

Low-Dose CT Image denoising and Restoration using General Adversarial Networks

Kevin Freire | kfreirea@ryerson.ca
 Walter Freire | andrei.freire@ryerson.ca
 Department of Electrical and Computer Engineering
 Ryerson University, Toronto, Canada.

Abstract—

Index Terms—Image restoration, Image Denoising, General Adversarial Networks, GANs, Low-Dose CT, Medical Imagery, Deep Neural Networks.

I. INTRODUCTION

Computerize Tomography (CT) has enabled direct imaging on the 3-dimensional structure of different organs and tissues inside the human body in a non-invasive manner. CT scans are constructed by combining the X-ray scans taken on several angles and orientations. It has several utilities but very useful in detecting lesions, tensions, tumours, and metastasis. It can reveal their presence and the spatial location, size and extent of the tumour. CT imaging has become a frequent tool for cancer diagnosis, angiography, and detecting internal injuries. However, despite the evidence of its utility for diagnosis and patient management, the potential risk of radiation-induced malignancy exists [1]. Studies found that CT alone contributes to almost half of the total radiation exposure from medical use alone. Recent studies reveal that as much as 1.2 - 2% of cancers may eventually be caused by the radiation dose conceived by the patient while undergoing CT scans [2]. To reduce the risk, the principles of ALARA (As Low As Reasonably Achievable) is now a profound practice predicted in CT imaging [3]. In regards to this practice, Low Dose CT (LDCT) is a promising solution in reducing radiation exposure [4]. In low dose CT, radiation exposure is decreased by lowering the tube current, or voltage. However, by reducing the tube voltage or current it introduces several artefacts and lowers the diagnostic quality of the LDCT image [5]. In order to boost the quality of an LDCT image, the reconstruction of LDCT has become a primary research. There are various methods that can be classified under three categories: (a) iterative reconstruction, (b) sinogram filtration based techniques, and (c) image post-processing based technique. In recent times, researchers were trying to develop new iterative algorithms (IR) for LDCT image reconstruction. Iterative reconstruction algorithms considerably suppresses the image noise, but still lose some details and suffer from remaining artefacts. Other disadvantages with IR techniques is the high computational cost, which is a bottle neck in practical utilization. Sinogram filtration on the other hand, directly works on the projection data before reconstructing the image and is more computationally economical than the IR technique. However, the data

of commercial scanners are not readily available to users, and this method suffers from edge blurring and the resolution loss. Many efforts were made in the image domain to reduce LDCT noise and suppress artefacts.

With the explosive evolution of deep neural networks, the LDCT denoising task is now dominated by deep neural networks. However, the research on deep learning-based LDCT denoising is confined to designing a network architecture based on the vanilla convolution operation. There has been a surge in interest in desinging General Adversarial Networks (GAN), mainly for image generation and becoming popular in denoising medical images as shown in [6], [7], [8]. In this paper we explore the possibility of applying GAN to the task of LDCT desnoising.

In many image-related reconstruction tasks, it is known that minimizing the per-pixel loss between the output image and the ground truth alone generate either blurring or makes the result visually unappealing [9]. The same effect was observed in the traditional neural network-based CT denoising works [10], [11]. The adversarial loss introduced by GAN can be treated as a driving force to push the generated image to look as close to as the ground truth image or in this case the Normal Dose CT image (NDCT) which also reduces the blurring effect. Further more, an additional perceptual loss was also introduced to measure the feature of the denoised image, with a focus on areas that the human eye cannot see.

Generative Adversarial Network was first introduced in 2014 by Goodfellow *et al.* [12]. It is a genrative model trying to generate real world images by employing a min-max optimization framework where two networks (Generator G and Discriminator D) are trained against each other. G tries to synthesize real appearing images from random noise wheras D is trying to distiguish between the generated and real images. If the Generator G get sufficiently well trained, the Discriminator D will eventually be unable to tell if the generated image is fake or not.

