGAN (47.17 dB and 0.9437). (h) GAN-CIRCLE (47.01 dB and 0.9435). (Display window [-1350 150]HU)

256×256 images. The baseline PSNR and SSIM of LDCT images averaged over ten patients are 40.88 dB(1.22) and 0.8710(0.0387), respectively.

1. **Influence of the Number of Training Patches**: As the patches extracted from each single 512×512 image increase from 1 to 50, the total training patches increase from about 5000 to about 268000. The average PSNR and SSIM values for ten trails are plotted in Figure 4.5 with error bars denoting the standard deviation. These values are significantly greater than the original LDCT values and indicate effective denoising. As can be seen, both PSNR and SSIM values increase with the number of training patches, except for from 4 patches/image to 7 patches/image. It implies that the larger the number of training samples, the better the denoising performance. However, the gain from the increased

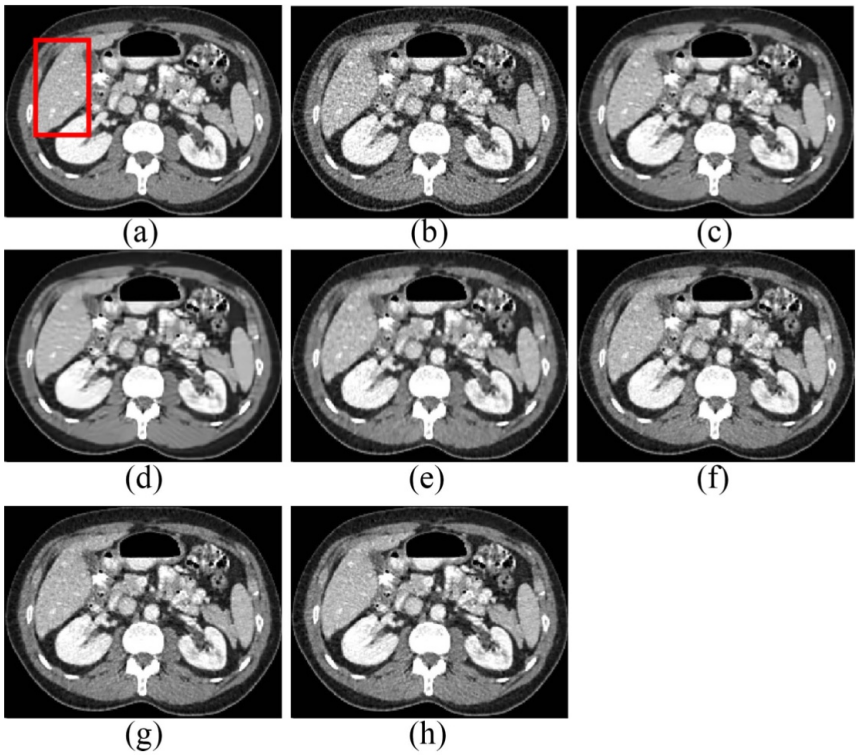


Figure 4.3: Abdominal transverse slice for different methods. (a) FDCT. (b) LDCT (PSNR: 43.23 dB and SSIM: 0.9149). (c) KSVD (45.82 dB and 0.9534). (d) BM3D (46.31 dB and 0.9572). (e) RED-CNN (46.23 dB and 0.9606). (f) CycleGAN (45.70 dB and 0.9554). (g) IdentityGAN (45.38 dB and 0.9523). (h) GAN-CIRCLE (46.21 dB and 0.9611). (Display window [-160 240]HU)

number of patches per image diminishes as it goes beyond 25 patches/image. The use of 25 patches/image seems to have a good balance between the denoising performance and computational burden.

