

Hybrid-Collaborative Noise2Noise Denoiser for Low-Dose CT Images

Ahmed M. Hasan^{ID}, Mohammad Reza Mohebbian, Khan A. Wahid^{ID}, *Senior Member, IEEE*, and Paul Babyn

Abstract—Low-dose computed tomography (LDCT) has yet to get its full potential benefit due to excessive quantum noise. Although learning to restore an image using two noisy images in the Noise2Noise (N2N) model has shown good promise for different noise models, it does not perform well in LDCT. In this article, we have introduced a collaborative technique to train multiple N2N generators simultaneously and learn the image representation from LDCT images. We have presented three models using this collaborative N2N (CN) principle employing CN two generators (CN2G), CN three generators (CN3G), and hybrid CN3G (HCN3G). The CN3G model has shown better performance than the CN2G model in terms of denoised image quality at the expense of an additional LDCT image. The HCN3G model has taken the advantages of both these models by managing to train three collaborative generators using only two LDCT images by leveraging one previous work called blind source separation (BSS) with block matching 3-D (BM3D). To make the collaboration among different generators more efficient, we have introduced collaborative loss terms among the generators. All three methods have shown improved performance in terms of peak signal-to-noise ratio and structural similarity index metrics compared to similar benchmark methods.

Index Terms—Bland–Altman analysis, collaborative loss, hybrid model, Noise2Noise (N2N) peak signal-to-noise ratio (PSNR), structural similarity index metrics (SSIM).

I. INTRODUCTION

CT IMAGING has been used in medical imaging for diagnosis for more than 40 years. Initially, the focus was more about getting quality images from CT scanners, but now, there has been a growing awareness of the associated radiation

Manuscript received August 23, 2019; revised January 7, 2020, February 21, 2020, and April 16, 2020; accepted June 2, 2020. Date of publication June 12, 2020; date of current version March 3, 2021. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), in part by the Canada Foundation for Innovation (CFI), and in part by the Western Economic Diversification Canada. (*Corresponding author: Ahmed M. Hasan*.)

Ahmed M. Hasan is with the Biomedical Engineering Division, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: hasan.ahmedmonjurul@usask.ca).

Mohammad Reza Mohebbian is with the Electrical Engineering Department, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: moremohebian@gmail.com).

Khan A. Wahid is with the Electrical Engineering Department, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada, and also with the Biomedical Engineering Division, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: khan.wahid@usask.ca).

Paul Babyn is with the Department of Medical Imaging, Saskatchewan Health Authority, Saskatoon, SK S7K 0M7, Canada (e-mail: paul.babyn@saskhealthauthority.ca).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TRPMS.2020.3002178

dose and the need to reduce the overall radiation dosage [1]. As there is an inherent tradeoff between the noise level and radiation dose of CT images, different CT manufacturers have developed algorithms to get high-quality images from low-dose CT scanning. This is made possible due to the “as low as reasonably achievable” (ALARA) principle [2]. However, it still remains an open challenge of how to get diagnostically useful images, especially with low-contrast areas using low-dose scanner settings.

Denoising is a process of reconstructing the original image by suppressing noise. The denoising methods for CT images can be broadly classified into three groups depending on where the denoising takes place, such as whether on the raw projection data, during iterative reconstruction (IR) or during image postprocessing. In the first method, the projection data or sinogram is modified before the filtered back projection (FBP) is undertaken using sinogram smoothing [3]. In IR, images are reconstructed using projection data for iterative forward and backward projection operations [4]–[7]. One recent study shows a promise in a hybrid approach between IR and deep learning method [8]. Major CT vendors have integrated IR algorithms in their image processing pipeline. Although these IR methods show improved performance, the main drawback is their expensive computation and longer processing time. In the image processing techniques, denoising happens in the image domain rather than the raw projection domain. However, noise levels in a CT image are not uniform and can be approximately modeled as the additive colored Gaussian noise [9], [10]. As a result, it is difficult to denoise the CT image effectively with traditional denoising algorithms. Some image processing algorithms provided a better performance in such denoising task, such as the K-SVD [11], MAP algorithm [12], nonlocal means filter [13], diffusion filter [14], and reconstruction of difference [15].

Recently, deep learning has gained popularity in several image restoration tasks, including contour detection [16], image recognition [17], image captioning [18], image segmentation [19], and radiomics [20]. We have seen a lot of studies on low-dose CT image denoising that are based on training with low-dose and normal-dose image pairs [21]–[23] and adjacent frames [24]. The main challenge in training low-dose and normal-dose image pairs is to obtain clean training targets. Because it is easier to get two noisy images than one clean target, a Noise2Noise (N2N) model was proposed in [25]. This method suggested that images can be restored by only looking at corrupted example pairs for different noise distributions, such as the additive Gaussian noise, Poisson noise, or multiplicative Bernoulli noise. Inspired by

the N2N model, several other models were proposed to learn image representation using a single noisy image, such as Noise2Self [26] and Noise2Void [27]. However, the authors of these papers used synthetic images or simulated datasets only. This motivates the investigation of the collaborative N2N (CN) model using phantom low-dose computed tomography (LDCT) images and rises the question of whether multiple generators can be trained to learn collaboratively image representation from these LDCT images.

Competition and collaboration are widely used in the multiagent environment [28]. Competition is analogous to the adversary principle between the generator and discriminator in various generative adversarial network (GAN) models and it has found a few applications in CT image denoising tasks [29]–[32]. GAN models use the min–max optimization framework, where the generator tries to reconstruct fake images from random noise; the discriminator works as a classifier to distinguish between the real and fake images. As the generator gets sufficiently trained, it starts to produce realistic images as fake images, and then the discriminator finds it difficult to distinguish between real and fake images. A potential collaboration between different generators has not been properly explored for different image processing tasks. We explore how collaboration between different N2N generators can be established for low-dose CT denoising tasks. We have initially tested two different models: 1) CN with two generators (CN2G) and 2) CN with three generators (CN3G) that work on two low-dose image sets and three low-dose image sets, respectively. The CN3G model has shown better performance than the CN2G model in terms of denoised reconstructed image quality at the expense of an additional low-dose CT image. So, we have proposed another model called hybrid CN three generators (HCN3G) using one of our previous works with blind source separation (BSS) with a block matching 3-D (BM3D) filter [33]. The HCN3G model can take the advantages of both previous models to train a three-generator model using only two low-dose CT image pairs. In order to make the collaboration among different generators more efficient, we have introduced collaborative loss terms among the generators in all our models. Refer to Section II for more details about the architectures of our three models. Our goal is to reduce the dose to the most vulnerable population, i.e., children, so we have used an appropriate phantom, CATPHAN 600 in this experiment. Refer to Section III for more details on that and other CT data acquisition process. These collaborative models help to reduce the radiation dose required, as well as finding low-contrast and high-resolution areas from the low-dose images.

II. COLLABORATIVE NOISE2NOISE GENERATOR

In a collaborative environment, multiple agents work together to achieve a single goal. In this article, we explore this principle to train multiple generators simultaneously in different configurations.

