

PIMA-CT: Physical Model-Aware Cyclic Simulation and Denoising for Ultra-low-dose CT Restoration

Peng Liu¹, Linsong Xu², Garrett Fullerton¹, Yao Xiao^{1,3}, James-Bond Nguyen¹, Zhongyu Li², Izabella Barreto⁴, Catherine Olguin⁴, and Ruogu Fang^{1,5,6,*}

¹*J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL, USA*

²*School of Software Engineering, Xi'an Jiaotong University, Xi'an, China*

³*Department of Radiation Physics, University of Texas MD Anderson Cancer Center, Houston, TX, USA*

⁴*Department of Radiology, University of Florida, Gainesville, FL, USA*

⁵*Center for Cognitive Aging and Memory, McKnight Brain Institute, University of Florida, Gainesville, FL, USA*

⁶*Department of Electrical and Computer Engineering, Herbert Wertheim College of Engineering, University of Florida, Gainesville, FL, USA*

Correspondence*:

Ruogu Fang
ruogu.fang@bme.ufl.edu

2 ABSTRACT

3 A body of studies has proposed to obtain high-quality images from low-dose and noisy Computed
4 Tomography (CT) scans for radiation reduction. However, these studies are designed for
5 population-level data without considering the variation in CT devices and individuals, limiting the
6 current approaches' performance, especially for ultra-low-dose CT imaging. Here, we incorporate
7 a physical anthropomorphic phantom model PIMA-CT with an unsupervised learning framework,
8 using a novel deep learning technique Cyclic Simulation and Denoising (CSD), to address the
9 limitations. We first acquired low-dose and standard-dose paired phantom CT scans and then
10 developed two generative neural networks: noise simulator and denoiser. The simulator extracts
11 real low-dose noise and tissue features from two separate image spaces (e.g., low-dose phantom
12 model scans and standard-dose patient scans) into a unified feature space. Meanwhile, the
13 denoiser provides feedback to the simulator on the quality of the generated noise. In this way, the
14 simulator and denoiser cyclically interact to optimize network learning and ease the denoiser to
15 simultaneously remove noise and restore tissue features. We thoroughly evaluate our method
16 for removing both real low-dose and Gaussian simulated low-dose noise. The results show that
17 CSD outperforms one of the state-of-the-art denoising algorithms without using any labeled data
18 (actual patients' low-dose CT scans) nor simulated low-dose CT scans. This study may shed light
19 on incorporating physical models in medical imaging, especially for ultra-low level dose CT scan
20 restoration.

21 **Keywords:** Physical model, Medical image denoising, Low-dose CT, Noise removal, Dose reduction

1 INTRODUCTION

The quality of medical imaging is critical for diagnosis and treatments. However, medical imaging often suffers from noise produced at either the image reconstruction or post-imaging stages. Medical physicists in radiology play several essential roles in maintaining imaging quality and stability for imaging machines, such as Computed tomography (CT). They usually adopt an anthropomorphic physical model to facilitate the assessment of imaging quality and the adjustment of the imaging machines' parameters before performing on real patients. Motivated by this, we hypothesized that a physical model could also help restore high-quality images for the cases in the post-imaging stage, such as radiation reduction in CT imaging.

Reducing radiation dose during imaging is a low-cost approach to release concerns about causing cancer or other negative health conditions using CT scanning (1), but this method introduces noise into CT scans, hindering the diagnostic effectiveness of such scans. Several studies (2, 3) have been proposed to address this problem by removing the noise from low-dose CT scanned images. However, these studies are designed based on Gaussian noise simulation for populations without considering the variation in CT devices and individuals, limiting the current approaches' performance, especially for ultra-low-dose CT imaging (see Figure 1A).

Moreover, most of the success of deep learning-based approaches for low-dose CT image restoration (4, 5) much relies on a large number of labeled images. However, obtaining the real low-dose CT scans is not available in practice. Accessing real image noise is critical for the development of any practical imaging algorithm. Also, real noise properties significantly vary among different CT machines and individuals. Thus, the Gaussian noise assumption is not always guaranteed in practical scenarios and significantly limits the existing approaches for ultra-low-dose CT imaging.