2. **Influence of the Number of Patients for Training**: In this part, the number of patches per image is fixed as 25. As the number of patients for training increases from one to nine, both average PSNR and SSIM values increase as shown in Figure 4.6. Note that the data of every patient are served as the test data once for each condition of the number of training patients (1, 2, 4, 6, and 9). The error bar denotes the standard deviation of ten trials. The largest

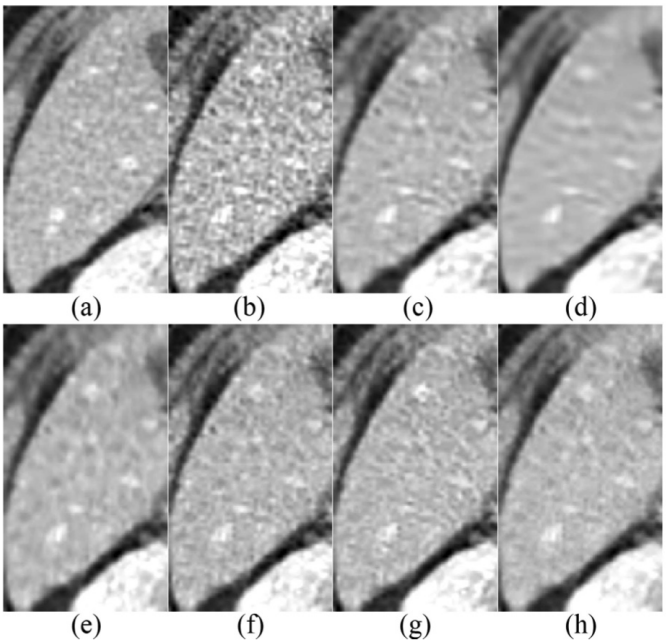


Figure 4.4: Zoomed-in liver region for different methods (the red rectangle in Fig. 4.3(a)). (a) FDCT. (b) LDCT (PSNR: 41.96 dB and SSIM: 0.6354). (c) KSVD (45.80 dB and 0.7404). (d) BM3D (46.60 dB and 0.7177). (e) RED-CNN (45.92 dB and 0.7363). (f) CycleGAN (45.13 dB and 0.7300). (g) IdentityGAN (44.83 dB and 0.7162). (h) GAN-CIRCLE (46.07 dB and 0.7423). (Display window [-160 240]HU)

increase is from one patient to two patients. Afterward, the increase becomes marginal. It demonstrates that the sufficient patient variety can significantly improve deep learning-based LDCT denoising. However, this improvement measured by PSNR and SSIM may reach a plateau at a certain number of training patients, e.g., four patients for the current dataset.

4.2.4 Totally Non-Overlapping Training

For non-overlapping patient training, the PSNR and SSIM are 43.31 dB and 0.9259, respectively. For non-overlapping organ training, the PNSR and SSIM are 43.47 dB and 0.9262, respectively. Even though there are no pairs of FDCT images and LDCT images in training, GAN-CIRCLE is able to extract useful features to