A. Collaborative Noise2Noise With Two Generators

In this configuration, we have two generators, G_A and G_B . The input and target pairs in these two generators are two

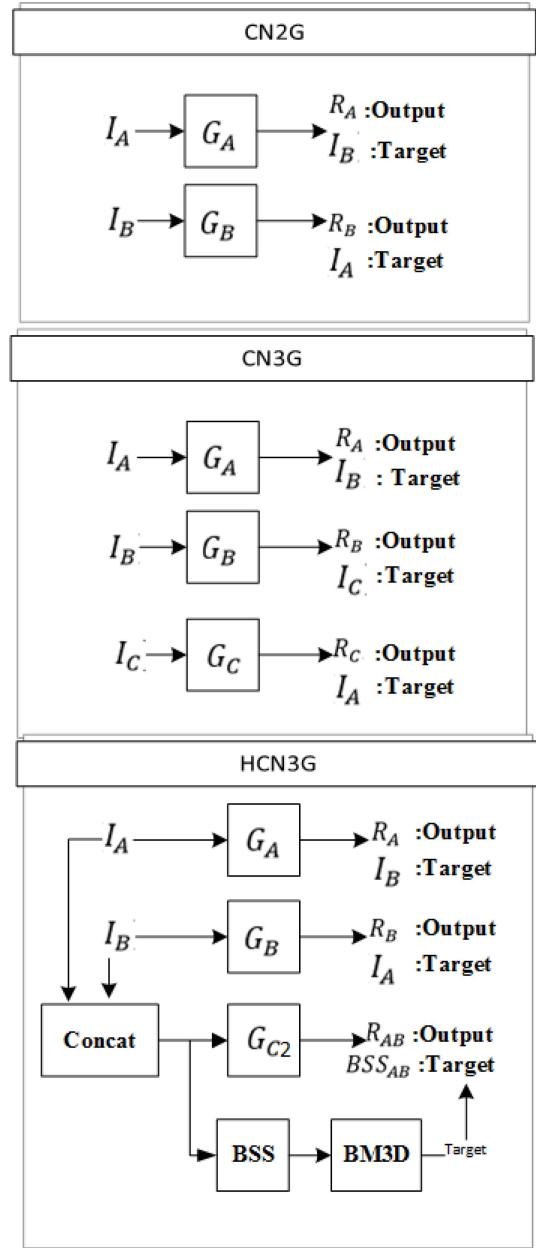


Fig. 1. Structures of different CN generators. I_X and R_X are the low-dose image and reconstructed image, respectively.

low-dose image pairs, (I_A, I_B) and (I_B, I_A) , and the reconstructed images are R_A and R_B , respectively. Fig. 1 shows the CN2G configuration. Each generator has an associated loss term, which is an L_2 loss that calculates the sum of the square difference between the target and the reconstructed images. In this regard, L_2 losses in generators G_A and G_B are given as

$$G_{A\text{Loss}} = L_2(R_A, I_B) \quad (1)$$

$$G_{B\text{Loss}} = L_2(R_B, I_A). \quad (2)$$

We have introduced a collaborative loss (C_{Loss}) between two generators using another L_2 loss function, which calculates the sum of the squared difference between two reconstructed images. The collaborative loss (C_{Loss}) is given as

$$C_{\text{Loss}} = L_2(R_A, R_B). \quad (3)$$

This collaborative loss is added to the losses of the individual generators using a weighting factor. The weighting factor controls the desired denoising effect. Therefore, the total loss (T_{Loss}) in this configuration is the summation of all three losses as given in the following:

$$T_{\text{Loss}_{\text{CN}2G}} = G_{A_{\text{Loss}}} + G_{B_{\text{Loss}}} + W \times C_{\text{Loss}} \quad (4)$$

where W is a weighting factor that controls the denoising effect. We have used a unity weighting factor to ensure an equal agreement among different loss terms.

B. Collaborative Noise2Noise With Three Generators

In this configuration, we have three generators G_A , G_B , and G_C . The input and target pairs in these three generators are three low-dose image pairs (I_A, I_B) , (I_B, I_C) , and (I_C, I_A) and the reconstructed images are R_A , R_B , and R_C , respectively. The structure of this configuration is also shown in Fig. 1. Here, L_2 losses in Generators G_A , G_B , and G_C are given in the following:

$$G_{A_{\text{Loss}}} = L_2(R_A, I_B) \quad (5)$$

$$G_{B_{\text{Loss}}} = L_2(R_B, I_C) \quad (6)$$

$$G_{C_{\text{Loss}}} = L_2(R_C, I_A). \quad (7)$$

The difference between any two generated outputs is termed *collaborative loss*. There are three collaborative loss components in this configuration as given in the following:

$$C_{\text{Loss}_{AB}} = L_2(R_A, R_B) \quad (8)$$

$$C_{\text{Loss}_{BC}} = L_2(R_B, R_C) \quad (9)$$

$$C_{\text{Loss}_{CA}} = L_2(R_C, R_A). \quad (10)$$

The total loss in this configuration is the summation of all six losses as given in the following:

$$T_{\text{Loss}_{\text{CN}3G}} = G_{A_{\text{Loss}}} + G_{B_{\text{Loss}}} + G_{C_{\text{Loss}}} + W \times (C_{\text{Loss}_{AB}} + C_{\text{Loss}_{BC}} + C_{\text{Loss}_{CA}}) \quad (11)$$

where W is a weighting factor that controls the denoising effect. Like before, a unity weighting factor is used to ensure an equal agreement among the loss terms.

C. Hybrid-Collaborative Noise2Noise With Three Generators

This HCN3G configuration utilizes two low-dose frames in three generators, which is inspired by our previous work [33]. In this strategy, we have separated the signal and noise components from two low-dose frames I_A and I_B by BSS; then the signal component is denoised by another denoiser, the block matching 3-D or BM3D, to get BSS_{AB} . The noise scale of BM3D is adaptively determined from the noise component of the BSS block. In the HCN3G model, three generators G_A , G_B , and G_{C2} are used with only two low-dose frames I_A and I_B . G_A and G_B have the same input and target image pairs as our first model (CN2G). G_{C2} uses the concatenation of low-dose frames I_A and I_B as input and BSS_{AB} as the target. The difference between G_C in CN3G and G_{C2} in HCN3G is that G_C accepts one low-dose image as a single-channel image, while G_{C2} accepts the concatenation of two low-dose images

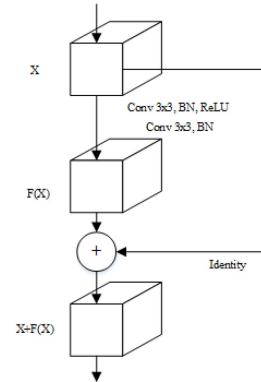


Fig. 2. Residual block which tries to learn the residue $F(x)$ due to the identity connection.

as a two-channel image. Fig. 1 illustrates this structure. $G_{A_{\text{Loss}}}$ and $G_{B_{\text{Loss}}}$ are same as (5) and (6); L_2 loss for G_{C2} is given as

$$G_{C_{\text{Loss}}} = L_2(R_{AB}, BSS_{AB}) \quad (12)$$

where R_{AB} is the reconstructed image of G_{C2} and BSS_{AB} is the denoised image from our previous BSS + BM3D method. There are three collaborative loss components in this configuration, which are defined as

$$C_{\text{Loss}_{AB}} = L_2(R_A, R_B) \quad (13)$$

$$C_{\text{Loss}_{BC}} = L_2(R_B, R_{AB}) \quad (14)$$

$$C_{\text{Loss}_{CA}} = L_2(R_{AB}, R_A). \quad (15)$$

Finally, the total loss is expressed in the following:

$$T_{\text{Loss}_{\text{HCN}3G}} = G_{A_{\text{Loss}}} + G_{B_{\text{Loss}}} + G_{C_{\text{Loss}}} + W \times (C_{\text{Loss}_{AB}} + C_{\text{Loss}_{BC}} + C_{\text{Loss}_{CA}}). \quad (16)$$

D. Generator Architecture

Several image-to-image generators were proposed in the literature, including the encoder-decoder structure [34], the U-Net structure [35], [36], the residual net (resnet) structure [17], and Derain [37]. In our experiment, we have used the resnet structure, because it is better in handling the vanishing gradient issue of the training a deep network [38], [39]. The basic component of the resnet structure is a residual block which uses an identity shortcut connection to skip two convolution layers as shown in Fig. 2. We can simply stack multiple residual blocks together as they are identity mappings which guide the gradient flow from the deep layer to the shallow layer [40], [41]. All generators G_A , G_B , and G_C use this same architecture, except G_{C2} which accepts a two-channel image as input instead of a one-channel input image.

