Study of Low-dose to High-dose CT using Supervised Learning with GAN and Virtual Imaging Trials

Fakrul Islam Tushar
Department of Electrical and Computer Engineering, Pratt School of Engineering
Duke University
Durham, NC
fakrulislam.tushar@duke.edu

Abstract

Computed tomography (CT) is one of the most widely used radiography exams worldwide for different diagnostic applications. However, CT scans involve ionizing radiational exposure, which raises health concerns. Counter-intuitively, lowering the adequate CT dose level introduces noise and reduces the image quality, which may impact clinical diagnosis. This study analyzed the feasibility of using a conditional generative adversarial network (cGAN) called pix2pix to learn the mapping from low dose to high dose CT images under different conditions. This study included 270 three-dimensional (3D) CT scan images (85,050 slices) from 90 unique patients imaged virtually using virtual imaging trials platform for model development and testing. Performance was reported as peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). Experimental results demonstrated that mapping a single low-dose CT to high-dose CT and weighted two low-dose CTs to high-dose CT have comparable performances using pix2pix CGAN and applicability of using VITs.

1 Instruction

Different non-invasive medical imaging modalities such as radiography, computer tomography (CT), mammography, and ultrasounds are highly integrated into the diagnostic workflow. CT scans images have been proven to be highly effective in different chest, abdomen, and pelvis diagnosis task due to distances contrast characteristics between other tissues, bone, and organs. However, to acquire high-quality CT scans images, there is a risk of high exposure to the radiation, which raises health concerns. Lowe-dose CT can significantly reduce the radiational health concerns with a tradeoff with the image quality, affecting diagnostic performances.

In recent years, several applications have been using deep-learning-driven models for different CT scans denoising tasks [1][2][3]. Yi et al. used a network trained adversarially in combination with a sharpness detection network to denoising the images [1]. Kang et al. proposed a denoising network using a residual wavelet network [2]. Ding et al. proposed a fidelity-embedded GAN to learn the conversion between low to standard-dose CT [3].

In this study we have utilized a cGAN called pix2pix to learn the mapping from low dose to high dose CT images under different conditions using simulated CT scans images generated using virtual imaging trial (VIT) platform [4][5][6][7][8].

2 Dataset

This study included 270 chest CT images (85,050 CT slices) generated using a VIT framework [4][8]. VIT is a process of simulating imaging evaluations with varying factors such as computational

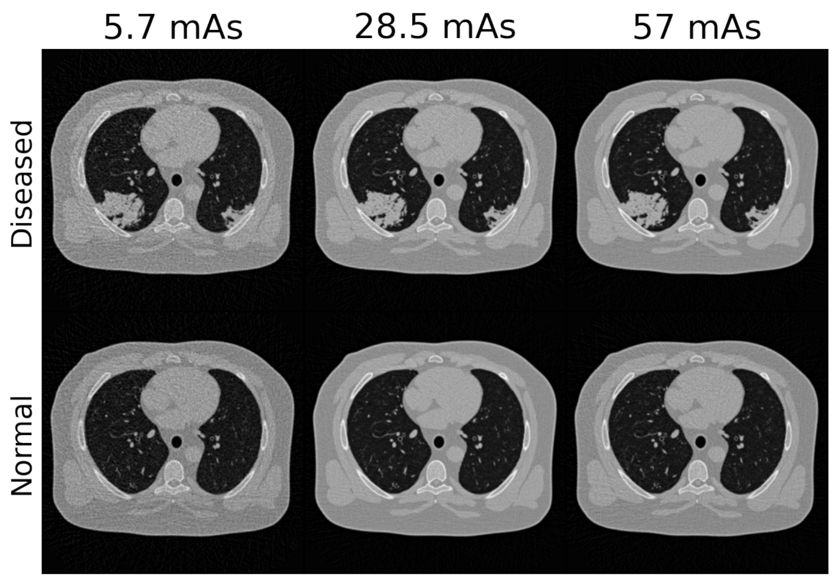

Figure 1: Simulated CT sample images. The top and bottom rows show diseased and normal slices from the same patients, respectively, different dose levels (left to right).

