Dual-task Learning For Low-Dose CT Simulation and Denoising

Mingqiang Meng[a,b], Yongbo Wang[a,b], Manman Zhu[a,b], Xi Tao[b,b], Zhaoying Bian[a,b], Dong Zeng[a,b], and Jianhua Ma[a,b]

[a]School of Biomedical Engineering, Southern Medical University, Guangdong, China
[b]Guangdong Artificial Intelligence and Digital Economy Laboratory (Guangzhou), Guangdong, China

ABSTRACT

Deep learning (DL) are being extensively investigated for low-dose computed tomography (CT). The success of DL lies in the availability of big data, learning the non-linear mapping of low-dose CT to target images based on convolutional neural networks. However, due to the commercial confidentiality of CT vendors, there are very few publicly raw projection data available to simulate paired training data, which greatly reduces the generalization and performance of the network. In the paper, we propose a dual-task learning network (DTNet) for low-dose CT simulation and denoising at arbitrary dose levels simultaneously. The DTNet can integrate low-lose CT simulation and denoising into a unified optimization framework by learning the joint distribution of low-dose CT and normal-dose CT data. Specifically, in the simulation task, we propose to train the simulation network by learning a mapping from normal-dose to low-dose at different levels, where the dose level can be continuously controlled by a noise factor. In the denoising task, we propose a multi-level low-dose CT learning strategy to train the denoising network, learning many-to-one mapping. The experimental results demonstrate the effectiveness of our proposed method in low-dose CT simulation and denoising at arbitrary dose levels.

Keywords: Computed tomography, dual-task learning, denoising network, simulation network

1. INTRODUCTION

X-raycomputed tomography (CT) is widely used in clinical diagnosis and treatment because of its ability to image the body's three-dimensional anatomy in a non-invasive manner. However, ionizing radiation generated during CT scanning will accumulate in the human body, and high radiation doses will induce the risk of cancer in human tissues and organs.[1] Given these risks, efforts have been made on reducing the radiation dose and the principles of As Low As Reasonably Achievable (ALARA) is profoundly practiced in clinical CT imaging.[2] However, reducing the radiation dose inevitably increases the noise and artifacts of reconstructed CT images, which compromises the diagnostic performance. Consequently, improving the image quality of low-dose CT (LDCT) has become a hot topic in medical imaging over the past decade.

Recently, with the rapid development of deep learning (DL) technology, the LDCT imaging algorithm is dominated by convolutional neural network and has achieved unprecedented success. DL-based algorithm learns the mapping from LDCT projection/image to normal-dose CT (NDCT) ones by designing an elaborate convolutional neural network (CNN), such as RED-CNN,[3] Wavelet networks,[4] and Tensor-Net.[5] A key factor in the success of these supervised algorithms is the availability of big data, that is, a large amount of paired LDCT and NDCT images.[6,7] Despite its superior denoising results, some issues still must be resolved before the DL models can be widely deployed in clinic. First, given the increase in total radiation dose, matched LDCT and NDCT data cannot be obtained in clinical practice. As a result, true NDCT and LDCT paired data are not available. Second, traditional LDCT simulation methods[8] usually insert noise into the raw projection data, however, very few raw data are publicly available to simulate paired training data, which degrades the generalization performance of

Further author information: (Send correspondence to J.M.)
J.M.: E-mail: jhma@smu.edu.cn

7th International Conference on Image Formation in X-Ray Computed Tomography, edited by Joseph Webster Stayman, Proc. of SPIE Vol. 12304, 123042O
© 2022 SPIE · 0277-786X · doi: 10.1117/12.2646640

the network. Third, most of DL-based models are designed for specific dose levels but perform poorly at lower doses.

To solve above problems, in this work, we propose a dual-task learning network (DTNet) for low-dose CT simulation and denoising at arbitrary dose levels. The presented DTNet can integrate LDCT simulation and denoising into a unified framework by learning the joint distribution of LDCT and NDCT data. Specifically, we first propose to train the simulation network by learning a mapping from NDCT to LDCT at different levels, where the dose level can be continuously controlled by a noise factor. In the denoising task, we present a multi-level LDCT learning strategy, which uses LDCT data of different levels to train the denoising network. Once trained, DTNet can be used for multi-level LDCT images simulation and denoising simultaneously, which greatly improves the applicability and generalization performance of the model.