The original setup of GAN does not contain any constraints to control what modes of data it can generate. However, the auxiliary information were provided during the generation, GAN can be driven to output images with specific modes. In this scenario, GAN is usually referred to as conditional GAN (cGAN) since the output is conditioned on additional information. Bera *et al.* [7] proposed a discriminator function for CT denoising task, which leverages self-attention and pixel-wise GANs for restoring the diagnostic quality of LDCT

images. Yi *et al.* [13], proposed an adversarial trained network and a sharpness detection network to denoise LDCT images. Both [6] and [8] used Wasserstein GAN with Perceptual Loss and Wasserstein Distance to generate a denoised image of an LDCT image.

In this paper, we make the following contributions:

- 1) We built a GAN architecture similar to Alsaiani *et al.* [14], with the exception of training it specifically on LDCT images.
- 2) Computed similar loss functions as [14] with the exception of using the Mean Squared Error Loss function instead of computing L2-norm.
- 3) Present the output results of the network after training.

The remainder of this paper is organized as follows. The proposed method based on [14], with perceptual loss, pixel-pixel loss, smooth loss and adversarial loss are presented in II. Then, experiments and analysis of the results are presented in section III and IV. Lastly, Section VI gives a summary of this paper and looks forward to some possible future research directions.

II. METHOD

We built a GAN network structure for image denoising, which like, [14], is based on ResNet [15]. A generator network is trained to generate noise free images through competition with the discriminator network, using ground truth NDCT images to improve the quality of the generated images. Our network makes use of residual blocks, skip connections, and batch normalization. Due to training limitations, we used three residual blocks. However, having a larger number of residual blocks would increase the training accuracy significantly, but with a higher computational complexity and longer training times.

Suppose that $z \in \{z_k\}_1^N$ denotes the LDCT image, $x \in \{x_k\}_1^N$ denotes the NDCT image, and p_l and p_r represent the distribution of LDCT and NDCT images, respectively. The generator G learns a mapping $G : z \rightarrow \hat{x}$, where z is the LDCT the generator is conditioned upon. \hat{x} is the denoised CT that is expected to be as close to as the NDCT image x . The discriminator D , is to differentiate the generated image z from the real x .

A. Refined Loss Function

In the original GAN [12], D and G are trained by solving the following minmax problem

$$\min_G \max_D L_{GAN}(D, G) = -E_{x \sim p_r}[D(x)] + E_{z \sim p_l}[1 - D(G(z))] \quad (1)$$

where $E(\cdot)$ denotes the expectation operator, $G(\cdot)$ and $D(\cdot)$ represents the outputs of G and D . The generator G transforms a noisy sample to mimic a real sample, which defines a data distribution, denoted by p_g . Then D is trained to become an optimal discriminator for a fixed G , the minimization search for G is equivalent to minimizing the Jensen-Shannon (JS) divergence of p_r and p_g , which will lead to a vanished gradient

on the Generator G [6] and G will stop updating as the training continues.

In the process of using CT image training, the feature difference between the LDCT and NCT image will have a certain probability of structural loss in denoising results. In order to obtain a good quality image, we include the Minimum Square Error Loss (MSE) between the Generated image \hat{x} and the ground truth x . The function is expressed as follows:

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

Where $\|\cdot\|^2$ is the square L2-norm (MSE Loss). A lower difference denotes a higher quality of the generated image. However, the MSE loss can potentially generate blurry images and cause the distortion or loss of details. Thus, we also applied the perceptual loss.

With perceptual loss, the rational behind it is two-fold. First, when a person compares two images, the perception is not performed pixel-by-pixel. The human vision extracts and compares features from images [16]. With perceptual loss, it guides the model towards image style conversion [17], image denoising [6], and other tasks. The perceptual loss function is introduced to learn feature distribution of NDCT images from the feature space to guide the denoising task of LDCT images. Therefore, we included a pre-trained deep CNN (VGG16) for feature extraction and compared the denoised output against the ground truth in terms of the extracted features. The perceptual loss function is defined in a feature space:

$$L_{Perceptual}(G) = E_{(x,z)} \left[\frac{1}{whd} \|\Phi(G(z)) - \Phi(x)\|^2 \right], \quad (3)$$

where Φ is the feature map obtained from the output of the second convolutional layer of the VGG-16 pretrained network, $G(z)$ is the denoising result of the LDCT image, w, h, d represents width, height and depth of the feature map, respectively. It should be noted that we used the VGG-16 network as the feature extractor. The input of the network is a colour image, including three channels, while the CT images is greyscale image. Thus, before using the network for feature extraction, we duplicated the CT image to make RGB channels.

