

Low Dose CT Image Denoising Using a Generative Adversarial Network with Perceptual Loss

Ravina Gadkar¹, Prof. Pravin Kalayankar²

¹M.E. Student, Department of Computer Engineering, TPCT'S
College of Engineering Osmanabad, Maharashtra, India

²Associate Professor, Department of Computer Engineering, TPCT'S
College of Engineering Osmanabad, Maharashtra, India

Abstract: The incessant development and universal use of CT in medical practice has raised a public concern over the associated radiation dose to the patient. Plummeting the radiation dose may lead to enlarged noise and artifacts, which can adversely affect the radiologist's judgment and confidence. Hence, advanced image restoration from low-dose CT data is needed to recover the diagnostic performance, which is an inspiring problem due to its ill-posed nature. Over the past years, various low-dose CT methods have produced exciting results. However, most of the algorithms developed for this application, including the recently disseminated deep learning performances, aim for minimizing the mean-squared-error (MSE) among a denoised CT image and the ground truth below generic penalties. Although the peak signal-to-noise ratio (PSNR) is developed, MSE- or weighted-MSE-based methods can compromise the perceptibility of important structural details after aggressive denoising. This paper introduces a new CT image denoising method based on the generative adversarial network (GAN) with Wasserstein distance and perceptual correspondence. The Wasserstein distance is a key perception of the optimal transport theory, and promises to improve the performance of GAN. The perceptual loss overpowers noise by comparing the perceptual features of a denoised output in contradiction of those of the ground truth in an established feature space, while the GAN focuses more on migrating the numbers noise distribution from tough to weak statistically. Therefore, our proposed method transfers our knowledge of visual awareness to the image denoising task and is capable of not only reducing the image noise level but also trying to keep the critical information at the same time. Promising results have been obtained in our experiments with clinical CT images.
Keywords: Noise Reduction Model, Perceptual Loss, Denoising Image

genetic damage and induce cancer in a likelihood interrelated to the radiation dose [1], [2]. Lowering the radiation dose intensifications the noise and artifacts in reassembled images, which can compromise diagnostic evidence. Hence, extensive efforts have been completed to design better image reconstruction or image processing methods for low-dose CT (LDCT). These methods commonly fall into three categories: (a) sinogram filtration beforehand reconstruction [3], (b) iterative reconstruction [5], [7], and (c) image postprocessing subsequently reconstruction [8]. Over the previous decade, researchers were dedicated to developing new iterative algorithms (IR) for LDCT image reconstruction. Commonly, those algorithms optimize an objective occupation that incorporates an precise system model [11], a statistical noise model [13] and prior information in the image domain. Popular image priors include total variation (TV) and its variants [16], as well as dictionary learning [19]. These iterative reconstruction algorithms impressively improved image quality but they may still lose particular details and suffer from outstanding artifacts. Also, they require a high computational asking price, which is a bottleneck in practical submissions. On the other hand, pre-filtration and image post processing are computationally efficient associated to iterative reconstruction. Noise characteristic was well modeled in the sinogram domain for sinogram-domain filtration. However, sinogram statistics of commercial scanners are not readily available to manipulators, and these methods may suffer from resolution loss and edge obscuring. The block-matching 3D (BM3D) algorithm was secondhand for image restoration in several CT imaging responsibilities [21]. With such image post-processing, image quality enhancement was indistinct but over-smoothing and/or residual errors were often observed in the processed images. These problems are difficult to address, given the nonuniform dissemination of CT image noise. The recent explosive improvement of deep neural networks suggests

I. INTRODUCTION

X-RAY computed tomography (CT) is one of the furthestmost significant imaging modalities in modern hospitals and clinics. However, in attendance is a potential radiation risk to the persevering, since x-rays could cause

new thinking and huge potential for the medical imaging pitch [22]. Despite the impressive denoising results with these innovative complex structures, they fall into a category of an end-to-end network that typically procedures the mean squared error (MSE) amongst the network output and the ground truth as the loss occupation. As revealed by the contemporary work, this perpixel MSE is often associated with over-smoothed edges and loss of information. As an algorithm tries to decrease per-pixel MSE, it overlooks subtle image textures/signatures critical for human perception. It is reasonable to assume that CT images distribute finished some manifolds. To tackle the above problems, here we propose to use a generative adversarial network (WGAN) with the Wasserstein distance as the inconsistency measure between distributions and a perceptual loss that calculates the difference between images in an established mouth space. The use of WGAN is to encourage that denoised CT images share the same dissemination as that of normal dose CT (NDCT) images. In the GAN framework, a generative network G and a discriminator network D are coupled forcefully and trained simultaneously. While the G network is trained to produce realistic images $G(z)$ from a random vector z , the D network is trained to discriminate between real and generated images. GANs have been used in many applications such as single image super-resolution, art creation, and image transformation. In the playing field of medical imaging, Nie *et al.* anticipated to use GAN to estimate CT image from its equivalent MR image. Wolterink *et al.* Are the paramount to apply GAN network for cardiac CT image denoising. And Yu *et al.* Despite its success in these areas, GANs still suffer from. In the original GAN, D and G are competent by solving the following minimax problem $\min_{G} \max_{D} L_{GAN}(D, G) =$