We have built our CN generators using a stacked residual block of size $k = 9$ between the downsampling and upsampling layers as shown in Fig. 3.

E. Evaluation Metrics

Along with visual assessment, we have used the following metrics for performance evaluation.

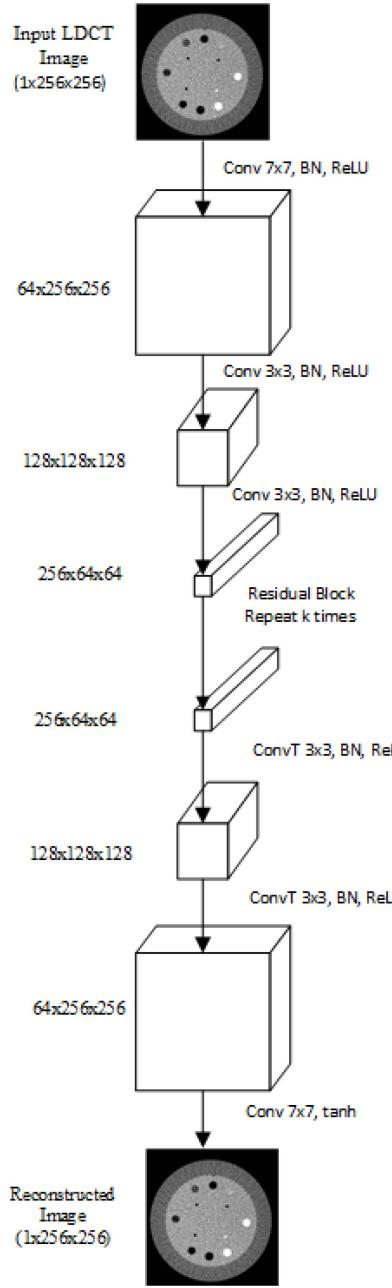


Fig. 3. CN generator architecture. Conv, ConvT, and BN represent convolution, transposed convolution, and batch normalization layers, respectively.

Peak Signal-to-Noise Ratio (PSNR): This metric will indicate the reconstruction accuracy or closeness to reference image; so, a higher value indicates a superior image. It can be calculated as follows:

$$\text{PSNR(dB)} = 10 \log_{10} \frac{\text{Peak}^2}{\text{MSE}}. \quad (17)$$

Structural Similarity Index Metrics (SSIM): This metric is consistent with our visual perception. The values of SSIM can vary from 0 to 1. A higher SSIM denotes a better image quality. If x and y are two local image windows selected from the same position of the two input images, then SSIM can be calculated as a function of the luminance comparison value, contrast comparison value, and structural comparison

TABLE I
CT SCANNER SETTINGS

Scan mode	Axial	Recon kernel	Standard
kVp	120	Image Size	512×512
Scan speed	Normal (1 second)	Beam hardening correction	NO
Collimation (mm)	20	DFOV(cm)	25
Number of views	Normal (984)	Slice thickness (mm)	0.625

value [42] as follows:

$$\text{SSIM}(x, y) = f(l(x, y), c(x, y), s(x, y)) \quad (18)$$

where

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (19)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (20)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}. \quad (21)$$

Here, μ_x and μ_y are the averages, σ_x and σ_y are the standard deviations, σ_{xy} is the covariance of the local windows, and C_1 , C_2 , and C_3 are constants.

Bland–Altman Analysis: The Bland–Altman analysis shows the difference of intensity $\text{Diff}(I)$ of sample points from different target materials between the original (or reference) and reconstructed images, as given in the following:

$$\text{Diff}(I) = |I_{\text{Ref}} - I_{\text{Rec}}| \quad (22)$$

where I_{Ref} and I_{Rec} are the intensity values of m sample points from the original and reconstructed images, respectively.

F. Modulation Transfer Function

The modulation transfer function (MTF) can measure the contrast difference of an image between a pair of white lines as follows:

$$\text{MTF} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \times 100\% \quad (23)$$

where I_{\max} and I_{\min} are maximum and minimum intensities, respectively, between a pair of white lines.

III. CT DATA ACQUISITION

Experimental data were taken using a GE healthcare multi-slice CT scanner (Discovery CT750 HD) from a CATPHAN 600 Phantom [43]. A 20-cm diameter image quality phantom was chosen because it has five different modules for clinical CT image assessment with respect to noise level, spatial resolution, low-contrast resolution, and the linearity of CT numbers [44]. This is a standard phantom often used for children and skinny adults [45]. The following settings are used for data acquisition, as shown in Table I.

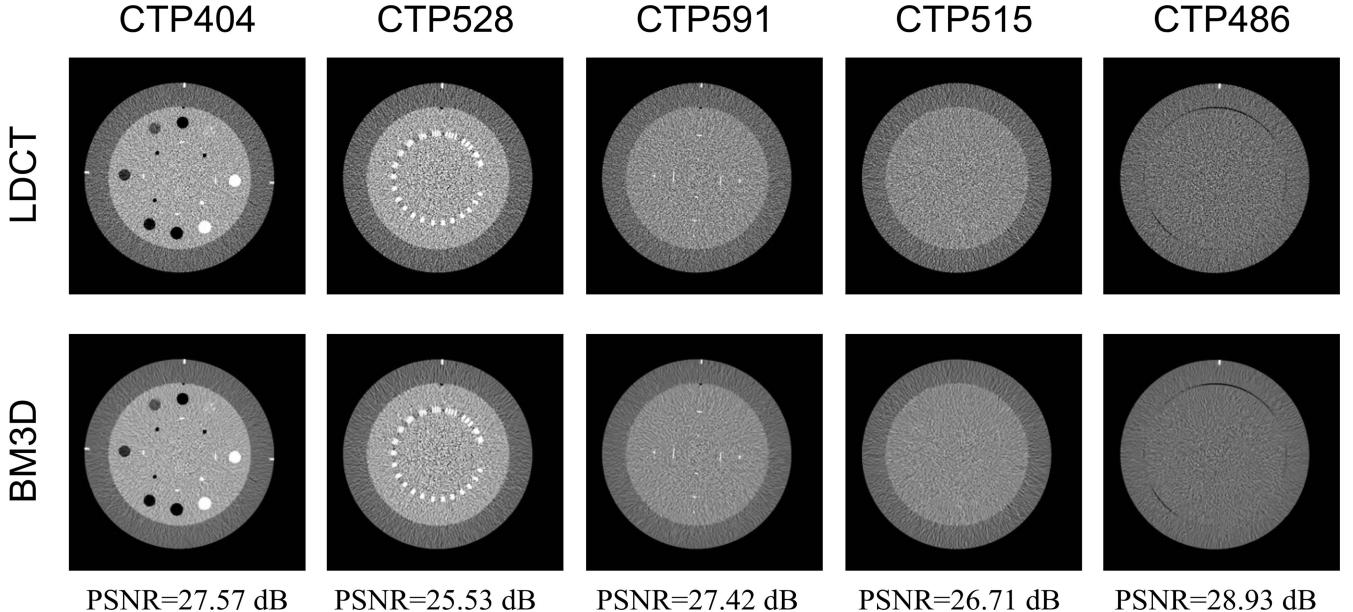


Fig. 4. Top row: input low-dose CT images from various modules of Catphan. Bottom row: denoised images from various modules of Catphan 600 using a BM3D filter with $\sigma = 30$. Window settings (WW/WL = 400/40) are used in all images.

