



Published in final edited form as:

IEEE Trans Med Imaging. 2022 December ; 41(12): 3587–3599. doi:10.1109/TMI.2022.3189759.

DuDoUFNet: Dual-domain under-to-fully-complete progressive restoration network for simultaneous metal artifact reduction and low-dose CT reconstruction

Bo Zhou [Student Member, IEEE],

Department of Biomedical Engineering, Yale University, New Haven, CT, 06511, USA.

Xiongchao Chen [Student Member, IEEE],

Department of Biomedical Engineering, Yale University, New Haven, CT, 06511, USA.

Huidong Xie [Student Member, IEEE],

Department of Biomedical Engineering, Yale University, New Haven, CT, 06511, USA.

S. Kevin Zhou [Fellow, IEEE],

School of Biomedical Engineering & Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou, China and the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 100190, China.

James S. Duncan [Life Fellow, IEEE],

Department of Biomedical Engineering and the Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT, 06511, USA.

Chi Liu [Senior Member, IEEE]

Department of Biomedical Engineering and the Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT, 06511, USA.

Abstract

To reduce the potential risk of radiation to the patient, low-dose computed tomography (LDCT) has been widely adopted in clinical practice for reconstructing cross-sectional images using sinograms with reduced x-ray flux. The LDCT image quality is often degraded by different levels of noise depending on the low-dose protocols. The image quality will be further degraded when the patient has metallic implants, where the image suffers from additional streak artifacts along with further amplified noise levels, thus affecting the medical diagnosis and other CT-related applications. Previous studies mainly focused either on denoising LDCT without considering metallic implants or full-dose CT metal artifact reduction (MAR). Directly applying previous LDCT or MAR approaches to the issue of simultaneous metal artifact reduction and low-dose CT (MARLD) may yield sub-optimal reconstruction results. In this work, we develop a dual-domain under-to-fully-complete progressive restoration network, called DuDoUFNet, for MARLD. Our DuDoUFNet aims to reconstruct images with substantially reduced noise and artifact by progressive sinogram to image domain restoration with a two-stage progressive restoration network design. Our experimental results demonstrate that our method can provide high-quality

reconstruction, superior to previous LDCT and MAR methods under various low-dose and metal settings.

Keywords

Low-dose CT; Metal artifact reduction; Dual-domain learning; Progressive restoration network

I. INTRODUCTION

X-RAY computed tomography (CT) is a non-invasive imaging technique that visualizes a patient's internal structures and has become one of the common examinations for medical diagnosis. Due to the increasing use of CT in clinical practice, concerns have been raised on the overall radiation dose to a patient. Thus, it is desirable to reduce the CT radiation dose. The most common ways include decreasing the operating current/voltage and shortening the exposure time of the x-ray tube [1], [2]. However, a weaker level of x-ray flux results in sinogram with higher Poisson noise, thus producing a reconstruction image with a low signal-to-noise ratio (SNR) [3], [4]. For patients with metallic implants, such as spinal implants [5] and hip prostheses [6], the image quality is further degraded due to the beam hardening effects caused by these high attenuation objects where the low-energy x-ray photons are attenuated more easily than the remaining high-energy photons, given a polychromatic x-ray spectrum in CT. Specifically, the projection data in the metal trace region in the sinogram does not follow simple exponential decay as with a monochromatic x-ray, thus results in additional metal artifacts in the reconstruction [3], [7]. The non-ideal image quality not only seriously affects the image quality for diagnostic purposes but also impacts other medical procedures that rely on CT, such as dose calculation in radiation therapy [8] and PET/SPECT attenuation corrections [9]. An example of low-dose CT with metallic implants is shown in Figure 1. With the increasing use of metallic implants and interest in reducing the CT radiation dose, how to reconstruct high-quality CT images for patients with metallic implants under low-dose settings is an important research direction in CT imaging.

For the low-dose CT (LDCT) reconstruction, there are many previous reconstruction approaches and can be summarized into two general categories, namely model-based iterative reconstruction (MBIR) and deep learning-based reconstruction (DLR). Previous MBIR methods reconstruct images from the noisy sinogram by iteratively minimizing a unified objective function combined from the prior information in the image domain and the statistical properties of data in the sinogram domain. Common choices of image priors include total variation (TV) and its variants [10], non-local mean [11], and dictionary learning [12]. Even though MBIR methods generate much-improved reconstruction quality, they rely on iterative forward and back-projection operations that not only require knowledge of vendor-specific scanning geometry but also are computationally heavy, thus suffering from a long reconstruction time.

On the other hand, DLR methods that have been developed for LDCT show promising reconstruction quality [13]. Chen et al. [14] first proposed to use a two-layer convolutional

neural network (CNN) for LDCT image denoising. Later, Chen et al. [15] further advanced to a RED-CNN design and show improved LDCT image denoising performance. Kang et al. [16] followed a similar approach but applied the CNN to a directional wavelet transform of the CT image. Yi et al. [17] proposed to add adversarial learning [18] in the denoising network training to further improve the sharpness in the denoised image. Similarly, Yang et al. [19] proposed to add Wasserstein distance-based adversarial learning and a perceptual loss [20] to further improve the performance. Other than using DLR in the image domain alone, Yin et al. [21] also proposed to use a sinogram domain network followed by an image domain network for domain progressive LDCT denoising. However, none of the previous LDCT algorithms considers the scenario when a patient has metallic implants, which would further degrade the image quality. Directly applying the previous LDCT approach to LDCT with metallic implants may result in sub-optimal reconstruction performance.

For metal artifact reduction (MAR), there are many previous methods for MAR in full-dose CT setting, and can be summarized into two general categories too, namely traditional sinogram-based correction methods and deep learning-based methods. As the metal artifacts are non-local in the image due to the beam hardening in the metal-affected sinogram regions, traditional MAR methods either correct the metal-affected sinogram regions by modeling the physical effects of x-ray imaging [22] or replace the metal-affected sinogram regions by estimated values [23]–[27]. Substituting the metal-affected sinogram regions by linear interpolation of its neighboring unaffected sinograms is one typical solution [23]. However, the inconsistency between the interpolated values and unaffected values often leads to new artifacts in the reconstructions. For improved estimation of the metal-affected region, previous works also attempted to utilize the forward projection of synthesized prior images [24], [27]–[30]. In general, these methods aim to first estimate an artifact-reduced prior image or sinogram based on physical properties of CT, and then use this prior signal to guide the sinogram completion. For instance, NMAR [24] generates a prior image by a multi-threshold segmentation of the initial reconstructed image [28]. Then, the forward projection of the prior image is used to normalize the sinogram before linear interpolation, thus improving the value estimation in the metal-affected sinogram regions.

With the recent advances in CNN for medical image reconstruction, deep learning-based MAR methods have also been proposed and demonstrated promising performance. Park et al. [31] proposed to use UNet [32] to correct the sinogram inconsistency caused by beam hardening. Gjestebj et al. [33] further improved the NMAR [24] by utilizing a CNN with three convolutional layers for improved sinogram data correction. Similarly, Zhang et al. [34] used a CNN with five convolutional layers to generate an artifact-reduced image from initial artifact-reduced images, and then the generated image is forward projected to aid the metal-affection sinogram regions' corrections. Besides addressing the MAR in the sinogram domain, there are also many previous works on reducing metal artifacts by image post-processing. Huang et al. [35] used a CNN with a residual connection between the input and output for MAR in cervical CT. Wang et al. [36] utilized a 3D conditional adversarial network [18] for MAR in ear CT. Gjestebj et al. [37] proposed to further improve the MAR performance by adding a perceptual loss [20]. Combining the sinogram domain and image domain, dual-domain restoration methods have been developed. Lin et al. [38] first proposed DuDoNet that combines the sinogram domain correction and image

domain correction into a single network. Lyu et al. [39] further improved the DuDoNet by adding the metal mask information in the dual-domain network. Similarly, Yu et al. [40], [41] found that dual-domain learning with deep sinogram completion by first estimating a prior image using CNN and then correct sinogram using another CNN yields further improved MAR performance. However, the aforementioned MAR algorithms only address the MAR in full-dose CT. With MAR in LDCT, the image quality is heavily degraded by both noise and metal artifacts. While the previous image-to-image based MAR algorithms [35]–[37] could be adapted to the simultaneous metal artifact reduction and low-dose CT reconstruction (MARLD) problem, the performance may be sub-optimal since the image quality is heavily degraded by both noise and metal artifacts, and image-domain alone correction may be insufficient. Other MAR algorithms are hard to be directly adapted to the MARLD problem since they generally assume only the metal-affected sinogram needs to be corrected.

As the reconstruction quality is seriously degraded when these two conditions are present simultaneously, the stand-alone LDCT methods and MAR methods may not be suitable for MARLD. To address these issues, we develop a dual-domain under-to-fully-complete progressive restoration network (DuDoUFNet) for MARLD. The pipeline of our DuDoUFNet is illustrated in Figure 2. Specifically, we propose progressive restoration in terms of both domain and network design. First, we use progressive dual-domain restoration. Since the image noise and non-local metal artifacts are caused by noisy and inconsistent projection data in the sinogram domain, we propose to progressively restore information on the sinogram domain and then on the image domain. Second, we propose to use a under-to-fully-complete progressive restoration network (UFNet) for the dual-domain restorations. Please note that all of the aforementioned previous deep learning-based LDCT and deep learning-based MAR methods relied on UNet or its variants [32] for restorations. While UNet with downsampling and upsampling operations can recover satisfactory general image content, it is not suitable for recovering fine details. Thus, it is neither suitable for sinogram domain restoration since a sinogram with a loss of fine details may lead to secondary artifacts in the reconstructed image, or suitable for image domain restoration since the final restored image may lose fine details. Inspired by the recent works in image restoration [42]–[47], we combined an under-complete restoration network and a fully-complete restoration network with cross stage connections for progressive restoration, where the multi-scale features from the under-complete network helps enrich the original resolution fine features from fully-complete restoration for final restorations. The proposed network is used for both sinogram and image domain restorations. Our DuDoUFNet is trained in an end-to-end fashion with losses supervising in both domain progressive outputs and network progressive outputs.