$$E_{x \sim P_r}[\log D(x)] + E_{z \sim P_z}[\log(1 - D(G(z)))] \dots (1)$$

where $E(\cdot)$ designates the expectation operator; P_r and P_z are the real data distribution and the noisy documents distribution. The generator G transforms a noisy sample to mimic a material sample, which defines a documents distribution, denoted by P_g . When D is trained to become an optimal discriminator for aimmovable G , the minimization examination for G is equivalent to minimizing the Jensen-Shannon (JS) divergence of P_r and P_g , which will central to vanished gradient on the generator G and G will discontinue updating as the training continues.

The foundation behind the perceptual loss is two-fold. First, when a person associates two images, the perception is not performed pixel-by-pixel. Human vision essentially

extracts and associates features from images. Therefore, instead of using pixel-wise MSE, we employment another pre-trained deep CNN (the famous VGG) for feature extraction and associate the denoised production against the ground certainty in terms of the mined features. Second, from a mathematical point of view, CT images are not consistently distributed in a high-dimensional Euclidean space. They reside more likely in a low-dimensional manifold. With MSE, we are not calculating the intrinsic similarity concerning the images, but just their superficial changes in the brute-force Euclidean distance. By comparing images according their intrinsic structures, we should project them against a manifold and estimate the geodesic distance instead. Therefore, the usage of the perceptual loss for WGAN should enable producing results with not only lower sound but also sharper information.

II. METHODS

A. Noise Reduction Model

Let $z \in \mathbb{R}^{N \times N}$ denote a LDCT image and $x \in \mathbb{R}^{N \times N}$ represent the corresponding NDCT image. The goal of the denoising process is to seek a occupation G that maps LDCT z to NDCT x :

$$G : z \rightarrow x \dots (2)$$

On the other hand, we can also take z as a sample from the LDCT image distribution P_L and x from the NDCT distribution or the real distribution P_r . The denoising function G maps samples from P_L into a certain distribution P_g . By fluctuating the function G , we aim to change P_g to make it close to P_r . In this way, we treat the denoising operation as moving one data dissemination to another. Typically, noise in x-ray photon dimensions can be simply modeled as the amalgamation of Poisson quantum noise and Gaussian electronic noise. On the contrary, in the reconstructed images, the noise prototypical is usually complicated and nonuniformly distributed across the whole image. Thus there is no strong clue that indicates how data distributions of NDCT and LDCT images are connected to each other, which makes it difficult to denoise LDCT images expending traditional methods. However, this uncertainty of noise model can be discounted in deep learning denoising because a deep neural system itself can efficiently learn high-level features and a demonstration of data distribution commencing modest sized image patches through a neural network.

B. WGAN

Related to the original GAN network, WGAN procedures the Wasserstein distance instead of the JS divergence to associate data distributions. It resolves the following minimax problem to obtain both D and G [41]: $\min_{G} \max_{D} L_{WGAN}(D, G) =$

$$-E_x[D(x)] + E_z[D(G(z))]GD + \lambda E_x[(\|\nabla_x D(x)\|_2 - 1)^2]$$

(3)

where the principal two terms perform a Wasserstein distance estimation; the preceding term is the gradient penalty term aimed at network regularization; x^{\wedge} is uniformly sampled along straight lines connecting couples of generated and real samples; and λ is a constant weighting parameter. Associated to the original GAN, WGAN removes the log function in the wounded and also drops the last sigmoid coating in the implementation of the discriminator D . specifically, the networks D and G are trained on the other hand by fixing one and informing the supplementary.

C. Perceptual Loss

While the WGAN network encourages that the generator transforms the data dissemination from high noise to a low noise version, additional part of the loss function is added for the network to keep image details or material content. Typically, a mean squared error (MSE) loss occupation is used, which tries to minimize the pixel-wise error flanked by a denoised patch $G(z)$ and a NDCT image patch x as [25], [26]

$$L_{MSE}(G) = E_{(x,z)} \left[\frac{1}{N^2} \|G(z) - x\|_F^2 \right], \quad (4)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. However, the MSE loss can potentially produce blurry images and cause the distortion or loss of details. Thus, as an alternative of using a MSE measure, we apply a perceptual loss function demarcated in a feature space

$$L_{Perceptual}(G) = E_{(x,z)} \left[\frac{1}{whd} \|\phi(G(z)) - \phi(x)\|_F^2 \right], \quad (5)$$

where ϕ is a feature extractor, and w , h , and d stand for the width, height and depth of the feature space, respectively. In our application, we adopt the well-known pre-trained VGG-19 network as the feature extractor. Since the pretrained VGG network takes color descriptions as input while CT images are in grayscale, we duplicated the CT images to make RGB channels previously they are fed into the VGG network. The VGG-19 network comprehends 16 convolutional layers followed by 3 fully-connected layers. The output of the 16th convolutional layer is the feature

removed by the VGG network and used in the perceptual loss occupation,

$$L_{VGG}(G) = E_{(x,z)} \left[\frac{1}{whd} \|VGG(G(z)) - VGG(x)\|_F^2 \right] \dots\dots\dots(6)$$

For opportuneness, we call the perceptual loss computed by VGG network VGG loss.