TABLE II
PSNR IN dB FOR VARIOUS METHODS AND MODULES

Module	Input LDCT	Benchmark		Our methods		
		N2N	BSS+ BM3D	CN2G	CN3G	HCN3G
CTP404	24.53 ± 0.07	31.12 ± 0.08	31.29 ± 0.29	33.38 ± 0.11	34.06 ± 0.26	34.61 ± 0.12
CTP528	22.38 ± 0.12	30.12 ± 0.16	29.66 ± 0.15	31.46 ± 0.45	33.30 ± 0.45	33.89 ± 0.02
CTP591	24.43 ± 0.02	31.55 ± 0.05	31.48 ± 0.30	33.13 ± 0.33	33.94 ± 0.34	34.62 ± 0.07
CTP515	23.54 ± 0.40	31.78 ± 0.23	30.19 ± 0.45	33.54 ± 0.20	34.51 ± 0.39	33.76 ± 0.35
CTP486	25.12 ± 0.17	32.08 ± 0.04	32.37 ± 0.27	34.25 ± 0.29	35.11 ± 0.33	35.78 ± 0.13

The cathode current was varied to change the radiation dosages ($CTDI_{vol}$). In a low-dose setting, 10-mA cathode current was used, which corresponded to 0.88-mGy dosage. In a normal-dose setting, 300-mA cathode current was used, which corresponded to a 26.47-mGy dosage.

Due to the limitations of the computing resource used in the experiments, we have used CT images of 256×256 resolution for training (instead of the original 512×512 size), as done in the original N2N work [25].

IV. RESULTS

This section shows the comparison among the generated images from three of our models with two benchmark models and reference images from five different modules of the CATPHAN 600 phantom. Our models correspond to CN2G, CN3G, and HCN3G models. We have used Noise2Noise [25] and our previous work using blind source separation (BSS + BM3D) [33] as the benchmark methods. We have used two low-dose frames in CN2G and HCN3G models and three low-dose frames in the CN3G model. Since each low-dose image has 3.33% dosage compared to that of

TABLE III
SSIM FOR VARIOUS METHODS AND MODULES

Module	Input LDCT	Benchmarks		Our methods		
		N2N	BSS+ BM3D	CN2G	CN3G	HCN3G
CTP404	0.64 ± 0.002	0.77 ± 0.003	0.77 ± 0.004	0.82 ± 0.002	0.83 ± 0.01	0.85 ± 0.002
CTP528	0.61 ± 0.002	0.76 ± 0.002	0.76 ± 0.003	0.79 ± 0.01	0.82 ± 0.01	0.84 ± 0.001
CTP591	0.60 ± 0.001	0.76 ± 0.000	0.76 ± 0.001	0.79 ± 0.01	0.81 ± 0.01	0.83 ± 0.002
CTP515	0.59 ± 0.001	0.76 ± 0.004	0.76 ± 0.005	0.79 ± 0.01	0.82 ± 0.01	0.83 ± 0.001
CTP486	0.61 ± 0.003	0.75 ± 0.000	0.76 ± 0.008	0.81 ± 0.01	0.83 ± 0.01	0.85 ± 0.003

the normal-dose image, the two and three low-dose frames in an image sequence correspond to 6.7% and 10% dosage, respectively.

Tables II and III show the PSNR and SSIM indices for different images concerning the normal-dose image from the respective modules. For all methods, we have shown the average values with the standard deviation of PSNR and SSIM of different generators. In addition to these performance metrics, we also used a modified Bland–Altman analysis to measure material noise characteristics in the CTP404 module and contrast difference between white line pairs in terms of MTF in the CTP528 module.

Since one of our benchmark methods used a hybrid method of BSS combined with BM3D, as an initial check, we have applied a BM3D filter only to denoise the low-dose CT images from different modules of the CATPHAN 600 phantom. The BM3D filter uses a noise scale parameter σ to control the denoising effect [46]. Fig. 4 shows the denoised images from different modules of the CATPHAN 600 phantom using a BM3D filter with $\sigma = 30$. As we increase the value of σ , we get more denoising on the images with an increasing blurring effect.