Table 1: Simulated diseased and normal CT images with different CT dose levels.

	Number of volumes (number of slices)	
Dose-level (mAs)	Diseased	Normal
5.7	50 (15,750)	40 (12,600)
28.5	50 (15,750)	40 (12,600)
57	50 (15,750)	40 (12,600)
Total	150 (47,250)	120 (37,800)

human phantoms, imaging scanner systems, and virtual readers [4][5]. Simulating the CT images involved four main steps. First, 50 CT images from 50 unique patients with lung abnormalities were acquired, and disease morphology was simulated. The second and third act incorporates the simulated appearance of the diseased region created in the previous step to 4D extended cardiac-torso (XCAT) phantoms [4]. Finally, CT scans of these virtual human patient models were imaged using a radiographic simulator called DukeSim [9]. We have imaged 50 diseased and 40 normal CT scans with three different dose levels.

Prior to model development, all CT volumes were resampled to voxel size of $2 \times 2 \times 1$ mm ($height \times width \times depth$) by B-spline interpolations, clipped to intensity range in Hounsfield units $(-1000, 500)$, normalized to 0 mean and 1 standard deviation and resized the slices to 256×256 ($height \times width$). Table 1 detailed the number of volumes and the different dose levels of the virtual images of CT images, and fig. 1 shown CT slices from simulated CT images.

3 Method

3.1 pix2pix cGAN

GANs have shown promising results in different image generation and transformation tasks [7]. This study has adopted the pix2pix cGAN proposed by Isola et al. to perform low-dose to high-dose CT image conversion using 2D CT slices [7]. The pix2pix cGAN consists of a generator and discriminator, and a modified U-Net has been used as a generator network. The generator weight was optimized using a weighted sigmoid cross-entropy and mean absolute error loss [7]. The discriminator network is a convolutional PatchGAN classifier trained to discriminate the generated and actual patches of the

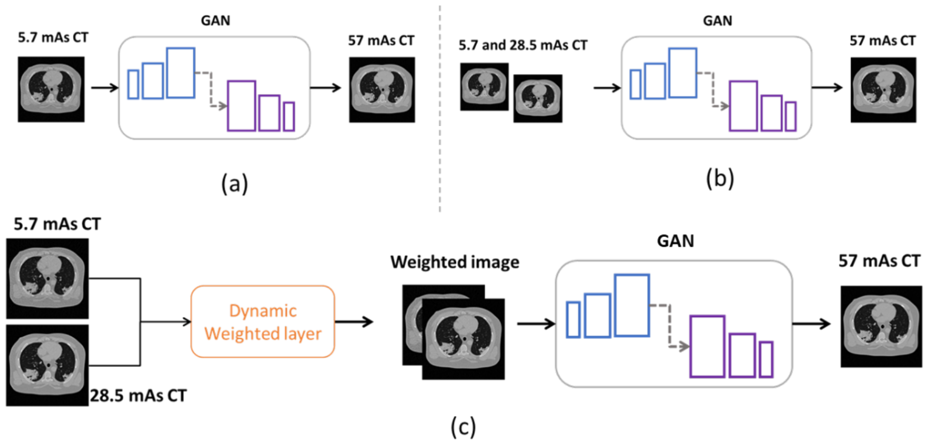

Figure 2: Overall workflow of different experiments utilizing single and multiple low-dose CT generate high-dose CT slices using (a).$cGAN$-SD (b)$cGAN$-DD and (c)$cGAN$-Dy.

images [7]. The Discriminator network was optimized with the combination of a real and generated sigmoid cross-entropy loss [7].

3.2 Experiments

The effectiveness of pix2pix cGAN was analyzed and evaluated with three different kinds of experimental setups detailed in sections 3.2.2,3.2.2 and 3.2.3.