Combining Eqs. (3) and (6) together, we get the overall joint loss occupation expressed as

$$\min_{G} \max_{D} L_{WGAN}(D, G) + \lambda_1 L_{VGG}(G) \quad (7)$$

where λ_1 is a weighting consideration to control the trade-off between the WGAN adversarial loss and the VGG perceptual loss.

D. Network Structures

The complete view of the proposed network structure. For convenience, we name this network WGANVGG. It comprises three parts. The first part is the generator G , which is a convolutional neural network (CNN) of 8 convolutional layers. Following the communal practice in the deep learning community, small 3×3 kernels were rummage-sale in each convolutional layer. Due to the stacking structure, such a network can cover a great enough receptive field efficiently. Every single of the first 7 hidden layers of G have 32 filters. The last layer generates only one feature map with a single 3×3 filter, which is also the output of G . We use Rectified Linear Unit as the activation function.

The second part of the network is the perceptual loss calculator, which is realized by the pre-trained VGG network. A denoised output image $G(z)$ from the generator G and the ground truth image x are fed into the pre-trained VGG network for feature extraction. Then, the objective loss is calculated using the extracted features from a specified layer according to Eq. (6).

The third part of the network is the discriminator D . As shown in Fig. 2, D has 6 convolutional layers with the construction inspired by others' work [29]. The first two convolutional layers have 64 filters, then shadowed by two convolutional layers of 128 filters, and the last two

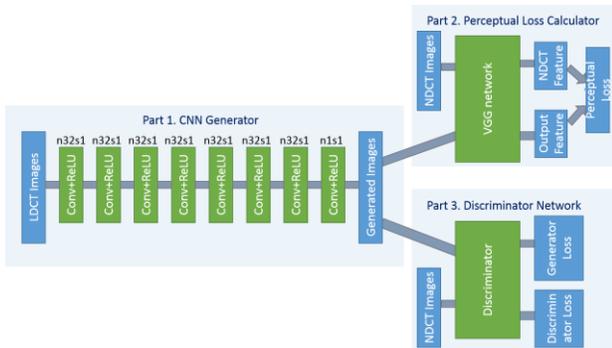


Fig. 2.1. The overall structure of the proposed WGAN-VGG network. In Part 1, n stands for the number of convolutional kernels and s for convolutional stride. So, $n32s1$ means the convolutional layer has 32 kernels with stride 1.

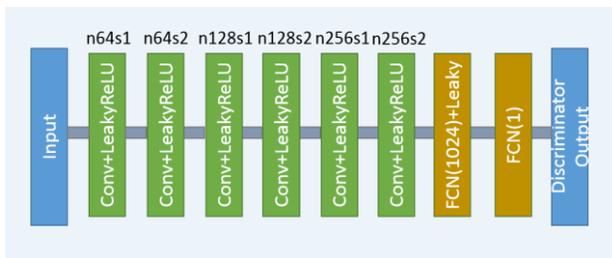


Fig. 2.2. The structure of the discriminator network. n and s have the same meaning as in Fig. 1

convolutional layers have 256 filters. Following the same logic as in G , all the convolutional layers in D have a small 3×3 kernel size. After the six convolutional layers, there are two fully-connected layers, of which the first has 1024 outputs and the other has a single output. there is no sigmoid cross entropy layer at the end of D .

The network is trained using image patches and applied on entire images. The details are provided in Section III on experiments.

E. Other Networks

For comparison, we also trained four other networks.

- CNN-MSE with only MSE loss
- CNN-VGG with only VGG loss
- WGAN-MSE with MSE loss in the WGAN framework
- WGAN with no other additive losses

- Original GAN

III SYSTEM ARCHITECTURE

A.EXISTING SYSTEM In existing approach, Over the past years, various low-dose CT methods have produced impressive results. However, most of the algorithms developed for this application, including the recently popularized deep learning techniques, aim for minimizing the mean-squared-error (MSE) between a denoised CT image and the ground truth under generic penalties. Although the peak signal-to-noise ratio (PSNR) is improved, MSE- or weighted-MSE-based methods can compromise the visibility of important structural details after aggressive denoising. This paper introduces a new CT image denoising method based on the generative adversarial network (GAN) with Wasserstein distance and perceptual similarity. The Wasserstein distance is a key concept of the optimal transport theory, and promises to improve the performance of GAN. The perceptual loss suppresses noise by comparing the perceptual features of a denoised output against those of the ground truth in an established feature space, while the GAN focuses more on migrating the data noise distribution from strong to weak statistically.

Disadvantages: These patches cannot be captured well by any compact quantized representation. It may not appear in another image. The drawback is that many of these methods require a large number of training samples which is not always available in practice. It relies on the ground-truth PatchSNR at each patch - information which is generally not available. It is not clear if these models actually capture the full statistical structure of natural images.