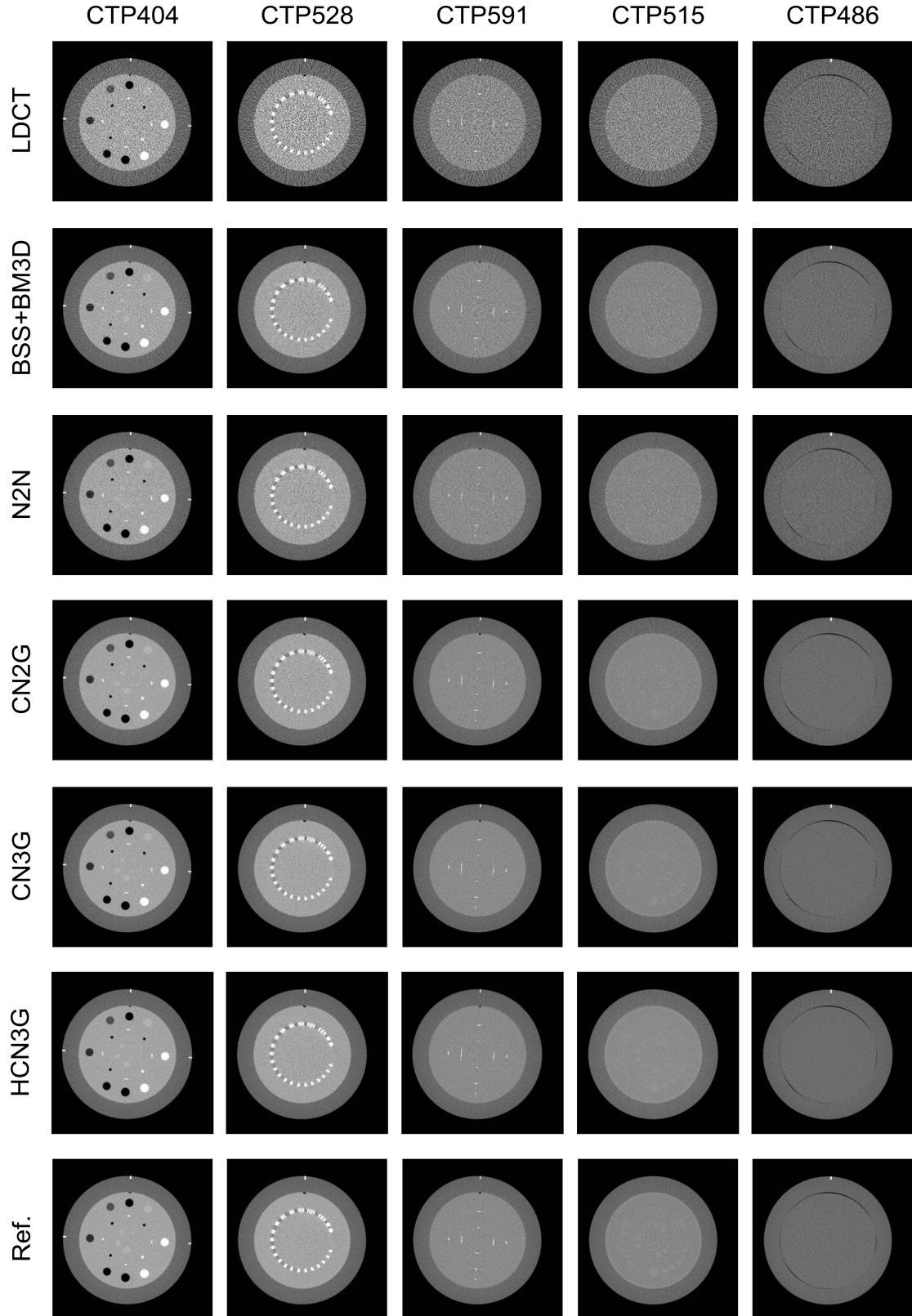


Fig. 5. Top row: input low-dose CT images from various modules of Catphan. Middle five rows: generated images from various methods and modules of Catphan. Row two and three show the benchmark methods (BSS + BM3D and N2N) and the next three rows show our three methods. Bottom row: reference or normal-dose CT (NDCT) images from various modules of Catphan. Window settings (WW/WL = 400/40) are used in all images.

Fig. 5 shows the visual comparison among the generated images from three of our models—CN2G, CN3G, and HCN3G with two benchmark models BSS + BM3D and N2N and reference images from the five different modules of the CATPHAN 600 phantom.

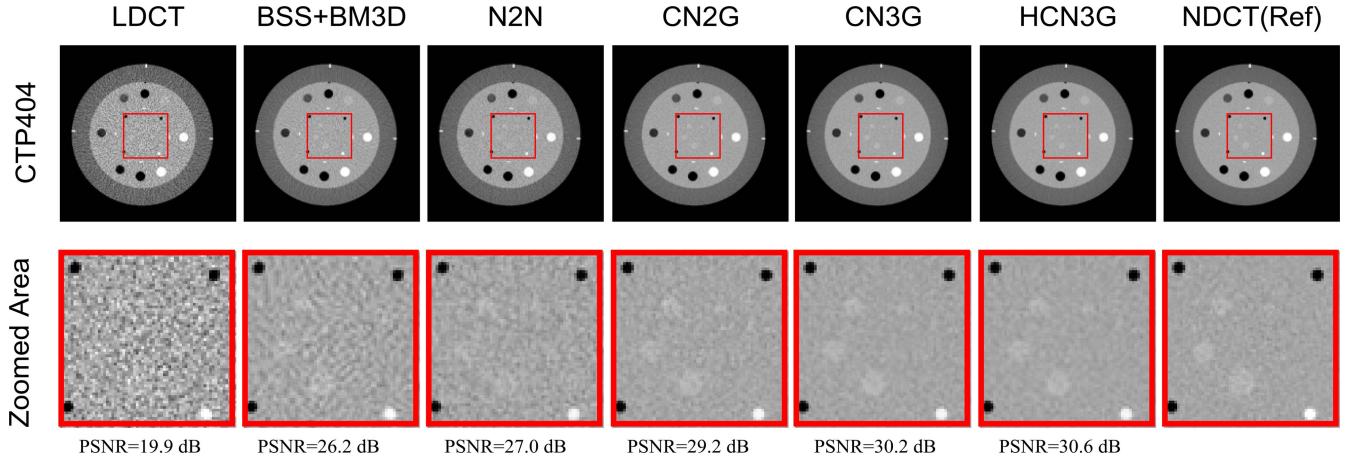


Fig. 6. Visual comparison of sensitometry targets in the CTP404 module. An area is zoomed in the bottom row to show the different diameters contrast targets in various methods. Window settings ($WW/WL = 400/40$) are used in all images. PSNR of the zoomed section is calculated with respect to NDCT (reference image).

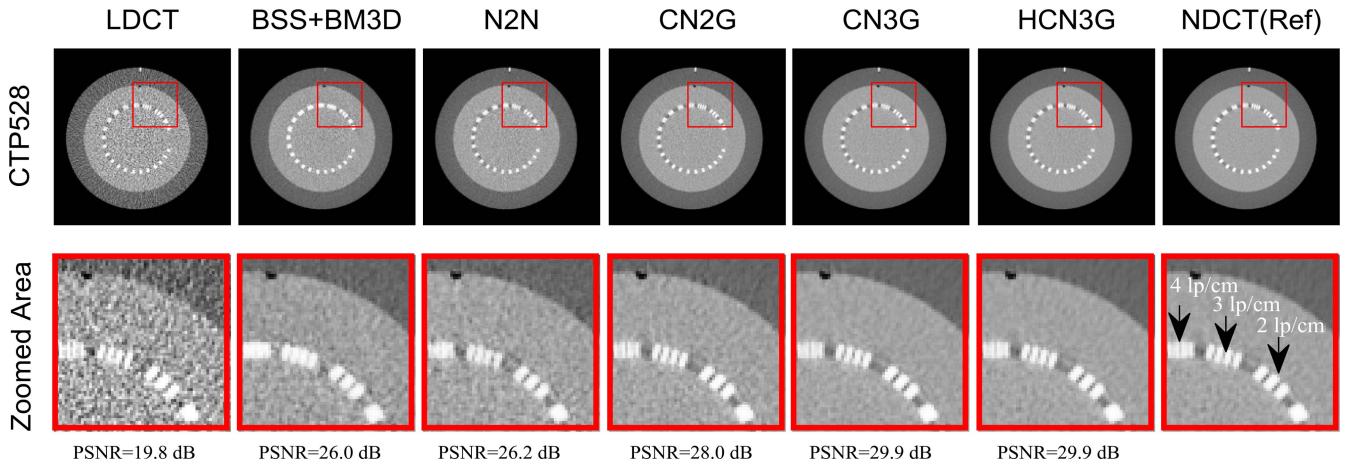


Fig. 7. Visual comparison of the spatial resolution in the CTP528 high-resolution module. An area is zoomed in the bottom row to show the 2, 3, and 4 line pairs per centimeter sections in various methods. Window settings ($WW/WL = 400/40$) are used in all images. PSNR of the zoomed section is calculated with respect to NDCT (reference image).