3.2.1 cGAN-SD

First, the model was trained to map one low-dose (5.7 or 28.5 mAs) CT image to a high-dose/standard-dose (57 mAs) CT image denoted as cGAN Single dose ($cGAN$-SD). Two models were developed: model $cGAN$-$SD_{5.7}$ and $cGAN$-$SD_{28.5}$ were trained to map 5.7 or 28.5 mAs CT slices to high-dose/standard-dose (57 mAs) CT slices (shown in fig.2(a)).

3.2.2 cGAN-DD

For second expriment, a model was trained with multiple low dose-level (5.7 or 28.5 mAs) CT to map to high dose (57 mAs) CT slices and denoted as cGAN double dose ($cGAN$-DD) (shown in fig.2(b)).

3.2.3 cGAN-Dy

Finally, the third experiment utilized two different low-dose CT slices and weighted the low-dose CT slices using a dynamic weighting layer (1×1 convolution layer) before feeding the $cGAN$ to generate high-dose/standard-dose (5.7 mAs) CT slices demoted as cGAN dynamic ($cGAN$-Dy) (shown in fig.2(c)).

3.2.4 Training and Implementation

The dataset was divided by patient for each dose-levels, 50 patients ($15,750$ CT slices) for training, 20 patients ($6,300$ CT slices) for validation and 20 patients ($6,300$ CT slices) for testing. Generator and discriminator weights were optimized using Adam optimizer (learning rate of $2e-4$). All the models were implemented using TensorFlow Python library (version2.6). All the codes and model weights area available at https://github.com/fitushar/.

Table 2: Comparison among different pix2pix cGANs on the test-set. These two matrics are peak signal-to-noise ratio(PSNR) and structure similarity (SSIM).The results are shown as $mean \pm std$ on all the CT slices in the test-set.

Model	PSNR ($mean \pm std$)	SSIM ($mean \pm std$)
$cGAN\text{-}SD_{5.7}$	30.18 ± 5.40	0.82 ± 0.09
$cGAN\text{-}SD_{28.5}$	30.58 ± 5.18	0.82 ± 0.09
$cGAN\text{-}DD$	30.12 ± 5.33	0.81 ± 0.09
$cGAN\text{-}Dy$	30.48 ± 5.09	0.82 ± 0.09

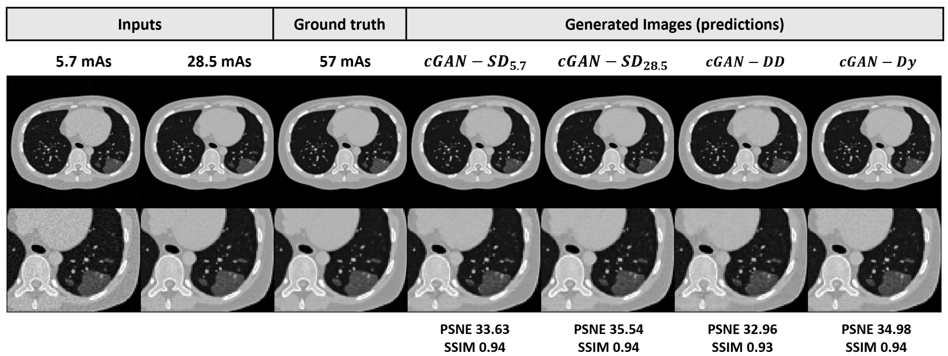

Figure 3: Shown the visual results of generated high-dose (57 mAs) CT Slices from low-dose CT slice.First two colums are inputs (low-dose CT), 3rd column is ground truth, column four to sever represents the generated high-dose CT slices using $cGAN\text{-}SD_{5.7}$,$cGAN\text{-}SD_{28.5}$,$cGAN\text{-}DD$ and $cGAN\text{-}Dy$,respectively. 2nd row shown zoom regions around the lesion.

4 Results

The performance was evaluated using two metrics, peak signal-to-noise ratio(PSNR) and structure similarity (SSIM).The evaluation was performed on 6300 CT slices of the test dataset and reported as the $mean \pm std$. Table 2 shown the performance of the four developed models on the testing dataset. All the model's performance is identical except ($cGAN\text{-}DD$) which performed with a SSIM of 0.82 ± 0.09, lower than all other models.