II. METHODS

A. Overview

The framework of our DuDoUFNet is illustrated in Figure 2. The DuDoUFNet aims to simultaneously reduce the noise and metal artifacts in low-dose CT by sinogram domain and image domain progressive restoration learning. Specifically, the DuDoUFNet consists of a

sinogram domain progressive restoration network G_{sino} and an image domain progressive restoration network G_{img} , where G_{sino} and G_{img} share the same architecture (Figure 3) while configuring with a different number of channel inputs. The progressive restoration network is a two-stage network with under-complete and fully-complete restoration outputs. The G_{sino} and G_{img} are connected by a Filtered Back Projection (FBP) layer [45] to enable the dual-domain restoration learning.

Since the noise and metal artifacts in the image domain are highly correlated caused by the noise and missing projection data in the sinogram domain, we propose to use sinogram restoration learning followed by image restoration learning. Given the low-dose metal-affected sinogram $S_{Idma} \in \mathbb{R}^{L \times V}$ and the metal trace $M_{project} \in \mathbb{R}^{L \times V}$ that forward projected from the metal mask $M \in \mathbb{R}^{H \times W}$, we first concatenate S_{Idma} and M_{proj} and input the two-channel sinogram into the G_{sino} for progressive sinogram restoration, where M_{proj} indicates the metal size and location information in the sinogram domain. G_{sino} outputs both the under-complete restoration output S_u and fully-complete restoration output S_o . Then, the finest restored sinogram S_o is converted into an initial restored image $X_o \in \mathbb{R}^{H \times W}$ via the FBP layer [45] for further image domain restoration. Since the FBP image X_{Idma} reconstructed from S_{Idma} still provides correct anatomical outlines in non-metal regions (Figure 1), we concatenate X_{Idma} and M with X_o and input the 3-channel image into the G_{img} for progressive image restoration. Similar to the sinogram restoration, G_{img} outputs both the under-complete restoration output X_u and fully-complete restoration output X_{final} and the image X_{final} is our final reconstruction outputs. During training, supervision is provided for S_o , S_u , X_o , X_u , and X_{final} .

B. Under-to-fully-complete Restoration Network

The backbone restoration network is a two-stage network, called under-to-fully-complete progressive restoration network (UFNet), consisting of an under-complete restoration subnetwork (UnNet), a fully-complete restoration subnetwork (FuNet), and cross-stage connections. The architecture of UFNet is illustrated in Figure 3.

The UnNet is based on a U-shape network [32] for the stage-1 restoration, as illustrated in the bottom part of Figure 3. Specifically, we propose to use a residual squeeze-and-excitation block (RSEB) for feature extraction. Given an input feature F_{in} for RSEB, the output can be written as:

$$F_{out} = F_{in} + P_{se} [P_{ex}(F_{in})], \quad (1)$$

where P_{ex} consists of two convolutional layers for feature extraction, and P_{se} is the squeeze-and-excitation layer [48], generating channel-attention from the input feature for channel-wise feature re-calibration. Given an input X_{in} in UnNet, we first use a convolutional layer with an RSEB for initial feature extraction. Then, two consecutive RSEBs are employed for feature extraction at 3 different levels of the UnNet, generating encoder features of F_{e1} , F_{e2} , F_{e3} and decoder features of F_{d3} , F_{d2} , F_{d1} which contain multi-scale features suitable for general appearance restoration and are used in the stage-2 restoration. The final output of UnNet then can be written as:

$$X_u = X_{in1} + P_{fu}(F_{d1}), \quad (2)$$

where P_{fu} is a convolutional layer for reducing the number of feature channel to 1, and X_{in1} is the first-channel data of X_{in} .

To utilize the final supervised feature of UnNet for the stage-2 restoration, we also proposed a self-attention connection, as illustrated by the pink connections in Figure 3. Specifically, the self-attention feature can be computed via:

$$F_{att} = F_{d1} + P_1(F_{d1}) \odot \sigma(P_2(X_u)), \quad (3)$$

where P_1 is a 3×3 convolutional layers for generating the unweighted feature, and P_2 is another 3×3 convolutional layers with input of X_u for generating spatial-wise attention weights. The attention weight is normalized by the sigmoid function σ , such that the attention weights lie between 0 and 1. The spatial attention generated from the final output of UnNet helps the spatial-wise re-calibration of the UnNet finest scale's feature which is used as part of the initial feature in the stage-2 restoration.

In the stage-2 restoration with FuNet, as illustrated in the top part of Figure 3, given the same input X_{in} and the self-attention feature F_{att} , an initial feature can be built via:

$$F_{init} = \{F_{att}, P_{init}(X_{in})\}, \quad (4)$$

where $\{ \}$ denotes channel-wise concatenation operation, and P_{init} consists of a convolutional layer followed by an RSEB. Then, F_{init} is input into three consecutive fully-complete restoration blocks (FRB) for restoration. The intermediate outputs of FRBs also fuse with the coarse-to-fine features from UnNet for restoration, such that the multi-scale features of UnNet is utilized to aid the original resolution feature recovery. The process can be formulated as:

$$F_{o1} = P_{o1}(F_{init}) + P_{e3}(F_{e3}) + P_{d3}(F_{d3}), \quad (5)$$

$$F_{o2} = P_{o2}(F_{o1}) + P_{e2}(F_{e2}) + P_{d2}(F_{d2}), \quad (6)$$

$$F_{o3} = P_{o3}(F_{o2}) + P_{e1}(F_{e1}) + P_{d1}(F_{d1}), \quad (7)$$

where P_{o1} , P_{o2} , and P_{o3} are FRB which consists of multiple RSEB with residual connections between the block's input and output. P_{e3} and P_{d3} consists of a $\times 4$ upsampling operation followed by a convolutional layer (yellow lines in Figure 3), while P_{e2} and P_{d2} consists of a $\times 2$ upsampling operation followed by a convolutional layer (green lines in Figure 3). P_{e1} and P_{d1} only contains one convolutional layer (blues lines in Figure 3). Finally, The output of FuNet can be written as:

$$X_{final} = X_{in1} + P_{fo}(F_{o3}), \quad (8)$$

where P_{f0} is a convolutional layer for reducing the number of feature channels to 1. The residual connection is added to reduce the vanishing gradient issue. Here, we set the number of RSEB to 6, where FuNet contains approximately the same number of RSEB as the UnNet, such that both subnetworks contain similar amounts of parameters.

Since the FuNet does not use any downsampling operation, it is able to maintain high-resolution features. The multi-scale feature from stage-1 helps enrich the feature in stage-2, thus aiding the final restoration with fine details.

C. Dual-domain Learning Objective

Our DuDoUFNet learns to restore the low-dose image with metal artifacts in both sinogram and image domains. Our loss function consists of three parts, including sinogram progressive restoration loss, intermediate image loss, and image progressive restoration loss. As shown in Figure 2, the sinogram domain loss directly supervises the under-complete and fully-complete restoration outputs from the UFNet by:

$$\mathcal{L}_{sino} = \|S_o - S_{gr}\|_1 + \|S_u - S_{gr}\|_1, \quad (9)$$

where S_u and S_o are the under-complete and fully-complete restored sinograms. S_{gr} is the ground-truth full-dose sinogram without metallic implants. Then, the finest restored sinogram S_o is converted to the intermediate restored image X_o using a FBP layer. The intermediate image loss is computed via:

$$\mathcal{L}_{int} = \|(1 - M) \odot (X_o - X_{gr})\|_1, \quad (10)$$

where M is the metallic implants segmentation mask in the image domain, and X_{gr} is the ground-truth full-dose image without metallic implants. With X_o fed into image domain UFNet, the image domain loss directly supervise the under-complete and fully-complete restoration outputs by:

$$\mathcal{L}_{img} = \|(1 - M) \odot (X_{final} - X_{gr})\|_1 + \|(1 - M) \odot (X_u - X_{gr})\|_1, \quad (11)$$

where X_{final} is the final restoration output of our DuDoUFNet. In the image domain losses, multiplying $1 - M$ in the loss allows the network to learn the restoration in the non-metal region since the metal is not our focus. Finally, the total loss function can be written as:

$$\mathcal{L}_{tot} = \alpha_1 \mathcal{L}_{sino} + \alpha_2 \mathcal{L}_{int} + \alpha_3 \mathcal{L}_{img} \quad (12)$$

where we empirically set $\alpha_1 = 5$ and $\alpha_2 = \alpha_3 = 1$ to achieve optimal performance. Higher weights are used for the sinogram domain restoration loss as compared to image domain restoration loss, since the dual-domain framework follows a sinogram to image domain progressive restoration, the image domain restoration relies on a reasonably restored sinogram from the sinogram restoration network. Our DuDoUFNet is trained in an end-to-end fashion by optimizing the \mathcal{L}_{tot} .

D. Data Preparation

We used realistically simulated low-dose CT images with metallic implants for training and evaluation of our method. Specifically, we used a similar data preparation procedure as the previous MAR works [38]–[40], where we randomly picked 1200 full-dose CT 2D images with a size of 512×512 pixels from the DeepLesion dataset [49] and collected 100 manually segmented metal implants with various locations, shapes, and sizes from [34]. Then, we randomly selected 1000 CT images and 90 metal masks to synthesize the training data. The remaining 10 metal masks were paired with the remaining 200 CT images to generate 2000 combinations for evaluation.

We used a similar x-ray projection protocol as in [15], [38]–[40] to simulate the low-dose metal-affected sinogram and the corresponding reconstruction images, by inserting metallic implants into clean CT images. More specifically, we considered an equiangular fan-beam projection geometry with a 120 kVp polyenergetic x-ray source. We simulated two low-dose CT scenarios with Poisson noise in the sinogram, where we used the incident x-ray containing 2×10^5 photons for 1/2 dose level, and 1×10^5 photons for 1/4 dose level, respectively. For each image, the sinogram was generated via 360 projection views uniformly spaced between 0 – 360 degrees. The CT images were resized to 416×416 before the simulation, thus resulting in the sinogram with the size of 641×640 .

E. Implementation Details

The DuDoUFNet is implemented using Pytorch¹, and it is trained in an end-to-end manner with a differential filtered back-projection layer proposed in [45]. The Adam solver [50] was used to optimize our network with the parameters $(\beta_1, \beta_2) = (0.5, 0.999)$ and a learning rate of $1e-4$. We trained 600 epochs with a batch size of 3 on an NVIDIA Quadro RTX 8000 GPU with 48GB memory.