B. PROPOSED SYSTEM

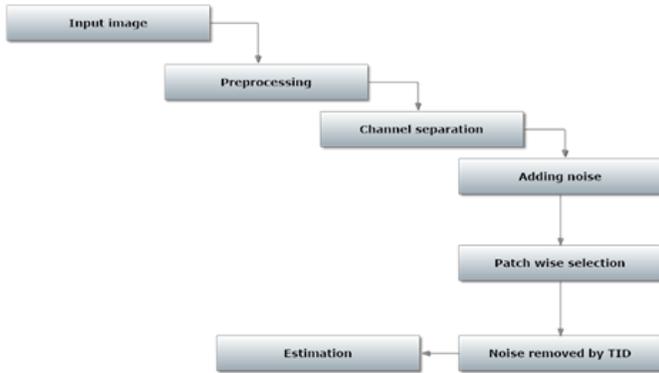


Fig 3.1 System Architecture

The proposed method, Typically, noise in x-ray photon measurements can be simply modeled as the combination of Poisson quantum noise and Gaussian electronic noise. On the contrary, in the reconstructed images, the noise model is usually complicated and non-uniformly distributed across the whole image. Thus there is no clear clue that indicates how data distributions of NDCT and LDCT images are related to each other, which makes it difficult to denoised LDCT images using traditional methods. However, this uncertainty of noise model can be ignored in deep learning denoising because a deep neural network itself can efficiently learn high-level features and a representation of data distribution from modest sized image patches through a neural network. Compared to the original GAN network, WGAN uses the Wasserstein distance instead of the JS divergence to compare data distributions. It solves the following minimax problem to obtain both D and G. In contrast to the existing methods, the proposed algorithm requires only a few targeted images in the database. Moreover, the proposed algorithm offers two new insights into the denoising problem. First, we show that when designing a linear denoising filter, the basis matrix can be learned by solving a convex optimization involving group sparsity, and the solution is the classical eigen-decomposition. This provides justifications of many well-known denoising algorithms in which PCA is used as a learning step. Second, we show that when estimating the spectral components of the denoising filter, a localized prior can be used and the denoising quality is improved by minimizing the associated Bayesian mean squared error. We show how such Bayesian treatment leads to a simple and effective denoising algorithm. This leads to a state-of-the-art denoising performance, equivalent and sometimes surpassing recently published leading alternative denoising methods. The employed 3-D transform there is a separable composition of the (2-D) SA-

DCT and a 1-D orthonormal transform. This transform is applied on 3-D groups that are generalized cylinders with adaptive-shape cross sections (as opposed to square prisms in BM3D). The magnified areas indicate that our proposed method removes the noise significantly and better reconstructs some fine details. In Figure, we plot and compare the average PSNR values on 4 test images over a range of noise levels. The results show that TID is consistently better than its competitors. Our first experiment considers denoising a text image with the help of other similar but non-identical texts. This is a simplified setup for problems such as hand writing, bar codes and license plates. To prepare the experiment, we add noise to a randomly chosen document and use 9 other clean documents (of the same font size) for denoising. We add zero-mean Gaussian noise with standard deviations from $\sigma = 20$ to $\sigma = 80$ to the test images. The patch size is set as 8×8 (i. e., $d = 64$). Two quality metrics, namely Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) are used to evaluate the objective quality of the denoised images. Among all the methods, TID yields the highest PSNR and SSIM values. The PSNR is 5 dB better than the benchmark BM3D (internal) denoising algorithm. Some existing learning-based methods, such as EPLL, do not perform well due to the insufficient training samples from the targeted database. Compared to other external denoising methods, TID shows a better utilization of the targeted database. The results show that TID is consistently better than its competitors. Since the default search window size for internal BM3D is only 39×39 , we conduct an experiment to explore the effect of different search window sizes for BM3D. The PSNR results are shown.

Flow Diagram of system

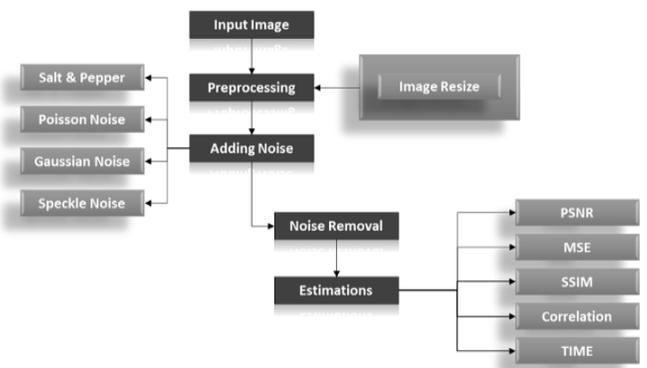


Fig 3.2 Flow Chart

Advantages:

The denoising is independently performed on each patch, GPU can be used to parallelize the computation. It is used to evaluate the objective quality of the denoised images. Compared to other external denoising methods, TID shows a better utilization of the targeted database.