B. Generator Network

The goal of a single image denoising is to generate a realistic image with high quality. The

C. Discriminator Network

III. EXPERIMENT SETTINGS

IV. RESULTS AND DISCUSSION

V. COMPUTATIONAL COMPLEXITY

VI. CONCLUSION

[6] Yang *et al* - Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss

- [7] Bera *et al* - Noise Conscious Training of Non Local Neural Network Powered by Self Attentive Spectral Normalized Markovian Patch GAN for Low Dose CT Denoising
- [8] Yin *et al* - Unpaired Image Denoising via Wasserstein GAN in Low-Dose CT Image with Multi-Perceptual Loss and Fidelity Loss
- [18] Radford *et al.* - UNSUPERVISED REPRESENTATION LEARNING WITH DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS
- [10] Chen *et al.* - Low-dose ct via deep neural network
- [11] Chen *et al* - Low-dose CT with a residual encoder-decoder convolutional neural network
- [9] Huang *et al.* - Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis
- [12] Goodfellow *et al.* - Generative adversarial nets
- [13] Yi *et al.* - Sharpness-aware low-dose CT denoising using conditional generative adversarial network
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [16] M. Nixon and A. Aguado, *Feature extraction and image processing for computer vision*. Academic press, 2019.
- [17] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European conference on computer vision*, pp. 694–711, Springer, 2016.
- [18] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

REFERENCES

- [1] D. J. Brenner and E. J. Hall, "Computed tomography—an increasing source of radiation exposure," *New England journal of medicine*, vol. 357, no. 22, pp. 2277–2284, 2007.
- [2] D. A. Schauer and O. W. Linton, "National council on radiation protection and measurements report shows substantial medical exposure increase," 2009.
- [3] R. Protection, "Icrp publication 103," *Ann ICRP*, vol. 37, no. 2.4, p. 2, 2007.
- [4] S. Trattner, G. D. Pearson, C. Chin, D. D. Cody, R. Gupta, C. P. Hess, M. K. Kalra, J. M. Kofler Jr, M. S. Krishnam, and A. J. Einstein, "Standardization and optimization of ct protocols to achieve low dose," *Journal of the American College of Radiology*, vol. 11, no. 3, pp. 271–278, 2014.
- [5] F. E. Boas, D. Fleischmann, *et al.*, "Ct artifacts: causes and reduction techniques," *Imaging Med*, vol. 4, no. 2, pp. 229–240, 2012.
- [6] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1348–1357, 2018.
- [7] S. Bera and P. K. Biswas, "Noise conscious training of non local neural network powered by self attentive spectral normalized markovian patch gan for low dose ct denoising," *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3663–3673, 2021.
- [8] Z. Yin, K. Xia, Z. He, J. Zhang, S. Wang, and B. Zu, "Unpaired image denoising via wasserstein gan in low-dose ct image with multi-perceptual loss and fidelity loss," *Symmetry*, vol. 13, no. 1, p. 126, 2021.
- [9] R. Huang, S. Zhang, T. Li, and R. He, "Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis," in *Proceedings of the IEEE international conference on computer vision*, pp. 2439–2448, 2017.
- [10] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang, "Low-dose ct via convolutional neural network," *Biomedical optics express*, vol. 8, no. 2, pp. 679–694, 2017.
- [11] 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 transactions on medical imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.
- [12] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [13] X. Yi and P. Babyn, "Sharpness-aware low-dose ct denoising using conditional generative adversarial network," *Journal of digital imaging*, vol. 31, no. 5, pp. 655–669, 2018.
- [14] A. Alsaiaari, R. Rustagi, M. M. Thomas, A. G. Forbes, *et al.*, "Image denoising using a generative adversarial network," in *2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT)*, pp. 126–132, IEEE, 2019.