The CTP404 module has sensitometry targets made from teflon, delrin, acrylic, polystyrene and low-density polyethylene (LDPE), polymethylpentene (PMP), and air. It has 10, 8, 6, 4, and 2-mm acrylic spheres in the center. The CTP528 module has high-resolution sections from 1–21 lines pairs per cm. In Fig. 5 and Tables II and III, our proposed methods show higher PSNR and SSIM values than other benchmark methods. Fig. 6 shows the zoomed areas in the CTP404 module to highlight the acrylic spheres of different sizes. In the low-dose CT image, these acrylic spheres are completely invisible in the zoomed section. In our methods, the four acrylic spheres are quite conspicuous. The zoomed image in HCN3G looks the best among our three methods. The PSNR value of the zoomed section in the HCN3G method has the best value among different methods. Fig. 7 shows the zoomed areas in the CTP528 module to highlight two, three, and four line pairs per centimeter sections. In the low-dose CT image, we can see these line pairs; however, due to noise, the edges do not look sharp. In our methods, the line pairs appear clear and sharp. The PSNR value of the zoomed section in

the HCN3G method has the highest value compared to other methods.

The CTP515 module has some low-contrast subslice targets, which pose the hardest challenge to any denoising algorithms. Fig. 5 and Tables II and III show that our present proposed methods demonstrate better performance than the two benchmark methods. Fig. 8 shows the zoomed areas in this module to highlight the low-contrast targets of different diameters in various methods. In the low-dose CT image, these acrylic spheres are completely invisible in the zoomed section. In our methods, several different size acrylic spheres are quite noticeable. The zoomed image in CN3G appears the best compared to our methods. The PSNR value of the zoomed section in using the CN3G method has also shown the highest value among different methods.

Finally, the CTP591 module contains two coarse pairs and one fine pair of opposed ramps. The CTP486 module is used for spatial uniformity and noise precision. Like previous cases, our proposed method performs better for these two modules than other methods, as shown in Fig. 5 and Tables II and III.

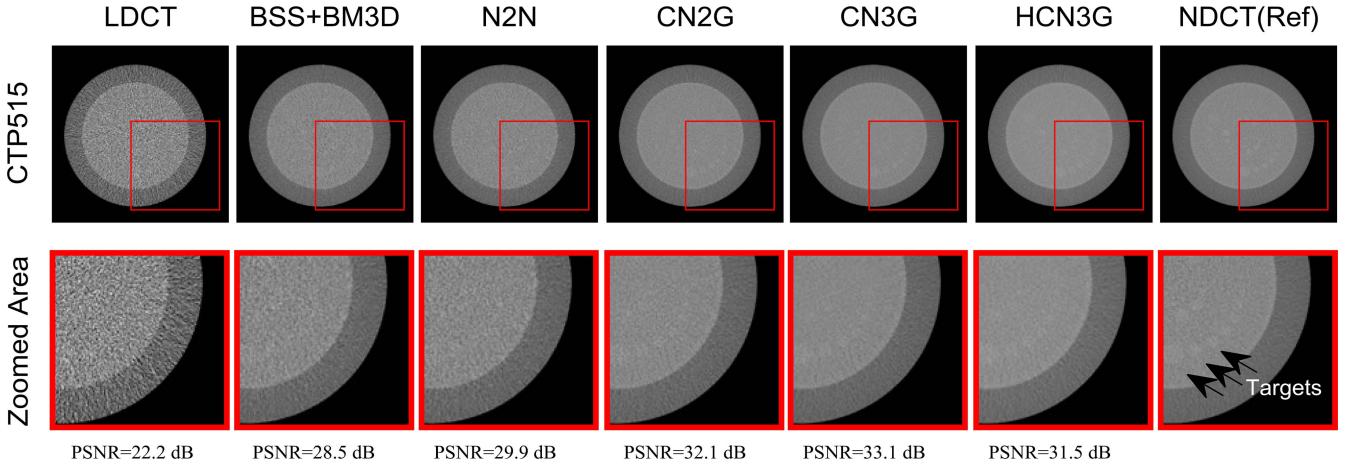


Fig. 8. Visual comparison of low-contrast targets in the CTP515 module. An area is zoomed in the bottom row to show the 1% contrast targets of different diameters in various methods. Window settings ($WW/WL = 400/40$) are used in all images. PSNR of the zoomed section is calculated with respect to NDCT (reference image).

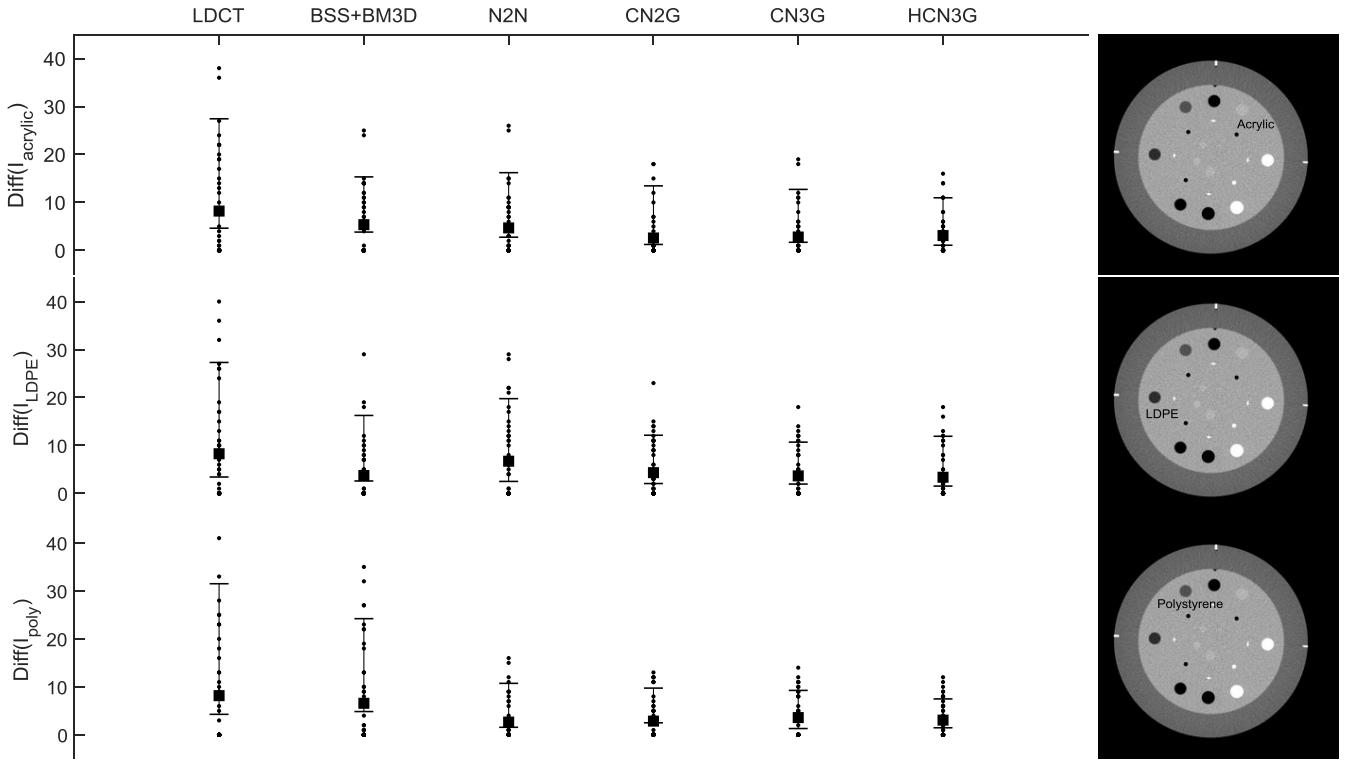


Fig. 9. Intensity difference in acrylic, LDPE, and polystyrene using a modified Bland–Altman analysis in the CTP404 module.