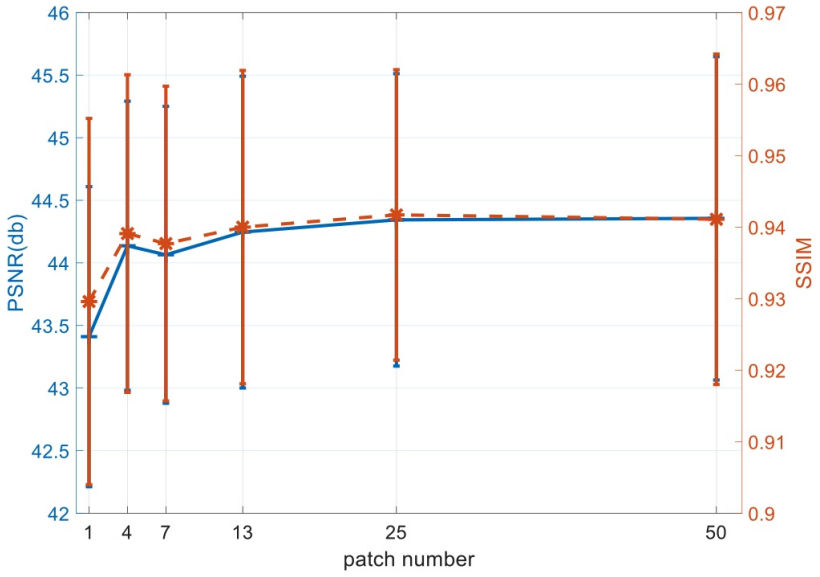


Figure 4.5: Influence of the number of training patches per image on denoising performance. The horizontal axis denotes the number of patches per image that are randomly extracted from each 512×512 image. Each of ten patients was used as the test data once and the results were averaged over ten trials for each data point.

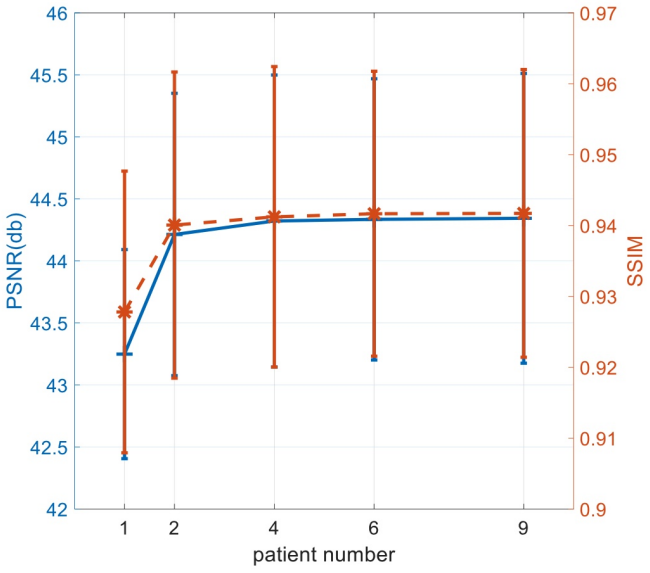


Figure 4.6: Influence of the number of patients in training (the horizontal axis on denoising performance. Twenty-five patches are randomly extracted from each 512×512 image. Each of ten patients was used as the test data once and the results were averaged over ten trials for each data point.

Table 4.4: Computation Time for Different Methods (Training for Each Epoch and Test for Each Image)

	RED-CNN	Cycle-GAN	IdentityGAN	GAN-CIRCLE
Train	1.20 min	21.12 min	119.78 min	5.95 min
Test	0.01 sec	0.05 sec	0.04 sec	0.02 sec

effectively remove noise. These two values are better than the original LDCT values (PSNR: 40.08 dB; SSIM: 0.8541), mixed training with one patient (PSNR: 43.25 dB; SSIM: 0.9236), and comparable to mixed training with four patients (PSNR: 43.50 dB; SSIM: 0.9268). The mixed training means that the FDCT and LDCT images of the same patient are used in the training although the one-to-one correspondence is scrambled and not utilized. One abdominal slice is shown in Figure 4.7. Even without pairs of of FDCT and LDCT images, non-overlapping training of GAN-CIRCLE effectively suppresses noise and achieves similar performance to the mixed training with four patients. This result provides preliminary evidence that large patient data without one-to-one correspondence, i.e., no need of registration or alignment, could be utilized for effective LDCT denoising through the cycle consistent adversarial learning.