2. METHODOLOGY

2.1 Low-dose CT Simulation Network

Let $x \in \mathbb{I}^{H\times W}$ denotes a NDCT image and $y \in \mathbb{I}^{H\times W}$ denotes the corresponding LDCT image. Low-dose CT simulation is an inverse process of LDCT denoising, which learns the opposite mapping from x to y. To precisely control the noise level of the generated LDCT image, we propose a noise control factor guided LDCT simulation scheme, as illustrated in Fig. 1 (a). Specifically, given a NDCT image and a mask image with value of 1, the simulator network encodes them into high-dimensional features for coupling and progressively decoding to reconstruct LDCT image of different levels, where the dose level can be continuously controlled by a noise factor. In addition, instead of directly applying the pixel-wise loss to the target image, we use a gaussian filter to extract the first-order statistics information of simulation and reference LDCT noise, and then constraint it with the MAE loss function, which can be formulated as:

$$L_S = arg\min_{\theta_S} \sum_{i=1}^{N} ||GF(S(x_i, m \cdot k_j) - x_i) - GF({y_i}^j - x_i)||_1. \tag{1}$$

Here, m is the mask image with value of 1, $k_j = 1, ..., j$ is the noise factor that controls the simulated dose level j. θ_S represents the parameters of simulation network S. GF is a 2d Gaussian filter convolution kernel with a size of 5×5.

2.2 Low-dose CT Denoising Network

Typically, the DL-based denoising problem is to build a prediction network $R(\cdot)$ that learns the non-linear mapping from y to x, i.e., $R : y \to x$. In clinical, CT images of various dose levels may be obtained to meet the clinical diagnosis demands. In Fig. 1 (b), we propose a multi-level low-dose CT learning strategy to train the denoising network with different levels low-dose CT data by minimizing the mean absolute error (MAE) loss function, which is expressed as:

$$L_R = arg\min_{\theta_R} \sum_{i=1}^{N} ||R({y_i}^j) - x_i||_1. \tag{2}$$

Here, ${y_i}^j, i = 1, ..., N, j = 1, ..., M$ represents the LDCT images, where N is the total number of LDCT images and M is number of dose levels. θ_R represent the parameters of denoising network R. Once trained, the denoising model can be applied to LDCT reconstruction at different dose levels in the clinic.

2.3 Dual-task Learning Network

To jointly optimize low-dose CT simulation and denoising tasks, we adopt the dual-task learning network (DT-Net), which uses a joint discriminator to alternately optimize the simulation and denoising network by learning the joint distribution $p(x, y^j)$ of LDCT and NDCT data, as shown in Fig. 1 (c). Let $p_S(x, y^j)$ and $p_R(x, y^j)$