Fig. 9 shows a modified Bland–Altman analysis that illustrates the difference of pixel intensities for different methods with reference image $\text{Diff}(I)$ for three target materials acrylic, polystyrene, and LDPE in the CTP404 module. We have chosen $m = 49$ sample points from a square region of 7×7 from each of these target materials to calculate $\text{Diff}(I)$. Table IV shows the 95% confidence limit for various methods and target materials: acrylic, LDPE, and polystyrene. Other target materials in the CTP404 module were excluded in this analysis because they showed less variation of pixels intensities for the Window settings ($WW/WL = 400/40$). For polystyrene, acrylic, and LDPE, due to the material characteristics, the

analysis shows a skewed distributed $\text{Diff}(I)$ for all methods. However, the overall variances of $\text{Diff}(I)$ are quite low in our three methods, compared to benchmark methods for target materials—LDPE, polystyrene, and acrylic. HCN3G performed the best reduction in the noise variations for the target materials—acrylic and polystyrene. The CN3G model showed the lowest 95% confidence limit for the target material LDPE. For acrylic, the mean $\text{Diff}(I)$ values of our methods are significant at $p < 0.01$, and for LDPE and polystyrene, the mean $\text{Diff}(I)$ values of our methods are significant at $p < 0.05$.

Fig. 10 shows the line profile along a line drawn across four, three, and two-line pairs per centimeter. The distance

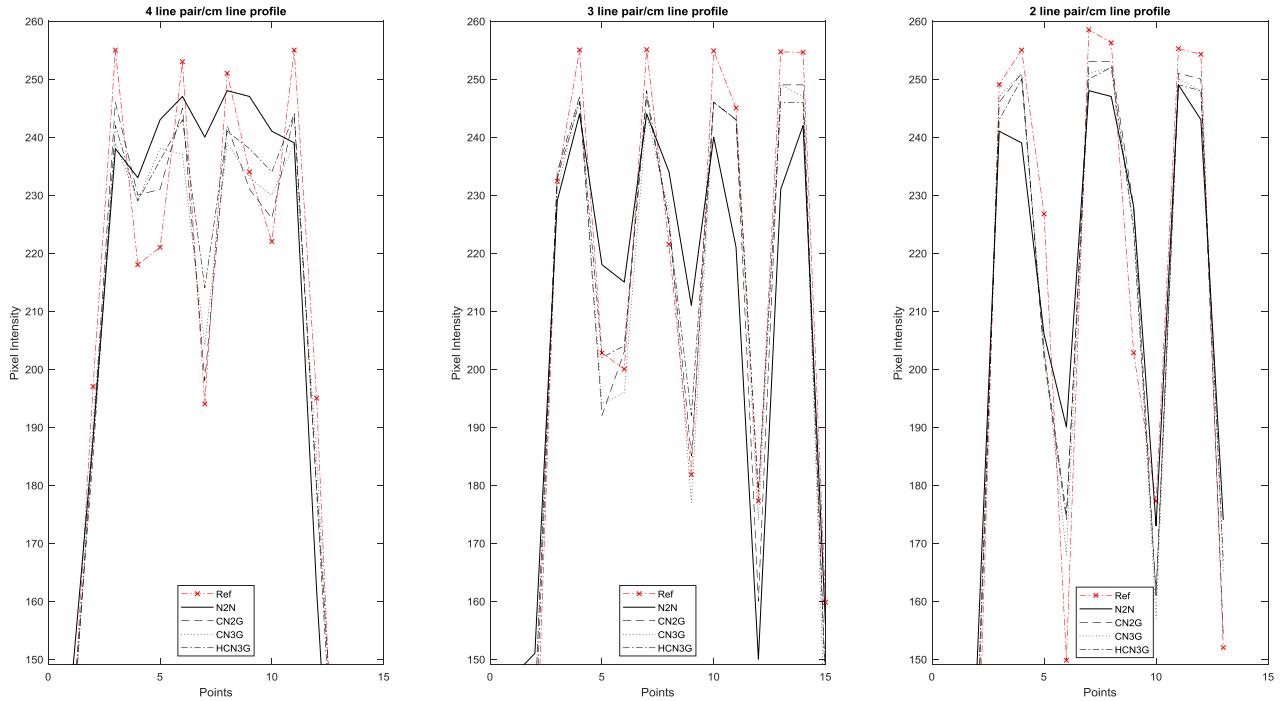


Fig. 10. Line profiles for the 4, 3, and 2 line pair per centimeter sections of different methods in the CTP528 module.

TABLE IV

PIXEL INTENSITY DIFFERENCES IN TERMS OF 95% CONFIDENCE LIMIT FOR VARIOUS METHODS AND MATERIALS IN THE CTP404 MODULE

Material	Input LDCT	Benchmarks		Our methods		
		N2N	BSS+ BM3D	CN2G	CN3G	HCN3G
Acrylic	22.81	13.45	11.72	12.22	11.03	9.89
LDPE	23.88	17.26	9.81	10.07	8.75	10.38
Polystyrene	27.23	9.16	15.89	7.23	7.95	6.02

TABLE V

CONTRAST DIFFERENCES IN TERMS OF MTF FOR VARIOUS METHODS AND LINE PAIRS IN THE CTP528 MODULE

Line pairs (lp/cm)	Benchmarks		Our methods		
	Ref	N2N	CN2G	CN3G	HCN3G
2 lp/cm	26.2%	13.4%	22.2%	19.8%	22.0%
3 lp/cm	16.9%	7.25%	14.3%	15.8%	12.7%
4 lp/cm	13.2%	2.00%	10.6%	7.90%	6.30%

between two white lines varies with different line pairs per centimeter. For two, three, and four line pairs per centimeter sections, the gap between two white lines are 0.25, 0.167, and 0.125 cm, respectively [43]. We have calculated the contrast difference between a white line pair of an image using MTF, as shown in Table V. In clinical applications, an MTF below 5%–10% cannot be detected by human vision [47]. In our experiment, we have set the cut-off MTF at 10%. For two and three line pairs per centimeter, our methods show a slightly lower contrast difference in terms of MTF than the reference. Since all these MTF values are above the cut-off value of 10%, these line pairs are quite detectable. But for four line pairs per centimeter, it has become difficult to detect line pairs for most

of the methods. Amongst our three methods, CN2G shows better contrast difference in terms of MTF than CN3G and HCN3G for four line pairs per centimeter.

V. CONCLUSION

In this article, we have introduced a collaborative technique to train multiple N2N generators using three different models: 1) CN2G; 2) CN3G; and 3) HCN3G. By using this technique, we can reduce the radiation dosage quite significantly without losing significant image details, especially for low-contrast areas. Besides, they overcome some problems with single N2N generators in the low-dose CT image denoising task due to its nonuniform noise model. Amongst our three methods, the HCN3G model performs the best in terms of PSNR, SSIM, and material noise characteristics, while CN2G and CN3G perform better in terms of contrast difference. In addition, we have introduced collaborative loss terms which work very well with the L_2 losses of individual generators of our model. In our experiment, we have not tested all the different configurations of CN models using different weighting factors exhaustively. Our future work can be directed to investigate the other configurations of the CN generators to get even better performance at higher resolutions.