4.2.5 Network Complexity and Computation Time

Three variants, CycleGAN [41], GAN-CIRCLE [42], and IdentityGAN [43], have been proposed for LDCT denoising or super resolution. Their network structures are at different complexity levels: 11.378 million parameters for one generator in CycleGAN, 0.168 million in GAN-CIRCLE, and 6.477 million for IdentityGAN. The training of deep learning-based methods for LDCT denoising is time consuming. The training time (min per epoch) and the test time (sec per slice) for the whole 256×256 image training are listed in Table IV using NVIDIA Tesla V100. For patch-based

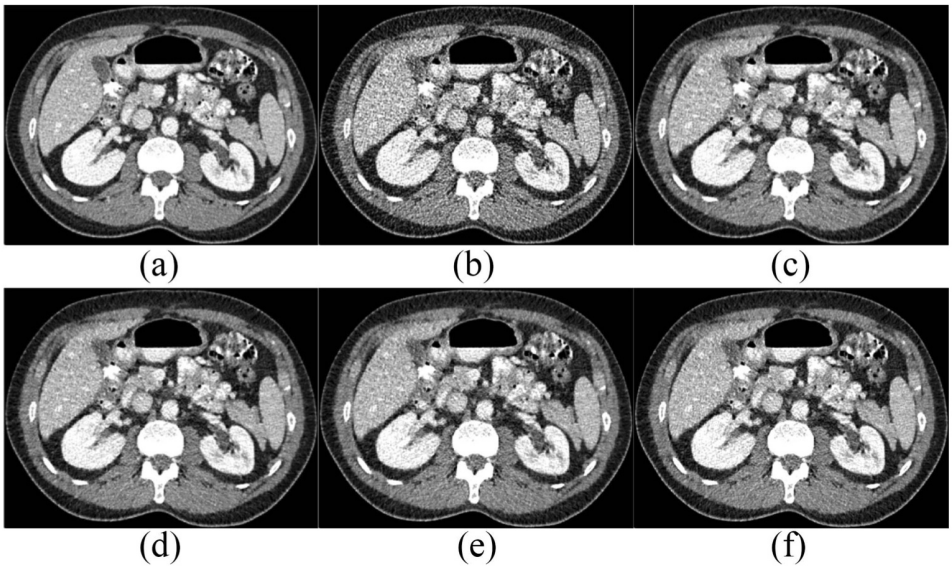


Figure 4.7: Abdominal transverse slice for (a) FDCT; (b) LDCT (PSNR: 38.28 dB and SSIM: 0.7960); (c) 1-patient mixed training (41.47 dB and 0.8891); (d) 4-patient mixed training (41.81 dB and 0.8952); (e) non-overlapping patient training (41.75 dB and 0.8924); and (f) non-overlapping organ training (41.74 dB and 0.8928). (Display window [-160 240] HU)

training, GAN-CIRCLE (13440064 $\times$ 64 patches by extracting 25 patches per image and a total of nine patients) takes about 23.5 h for 150 epochs with a batch size of 40 on the NVIDIA Tesla V100. The computation time linearly increases with the number of training patches. Once the training is done, the deep learning-based methods are much faster than KSVD and BM3D to denoise images (see test time in Table 4.4).

CHAPTER 5

CONCLUSION

In this work, we developed a novel image-to-image translation method based on CycleGAN and compared it with two other variants in LDCT image denoising. These cycle-consistent based GAN methods do not require aligned image pairs and yet provide the denoising performance comparable to the paired training, RED-CNN, which makes them particularly valuable to utilize abundant unpaired training samples. Among three variants, CycleGAN, IdentityGAN, and GAN-CIRCLE, the latter uses a lightweight network and achieves the best denoising performance for the NIH-AAPM-Mayo Clinic Low Dose CT data. The much more complicated network structures in CycleGAN and IdentityGAN do not improve the denoising performance, but demand much more computational resources. Although the LDCT data used in this study were synthesized using a validated noise insertion method, further investigation using real LDCT data would be preferred to verify the findings in this work and to improve unpaired deep learning methods. However, these studies are out of the scope of this work and will be investigated in the future.