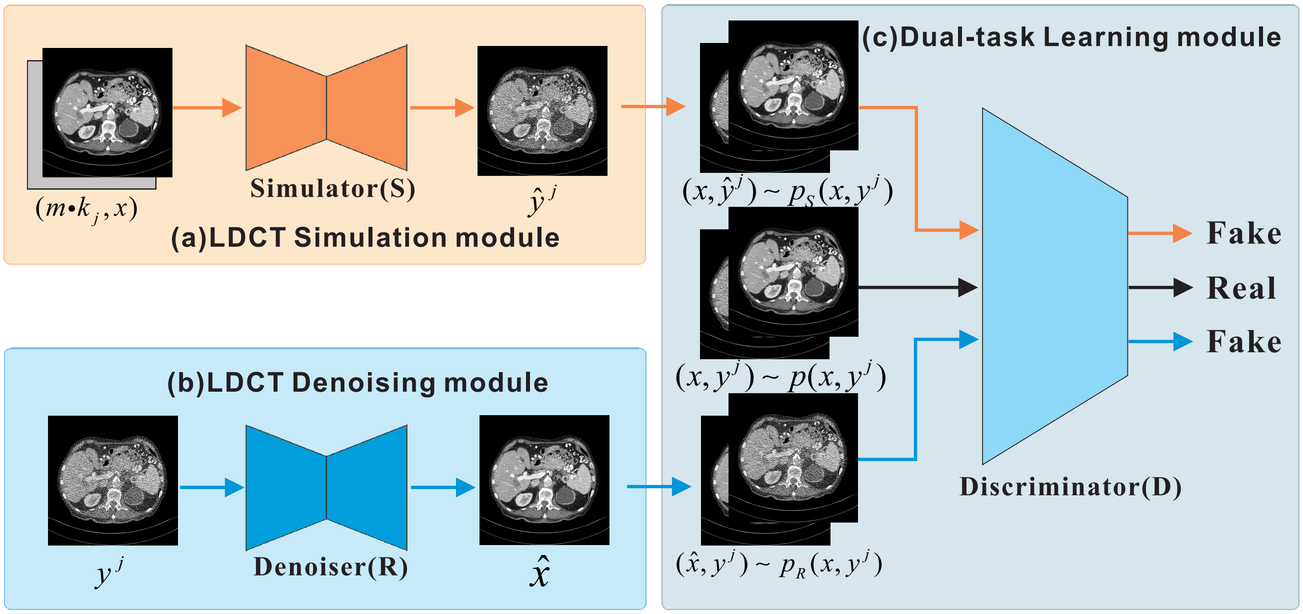

Figure 1. The overall structure of the proposed DTNet framework. (a) Low-Dose CT simulation module, (b) Low-Dose CT denoising module, (c) Dual-task learning module. m is the mask image with value of 1. k_j is the noise scale factor.

represent the pseudo joint distribution of simulation and denoising task, respectively. The dual adversarial loss can be defined as follows:

$$\begin{aligned} \min_{S,R} \max_{D}(S, R,, D) = & E_{(x,y^j)\backsim p(x,y^j)}[D(x, y^j)] \\ & - \lambda_S \cdot E_{(x,\hat{y}^j)\backsim p_S(x,y^j)}[D(x, \hat{y}^j)] \\ & - \lambda_R \cdot E_{(\hat{x},y^j)\backsim p_R(x,y^j)}[D(\hat{x}, y^j)]. \end{aligned} \tag{3}$$

Here, $E[\cdot]$ denotes the expectation operator, D represents the discriminator, which is used to receive the image pair (x, y^j), $(\hat{x}, y^j)$, $(x, \hat{y}^j)$ and distinguish them as real or fake samples. The hyper-parameters λ_R and λ_S controls the weight of GAN loss.

The final objective function for optimizing DTNet can be formulated as:

$$\min_{S,R} \max_{D}(S, R, D) + \alpha \cdot L_S + \beta \cdot L_R, \tag{4}$$

where α and β are the weight parameters that control the trade-off between adversarial loss and fidelity loss of simulation and denoising tasks.

3. EXPERIMENTS

3.1 Dataset

The 2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge dataset[9] published by Mayo Clinic is used to evaluate the effectiveness of the proposed DTNet model. This dataset contains 10 anonymized patient normal-dose raw projection data, which were acquired using the Somatom Definition AS+ CT system under 100kV or 120kV and automatic exposure controlling mode, and simulated quarter-dose projection data. In order to obtain LDCT data at different dose levels, we re-simulated 1/4, 1/8, 1/16, 1/20-dose projection data using the corresponding simulation algorithm[8] to insert quantum and electronic noise into the normal-dose projection data. In this study, we selected seven patients with a total of 17056 image data pairs for training. Specifically, 4568 image pairs collected from two patients are used to validate the performance of DTNet, and 2100 image pairs from the remaining one patient are selected as the testing set.