ACKNOWLEDGMENT

The research is conducted at the Multimedia Processing and Prototyping Lab in the Department of Electrical and Computer Engineering at the University of Saskatchewan. The authors would like to thank Troy Anderson from Royal University Hospital, Saskatoon, Canada for setting up and collecting slice images using the CATPHAN 600 phantom.

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, 2007.
- [2] N. Savage, "Medical imagers lower the dose," *IEEE Spectr.*, vol. 47, no. 3, pp. 14–16, Feb. 2010.
- [3] T. Li *et al.*, "Nonlinear sinogram smoothing for low-dose X-ray CT," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 5, pp. 2505–2513, Oct. 2004.
- [4] Y. Chen *et al.*, "Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means," *Phys. Med. Biol.*, vol. 57, no. 9, p. 2667, 2012.
- [5] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1682–1697, Sep. 2012.
- [6] P. Prakash *et al.*, "Reducing abdominal CT radiation dose with adaptive statistical iterative reconstruction technique," *Invest. Radiol.*, vol. 45, no. 4, pp. 202–210, 2010.
- [7] N. Negi *et al.*, "Hepatic CT perfusion measurements: A feasibility study for radiation dose reduction using new image reconstruction method," *Eur. J. Radiol.*, vol. 81, no. 11, pp. 3048–3054, 2012.
- [8] J. He *et al.*, "Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 371–382, Feb. 2019.
- [9] E. Bar-Aviv, Z. Devir, E. Dahan, T. Kenig, and G. Rosman, "Denoising medical images," U.S. Patent 8 605 970, Dec. 10, 2013.
- [10] A. Britten, M. Crotty, H. Kiremidjian, A. Grundy, and E. Adam, "The addition of computer simulated noise to investigate radiation dose and image quality in images with spatial correlation of statistical noise: An example application to X-ray CT of the brain," *Brit. J. Radiol.*, vol. 77, no. 916, pp. 323–328, 2004.
- [11] Y. Chen *et al.*, "Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing," *Phys. Med. Biol.*, vol. 58, no. 16, p. 5803, 2013.
- [12] R. J. Geraldo, L. M. V. Cura, P. E. Cruvinel, and N. D. A. Mascarenhas, "Low dose CT filtering in the image domain using map algorithms," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 1, no. 1, pp. 56–67, Jan. 2017.
- [13] Z. Li *et al.*, "Adaptive nonlocal means filtering based on local noise level for CT denoising," *Med. Phys.*, vol. 41, no. 1, 2014, Art. no. 011908.
- [14] A. M. Mendrik, E.-J. Vonken, A. Rutten, M. A. Viergever, and B. van Ginneken, "Noise reduction in computed tomography scans using 3-D anisotropic hybrid diffusion with continuous switch," *IEEE Trans. Med. Imag.*, vol. 28, no. 10, pp. 1585–1594, Oct. 2009.
- [15] S. Seyyedi, E. Liapi, T. Lasser, R. Ivkov, R. Hatwar, and J. W. Stayman, "Low-dose CT perfusion of the liver using reconstruction of difference," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 3, pp. 205–214, Mar. 2018.
- [16] G. Bertasius, J. Shi, and L. Torresani, "Deepedge: A multi-scale bifurcated deep network for top-down contour detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 4380–4389.
- [17] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 2472–2481.
- [18] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3156–3164.
- [19] X. Yi and M. Eramian, "LBP-based segmentation of defocus blur," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1626–1638, Apr. 2016.
- [20] M. Hatt, C. Parmar, J. Qi, and I. El Naqa, "Machine (deep) learning methods for image processing and radiomics," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 2, pp. 104–108, Mar. 2019.
- [21] H. Chen *et al.*, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.
- [22] H. Chen *et al.*, "Low-dose CT via convolutional neural network," *Biomed. Opt. Exp.*, vol. 8, no. 2, pp. 679–694, 2017.
- [23] V. S. Kadimesetty, S. Gutta, S. Ganapathy, and P. K. Yalavarthy, "Convolutional neural network-based robust denoising of low-dose computed tomography perfusion maps," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 2, pp. 137–152, Mar. 2019.
- [24] D. Wu, H. Ren, and Q. Li, "Self-supervised dynamic CT perfusion image denoising with deep neural networks," *IEEE Trans. Radiat. Plasma Med. Sci.*, early access, May 21, 2020, doi: 10.1109/TRPMS.2020.2996566.
- [25] J. Lehtinen *et al.*, "Noise2Noise: Learning image restoration without clean data," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 2965–2974.
- [26] J. Batson and L. Royer, "Noise2Self: Blind denoising by self-supervision," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 524–533.
- [27] A. Krull, T. O. Buchholz, and F. Jug, "Noise2Void-learning denoising from single noisy images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2129–2137.
- [28] R. A. Flores-Mendez, "Towards the standardization of multi-agent system architectures: An overview," *ACM Crossroads Intell. Agents Assoc. Comput. Mach.*, vol. 5, no. 4, pp. 18–24, 1999.
- [29] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Işgum, "Generative—Adversarial networks for noise reduction in low-dose CT," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2536–2545, Dec. 2017.
- [30] Q. Yang *et al.*, "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1348–1357, Apr. 2018.
- [31] X. Yi and P. Babyn, "Sharpness-aware low-dose CT denoising using conditional generative adversarial network," *J. Digit. Imag.*, vol. 31, no. 5, pp. 655–669, 2018.
- [32] A. Bora, E. Price, and A. G. Dimakis, "Ambientgan: Generative models from lossy measurements," in *Proc. ICLR*, vol. 2, 2018, p. 5.
- [33] A. M. Hasan, A. Melli, K. A. Wahid, and P. Babyn, "Denoising low-dose ct images using multiframe blind source separation and block matching filter," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 4, pp. 279–287, Jul. 2018.
- [34] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [35] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Assist. Intervent.*, 2015, pp. 234–241.
- [36] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1125–1134.
- [37] H. Zhang, V. Sindagi, and V. M. Patel, "Image de-raining using a conditional generative adversarial network," *IEEE Trans. Circuits Syst. Video Technol.*, early access, Jun. 3, 2019, doi: 10.1109/TCSVT.2019.2920407.
- [38] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 694–711.
- [39] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2223–2232.
- [40] M. Drozdzial, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, "The importance of skip connections in biomedical image segmentation," in *Deep Learning and Data Labeling for Medical Applications*. Cham, Switzerland: Springer, 2016, pp. 179–187.
- [41] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 630–645.
- [42] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [43] D. Goodenough, *Catphan 500 and 600 Manual*. Greenwich, NY, USA: Phantom Lab, 2012.
- [44] N. V. Slavine, R. W. McColl, O. K. Oz, J. Guild, J. A. Anderson, and R. E. Lenkinski, "Phantom and preclinical studies for image improvement in clinical CT," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 1, pp. 96–102, Jan. 2019.
- [45] J. Santos, M. D. C. Batista, S. Foley, G. Paulo, M. F. McEntee, and L. Rainford, "Paediatric CT optimisation utilising Catphan 600 and age-specific anthropomorphic phantoms," *Radiat. Protect. Dosimetry*, vol. 162, no. 4, pp. 586–596, 2014.
- [46] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," in *Proc. SPIE*, 2006, pp. 354–365.
- [47] W. A. Kalender, *Computed Tomography: Fundamentals System Technology Image Quality Applications*, Erlangen, Germany: Publicis, 2011, pp. 118–119.