F. Evaluation Strategies and Baselines

For quantitative evaluation, we measured the reconstruction performance using three evaluation metrics, including Structural Similarity Index (SSIM), Root Mean Square Error (RMSE), and Peak Signal-to-Noise Ratio (PSNR) which are computed using the predicted reconstructions and the ground-truth reconstructions in the non-metal regions. SSIM focuses on the evaluation of structural recovery, while RMSE with a unit of Hounsfield unit (HU) and PSNR with a unit of dB stress the evaluation of intensity profile recovery. For comparative evaluation, we first compared our results against the classic deep learning-based LDCT method, called REDCNN [15]. For a fair comparison, the REDCNN was trained with a modified image domain loss, as defined in the first term of Eq. 11, so that REDCNN focuses on image recovery in the non-metal regions. Second, we also compared against two of the previous state-of-the-art deep learning-based MAR methods, including DuDoNet++ [39] and DSPNet [40]. Lastly, since we consider simultaneous metal artifact reduction and low-dose CT denoising in this work, we also combined previous deep learning-based LDCT methods and deep learning-based MAR methods as an end-to-end network for comparisons. Specifically, we created four combinations of denoising +

¹ <http://pytorch.org/>

MAR or MAR + denoising methods, including 1) REDCNN followed by DSPNet, 2) DSPNet followed by REDCNN, 3) REDCNN followed by DuDoNet++, and 4) DuDoNet++ followed by REDCNN. The combined methods are trained jointly as an integrated model, with simultaneous optimization of the denoising loss functions and the MAR loss functions defined in the original methods during the training. For both quantitative and qualitative evaluations, we evaluated the performance under different dose levels and metal insertion settings. For the evaluation under different metal size settings, the metal sizes were calculated and ordered based on the number of pixel in the metal masks.

III. RESULTS

A. Experimental Results

The visual comparison between our DuDoUFNet and previous LDCT and MAR methods under different low-dose settings are shown in Figure 4. As we can see from the first example in the 1/2 dose experiment (1st row in Figure 4), the pelvic CT image from FBP reconstruction suffers from both non-local noise and metal artifacts. Structural distortions with a higher noise level can be observed in near-metal regions, especially near the iliac artery region. REDCNN and DuDoNet++ can reduce the global noise in the reconstructions, but they cannot fully correct the metal artifacts. DSPNet neither reduces the noise or the metal artifacts. As compared to previous methods, our DuDoUFNet can produce reconstruction with structure and intensity best matching with the ground-truth image. Similar observations can be found for the second example in the 1/2 dose experiment (2nd row in Figure 4). With large metallic implants, the abdominal CT image from FBP reconstruction is heavily degraded by the metal artifacts along with noise, where the kidney structures are diminished due to these factors. While it is difficult for previous methods to recover these soft-tissue structures, our DuDoUFNet is able to recover general kidney structures that best match the ground truth. In the 3rd and last rows, we can see the noise and metal artifacts are further amplified with 1/4 dose level. As compared to other methods, our DuDoUFNet still achieves the best reconstruction performance where we are able to preserve the fine details in both bone and soft-tissue regions. For example, in the 3rd row of Figure 4, the soft-tissue structure of the pectoralis major muscle and the bone structure of the sternum are heavily distorted by previous methods, while our DuDoUFNet can maintain these subtle structures. Similarly, in the last row of Figure 4, the structural details of the spine are challenging for previous methods to reconstruct, while our DuDoUFNet can recover the bone signal with a high fidelity. Moreover, the visual comparison between our DuDoUFNet and the combined methods of denoising + MAR or MAR + denoising are shown in Figure 5. As we can observe, simply combining previous deep learning-based LDCT methods and MAR methods still generates reconstruction with residual artifacts and does not lead to superior structure recovery as compared to our DuDoUFNet.

The quantitative evaluation of different methods under two low-dose settings is shown in Table I. In the 1/2 low-dose experiment, compared to the original FBP reconstructions, our DuDoUFNet can significantly improve the image quality with SSIM increased from 0.425 to 0.978 and RMSE decreased from 186.22 to 20.18, as well as outperforming the best previous method's performance of DuDoNet++ and combined methods. Similar

observations can be found in the 1/4 low-dose experiment, where we maintain the SSIM over 0.970 and RMSE lower than 23.00 even with a reduced dose level.

Furthermore, we evaluated our DuDoUFNet's performance when different sizes of metallic implants were presented. Figure 6 visualizes our DuDoUFNet's reconstructions when metallic implants vary from small to large under both 1/2 and 1/4 low-dose settings. Our method can consistently reconstruct noise-reduced and artifact-suppressed high-quality images under different metal insertion and dose conditions. Table II outlines the quantitative results of DuDoUFNet's performance under 1/2 dose setting when different sizes of metallic implants were introduced. In general, patients with large metal implants were harder to reconstruct as compared to small implants, thus resulting in a lower SSIM and a higher RMSE. However, our method consistently maintains the SSIM over 0.973 and RMSE lower than 26.16 across a range of metallic implants with varying sizes, outperforming previous methods in each metallic implant setting.

B. Ablative Studies

1. The impact of dual-domain learning: Our DuDoUFNet progressively restores signal in the sinogram domain, and then in the image domain. To study the effectiveness of progressive dual-domain learning, we also trained a network with image domain restoration only. Specifically, we used a network that has the same design as in the green box of Figure 2, but without the sinogram restoration input X_O . An example of visual comparison is shown in Figure 7, where we can observe more residual artifacts when relying solely on image domain restoration, as compared to our DuDoUFNet with dual-domain restoration. The corresponding quantitative comparison is summarized in Table III. Consistent with visual observations, dual-domain-based learning achieves superior SSIM and RMSE performances as compared to single-domain-based learning method, increasing SSIM from 0.966 to 0.978 and decreasing RMSE from 24.36 to 20.18.
2. The impact of the sub-network structure: The UFNet deployed in the DuDoUFNet consists of two sub-networks, including an UnNet and a FuNet. To investigate the impact of individual sub-network structures, we analyzed the reconstruction performance when either only UnNet or only FuNet is used in our dual-domain framework. A visual comparison is shown in Figure 8. As we can observe, more residual artifacts are presented in the brachiocephalic vein and sternum regions when only UnNet or FuNet is used in the dual-domain framework, as compared to our DuDoUFNet with both subnetworks integrated. The corresponding quantitative comparison is summarized in Table IV. Using only UnNet in the dual-domain framework provides slightly better reconstruction performance as compared to when using only FuNet in the dual-domain frameworks. However, using UFNet with both UnNet and FuNet integrated still yields superior performance as compared to when only UnNet is used, decreasing RMSE from 21.96 to 20.18 and increasing PSNR from 42.92 to 43.62.

3. The impact of the two-stage supervision: In our DuDoUFNet, the UFNet is trained with two-stage supervision in both sinogram and image domains. To validate the impact of the two-stage supervision applied on fully-complete restoration and under-complete restoration in both domains, we analyzed the reconstruction performance when different two-stage supervision settings are used. Table V outlines the reconstruction performance when loss supervision is used and not used on under-complete restoration outputs in the sinogram or the image domain of DuDoUFNet. As we can observe, when no two-stage supervision of UFNet is applied on both domains (1st row), it only achieves SSIM of 0.974 and RMSE of 22.56. Even though adding two-stage supervision in the image domain (2nd row) boosts the reconstruction performance, using two-stage supervision in the sinogram domain (3rd row) alone achieves even better performance since the image restoration relies on accurate sinogram restoration. Finally, using two-stage supervision in both domains, we are able to achieve the best reconstruction performance, with SSIM=0.978 and RMSE=20.18.
4. The influence of metal mask segmentation: DuDoUFNet uses both metal mask and metal trace as additional input to aid the dual-domain progressive restoration. An accurate metal mask (or metal trace equally) is important for the robust performance of our method. To study the impact of metal mask segmentation, we took different variants of the ground-truth metal mask as DuDoUFNet's input and analyzed the reconstruction performance in the non-metal region, i.e. regions outside of the ground-truth metal mask. Specifically, we created four different variants of metal masks using erosion and dilation operations by applying these operations 1 and 2 times (with a disk structure with a radius equal to 1), such that over-segmentation and under-segmentation scenarios were simulated. Figure 9 shows an example of DuDoUFNet's reconstruction by taking the original ground-truth mask (a), the dilation metal masks (b & c), and the erosion metal masks (e & f) as input. As we can observe, while slight segmentation errors lead to perturbed bone structure recovery, the noise and metal artifacts in all reconstructions of DuDoUFNet are suppressed, generating reasonable reconstruction results. The quantitative analysis is summarized in Figure 10. While inaccurate metal masks generated by dilation and erosion operations degrade the image recovery performance, the SSIM only decreases from 0.9779 to 0.9724 when an over-segmentation is presented. Similarly, the SSIM only reduces from 0.9779 to 0.9748 when an under-segmentation is presented. Under small metal segmentation errors, the DuDoUFNet can still maintain reasonable reconstruction with SSIM over 0.97.

IV. DISCUSSION

In this work, we developed a novel network, called DuDoUFNet, for simultaneous metal artifact reduction and low-dose CT reconstruction. Specifically, we proposed progressive restoration in terms of both domain and network design, enabling satisfactory reconstruction when the conditions of low-dose acquisitions and metallic implants co-exist. First of all,

inheriting from the previous MAR method in a full-dose CT using dual-domain learning [38]–[40], we proposed to use progressive dual-domain restoration, where we first restore sinogram signal and then restore image signal. We first perform sinogram restoration because the reconstruction image noise and metal artifacts are primarily caused by the degraded projection data in the sinogram domain. A strong network that can restore sinogram signal maintaining the fine details is also a key for providing a good initial reconstruction for image domain restoration. Therefore, secondly, we proposed a progressive under-to-fully-complete restoration network, called UFNet, for each domain's restoration. The UFNet combines two subnetworks, including FuNet containing no downsampling or upsampling operations and UnNet with encoding-decoding operations. The UnNet with downsampling and upsampling can generate restoration features at multiple scales, while FuNet operates on original resolution can generate restoration features focusing on the level of the original resolution. Then, combining the UnNet and FuNet via cross-stage connections enables us to generate spatial-enriched features for the final restoration. As we can observe from Figure 4 and Table I, as compared to DuDoNet++ relying on UNet [32] for sinogram restoration, our DuDoUFNet based on UFNet with fine detail restoration in both domains achieves superior reconstruction performance.