We address these problems by incorporating an anthropomorphic physical phantom model into generative adversarial networks. The proposed framework is named Cyclic Simulation and Denoising (CSD). The physical model provides paired low-dose and standard-dose phantom CT scans before scanning the actual patients. These phantom scans can offer statistical noise prior, which is related to the specific CT machine for patient diagnosis, for CSD to precisely capture noise properties and remove real complex noise from CT scans. Our CSD is composed of noise simulation and denoising two networks. The simulation network facilitates the denoising network to learn real noise properties. The denoising network thus can access realistic noise through physical phantom CT scans. However, phantom scans lack tissue features (see Figure 1B). The missing tissue information prevents feasible phantom-based solutions for CT image restoration. As one can see in Figure 1C, the model trained with paired low-dose and standard-dose phantom scans fails to remove real noise from low-dose patient scans. To overcome this problem, we train CSD using normal-dose and phantom CT scans simultaneously to embrace realistic noise and tissue features into a unified learning framework without the access labeled or Gaussian noise simulated data.

We evaluate our CSD for removing both real low-dose and Gaussian simulated noise. The results show that CSD outperforms one of the state-of-the-art denoising algorithms for ultra low-quality medical image restoration. Our main contributions include that (1) we incorporate an anthropomorphic physical phantom model into generative adversarial learning to address the challenges of removing real noise from ultra-low-dose CT scans for radiation reduction; (2) we develop a unsupervised framework in the combination of phantom CT scans that can outperform one of the start-of-the-art methods without using any labeled or other noise simulation data; (3) to the best of our knowledge, this is the first study to incorporate physical model into deep learning for medical imaging.

2 MATERIALS AND METHODS

The problem of CT image denoising can be understood by $L = H + N$, where H is the clean, standard-dose CT image, L is the noisy, low-dose CT image, and N is additive image noise. Though an additive relationship does not completely represent the relationship between clean and noisy images, this formula provides a baseline for understanding the problem.

We utilize two deep networks in the framework. The first network G_s is the noise simulator and can be modeled by $L = G_s(H, \alpha)$, where α is the desired simulated dose level and implicitly indicated in training data. The second network G_d is the denoiser that can be modeled by $H = G_d(L)$, where G_d is the network generating a clean image from a given low-dose noisy input L .

2.1 Unsupervised Learning by Incorporating Physical Model

We use a head phantom model to obtain paired low-dose and standard-dose phantom CT scans, with which we combine the normal dose (standard-dose) patient CT scans to develop our CSD model. The phantom scans allow the model to access real noise properties and the patient scans offer the actual brain tissue features to the model. In this way, we eliminate the need for noisy low-dose CT scans from actual patients and even the Gaussian noise simulated low-dose CT scans to develop our mode (Figure 2A). Therefore, we present an *unsupervised learning* framework by incorporating an anthropomorphic physical phantom model.

2.2 Cyclic Simulation and Denoising (CSD)

2.2.1 Overview

We develop two deep networks to perform simulator and denoiser individually. To ease the network training, we first use paired low-dose and standard-dose phantom CT scans to pre-train the simulator and denoiser, separately. Then, we plug the simulator and denoiser pre-trained models into our CSD framework (Figure 2A). In particular, we start with noise simulation using both the phantom and patient CT scans to generate low-dose noisy patient CT images that simultaneously provide noise and tissue features for training the denoiser (Figure 2B). Meanwhile, CSD also allows the backward training process from denoiser to simulator. The denoiser takes phantom noisy scans and simulated noisy patient scans as input to learn how to remove realistic noise and restore tissue features simultaneously, while the simulator mainly plays as a regularizer to the denoiser for stabilizing the training (Figure 2B). The interaction between simulator and denoiser forms a dynamic data-driven framework, named Cyclic Simulation and Denoising (CSD), to address the challenges of low-dose CT image restoration.