$cGAN\text{-}SD_{28.5}$ which maps 5.7 mAs CT slice to 57 mAs almost similar to the $cGAN\text{-}SD_{5.7}$ which maps 10 times lower-dose 5.7 mAs CT slice to 57 mAs. For $cGAN\text{-}Dy$ which used dynamic weighted layer two utilized two low-dose level CT slices haven't shown any performance improvement compared to other models using single low-dose CT ($cGAN\text{-}SD_{28.5}$ and $cGAN\text{-}SD_{5.7}$)

Fig.3 shown the visual results of generated high-dose CT slices using different models.

5 Discussion and Conclusion

This study analyzed the feasibility of utilizing virtually generated CT images to generate high-dose CT from low-dose CT applying pix2pix cGANs. Different experiments were performed using single and multiple low-dose to high-dose CT mapping using 4 different pix2pix cGANs. Performance of all the model are comparable.

This study had some noticeable limitations. Hyper-parameter tuning wasn't performed due to the time constrain and high computational time of the models; an intensive hyper-parameter tuning is needed to improve the performance. Only one type of GANs was applied for the image generation. Future work will investigate existing denoising deep learning models and applications of the state-of-the-art generative algorithms in denoising tasks.

Overall, pix2pix cGANs have shown promising results in denoising tasks and proven the feasibility of using VITs. If scalable, this approach could reduce radiation risk and improve the diagnosis.

Acknowledgments

I am grateful for helpful discussions with Roarke Horstmeyer, PhD and Joseph Y. Lo, PhD.

References

[1] Xin Yi and Paul Babyn. Sharpness-aware low-dose ct denoising using conditional generative adversarial network. *Journal of digital imaging*, 31(5):655–669, 2018.

[2] Eunhee Kang, Jong Chul Ye, et al. Wavelet domain residual network (wavresnet) for low-dose x-ray ct reconstruction. *arXiv preprint arXiv:1703.01383*, 2017.

[3] Hyoung Suk Park, Jineon Baek, Sun Kyoung You, Jae Kyu Choi, and Jin Keun Seo. Unpaired image denoising using a generative adversarial network in x-ray ct. *IEEE Access*, 7:110414–110425, 2019.

[4] Ehsan Abadi, W Paul Segars, Hamid Chalian, and Ehsan Samei. Virtual imaging trials for coronavirus disease (covid-19). *AJR. American journal of roentgenology*, 216(2):362, 2021.

[5] Ehsan Abadi, William P Segars, Benjamin MW Tsui, Paul E Kinahan, Nick Bottenus, Alejandro F Frangi, Andrew Maidment, Joseph Lo, and Ehsan Samei. Virtual clinical trials in medical imaging: a review. *Journal of Medical Imaging*, 7(4):042805, 2020.

[6] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.

[7] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[8] Fakrul Islam Tushar, Ehsan Abadi, Saman Sotoudeh-Paima, Rafael B. Fricks, Maciej A. Mazurowski, W. Paul Segars, Ehsan Samei, and Joseph Y. Lo. Virtual vs. reality: external validation of COVID-19 classifiers using XCAT phantoms for chest computed tomography. In Karen Drukker, Khan M. Iftekharuddin, Hongbing Lu, Maciej A. Mazurowski, Chisako Muramatsu, and Ravi K. Samala, editors, *Medical Imaging 2022: Computer-Aided Diagnosis*, volume 12033, pages 18 – 24. International Society for Optics and Photonics, SPIE, 2022.

[9] Ehsan Abadi, Brian Harrawood, Shobhit Sharma, Anuj Kapadia, William P Segars, and Ehsan Samei. Dukesim: a realistic, rapid, and scanner-specific simulation framework in computed tomography. *IEEE transactions on medical imaging*, 38(6):1457–1465, 2018.