IV. EXPERIMENTAL RESULT

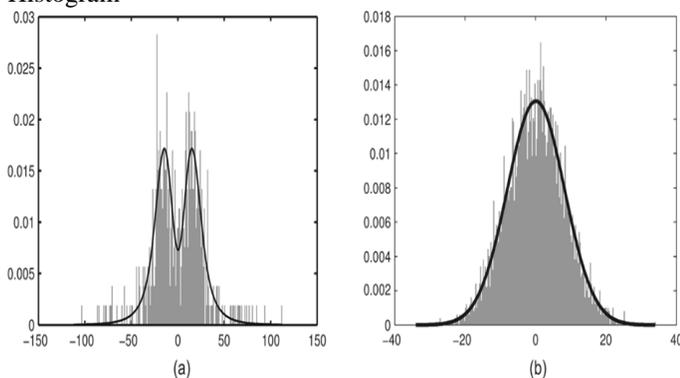
A. Experimental Datasets

We used a real clinical dataset authorized for “the 2016. NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge” by Mayo Clinic for the training and evaluation of the proposed networks [45]. The dataset contains 10 anonymous patients’ normal-dose abdominal CT images and simulated quarterdose CT images. When choosing the image patches, we excluded image patches that were mostly air. For comparison, we implemented a state-of-the-art 3D dictionary learning reconstruction technique as was performed from the LDCT projection data provided by Mayo Clinic.

B. Network Training

In our experiments, all the networks were optimized using Adam algorithm.

Fig. 4.1. Plots of validation (a) Noisy Histogram, (b) Denoising Histogram



Similar except that line 12 was changed to the corresponding loss function, and for CNN-MSE and CNN-VGG, lines 2-10 were removed and line 12 was changed according to their loss functions. The networks were implemented in Python with the library. A NVIDIA Titan XP GPU was used in this study.

C. Network Convergence

To visualize the convergence of the networks, we calculated the MSE loss and VGG loss over the 5,056 image patches for validation according to after each epoch shows the averaged MSE and VGG losses respectively versus the number of epochs for the five networks. Even though these two loss functions were not used at the same time for a given network, we still want to see how their values change during the training. In the two figures, both the MSE and VGG losses decreased initially, which indicates that the two metrics are positively correlated. However, the loss values of the networks in terms of MSE are increasing in the following order, CNN-MSE < WGAN-MSE < WGAN-VGG < CNN-VGG yet the VGG loss are in the opposite order. The MSE and VGG losses of GAN network are oscillating in the converging process. WGAN-VGG and CNN-VGG have very close VGG loss values, while their MSE losses are quite different. On the other hand, WGAN perturbed the convergence as measured by MSE but smoothly converged in terms of VGG loss. These observations suggest that the two metrics have different focuses when being used by the networks. The difference between MSE and VGG losses will be further revealed in the output images of the generators.

In order to show the convergence of WGAN part, we plotted the estimated Wasserstein values defined as $| - E[D(x)] + E[D(G(z))] |$ in Eq. (3). It can be observed that increasing the number of epochs did reduce the W-distance, although the decay rate becomes smaller. For the WGAN VGG curve, the introduction of VGG loss has helped to improve the perception/visibility at a cost of a compromised loss measure. For the WGAN and WGAN-MSE curves, we would like to note that what we computed is a surrogate for the W-distance which has not been normalized by the total number of pixels, and if we had done such a normalization the curves would have gone down closely to zero after 100 epochs.

D. Denoising Results

To show the denoising effect of the selected networks, we took two representative slices are the zoomed regions-of-interest (ROIs) marked by the red rectangles. All the networks demonstrated certain denoising capabilities. However, CNN-MSE blurred the images and introduced waxy artifacts as expected, which are easily observed in the zoomed ROIs. WGAN-MSE was able to improve the result of CNN-MSE by avoiding over-smooth but minor streak artifacts can still be observed especially compared to CNN-VGG and WGAN-VGG. Meanwhile, using WGAN or GAN alone generated stronger noise than the other networks enhanced a few white structures in the WGAN/GAN generated images, which are originated from the low dose

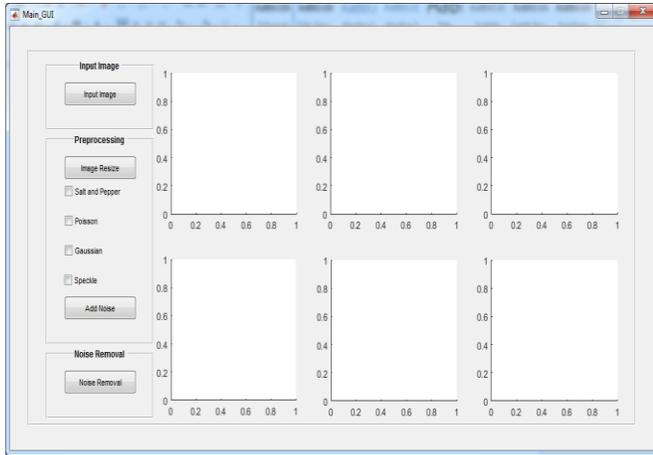


Fig.4.2 Input image

By using VGG loss, we transferred the knowledge of human perception that is embedded in VGG network to CT image quality evaluation. The performance of using WGAN or GAN alone is not acceptable because it only maps the data distribution from LDCT to NDCT but does not guarantee the image content correspondence. As for the lesion detection in these two slices, all the networks enhance the lesion visibility compared to the