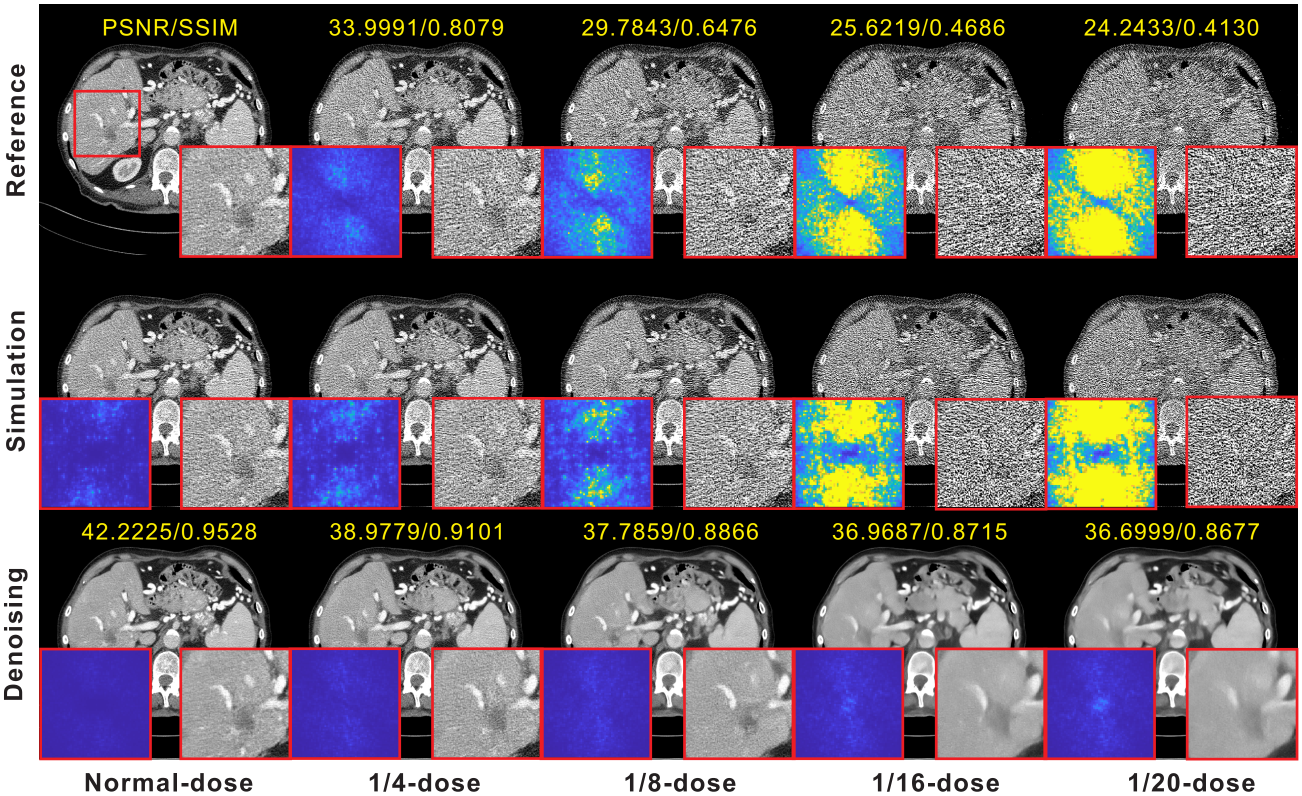

Figure 2. The simulation and denoising results at different dose levels of proposed DTNet. The display window of images and zoomed-in ROIs is [−140, 260] HU. The display window of NPS maps is [0 3000] HU^2mm^2.

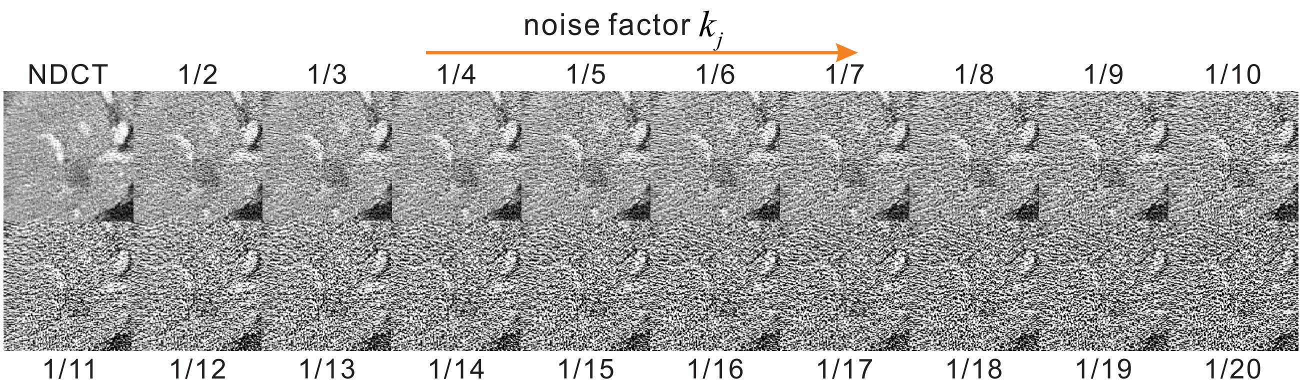

Figure 3. Arbitrary dose levels simulation results of proposed DTNet. From the left to the right column, the noise level of CT image gradually increases as the noise factor k_j increases. The display window is [−140, 260] HU

3.2 Implementation details

In our experiments, the proposed DTNet model consists of three sub-networks: simulator S, denoiser R and discriminator D. For S and R, we use the same generator network structure UNet,[10] which contains an encoder and decoder. The discriminator D has a similar structure to the PatchGAN.[11] The DTNet model is optimized in an alternating manner using Adam algorithm. The learning rates of S, R and D are set to $1e^{-4}$, $1e^{-4}$, and $2e^{-4}$, respectively. The hyper-parameters of loss function are selected to be $\alpha = 100$, $\beta = 10$, $\lambda_R = \lambda_S = 0.5$. During training, we randomly extracted 4 patches of size 128×128 as input in each iteration, D is updated three times while S and R are updated once. All networks are implemented using Pytorch and trained with an GeForce RTX 3090 GPU.

Table 1. PSNR and SSIM Quantitative comparison of DTNet denoising results at different dose levels.

Dose	RED-CNN	WGAN-VGG	DTNet
1/20	36.5572 ± 2.3307	33.3125 ± 2.5499	$\mathbf{36.6999 \pm 2.3983}$
	0.8665 ± 0.0573	0.7543 ± 0.0994	$\mathbf{0.8677 \pm 0.0583}$
1/16	36.8386 ± 2.3549	34.0276 ± 2.5127	$\mathbf{36.9687 \pm 2.4371}$
	0.8705 ± 0.0561	0.7795 ± 0.0908	$\mathbf{0.8715 \pm 0.0574}$
1/8	37.6845 ± 2.5182	35.8733 ± 2.5104	$\mathbf{37.7859 \pm 2.5292}$
	0.8839 ± 0.0533	0.8371 ± 0.0712	$\mathbf{0.8866 \pm 0.0515}$
1/4	38.6890 ± 2.6873	37.4532 ± 2.5651	$\mathbf{38.9779 \pm 2.7158}$
	0.9024 ± 0.0472	0.8789 ± 0.0557	$\mathbf{0.9101 \pm 0.0426}$

4. RESULTS

Fig. 2 shows the simulation and denoising results of abdominal CT image at five different dose levels: normal-dose, 1/4, 1/8, 1/16, 1/20-dose. In the simulation task, the corresponding simulation noise factors are set to 0, 1/4, 1/8, 1/16, and 1/20, respectively. Note that the noise factor equal to 0 means that it does not insert any noise into the NDCT image. It can be observed that the proposed method can simulate LDCT images of different dose levels and the learned noise intensity and characterization are similar to the reference images. In addition, we also calculate the noise power spectrum (NPS) maps of magnified ROIs to evaluate the statistical property of noise. We can observe that the NPS of generated LDCT images is very close to the reference images. When the noise factor is set to 0, only little noise is embedded in the output image, which can be seen from the ROI and the corresponding NPS. To verify the robustness of DTNet in simulating other dose levels which are not including in training data, as shown in Fig. 3. It can be seen that the noise intensity of simulated LDCT images continuously increases with the increase of noise factor k_j. This demonstrates that the proposed DTNet has the ability to simulate the realistic LDCT images and can control the noise level well.

In the denoising task, the last row in Fig. 2 shows that DTNet can efficiently suppress noise and artifacts at difference dose levels. In particular, for normal-dose images with a small amount of noise, we can also remove the noise without smoothing the image content. And for ultra-low doses, such as 1/20-dose, some small structures are completely drowned out by noise and are difficult to recover well. Therefore, ultra-low dose scanning can be used for special imaging tasks where anatomical details are not important, such as localization imaging. To quantitatively analyze the denoising performance of DTNet, we calculate the peak-to-noise ratio (PSNR) and structural similarity (SSIM), as summarized in Table 1. We can see that DTNet obtains the best quantitative values at different dose levels compared to RED-CNN[3] and WGAN-VGG,[12] which is consistent with the visual evaluation.

5. CONCLUSION

In this paper, we have presented a dual-task learning network (DTNet) for LDCT simulation and denoising tasks. In the simulation task, the simulation network encodes the NDCT image and mask image into high-dimensional features for coupling and decoding to generate LDCT images at different dose levels, where the dose levels can be controlled by a noise factor. In the denoising task, the multi-level LDCT learning strategy is used to train the denoising network, which can learn many-to-one end-to-end mapping. The presented DTNet integrates the LDCT simulation and denoising tasks into a unified optimization model. Both the quantitative and qualitative evaluation results have demonstrated the promising performance of DTNet in terms of LDCT simulation and denoising. In the feature, we will further improve the performance of DTNet by incorporating advanced network and prior information.

ACKNOWLEDGMENTS

This work was supported in part by the NSFC under Grant U21A6005 and Grant U1708261, the National Key R&D Program of China under Grant No. 2020YFA0712200, and Young Talent Support Project of Guangzhou Association for Science and Technology.

REFERENCES

[1] Smith-Bindman, R., Lipson, J., Marcus, R., Kim, K.-P., Mahesh, M., Gould, R., Berrington de Gonzlez, A., and Miglioretti, D. L., "Radiation Dose Associated With Common Computed Tomography Examinations an0d the Associated Lifetime Attributable Risk of Cancer," *Archives of Internal Medicine* **169**(22), 2078-2086 (2009).

[2] Valentin, J., "International Commission on Radiological Protection. The 2007 recommendations of the international commission on radiological protection," *Annals of the ICRP, ICRP Publication* 103, 2-4 (2007).

[3] Chen, H., Zhang, Y., Kalra, M. K., Lin, F., Chen, Y., Liao, P., Zhou, J., and Wang, G., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE transactions on medical imaging* **36**(12), (2524-2535) (2017).

[4] Kang, E., Min, J., and Ye, J. C., "A deep convolutional neural network using directional wavelets for lowdose Xray CT reconstruction," *Medical physics* **44**(10), e360-e375 (2017).

[5] Tao, X., Wang, Y., Lin, L., Hong, Z., and Ma, J., "Learning to Reconstruct CT Images from the VVBP-Tensor," *IEEE Transactions on Medical Imaging*, (2021).

[6] Wang, G., Ye, J. C., and De Man, B., "Deep learning for tomographic image reconstruction," *Nature Machine Intelligence* **2**(12), 737-748 (2020).

[7] Wang, G., Jacob, M., Mou, X., Shi, Y., and Eldar, Y. C., "Deep Tomographic Image Reconstruction: Yesterday, Today, and TomorrowEditorial for the 2nd Special Issue Machine Learning for Image Reconstruction," *IEEE transactions on medical imaging* **40**(11), 2956-2964 (2021).

[8] Yu, L., Shiung, M., Jondal, D., and McCollough, C. H., "Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols," *Journal of computer assisted tomography* **36**(4), 477-487 (2012).

[9] https://www.aapm.org/GrandChallenge/LowDoseCT/.

[10] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical image computing and computer-assisted intervention*, 234-241, (2015).

[11] Bera, S., and Biswas, P. K., "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*, (2021).

[12] Yang, Q., Yan, P., Zhang, Y., Yu, H., Shi, Y., Mou, X., Kalra, M. K., Zhang, Y., Sun, L., and Wang, G., "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE transactions on medical imaging* **37**(6), 1348-1357 (2018).