Using patch-based training where more training samples are generated from patches, we investigated the influence of the size of training samples and the variety of training patients on denoising performance of the unpaired deep learning method. Our results show that increasing the number of training patches and the number of patients can improve the denoising performance as shown in Figure 4.5 and 4.6. The increase of the number of training patches saturates the improvement for 25 random patches per image even though more patches per image are available. This is likely due to the

information carried by the training data (of nine patients) is sufficiently represented by these patches. With further increase of the patches, the redundant information has none or little help on better training. The extra information introduced by new patient data can also boost the performance of deep learning-based denoising as shown in Figure 4.6. However, benefits of additional patient data diminish after using more than four patients. It seems that the current network structures can use less than nine patient data to achieve a denoising performance close to the optimum. We attempted to add the layers of GAN-CIRCLE and obtained a similar PSNR performance. Whether the network structures can be further optimized to achieve better denoising performance or the upper bound of denoising for the current dataset has been reached is worth further investigation.

The semi-unsupervised deep learning methods (using unpaired images), such as CycleGAN, have a unique advantage over the supervised deep learning methods (using paired images). They include more training data as no alignment of images in two domains is needed. (Note that we call this unpaired learning as "semi-unsupervised" learning because it still needs labels to denote which domain the training data belong to, while the unpaired learning was simply called "unsupervised" learning in [42] where semi-unsupervised was used for mixed paired and unpaired learning.) Although the LDCT and FDCT data in training were scrambled to void direct use of image pairs, the LDCT and FDCT pairs did exist in the training data, which could provide information leading to spurious good performance. The additional non-overlapping experiment was conducted to demonstrate the effectiveness of CycleGAN for unpaired training without counterparts in two domains, i.e., LDCT and FDCT, in terms of both patient-wise non-overlapping and organ-wise non-overlapping (Figure 4.7). These promising results show that unpaired learning based on CycleGAN can provide a

powerful tool to use a large number of images in two domains without the requirement of alignment.

In future work, we will investigate how to systematically optimize network architectures and hyperparameters since the empirical trial and error is not only suboptimal but also is unbearably time consuming. The investigation of the upper bound of denoising performance for the supervised deep learning methods (using paired images) and the semi-unsupervised deep learning methods (using unpaired images) based on CycleGAN for LDCT denoising will be another interesting direction. With a sufficiently large number of patient data, we envision that the unpaired learning can approach and go beyond the performance of the supervised learning.

PUBLICATIONS

Li, Zeheng, Junzhou Huang, Lifeng Yu, Yujie Chi, and Mingwu Jin. "Low-Dose CT Image Denoising Using Cycle-Consistent Adversarial Networks." In 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), pp. 1-3. IEEE, 2019.

Li, Zeheng, Shiwei Zhou, Junzhou Huang, Lifeng Yu, and Mingwu Jin. "Investigation of low-dose CT image denoising using unpaired deep learning methods." IEEE Transactions on Radiation and Plasma Medical Sciences (2020).

REFERENCES

[1] D. J. Brenner and E. J. Hall, "Computed tomography—An increasing source of radiation exposure," New England J. Med., vol. 357, no. 22, pp. 2277–2284, Nov. 2007, doi: 10.1056/NEJMra072149.

[2] A. B. de Gonzalez and S. Darby, "Risk of cancer from diagnostic X-rays: Estimates for the U.K. and 14 other countries," Lancet, vol. 363, no. 9406, pp. 345–351, 2004.

[3] P. J. La Rivière, J. Bian, and P. A. Vargas, "Penalized-likelihood sino- gram restoration for computed tomography," IEEE Trans. Med. Imag., vol. 25, no. 8, pp. 1022–1036, Aug. 2006.

[4] G. Wang, D. L. Snyder, J. A. O'Sullivan, and M. W. Vannier, "Iterative deblurring for CT metal artifact reduction," IEEE Trans. Med. Imag., vol. 15, no. 5, pp. 657–664, Oct. 1996.

[5] P. J. La Riviere, "Penalized-likelihood sinogram smoothing for low- dose CT," Med. Phys., vol. 32, no. 6, pp. 1676–1683, Jun. 2005, doi: 10.1118/1.1915015.

[6] Y. Chen et al., "Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means," Phys. Med. Biol., vol. 57, no. 9, pp. 2667–2688, 2012.