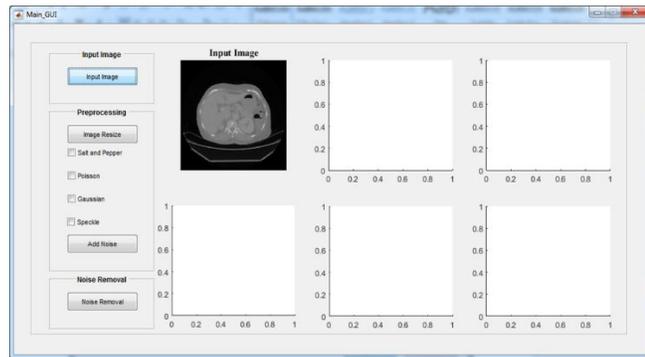


Fig.4.3 Display input image

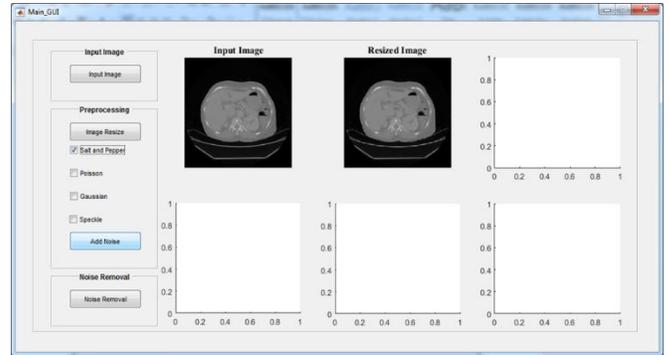


Fig.4.4 Resize image

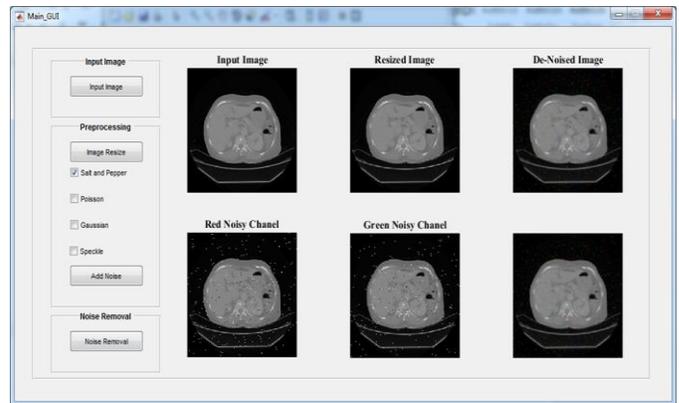


Fig.4.5 denoised Image

original noisy low dose FBP images as noise is reduced by the different approaches. As for iterative reconstruction technique, the reconstruction results depend greatly on the choices of the regularization parameters. As for the lesion detection in these two slices, all the networks enhance the lesion visibility compared to the original noisy low dose FBP images as noise is reduced by the different approaches. As for iterative reconstruction technique, the reconstruction results depend greatly on the choices of the regularization parameters.