From our experimental results, we demonstrated the feasibility of using our DuDoUFNet for simultaneous metal artifact reduction and low-dose CT reconstruction under different acquisition settings and a variety of metallic implants. First of all, as we can see from Figure 4 and Table I, our method can consistently outperform previous LDCT and MAR methods, including REDCNN, DuDoNet++, and DSPNet, where we are able to maintain SSIM over 0.974 and RMSE under 23.00, under different dose settings. Please note while DSPNet is considered the state-of-the-art method for full-dose MAR, it reconstructs non-ideal results (Table I and Figure 4) where it neither reduces the noise or the metal artifact. It is mainly due to the non-metal sinogram replacement operation in the last part of DSPNet's design, where they assume the non-metal sinogram is noiseless and should be kept, thus not suitable for MARLD. Moreover, while combining the previous deep learning-based MAR and LDCT methods also achieved reasonable reconstruction performance (Figure 5 and Table I), simply concatenating these methods together with complex models created, i.e. a large number of model parameters, does not necessary yielded better performances. For example, under the 1/2 low-dose setting (Table I), REDCNN→DuDoNet++ achieves PSNR=42.25 and DuDoNet++→REDCNN achieves PSNR=42.46, which are both better than REDCNN alone with PSNR=41.37. It implies that combining the previous denoising method and MAR method indeed helps the MARLD task when compared to the denoising method alone. However, both REDCNN→DuDoNet++ and DuDoNet++→REDCNN generate an even worse reconstruction performance than the DuDoNet++ alone. According to this observation, we believe using a simple image domain-based denoising method, such as REDCNN, may not be sufficient for a complex image restoration task, i.e. MARLD. Adding the MAR method, such as DuDoNet++, to the denoising method increase the model complexity, as well as enables the learning in the dual domain, thus providing better performance. However, the combined models with increased model complexity are also prone to overfitting which could degrade the reconstruction performance. Thus, DuDoNet++ alone with slightly lower model complexity and with learning in dual-domain could

potentially generate even better performance than the combined methods. On the other hand, our DuDoUFNet with a slightly reduced model complexity, as compared to the combined methods, still provided significantly better reconstruction results. Second, as we can observe from Table II and Table I, our DuDoUFNet's reconstruction performance is gradually degraded either when increasing the size of the metallic implant or reducing the x-ray dose level. For instance, the RMSE value increases from 16.80 to 26.16 when increasing the size of the metal implants (Table II). Similarly, the RMSE value increases from 20.18 to 22.96 when lowering the dose level from 1/2 to 1/4 (Table I). Even though our DuDoUFNet achieves the lowest RMSE values among the previous methods across different sizes of metal implants and different low-dose settings, we think a proper level of x-ray dose needs to be considered to balance the trade-off between image quality, radiation exposure, and metal size. For example, a higher dose level or even near-full-dose level should be considered when a large metal implant is known in a patient.

Our current work also has potential limitations. First, our DuDoUFNet consists of 4 subnetworks with 2 subnetworks in the sinogram domain and 2 subnetworks in the image domain, thus the computation complexity and the number of parameters increase nearly 4 times as compared to when using only one subnetwork for restoration. Even though the computation complexity and the number of network parameters are higher than previous stand-alone MAR and LDCT methods, the average inference time of our DuDoUFNet is < 0.5s, which still achieves reasonable speed for high-quality reconstruction, and the model complexity is still slightly lower than the combined method of MAR+LDCT with a better performance. Second, because it is infeasible to collect real metal-inserted low-dose CT and metal-free CT data for training and there is no public real projection data, we used a realistic CT simulation to produce synthesized training pairs from clinical metal-free full-dose CT images [49], which is a common data preparation method utilized in full-dose CT MAR [38]–[40] and LDCT [15]. However, the quality of the simulated data may affect the reconstruction performance on real data, due to factors such as limited variability of metal sizes/shapes/locations, varied low-dose x-ray spectrum, and inconsistent x-ray exposure settings, causing domain shifts. In the future, we will investigate the DuDoUFNet's performance and generalizability on real low-dose CT patients with metallic implants, as well as explore how to build a better simulation dataset to further improve the reconstruction performance on real patient data. Finally, our DuDoUFNet, similar to the previous works of DuDoNet++ [39] and DSPNet [40], requires the metal mask as one of the network inputs. In our ablation studies, we show that the DuDoUFNet can provide reasonable reconstruction under slight metal segmentation errors (Figure 9 and Figure 10). In real applications, since the HU values in the metal region often consist of high values, we believe it is feasible to use semi-automatic methods, such as thresholding with manual adjustments, to obtain a reasonably accurate metal mask. On the other hand, fully automatic metal segmentation methods could also be applied to assist the generation of metal masks [51], [52].

Our work also suggests a few interesting directions for future studies. First, the radiation dose in CT is commonly reduced by either decreasing the projection x-ray flux or decreasing the number of projections. In this work, we only investigated the simultaneous metal artifact reduction and low-dose CT using full-view sinogram with reduced projection x-ray flux. While there were previous works on simultaneous metal artifact reduction and sparse-view

CT reconstruction [53], combining the low-dose/low-current CT with sparse view settings could further reduce the radiation dose. Thus, investigation on how to simultaneously reduce metal artifacts and reconstruct low-current sparse-view CT would be an interesting direction. Second, CT is commonly acquired along with PET and SPECT for the purposes of attenuation correction and medical diagnosis. Radiation dose and artifact reduction methods have been extensively investigated for stand-alone modalities, while multi-modal low-dose reconstructions, such as low-dose PET-CT and low-dose SPECT-CT, remains under-explored. In fact, information from multi-modalities under low-dose settings could be mutually beneficial. Since our DuDoUFNet could be applied to different tomographic imaging modalities, investigations on how to extend our DuDoUFNet into a joint multi-modal low-dose reconstruction framework would also be an interesting direction. For example, FDG PET/CT scans for device-induced infection usually suffers from metal artifact-induced attenuation correction errors, which could be addressed by extending this work [54]. Finally, our DuDoUFNet is an open framework with flexibility in each subnetwork. Even though we used RSEB as our basic block in the UnNet and FuNet, we do not claim the optimality of the subnetwork design. Other feature extraction blocks, such as Swin Transformer block [55], could be adapted in our DuDoUFNet which may produce even better performance. As a matter of fact, image restoration networks based on the Swin transformer block [55] with U-shape design [56] and with no downsampling operation [57] have been proposed. Using them as our UnNet and FuNet may produce improved performance and could also be an interesting topic to investigate. Furthermore, adding other advanced fine detail recovery loss functions, such as SSIM loss [58] and gradient loss [59], may also help further improve the DuDoUFNet's reconstruction performance.

V. CONCLUSION

We proposed a dual-domain under-to-fully-complete progressive restoration network, called DuDoUFNet, for simultaneous metal artifact reduction and low-dose CT reconstruction. Our DuDoUFNet aims to reconstruct a noise- and artifact-reduced CT image by progressive sinogram to image domain restoration with a two-stage progressive restoration network. Our experimental results demonstrate that our method can provide high-quality reconstruction, superior to previous LDCT and MAR methods under a wide range of settings.

ACKNOWLEDGMENT

This work was supported by funding from the National Institutes of Health (NIH) under grant number R01EB025468.

REFERENCES

- [1]. McCollough CH, Primak AN, Braun N, Kofler J, Yu L, and Christner J, "Strategies for reducing radiation dose in ct," *Radiologic Clinics*, vol. 47, no. 1, pp. 27–40, 2009. [PubMed: 19195532]
- [2]. Kubo T, Lin P-JP, Stiller W, Takahashi M, Kauczor H-U, Ohno Y, and Hatabu H, "Radiation dose reduction in chest ct: a review," *American journal of roentgenology*, vol. 190, no. 2, pp. 335–343, 2008. [PubMed: 18212218]
- [3]. Boas FE, Fleischmann D et al. , "Ct artifacts: causes and reduction techniques," *Imaging Med*, vol. 4, no. 2, pp. 229–240, 2012.

- [4]. Goldman LW, "Principles of ct: radiation dose and image quality," *Journal of nuclear medicine technology*, vol. 35, no. 4, pp. 213–225, 2007. [PubMed: 18006597]
- [5]. Son SH, Kang YN, and Ryu M-R, "The effect of metallic implants on radiation therapy in spinal tumor patients with metallic spinal implants," *Medical Dosimetry*, vol. 37, no. 1, pp. 98–107, 2012. [PubMed: 21925861]
- [6]. Roth TD, Maertz NA, Parr JA, Buckwalter KA, and Choplin RH, "Ct of the hip prosthesis: appearance of components, fixation, and complications," *Radiographics*, vol. 32, no. 4, pp. 1089–1107, 2012. [PubMed: 22786996]
- [7]. Verburg JM and Seco J, "Ct metal artifact reduction method correcting for beam hardening and missing projections," *Physics in Medicine & Biology*, vol. 57, no. 9, p. 2803, 2012. [PubMed: 22510753]
- [8]. Giantsoudi D, De Man B, Verburg J, Trofimov A, Jin Y, Wang G, Gjestebj L, and Paganetti H, "Metal artifacts in computed tomography for radiation therapy planning: dosimetric effects and impact of metal artifact reduction," *Physics in Medicine & Biology*, vol. 62, no. 8, p. R49, 2017. [PubMed: 28323641]
- [9]. Konishi T, Shibutani T, Okuda K, Yoneyama H, Moribe R, Onoguchi M, Nakajima K, and Kinuya S, "Metal artifact reduction for improving quantitative spect/ct imaging," *Annals of Nuclear Medicine*, vol. 35, no. 3, pp. 291–298, 2021. [PubMed: 33460009]
- [10]. Fang R, Zhang S, Chen T, and Sanelli PC, "Robust low-dose ct perfusion deconvolution via tensor total-variation regularization," *IEEE transactions on medical imaging*, vol. 34, no. 7, pp. 1533–1548, 2015. [PubMed: 25706579]
- [11]. Green M, Marom EM, Kiryati N, Konen E, and Mayer A, "Efficient low-dose ct denoising by locally-consistent non-local means (lc-nlm)," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2016, pp. 423–431.
- [12]. Chen Y, Shi L, Feng Q, Yang J, Shu H, Luo L, Coatrieux J-L, and Chen W, "Artifact suppressed dictionary learning for low-dose ct image processing," *IEEE transactions on medical imaging*, vol. 33, no. 12, pp. 2271–2292, 2014. [PubMed: 25029378]
- [13]. Kulathilake KSH, Abdullah NA, Sabri AQM, and Lai KW, "A review on deep learning approaches for low-dose computed tomography restoration," *Complex & Intelligent Systems*, pp. 1–33, 2021.
- [14]. Chen H, Zhang Y, Zhang W, Liao P, Li K, Zhou J, and Wang G, "Low-dose ct denoising with convolutional neural network," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017, pp. 143–146.
- [15]. Chen H, Zhang Y, Kalra MK, 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*, vol. 36, no. 12, pp. 2524–2535, 2017. [PubMed: 28622671]
- [16]. Kang E, Chang W, Yoo J, and Ye JC, "Deep convolutional framelet denoising for low-dose ct via wavelet residual network," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1358–1369, 2018. [PubMed: 29870365]
- [17]. Yi X and Babyn P, "Sharpness-aware low-dose ct denoising using conditional generative adversarial network," *Journal of digital imaging*, vol. 31, no. 5, pp. 655–669, 2018. [PubMed: 29464432]
- [18]. Isola P, Zhu J-Y, Zhou T, and Efros AA, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.
- [19]. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, Kalra MK, 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*, vol. 37, no. 6, pp. 1348–1357, 2018. [PubMed: 29870364]
- [20]. Johnson J, Alahi A, and Fei-Fei L, "Perceptual losses for real-time style transfer and super-resolution," in *European conference on computer vision*. Springer, 2016, pp. 694–711.
- [21]. Yin X, Zhao Q, Liu J, Yang W, Yang J, Quan G, Chen Y, Shu H, Luo L, and Coatrieux J-L, "Domain progressive 3d residual convolution network to improve low-dose ct imaging," *IEEE transactions on medical imaging*, vol. 38, no. 12, pp. 2903–2913, 2019. [PubMed: 31107644]