[7] J. C. R. Giraldo, Z. S. Kelm, L. S. Guimaraes, and L. Yu, "Comparative study of two image space noise reduction methods for computed tomog- raphy: Bilateral filter and nonlocal means," in Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2009, pp. 3529–3532.

[8] Z. Li et al., "Adaptive nonlocal means filtering based on local noise level for CT denoising," Med. Phys., vol. 41, no. 1, 2014, Art. no. 11908.

[9] J. Ma et al., "Low-dose computed tomography image restoration using previous normal-dose scan," Med. Phys., vol. 38, no. 10, pp. 5713–5731, 2011.

[10] W. Xu, S. Ha, and K. Mueller, "Database-assisted low-dose CT image restoration," Med. Phys., vol. 40, no. 3, 2013, Art. no. 031109.

[11] R.J.Geraldo,L.M.Cura,P.E.Cruvinel,andN.D.Mascarenhas,"Low dose CT filtering in the image domain using MAP algorithms," IEEE Trans. Radiat. Plasma Med. Sci., vol. 1, no. 1, pp. 56–67, Jan. 2017.

[12] A. M. Hasan, A. Melli, K. A. Wahid, and P. Babyn, "Denoising low- dose CT images using multiframe blind source separation and block matching filter," IEEE Trans. Radiat. Plasma Med. Sci., vol. 2, no. 4, pp. 279–287, Jul. 2018.

[13] G. H. Chen, J. Tang, and S. Leng, "Prior image constrained com- pressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," Med. Phys., vol. 35, no. 2, pp. 660–663, Feb. 2008, doi: 10.1118/1.2836423.

[14] L. Ouyang, T. Solberg, and J. Wang, "Noise reduction in low-dose cone beam CT by incorporating prior volumetric image information," Med. Phys., vol. 39, no. 5, pp. 2569–2577, 2012.

[15] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," Phys. Med. Biol., vol. 53, no. 17, pp. 4777–4807, Sep. 2008, doi: 10.1088/0031-9155/53/17/021.

[16] J. W. Stayman, H. Dang, Y. Ding, and J. H. Siewerdsen, "PIRPLE: a penalized-likelihood framework for incorporation of prior images in CT reconstruction," Phys. Med. Biol., vol. 58, no. 21, pp. 7563–7582, 2013.

[17] J. B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh, "A three- dimensional statistical approach to improved image quality for multislice helical CT," Med. Phys., vol. 34, no. 11, pp. 4526–4544, Nov. 2007, doi: 10.1118/1.2789499.

[18] J. Wang, T. Li, H. Lu, and Z. Liang, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low- dose X-ray computed tomography," IEEE Trans. Med. Imag., vol. 25, no. 10, pp. 1272–1283, Oct. 2006, doi: 10.1109/42.896783.

[19] H. Yu, S. Zhao, E. A. Hoffman, and G. Wang, "Ultra-low dose lung CT perfusion regularized by a previous scan," Acad. Radiol., vol. 16, no. 3, pp. 363–373, Mar. 2009, doi: 10.1016/j.acra.2008.09.003.

[20] H. Zhang, J. Huang, J. Ma, and Z. Bian, "Iterative reconstruction for X-ray computed tomography using prior-image induced nonlocal regu- larization," IEEE Trans. Biomed. Eng., vol. 61, no. 9, pp. 2367–2378, Sep. 2014.

[21] Z. Tian, X. Jia, K. Yuan, T. Pan, and S. B. Jiang, "Low-dose CT recon- struction via edge-preserving total variation regularization," Phys. Med. Biol., vol. 56, no. 18, p. 5949, 2011.

[22] Y. Zhang, Y. Wang, W. Zhang, F. Lin, Y. Pu, and J. Zhou, "Statistical iterative reconstruction using adaptive fractional order regularization," Biomed. Opt. Exp., vol. 7, no. 3, pp. 1015–1029, 2016.