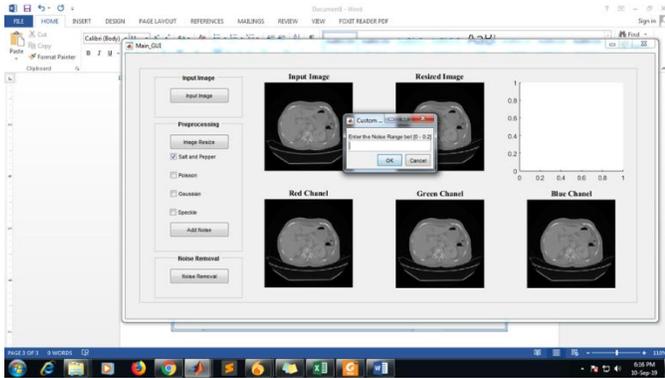


Fig.4.6 adding noise range

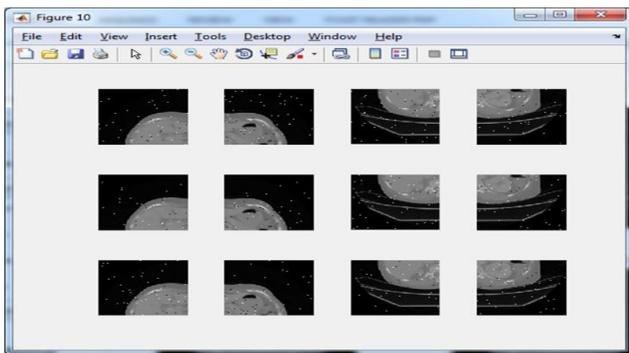


Fig.4.7 added noise range

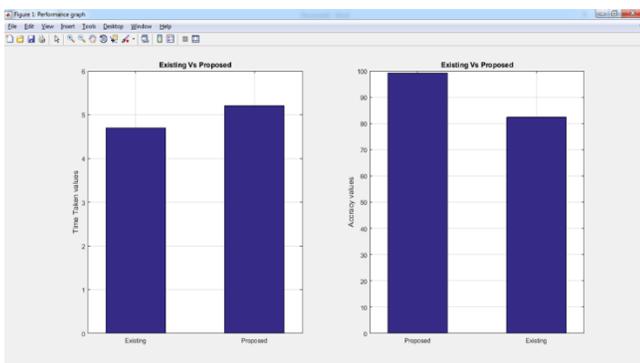


Fig.4.8 Performance Graph

The implemented dictionary learning reconstruction result gave the most aggressive noise reduction effect compared to the network outputs as a result of strong regularization. However, it over-smoothed some fine structures. The vessel pointed by the green arrow was smeared out while it is easily identifiable in NDCT as well as WGAN-VGG images. Yet, as an iterative reconstruction method, it has its advantage over postprocessing method. As pointed by the red arrow in Fig 8,

there is a bright spot which can be seen in NDCT images, but is not observable in LDCT and network. Since the WGAN-VGG image is generated from LDCT image, in which this bright spot is not easily observed, it is reasonable that we do not see the bright spot in the images processed by neural networks. In other words, we do not want the network to generate structure that does not exist in the original images. In short, the proposed WGAN-VGG network is a post-processing method and information that is lost during the FBP reconstruction cannot easily be recovered, which is one limitation for all the post-processing methods. On the other hand, as an iterative reconstruction method, algorithm generates images from raw data, which has more information than the post-processing methods.

E. Quantitative Analysis

CNN-MSE ranks the first in terms of PSNR, while WGAN is the worst. Since PSNR is equivalent to the per-pixel loss, it is not surprising that CNN-MSE, which was trained to minimize MSE loss, outperformed the networks trained to minimize other feature-based loss. It is worth noting that these quantitative results are in decent agreement, in which CNN-MSE has the smallest MSE loss and WGAN has the largest. In Wolterink *et al.* trained three networks, i.e. GAN, CNN-MSE, and GAN-MSE for cardiac CT denoising. Their results are also consistent with ours. Interestingly, despite the high PSNRs obtained by MSE-based networks, the authors in the two papers all claim that GAN and VGG loss based networks have better image quality and diagnostic information. To gain more insight into the output images from different approaches, we inspect the statistical properties by calculating the mean CT numbers (Hounsfield Units) and standard deviations (SDs) of two flat regions (marked by the blue rectangles). In an ideal scenario, a noise reduction algorithm should achieve mean and SD to the gold standard as close as possible. The NDCT FBP images were used as gold standard because they have the best image quality in this dataset. Both CNN-MSE produced much smaller SDs compared to NDCT, which indicates they over-smoothed the images and supports our visual observation. On the contrary, WGAN produced the closest SDs yet smaller mean values, which means it can reduce noise to the same level as NDCT but it compromised the information content. On the other hand, the proposed WGAN-VGG has outperformed CNN-VGG, WGAN-MSE and other selected methods in terms of mean CT numbers, SDs, and most importantly visual impression. In addition, we performed a blind reader study on 10 groups of images. Each group contains the same image slice but processed by different methods.

V. CONCLUSION

The most significant motivation for this paper is to method the gold standard NDCT images as much as possible. As described overhead, the feasibility and qualities of GAN has been investigated for this determination with the Wasserstein distance and the VGG loss. The transformation between using the MSE and VGG losses is rather substantial. Despite the fact that networks with MSE would proposal higher values for traditional figures of excellence, VGG loss based networks seem desirable for better visual image quality with more details and less artifacts. The experimental results have demonstrated that using WGAN helps develop image quality and statistical possessions. This could be a corporate problem for all the denoising networks. Instead of focusing on the design of a complex network organization, we have dedicated our effort to combine synergistic loss functions that guide the denoising development so that the resultant denoised results are as close to the gold customary as possible. Our experiment results with real clinical pictures have shown that the proposed WGAN-VGG network can effectively solve the well-known over-smoothing delinquent and generate images with concentrated noise and increased dissimilarity for improved lesion detection. In the future, we plan to incorporate the WGANVGG network with supplementary complicated generators such as the networks.