- [22]. Park HS, Hwang D, and Seo JK, "Metal artifact reduction for polychromatic x-ray ct based on a beam-hardening corrector," *IEEE transactions on medical imaging*, vol. 35, no. 2, pp. 480–487, 2015. [PubMed: 26390451]
- [23]. Kalender WA, Hebel R, and Ebersberger J, "Reduction of ct artifacts caused by metallic implants." *Radiology*, vol. 164, no. 2, pp. 576–577, 1987. [PubMed: 3602406]
- [24]. Meyer E, Raupach R, Lell M, Schmidt B, and Kachelrieß M, "Normalized metal artifact reduction (nmar) in computed tomography," *Medical physics*, vol. 37, no. 10, pp. 5482–5493, 2010. [PubMed: 21089784]
- [25]. Zhang Y, Pu Y-F, Hu J-R, Liu Y, and Zhou J-L, "A new ct metal artifacts reduction algorithm based on fractional-order sinogram inpainting." *Journal of X-ray science and technology*, vol. 19, no. 3, pp. 373–384, 2011. [PubMed: 21876286]
- [26]. Mehranian A, Ay MR, Rahmim A, and Zaidi H, "X-ray ct metal artifact reduction using wavelet domain $L_{\{0\}}$ sparse regularization," *IEEE transactions on medical imaging*, vol. 32, no. 9, pp. 1707–1722, 2013. [PubMed: 23744669]
- [27]. Zhang Y, Yan H, Jia X, Yang J, Jiang SB, and Mou X, "A hybrid metal artifact reduction algorithm for x-ray ct," *Medical physics*, vol. 40, no. 4, p. 041910, 2013. [PubMed: 23556904]
- [28]. Bal M and Spies L, "Metal artifact reduction in ct using tissue-class modeling and adaptive prefiltering," *Medical physics*, vol. 33, no. 8, pp. 2852–2859, 2006. [PubMed: 16964861]
- [29]. Prell D, Kyriakou Y, Beister M, and Kalender WA, "A novel forward projection-based metal artifact reduction method for flat-detector computed tomography," *Physics in Medicine & Biology*, vol. 54, no. 21, p. 6575, 2009. [PubMed: 19826202]
- [30]. Wang J, Wang S, Chen Y, Wu J, Coatrieux J-L, and Luo L, "Metal artifact reduction in ct using fusion based prior image," *Medical physics*, vol. 40, no. 8, p. 081903, 2013. [PubMed: 23927317]
- [31]. Park HS, Lee SM, Kim HP, Seo JK, and Chung YE, "Ct sinogram-consistency learning for metal-induced beam hardening correction," *Medical physics*, vol. 45, no. 12, pp. 5376–5384, 2018. [PubMed: 30238586]
- [32]. Ronneberger O, Fischer P, and Brox T, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [33]. Gjestebj L, Yang Q, Xi Y, Zhou Y, Zhang J, and Wang G, "Deep learning methods to guide ct image reconstruction and reduce metal artifacts," in *Medical Imaging 2017: Physics of Medical Imaging*, vol. 10132. International Society for Optics and Photonics, 2017, p. 101322W.
- [34]. Zhang Y and Yu H, "Convolutional neural network based metal artifact reduction in x-ray computed tomography," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1370–1381, 2018. [PubMed: 29870366]
- [35]. Huang X, Wang J, Tang F, Zhong T, and Zhang Y, "Metal artifact reduction on cervical ct images by deep residual learning," *Biomedical engineering online*, vol. 17, no. 1, pp. 1–15, 2018. [PubMed: 29310661]
- [36]. Wang J, Zhao Y, Noble JH, and Dawant BM, "Conditional generative adversarial networks for metal artifact reduction in ct images of the ear," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 3–11.
- [37]. Gjestebj L, Shan H, Yang Q, Xi Y, Claus B, Jin Y, De Man B, and Wang G, "Deep neural network for ct metal artifact reduction with a perceptual loss function," in *In Proceedings of The Fifth International Conference on Image Formation in X-ray Computed Tomography*, 2018.
- [38]. Lin W-A, Liao H, Peng C, Sun X, Zhang J, Luo J, Chellappa R, and Zhou SK, "Dudonet: Dual domain network for ct metal artifact reduction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 10512–10521.
- [39]. Lyu Y, Lin W-A, Liao H, Lu J, and Zhou SK, "Encoding metal mask projection for metal artifact reduction in computed tomography," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2020, pp. 147–157.
- [40]. Yu L, Zhang Z, Li X, and Xing L, "Deep sinogram completion with image prior for metal artifact reduction in ct images," *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 228–238, 2020. [PubMed: 32956044]

- [41]. Yu L, Zhang Z, Li X, Ren H, Zhao W, and Xing L, "Metal artifact reduction in 2d ct images with self-supervised cross-domain learning," *Physics in Medicine & Biology*, vol. 66, no. 17, p. 175003, 2021.
- [42]. Zhang Y, Tian Y, Kong Y, Zhong B, and Fu Y, "Residual dense network for image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.
- [43]. Zhang Y, Li K, Li K, Wang L, Zhong B, and Fu Y, "Image super-resolution using very deep residual channel attention networks," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 286–301.
- [44]. Zamir SW, Arora A, Khan S, Hayat M, Khan FS, Yang M-H, and Shao L, "Multi-stage progressive image restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 14821–14831.
- [45]. Zhou B, Zhou SK, Duncan JS, and Liu C, "Limited view tomographic reconstruction using a cascaded residual dense spatial-channel attention network with projection data fidelity layer," *IEEE transactions on medical imaging*, vol. 40, no. 7, pp. 1792–1804, 2021. [PubMed: 33729929]
- [46]. Zhou B and Zhou SK, "Dudornet: Learning a dual-domain recurrent network for fast mri reconstruction with deep t1 prior," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2020, pp. 298–313.
- [47]. Zhou B, Schlemper J, Dey N, Salehi SSM, Liu C, Duncan JS, and Sofka M, "Dsformer: A dual-domain self-supervised transformer for accelerated multi-contrast mri reconstruction," *arXiv preprint arXiv:2201.10776*, 2022.
- [48]. Hu J, Shen L, and Sun G, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.
- [49]. Yan K, Wang X, Lu L, and Summers RM, "Deeplesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning," *Journal of Medical Imaging*, vol. 5, no. 3, p. 036501, 2018. [PubMed: 30035154]
- [50]. Kingma DP and Ba J, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [51]. Hegazy MA, Cho MH, Cho MH, and Lee SY, "U-net based metal segmentation on projection domain for metal artifact reduction in dental ct," *Biomedical engineering letters*, vol. 9, no. 3, pp. 375–385, 2019. [PubMed: 31456897]
- [52]. Gottschalk TM, Maier A, Kordon F, and Kreher BW, "View-consistent metal segmentation in the projection domain for metal artifact reduction in cbct—an investigation of potential improvement," *arXiv preprint arXiv:2112.02101*, 2021.
- [53]. Zhou B, Chen X, Zhou SK, Duncan JS, and Liu C, "Dudodr-net: Dual-domain data consistent recurrent network for simultaneous sparse view and metal artifact reduction in computed tomography," *Medical Image Analysis*, p. 102289, 2021. [PubMed: 34758443]
- [54]. Ahmed FZ, James J, Tout D, Arumugam P, Mamas M, and Zaidi AM, "Metal artefact reduction algorithms prevent false positive results when assessing patients for cardiac implantable electronic device infection," *Journal of Nuclear Cardiology*, vol. 22, no. 1, pp. 219–220, 2015. [PubMed: 25112539]
- [55]. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, and Guo B, "Swin transformer: Hierarchical vision transformer using shifted windows," *arXiv preprint arXiv:2103.14030*, 2021.
- [56]. Wang Z, Cun X, Bao J, and Liu J, "Uformer: A general u-shaped transformer for image restoration," *arXiv preprint arXiv:2106.03106*, 2021.
- [57]. Liang J, Cao J, Sun G, Zhang K, Van Gool L, and Timofte R, "Swinir: Image restoration using swin transformer," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 1833–1844.
- [58]. Zhao H, Gallo O, Frosio I, and Kautz J, "Loss functions for image restoration with neural networks," *IEEE Transactions on computational imaging*, vol. 3, no. 1, pp. 47–57, 2016.
- [59]. Zhou B, Lin X, Eck B, Hou J, and Wilson D, "Generation of virtual dual energy images from standard single-shot radiographs using multi-scale and conditional adversarial network," *arXiv preprint arXiv:1810.09354*, 2018.