2.2.2 Pretrain Simulator and Denoiser ($H \rightarrow \hat{L}, L \rightarrow \hat{H}$)

We train the simulator with a u-shape encoder-decoder generative adversarial network by formulating the objective as an adversarial learning. We use a discriminator D_s to differentiate real low-dose CT images from fake samples generated by the simulator G_s . We illustrate the formulation of the simulation as below.

$$\mathcal{L}_{GAN}(G_s, D_s) = \mathbb{E}_{L \sim p(L)}[\log(\mathbf{D}_s(L))] + \mathbb{E}_{H_{phantom} \sim p(H)}[\log(1 - \mathbf{D}_s(\mathbf{G}_s(H_{phantom})))] \quad (1)$$

To encourage the output of the denoiser to match the clean phantom scans, we use an ℓ_1 loss between the output and the ground truth image.

$$\mathcal{L}_1(G_d) = \mathbb{E}_{L, H \sim p(L, H)} \|H - \mathbf{G}_d(L)\|_1 \quad (2)$$

99 Initializing the weights by pretraining can significantly ease the convergence of two interactive generators
 100 in both spatial and temporal space. However, the phantom scans still lack the essential features as scanning
 101 on a real patient.

102 2.2.3 Learn Simulation Interacting with Denoiser: S2D ($H \rightarrow \hat{L} \rightarrow \hat{H}$)

103 We start with noise simulation to provide both noise and tissue features for training denoiser. We apply a
 104 discriminator D_s to train the simulator G_s . We formulate the simulation objective as below.

$$\mathcal{L}_{GAN}^{S2D}(G_s, D_s) = \mathbb{E}_{L \sim p(L)} [\log(\mathbf{D}_s(L))] + \mathbb{E}_{H \sim p(H)} [\log(1 - \mathbf{D}_s(\mathbf{G}_s(H)))] \quad (3)$$

105 The simulator feeds its output into the denoiser during training. Thus, we formulate the denoising loss
 106 using a modified Equation 2 as below.

$$\mathcal{L}_1^{S2D}(G_d) = \mathbb{E}_{L, H \sim p(L, H)} \|H - \mathbf{G}_d(\mathbf{G}_s(H))\|_1 \quad (4)$$

107 Besides the discriminator D_s , we take advantage of the denoising performance as regularization feedback
 108 to indicate the quality of the simulation. As the simulation becomes better, the denoising is getting harder.

109 Furthermore, the simulator D_s in S2D takes the standard-dose scans from both phantom and patients as
 110 inputs. The phantom data applies a latent constrain to the D_s and stabilizes the training. Interacting with
 111 denoising encourages the simulator to generate realistic low-dose noise. Further, the denoise can benefit of
 112 taking the output of the simulator as additional training data, dynamically.

113 2.2.4 Learn Denoising in Simulator: D2S ($L \rightarrow \hat{H} \rightarrow \hat{L}$)

114 The development of the training process from denoising to simulation has two significant varies from the
 115 cycle consistency study (6) (see Figure 2(c)). We first enable supervised learning to train the denoiser G_d
 116 using the standard-dose and the corresponding low-dose CT images. Compared to the adversarial learning,
 117 supervised learning provides a stronger supervision signal to build an accurate denoiser. More importantly,
 118 the simulator in S2D produces the noise gradually close to the desired level during training. Thus, we can
 119 acquire various noise level images from the simulator, with which, the denoiser-self implicitly learns to
 120 restore clean CT scans for a range level of low-dose CT scans, rather than a specific noise level indicated in
 121 the training data. Therefore, the input to the denoiser G_d in D2S includes phantom low-dose and simulated
 122 patient low-dose images. We use a ℓ_1 loss to train the denoiser G_d . The ℓ_1 loss encourages a pixel-wise
 123 match to the ground-truth. We illustrate the ℓ_1 loss as below.

$$\mathcal{L}_1^{D2S}(G_d) = \mathbb{E}_{L, H \sim p(L, H)} \|H - \mathbf{G}_d(L)\|_1 \quad (5)$$