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.
- [2] A. B. De Gonzalez and S. Darby, "Risk of cancer from diagnostic xrays: estimates for the UK and 14 other countries," *The Lancet*, vol. 363, no. 9406, pp. 345–351, Jan. 2004.
- [3] J. Wang, H. Lu, T. Li, and Z. Liang, "Sinogram noise reduction for lowdose CT by statistics-based nonlinear filters," in *Med. Imag. 2005: Image Process.*, vol. 5747. International Society for Optics and Photonics, Apr. 2005, pp. 2058–2067.
- [4] A. Manduca, L. Yu, J. D. Trzasko, N. Khaylova, J. M. Kofler, C. M. McCollough, and J. G. Fletcher, "Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT," *Med. Phys.*, vol. 36, no. 11, pp. 4911–4919, Nov. 2009.
- [5] M. Beister, D. Kolditz, and W. A. Kalender, "Iterative reconstruction methods in x-ray ct," *Physica Medica: Eur. J. Med. Phys.*, vol. 28, no. 2, pp. 94–108, Apr. 2012.
- [6] J. Ma, J. Huang, Q. Feng, H. Zhang, H. Lu, Z. Liang, and W. Chen, "Low-dose computed tomography image restoration using previous normal-dose scan," *Med. Phys.*, vol. 38, no. 10, pp. 5713–5731, Oct. 2011.
- [7] Y. Chen, X. Yin, L. Shi, H. Shu, L. Luo, J.-L. Coatrieux, and C. Toumoulin, "Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing," *Phys. Med. Biol.*, vol. 58, no. 16, p. 5803, Aug. 2013.
- [8] P. F. Feruglio, C. Vinegoni, J. Gros, A. Sbarbati, and R. Weissleder, "Block matching 3D random noise filtering for absorption optical projection tomography," *Phys. Med. Biol.*, vol. 55, no. 18, p. 5401, Sep. 2010.
- [9] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Phys. Med. Biol.*, vol. 49, no. 11, p. 2463, May. 2004.
- [10] R. M. Lewitt, "Multidimensional digital image representations using generalized Kaiser–Bessel window functions," *J. Opt. Soc. Amer. A*, vol. 7, no. 10, pp. 1834–1846, Oct. 1990.
- [11] B. R. Whiting, P. Massoumzadeh, O. A. Earl, J. A. O'Sullivan, D. L. Snyder, and J. F. Williamson, "Properties of preprocessed sinogram data in x-ray computed tomography," *Med. Phys.*, vol. 33, no. 9, pp. 3290 – 3303, Sep. 2006.
- [12] I. A. Elbakri and J. A. Fessler, "Statistical image reconstruction for polyenergetic x-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 89–99, Feb. 2002.
- [13] S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical x-ray CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 677–688, Mar. 2012.
- [14] 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, p. 4777, Aug. 2008.
- [15] Y. Liu, J. Ma, Y. Fan, and Z. Liang, "Adaptive-weighted total variation minimization for sparse data toward low-dose x-ray computed tomography image reconstruction," *Phys. Med. Biol.*, vol. 57, no. 23, p. 7923, Nov. 2012.
- [16] Z. Tian, X. Jia, K. Yuan, T. Pan, and S. B. Jiang, "Low-dose CT reconstruction via edge-preserving total variation regularization," *Phys. Med. Biol.*, vol. 56, no. 18, p. 5949, Nov. 2011.
- [17] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose xray CT reconstruction via dictionary learning," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1682–1697, Sep. 2012.

- [18] Y. Zhang, X. Mou, G. Wang, and H. Yu, "Tensor-based dictionary learning for spectral CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 36, no. 1, pp. 142–154, Jan. 2017.
- [19] D. Kang, P. Slomka, R. Nakazato, J. Woo, D. S. Berman, C.-C. J. Kuo, and D. Dey, "Image denoising of low-radiation dose coronary CT angiography by an adaptive block-matching 3D algorithm," in *SPIE Med. Imag.* International Society for Optics and Photonics, Mar. 2013, pp. 86692G–86692G.
- [20] G. Wang, M. Kalra, and C. G. Orton, "Machine learning will transform radiology significantly within the next 5 years," *Med. Phys.*, vol. 44, no. 6, pp. 2041–2044, Mar. 2017.
- [21] G. Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, Nov. 2016.
- [22] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb 2016.
- [23] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524 – 2535, Dec. 2017.
- [24] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*. Springer, Nov. 2015, pp. 234 – 241.
- [25] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," 2016. [Online]. Available: arXiv:1609.04802
- [26] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances Neural Inform. Process. Syst.*, 2014, pp. 2672–2680.
- [27] J.-Y. Zhu, P. Krahenbuhl, E. Shechtman, and A. A. Efros, "Generative visual manipulation on the natural image manifold," in *Eur. Conf. Comput. Vision*. Springer, 2016, pp. 597–613.
- [28] M. Arjovsky and L. Bottou, "Towards principled methods for training generative adversarial networks," in *NIPS 2016 Workshop on Adversarial Training. In review for ICLR*, vol. 2016, 2017.
- [29] S. Srinivas, R. K. Sarvadevabhatla, K. R. Mopuri, N. Prabhu, S. S. Kruthiventi, and R. V. Babu, "A taxonomy of deep convolutional neural nets for computer vision," *Frontiers Robot. AI*, vol. 2, p. 36, 2016.
- [30] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, June 2009, pp. 248–255