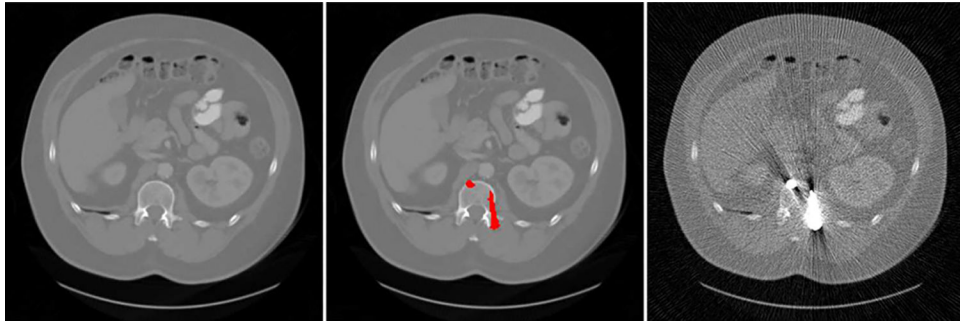


Fig. 1.
An example of CT image with metallic implants under low-dose acquisition condition. Left: metal-free full-dose CT image. Middle: full-dose CT image overlaid with metal segmentation (red mask) for simulation. Right: low-dose CT image with metallic implants (1/4 dose). The display window is $[-1000\ 1000]$ HU.

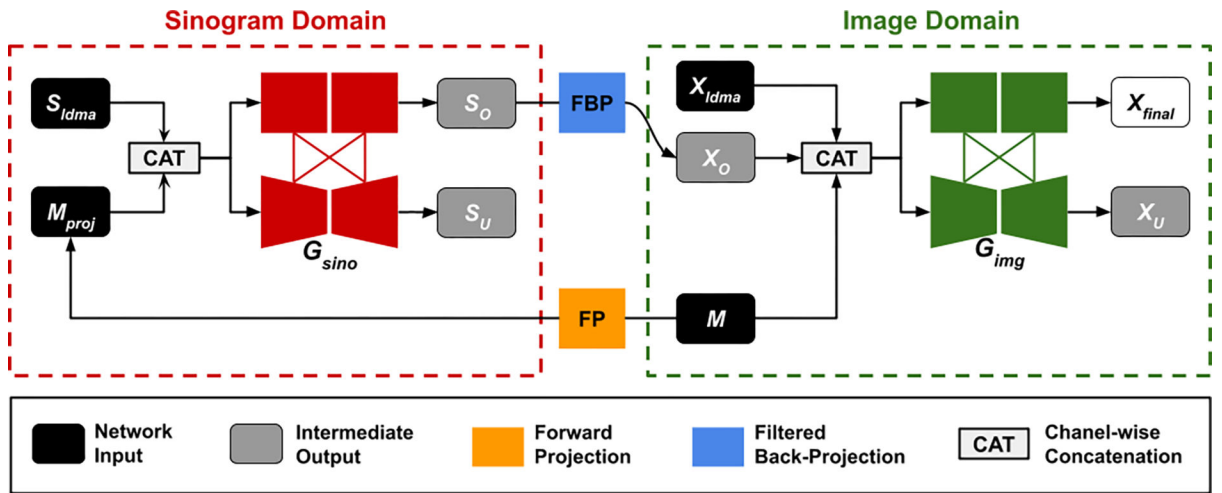


Fig. 2.

Framework of our dual-domain under-to-fully-complete progressive restoration network (DuDoUFNet) for simultaneous metal artifact reduction and low-dose CT reconstruction. Our DuDoUFNet consists of a sinogram progressive restoration network G_{sino} and an image progressive restoration network G_{img} . Given an input low-dose sinogram with metallic implants S_{ldma} , it is concatenated with metal trace M_{proj} and goes through G_{sino} to progressively restore signal in sinogram domain. The sinogram restoration output S_O is then converted to image X_O via the FBP layer, and concatenated with the low-dose metal artifact image X_{ldma} and metal mask M for progressive restoration in image domain via G_{img} .

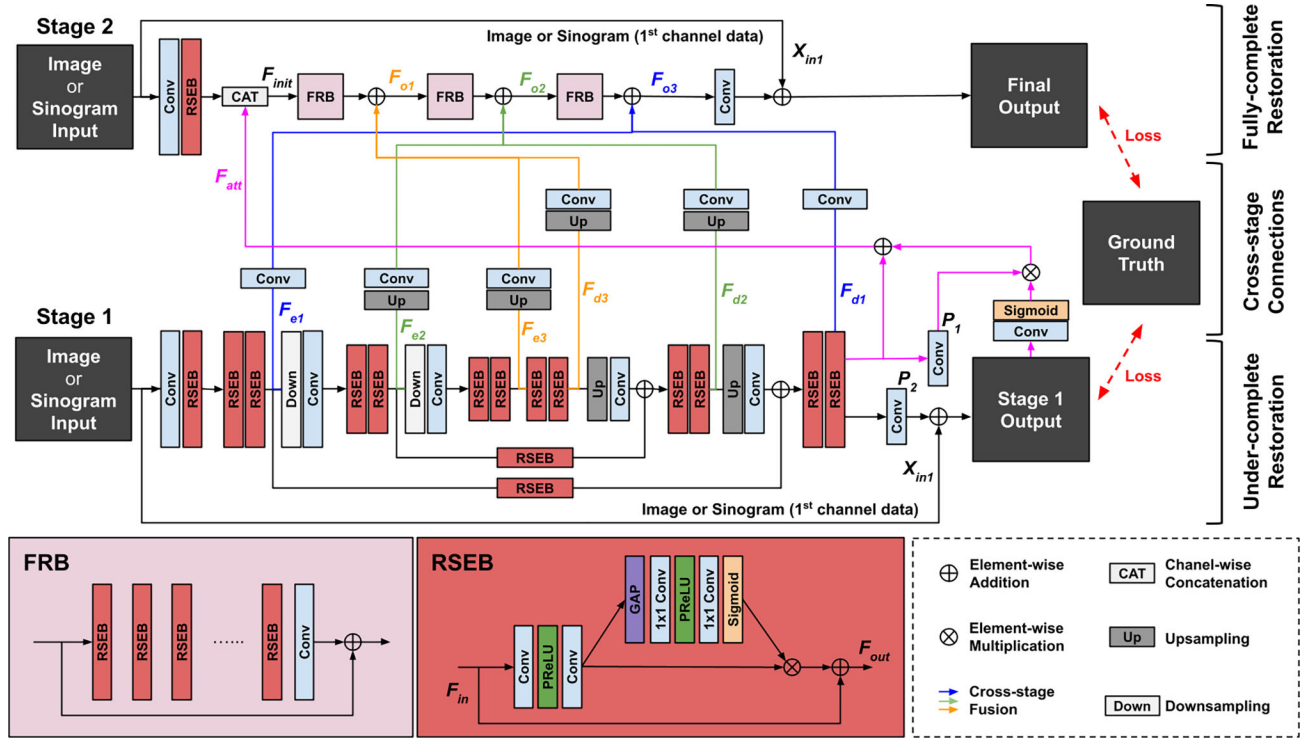


Fig. 3. The network architecture of under-to-fully-complete progressive restoration network (UFNet) used for the sinogram restoration network G_{sino} and the image restoration network G_{img} in DuDoUFNet (Figure 2). Our UFNet consists of a under-complete network (stage 1) and an fully-complete network (stage 2) with cross stage feature fusions (blue, yellow, and green lines). Supervision is applied on both stages' outputs.

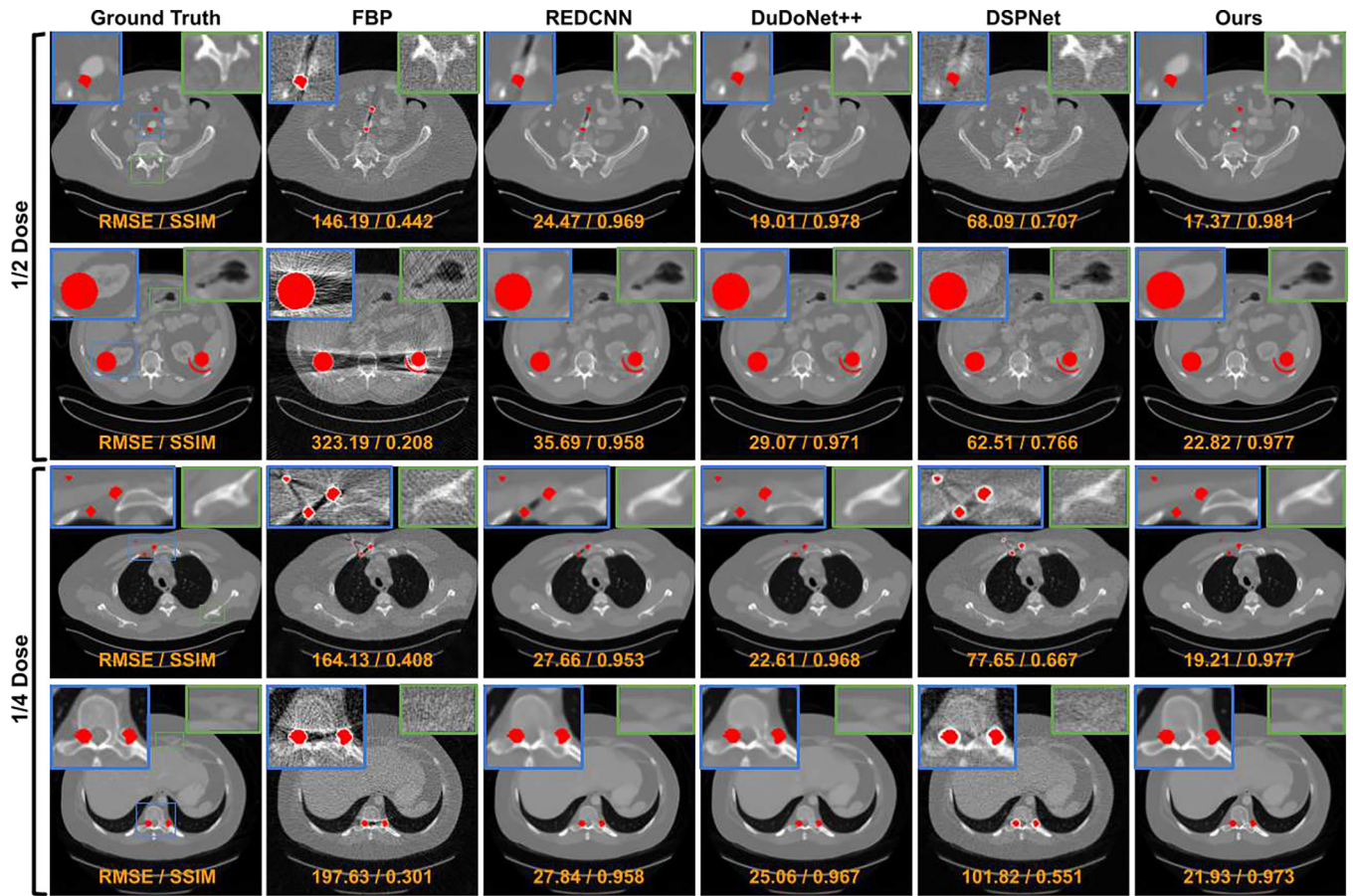


Fig. 4.