124 Besides, we use adversarial learning to train the simulator in D2S to match the desired noise distribution in
 125 the actual low-dose CT scans. The objective to this adversarial learning the distribution is written as below.

$$\mathcal{L}_{GAN}^{D2S}(G_s, D_s) = \mathbb{E}_{L \sim p(L)} [\log(\mathbf{D}_s(L))] + \mathbb{E}_{\hat{H} \sim p(H)} [\log(1 - \mathbf{D}_s(\mathbf{G}_s(\hat{H})))] \quad (6)$$

We develop the cyclic simulation and denoising training with regularizations in both directions and take advantage of both cycles $H \rightarrow \hat{L} \rightarrow \hat{H}$ and $L \rightarrow \hat{H} \rightarrow \hat{L}$. The total objective is illustrated as below.

$$G_s^*, G_d^* = \arg \min_{G_s, G_d} \max_{D_s} \lambda_1 \mathcal{L}_{GAN}^{S2D}(G_s, D_s) + \lambda_2 \mathcal{L}_1^{S2D}(G_d) + \lambda_3 \mathcal{L}_{GAN}^{D2S}(G_s, D_s) + \lambda_4 \mathcal{L}_1^{D2S}(G_d) \quad (7)$$

where λ indicates the weights of each loss. With these novel developments, the simulator and denoiser interact each other in a cyclic self-learning manner to enable realistic noise simulation and accurate denoising for low-dose CT image.

3 RESULTS

3.1 Datasets

We use *three* CT datasets during training and testing. The first dataset is obtained from the CT scanning on a single tissue-equivalent physical phantom model. This set contains various levels of low-dose series, scanned between 5 mAs and 95 mAs with 5 mAs intervals (see examples in the supplementary Figure S1). In this work, we simply use 20 mAs, 30 mAs, and 60 mAs low-dose phantoms for training noise simulation and evaluating the reality of various types of noise in Figure 1C. We also include the standard-dose (175 mAs) scans as the ground-truth. Each dose level of phantom series produces 138 CT scans. The second dataset is a public Retrospective Image Registration Evaluation (RIRE) dataset. This dataset includes 388 standard-dose CT scans. We use 80% for training the simulator and denoiser in the proposed CSD and also task 20% for demonstrating the advantages of CSD over end-to-end training a denoiser in Table 2, where we simulate the low-dose noise by adding Gaussian noise on normal dose CT scans. We compute the corresponding standard variation of Gaussian noise for a specific mAs by following (7). Additionally, we acquire a real patient dataset including paired standard-dose (190 mAs) and low-dose (20 mAs) in a total of 432 CT scans (see examples in the supplementary Figure S2). We use them for comparing various types of simulated noise in Figure 1C and evaluating real noise removal performance of our approach in Table 1, where 250 scans are used for training and 182 scans are used for testing. Moreover, we randomly select 373 scans from this dataset combining with 20% of the RIRE dataset, in total 449 scans included, to evaluate our CSD's generalizability in Table 2.

3.2 Evaluation Metric

We develop CSD with U-net (8) for the simulator network G_s and DnCNN (9) for the denoiser network G_d . We evaluate image denosing performance using Peak signal-to-noise ratio (PSNR) and image structural similarity index measure (SSIM).

3.3 Unsupervised Learning Performance on Real Low-dose Noise Removal

Here, we aim to demonstrate that the proposed CSD framework in a combination with phantom can remove the real low-dose noise effectively. We first take the start-of-the-art medical image denoising network (9) as a baseline and train it with Gaussian simulated low-dose CT scans at different noise levels. Then, we build the G_d in CSD using the baseline's architecture and train it with paired low-dose and standard-dose phantom CT scans at the same noise levels as Gaussian simulation. We test each model on 182 real low-dose CT scans at the noise level 20 mAs. The comparison results are shown in Figure 1C at 20 mAs and Table 1 at 30, 60 mAs noise levels. As one can see, the combination of the proposed CSD training framework and phantom simulation significantly outperforms the baseline with an average 1.56 dB improvement on PSNR

across three different noise levels. Furthermore, as one can see in Figure 1C, the baseline network, which is trained with paired low-dose and standard-dose phantom scans, performs much worse than the model trained with both our CSD phantom and Gaussian simulation, which may be due to the lack of critical tissue features in the phantom scans. Notably, these results may indicate that *CSD, in combination with phantom simulation, can encourage the denoiser to learn both real low-dose noise features from phantom and tissue image features from patient scans*, simultaneously, and leading to real low-dose noise removal with greater accuracy and precision.