[23] Y. Zhang, W. Zhang, Y. Lei, and J. Zhou, "Few-view image reconstruc- tion with fractional-order total variation," J. Opt. Soc. Amer. A Opt. Image Sci. Vision, vol. 31, no. 5, pp. 981–995, 2014.

[24] Y. Zhang, W. H. Zhang, H. Chen, M. L. Yang, T. Y. Li, and J. L. Zhou, "Few-view image reconstruction combining total variation and a high- order norm," Int. J. Imag. Syst. Technol., vol. 23, no. 3, pp. 249–255, 2013.

[25] G. L. Zeng and W. Wang, "Does noise weighting matter in CT iterative reconstruction?" IEEE Trans. Radiat. Plasma Med. Sci., vol. 1, no. 1, pp. 68–75, Jan. 2017.

[26] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

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

[28] J. Ma et al., "Iterative image reconstruction for cerebral perfusion CT using a pre-contrast scan induced edge-preserving prior," Phys. Med. Biol., vol. 57, no. 22, pp. 7519–7542, 2012.

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

[30] 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.

[31] H. Chen et al., "Low-dose CT denoising with convolutional neural network," in Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI), 2017, pp. 143–146.

[32] H. Chen et al., "Low-dose CT via convolutional neural network," Biomed. Opt. Exp., vol. 8, no. 2, pp. 679–694, 2017.

[33] H. Chen et al., "Low-dose CT with a residual encoder–decoder con- volutional neural network (RED-CNN)," IEEE Trans. Med. Imag., vol. 36, no. 12, pp. 2524–2535, Dec. 2017. [Online]. Available: https://128.84.21.199/pdf/1702.00288v1.pdf

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

[35] V. S. Kadimesetty, S. Gutta, S. Ganapathy, and P. K. Yalavarthy, "Convolutional neural network-based robust denoising of low-dose com- puted tomography perfusion maps," IEEE Trans. Radiat. Plasma Med. Sci., vol. 3, no. 2, pp. 137–152, Mar. 2019.

[36] I. Goodfellow et al., "Generative adversarial nets," in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 2672–2680.

[37] M. Mirza and S. Osindero. (2014). Conditional Generative Adversarial Nets. [Online]. Available: https://arxiv.org/abs/1411.1784

[38] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT," IEEE Trans. Med. Imag., vol. 36, no. 12, pp. 2536–2545, Dec. 2017.

[39] Q. Yang et al., "Low-dose CT image denoising using a generative adver- sarial network with Wasserstein distance and perceptual loss," IEEE Trans. Med. Imag., vol. 37, no. 6, pp. 1348–1357, Jun. 2018.

[40] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2017, pp. 1125–1134.

[41] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in Proc. IEEE Int. Conf. Comput. Vision, 2017, pp. 2223–2232.

[42] 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.

[43] E. Kang, H. J. Koo, D. H. Yang, J. B. Seo, and J. C. Ye, "Cycle- consistent adversarial denoising network for multiphase coronary CT angiography," Med. Phys., vol. 46, no. 2, pp. 550–562, 2019.

[44] J. Wang et al., "An experimental study on the noise properties of X-ray CT sinogram data in Radon space," Phys. Med. Biol., vol. 53, no. 12, p. 3327, 2008.

[45] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. Int. Conf. Learn. Represent., 2015, pp. 1–13.

[46] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classifica- tion with deep convolutional neural networks," in Proc. Adv. Neural Inf. Process. Syst., 2012, pp. 1097–1105.

[47] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks," in Proc. IEEE Int. Conf. Comput. Vision, 2017, pp. 2794–2802.

[48] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2016, pp. 770–778.

BIOGRAPHICAL STATEMENT

Zeheng Li received her B.S. degree in Electronic Engineering from Northeastern Polytechnical University in 2015. She also obtained her M.S degree in Information Engineering from Beihang University in 2018.