Visual comparison of low-dose reconstructions with metallic implants under 1/2 and 1/4 dose levels. The metal regions are overlaid with the red masks. The zoom-in regions (blue boxes for metal-proximal regions and green boxes for metal-distant regions) are annotated on the ground-truth images. RMSE and SSIM values are computed for individual images (orange). The display window is $[-1000\ 1000]$ HU.

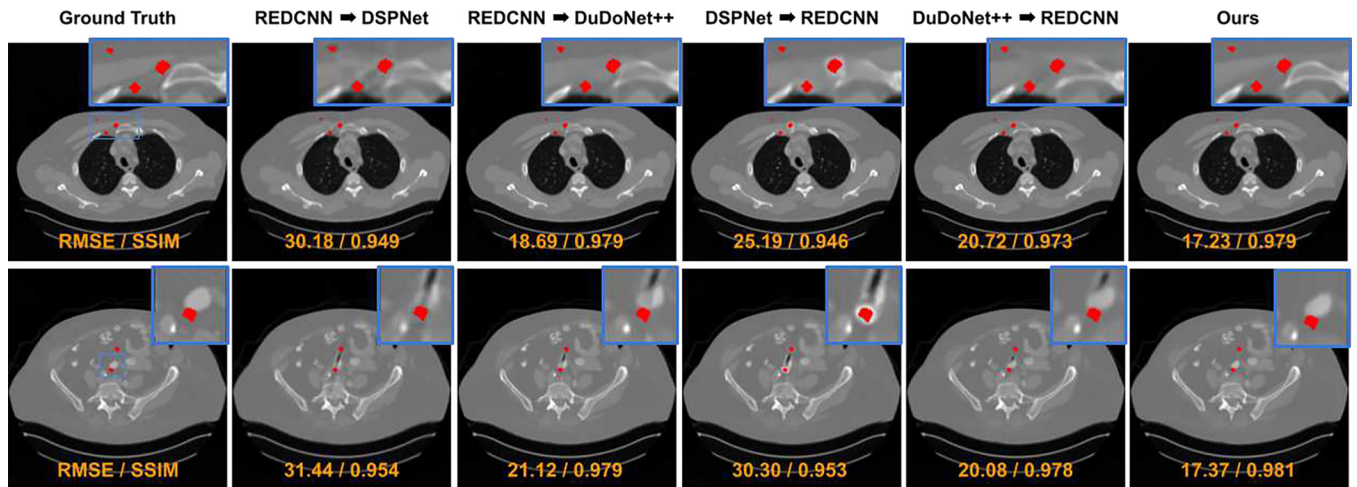


Fig. 5. Visual comparison of low-dose reconstructions with metallic implants under 1/2 dose level. The metal regions are overlaid with the red masks. The zoom-in regions (blue boxes) are annotated on the ground-truth images. RMSE and SSIM values are computed for individual images (orange). The display window is $[-1000 \ 1000]$ HU.

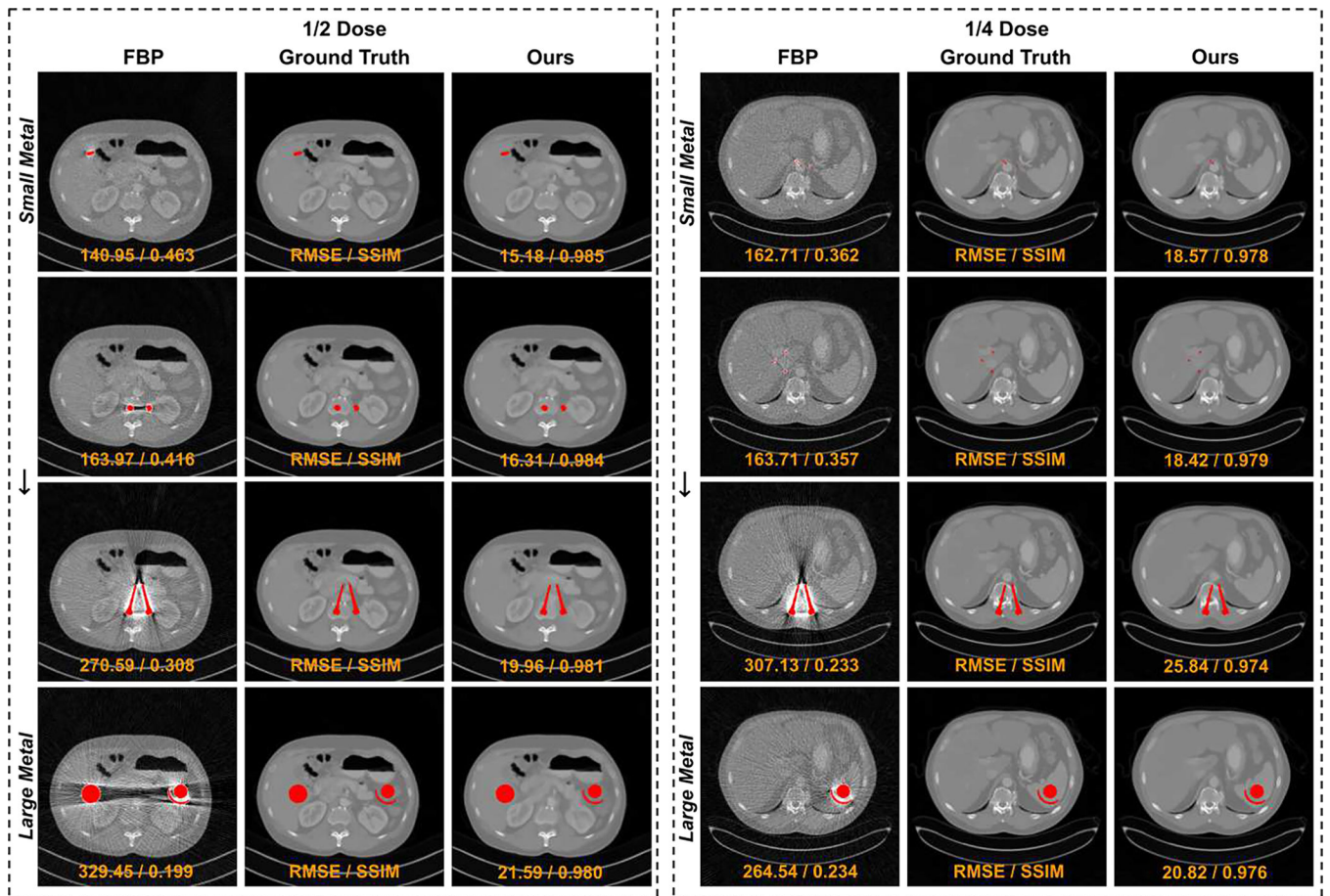


Fig. 6. Visual comparison of low-dose reconstructions with small to large metallic implants (top to bottom) under 1/2 and 1/4 low-dose settings. The metal regions are overlaid with the red masks. RMSE and SSIM values are computed for individual images (orange). The display window is $[-1000 \ 1000]$ HU.

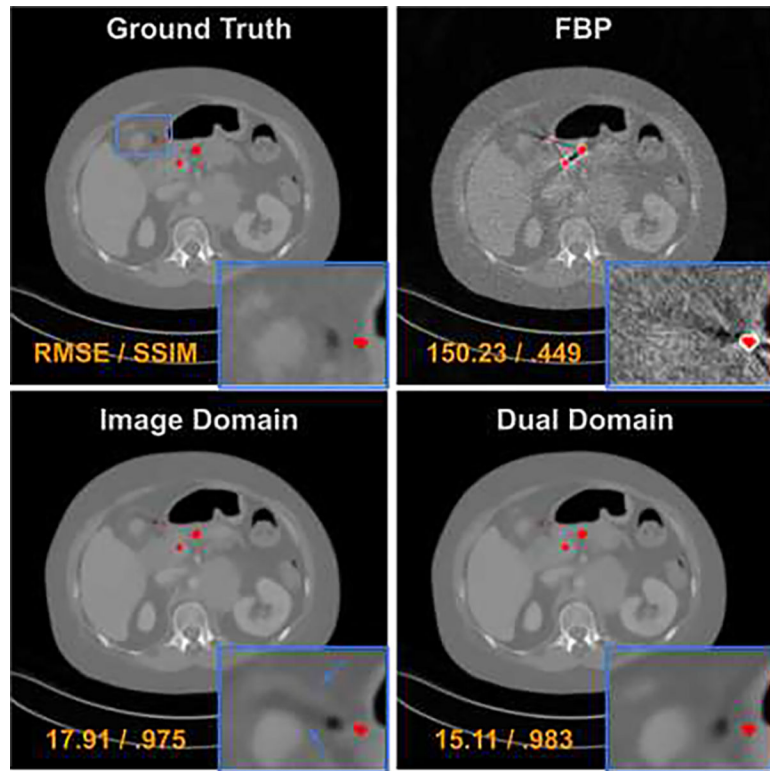


Fig. 7. Visual comparison of reconstructions under 1/2 low-dose setting. The bottom row compares the reconstructions from the image-domain-only network against our DuDoUFNet. The display window is $[-1000 \ 1000]$ HU.