3.4 Evaluate CSD's generalizability (Ablation without G_s)

Here, we further evaluate the proposed CSD's generalizability to train a denoiser targeting the general simulated low-dose noise, such as Gaussian simulation. We still use the same baseline network to conduct this study. We use the standard end-to-end manner and our CSD framework to train two networks with the same architecture as the baseline, separately. Notably, to have a fair comparison, we only use original noisy CT scans in training dataset as the input of the G_d in D2S cyclic training. Then, we compare the two networks to remove 30 and 60 mAs levels of Gaussian simulated low-dose noise from 449 CT scans. As one can see in Table 2, the model trained with our CSD can consistently outperform the one trained with end-to-end manner, with impressive average performance gain 0.355 dB for PSNR. In addition, we also show a visual result comparison in Figure 3. As one can see, the denoiser G_d trained with our CSD framework can produce more realistic CT scans from its low-dose noisy version. More visual results of low-dose simulation and denoising can be found in the supplementary Figure S3 and Figure S4. These results suggest that starting with simulation may *create a live environment* from which the denoiser can learn high-validity representations to achieve a better denoising performance. Theoretically, the simulator and denoiser in the CSD may play as a *regularizer to each other* to optimize the networks effectively.

4 CONCLUSION

This paper proposed incorporating an anthropomorphic physical phantom model with generative deep learning networks for medical imaging, with a focus on realistic low-dose CT image restoration. The combination of an anthropomorphic physical model with deep generative adversarial networks can eliminate the needs of both actual low-dose patients and even other low-dose simulation CT scans to build an unsupervised learning framework for low-dose CT image restoration. More importantly, an anthropomorphic physical model CT scanning can abstract the unique noise properties of a particular CT imaging machine for the deep learning model to take CT machine domain variation into account during training. Eventually, with the interaction between a noise simulation network and a denoising network in cyclic training processing, the proposed deep learning model embraces realistic noise from low-dose phantom CT scans and tissue features from normal-dose patient CT scan into a single unified framework for building a state-of-the-art method for real low-dose CT image restoration.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

197 PL, YX, and RF contributed to conception and design of the study. PL, GF, YX, and JN organized the
198 database. PL, LX, and ZL contributed to the software used in this study. PL performed the statistical
199 analysis. IB and CO provided digital scanning data of the anthropomorphic physical phantom. PL and
200 GF wrote the first draft of the manuscript. PL, GF, YX, and RF wrote sections of the manuscript. RF was
201 responsible for supervision, project administration, and funding acquisition. All authors contributed to
202 manuscript revision, read, and approved the submitted version.

FUNDING

203 This material is based upon work supported by the National Science Foundation under Grant No. (NSF
204 1908299).

SUPPLEMENTARY MATERIAL

205 Supplementary Material is available for this paper.

DATA AVAILABILITY STATEMENT

206 Publicly available datasets were analyzed in this study. This data can be found here: https://rire.insight-journal.org/download_training_data.html. Anonymized data are available from the corresponding authors
207 upon reasonable request. The code and software used in this study will be made publicly available upon
208 acceptance.
209

REFERENCES

- 210 1. Goldman A, Maldjian P. Reducing radiation dose in body CT: A practical approach to optimizing CT
211 protocols. *American Journal of Roentgenology* **200** (2013) 748–754.
- 212 2. Badretale S, Shaker F, Babyn P, Alirezaie J. Fully convolutional architecture for low-dose CT image
213 noise reduction. *IOP Conference Series: Materials Science and Engineering* (Hawaii, USA: International
214 Conference on Artificial Intelligence Applications and Technologies) (2017), vol. 261, 012012.
- 215 3. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, et al. Low-dose CT image denoising using a generative
216 adversarial network with wasserstein distance and perceptual loss. *IEEE transactions on medical*
217 *imaging* **37** (2018) 1348–1357.
- 218 4. Yang Q, Yan P, Kalra MK, Wang G. CT image denoising with perceptive deep neural networks. *ArXiv*
219 *abs/1702.07019* (2017).
- 220 5. Chen H, Zhang Y, Zhang W, Liao P, Li K, Zhou J, et al. Low-dose CT denoising with convolutional
221 neural network. *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* (IEEE)
222 (2017), 143–146.
- 223 6. Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial
224 networks. *Proceedings of the IEEE international conference on computer vision* (2017), 2223–2232.
- 225 7. Britten A, Crotty M, Kiremidjian H, Grundy A, Adam E. The addition of computer simulated noise to
226 investigate radiation dose and image quality in images with spatial correlation of statistical noise: an
227 example application to X-ray CT of the brain. *The British journal of radiology* **77** (2004) 323–328.

- 228 8 .Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation.
229 Navab N, Hornegger J, Wells WM, Frangi AF, editors, *International Conference on Medical Image*
230 *Computing and Computer-Assisted Intervention (MICCAI)* (Cham: Springer) (2015), 234–241.

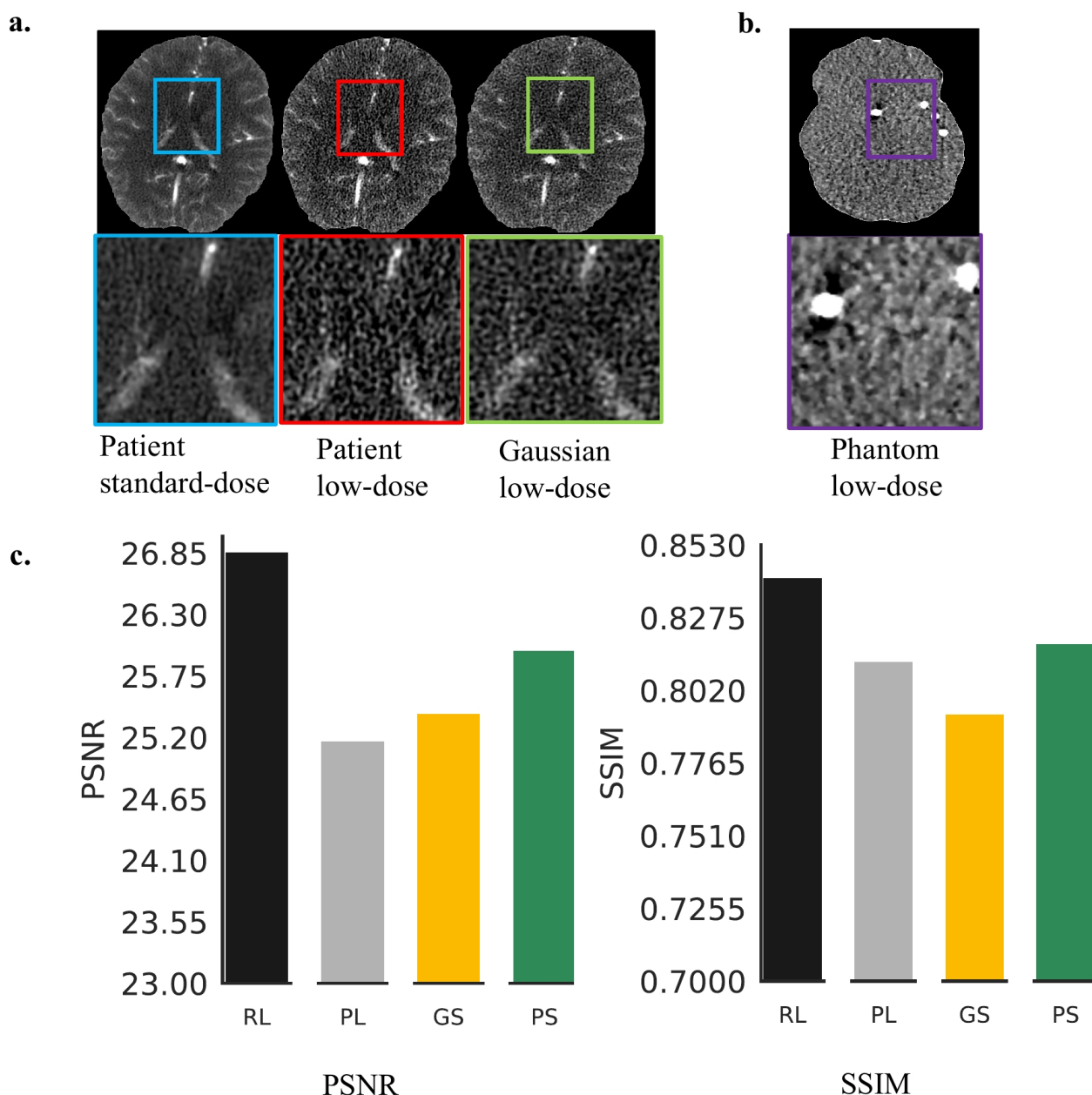
- 231 9 .Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a gaussian denoiser: Residual learning of deep
232 cnn for image denoising. *IEEE Transactions on Image Processing* **26** (2017) 3142–3155.

Table 1. The average real low-dose noise removal performance of a same deep neural network *trained* with Gaussian noise simulation and CSD + physical phantom noise simulation, separately. The best results are highlighted in bold.

PSNR (dB) / SSIM		
Noise level (mAs)	Trained with Gaussian	Trained with CSD + phantom
30	24.47/0.7555	26.10/0.8235
60	23.03/0.6960	25.51/0.7894

Table 2. The average Gaussian noise removal performance of the same deep neural network trained through the proposed CSD framework and the standard end-to-end manner, separately. The best results are highlighted in bold.

PSNR (dB) / SSIM		
Noise level (mAs)	End-to-end training	CSD training
30	31.93/0.9105	32.05/0.9124
60	33.33/0.9365	33.92/0.9429



A denoising model trained with:

RL: real low- and standard-dose paired patient CT scans

PL: real low- and standard-dose paired phantom CT scans

GS: Gaussian simulated low- and standard-dose paired CT scans

PS: our CSD incorporating with physical phantom model

Figure 1. Real low-dose has a different noise distribution from Gaussian noise and is hard to remove. (A) It shows a visual comparison of the standard-dose CT, real low-dose CT, and Gaussian simulated low-dose CT scans. (B) It shows a low-dose CT scanned by using a physical phantom model. (C) We trained four same structural deep neural networks(DNNs) using various types of low-dose noise with the same noise level (20 mAs radiation dose) and then compared the effectiveness of noise removal.

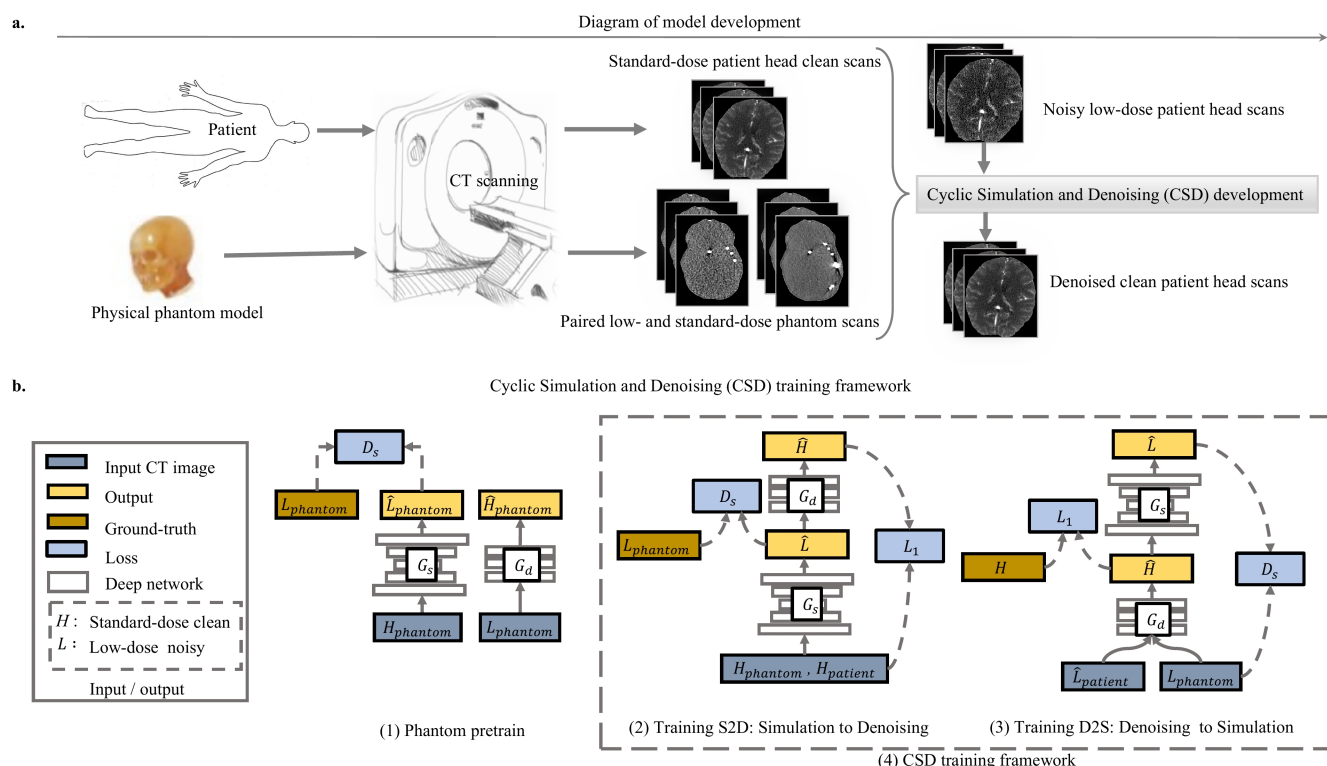


Figure 2. The overview of the model development in (A) and the proposed CSD training framework in (B). (A) demonstrates how we incorporate a physical phantom model into the proposed deep learning model CSD. (B) shows how our CSD is developed in detailed. Two training stages: first, we initialize the weights of simulator and denoiser by pretrain on physical phantom CT scan (1); second, the cycle-training from noise simulation to denoising (2) and another cycle-training from denoising to simulation (3) are developed, simultaneously. The G_s and G_d represent simulation and denoising, separately. During training, the two cycles interact with each other and are executed, alternatively.

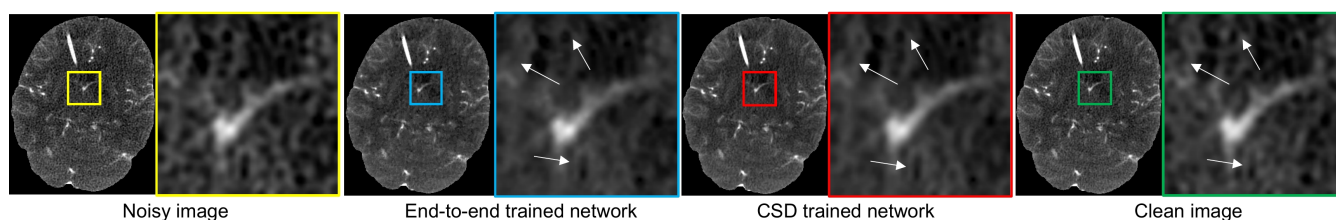


Figure 3. The visual comparison of the denoising performance between the network trained with end-to-end and the one trained with our CSD framework.