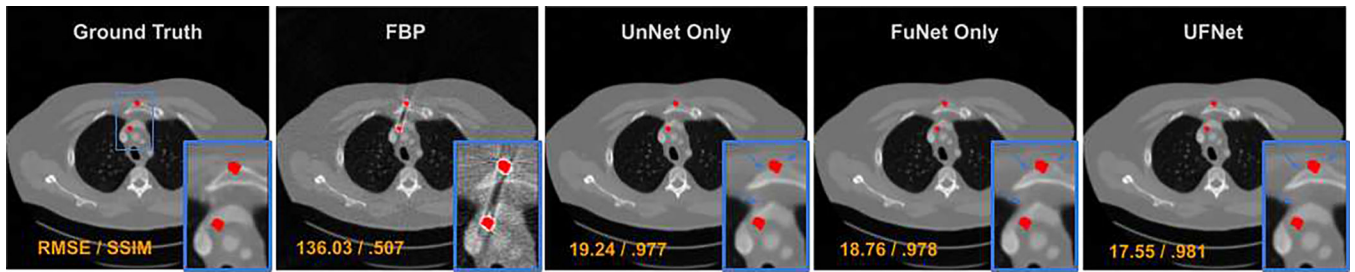


Fig. 8. Visual comparison of reconstructions under 1/2 low-dose setting when different sub-network is used in the DuDoUFNet. Our dual domain framework with UnNet only and FuNet only are compared against DuDoUFNet with UFNet. The display window is $[-1000\ 1000]$ HU.

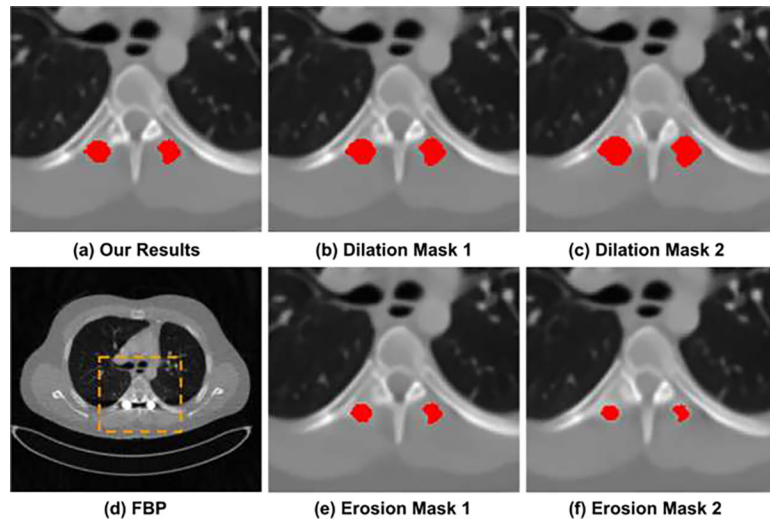


Fig. 9. Visual comparison of reconstructions from DuDoUFNet with different segmented metal masks as input (1/2 low-dose setting). The reconstruction with ground-truth segmentation (a) is compared against the reconstruction with dilated metal mask (b & c) and the reconstruction with eroded metal masks (e & f). The display window is $[-1000 \ 1000]$ HU.

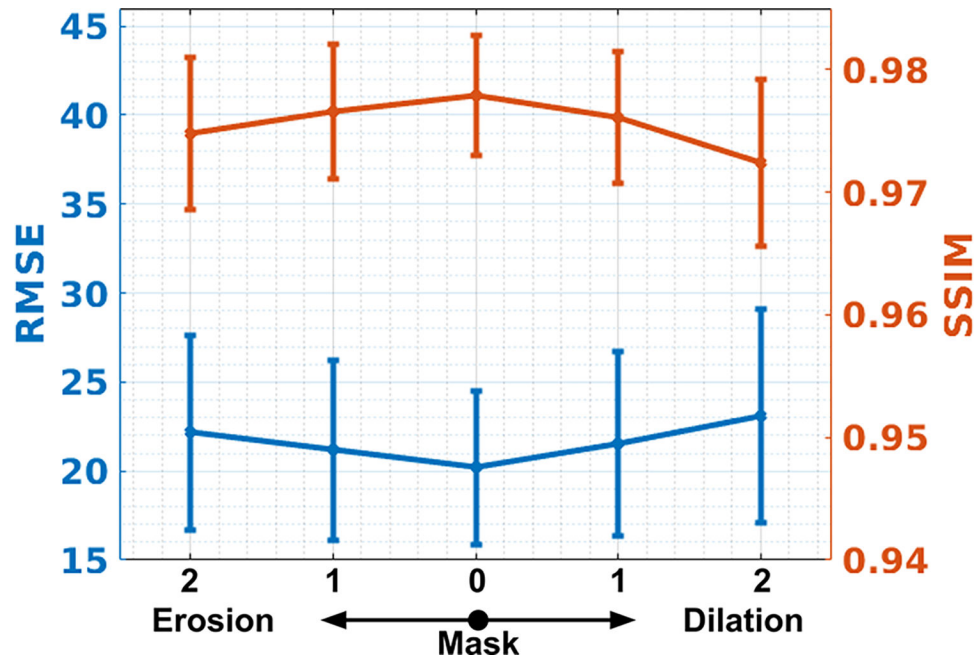


Fig. 10. Quantitative analysis of the impact of metal segmentation mask in DuDoUFNet. The number of metal mask erosion and dilation iteration is annotated on the x-axis. 1/2 low-dose setting is considered here.

TABLE I

Quantitative comparison of reconstructions under two different low-dose settings using SSIM, RMSE, and PSNR. Best results are marked in **bold**. Last column shows the number of trainable parameters of each methods.

Evaluation	1/2 Dose			1/4 Dose			Num of Param
	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	
FBP	.425 ± .098	186.22 ± 63.50	24.59 ± 2.72	.327 ± .083	214.89 ± 69.03	23.29 ± 2.55	-
REDCNN [15]	.964 ± .008	26.47 ± 7.07	41.37 ± 2.18	.956 ± .009	30.76 ± 7.14	39.99 ± 1.89	1.8M
DuDoNet++ [39]	.973 ± .006	22.73 ± 5.39	42.63 ± 1.94	.967 ± .007	26.62 ± 5.84	41.22 ± 1.80	3.6M
DSPNet [40]	.761 ± .043	61.18 ± 6.83	33.86 ± 0.95	.623 ± .055	90.17 ± 10.44	30.49 ± 0.99	3.6M
REDCNN [15] → DSPNet [40]	.948 ± .011	34.71 ± 7.39	38.91 ± 1.74	.931 ± .018	39.45 ± 8.01	37.78 ± 1.65	5.4M
REDCNN [15] → DuDoNet++ [39]	.975 ± .006	23.97 ± 6.70	42.25 ± 2.25	.969 ± .006	27.31 ± 7.42	41.10 ± 2.20	5.4M
DSPNet [40] → REDCNN [15]	.952 ± .006	28.14 ± 6.39	40.76 ± 1.85	.958 ± .008	29.95 ± 6.45	40.20 ± 1.78	5.4M
DuDoNet++ [39] → REDCNN [15]	.973 ± .006	23.08 ± 4.99	42.46 ± 1.78	.968 ± .006	26.32 ± 5.50	41.30 ± 1.71	5.4M
Ours	.978 ± .004 [‡]	20.18 ± 4.30 [‡]	43.62 ± 1.76 [‡]	.974 ± .005 [‡]	22.96 ± 4.77 [‡]	42.49 ± 1.72 [‡]	5.1M

[‡] means the difference between our DuDoUFNet and all the baseline methods are significant at $p < 0.05$.

Quantitative comparison of reconstructions with small to large metallic implants (right to left) under the 1/2 low-dose setting. Best results are marked in **bold**.

TABLE II

SSIM/RMSE	Large Metal ↔ Small Metal						Average
	.338/270.03	.408/203.55	.486/139.06	.505/137.49	.515/129.03	.425/186.22	
FBP	.338/270.03	.408/203.55	.486/139.06	.505/137.49	.515/129.03	.425/186.22	
REDCNN [15]	.961/30.65	.959/38.65	.966/23.08	.968/20.96	.968/20.55	.964/26.47	
DuDoNet++ [39]	.969/25.41	.969/27.92	.975/19.35	.976/18.62	.976/18.41	.973/22.73	
DSPNet [40]	.778/60.40	.765/66.79	.754/60.43	.748/61.17	.752/60.33	.761/61.18	
REDCNN [15] → DSPNet [40]	.939/40.70	.943/44.18	.953/30.84	.954/29.26	.956/28.87	.948/34.71	
REDCNN [15] → DuDoNet++ [39]	.971/27.49	.971/36.21	.978/20.35	.978/18.98	.979/18.73	.975/23.97	
DSPNet [40] → REDCNN [15]	.949/31.73	.949/38.50	.952/26.31	.954/22.53	.954/24.22	.952/28.14	
DuDoNet++ [39] → REDCNN [15]	.970/25.91	.969/28.94	.975/20.16	.976/19.31	.976/19.25	.973/23.08	
Ours	.974/23.68	.973/26.16	.980/17.58	.980/16.94	.981/16.80	.978/20.18	

TABLE III

Quantitative analysis of the dual-domain learning used in our DuDoUFNet. Image vs dual-domain-based learning is evaluated under 1/2 low-dose condition.

Setting	SSIM	RMSE	PSNR
Image-Domain Only	.966 ± .007	24.36 ± 5.82	42.03 ± 1.95
Dual-Domain	.978 ± .004	20.18 ± 4.30	43.62 ± 1.76

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE IV

Quantitative analysis of the network structure used in our DuDoUFNet. DuDoUFNet with unnet vs funet vs ufnet is evaluated under 1/2 low-dose condition.

Setting in DuDoUFNet	SSIM	RMSE	PSNR
UnNet	.975 ± .006	21.96 ± 5.09	42.92 ± 1.92
FuNet	.974 ± .005	22.19 ± 5.67	42.87 ± 2.08
UFNet	.978 ± .004	20.18 ± 4.30	43.62 ± 1.76

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE V

Ablation studies on utilizing different supervisory loss functions of under-complete network output in sinogram and image domains. 1/2 low-dose setting is considered here. \checkmark And \mathcal{X} means supervisory loss function(s) used and not used on under-complete restoration outputs in the sinogram or the image domain of DuDoUFNet, respectively.

Image	Sinogram	SSIM	RMSE	PSNR
\mathcal{X}	\mathcal{X}	$.974 \pm .006$	22.56 ± 5.72	42.73 ± 2.07
\checkmark	\mathcal{X}	$.976 \pm .005$	21.03 ± 4.67	43.28 ± 1.82
\mathcal{X}	\checkmark	$.976 \pm .005$	20.86 ± 4.62	43.35 ± 1.82
\checkmark	\checkmark	$.978 \pm .004$	20.18 ± 4.30	43.62 ± 1.76

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript