Deep Sinogram Completion With Image Prior for Metal Artifact Reduction in CT Images

Lequan Yu, *Member, IEEE*, Zhicheng Zhang, Xiaomeng Li, *Member, IEEE*,
and Lei Xing, *Member, IEEE*

Abstract—**Computed tomography (CT) has been widely used for medical diagnosis, assessment, and therapy planning and guidance. In reality, CT images may be affected adversely in the presence of metallic objects, which could lead to severe metal artifacts and influence clinical diagnosis or dose calculation in radiation therapy. In this article, we propose a generalizable framework for metal artifact reduction (MAR) by simultaneously leveraging the advantages of image domain and sinogram domain-based MAR techniques. We formulate our framework as a sinogram completion problem and train a neural network (SinoNet) to restore the metal-affected projections. To improve the continuity of the completed projections at the boundary of metal trace and thus alleviate new artifacts in the reconstructed CT images, we train another neural network (PriorNet) to generate a good prior image to guide sinogram learning, and further design a novel residual sinogram learning strategy to effectively utilize the prior image information for better sinogram completion. The two networks are jointly trained in an end-to-end fashion with a differentiable forward projection (FP) operation so that the prior image generation and deep sinogram completion procedures can benefit from each other. Finally, the artifact-reduced CT images are reconstructed using the filtered backward projection (FBP) from the completed sinogram. Extensive experiments on simulated and real artifacts data demonstrate that our method produces superior artifact-reduced results while preserving the anatomical structures and outperforms other MAR methods.**

Index Terms—**Metal artifact reduction, sinogram completion, prior image, residual learning, deep learning.**

I. INTRODUCTION

COMPUTED tomography (CT) systems have become an important tool for medical diagnosis, assessment, and therapy planning and guidance. However, the metallic implants within the patients, *e.g.*, dental fillings and hip prostheses, would lead to missing data in X-ray projections and cause strong star-shape or streak artifacts to the reconstructed CT images [1]. Those metal artifacts not only present undesirable visual effects in CT images with influencing diagno-

sis but also make dose calculation problematic in radiation therapy [2], [3]. With the increasing use of metallic implants, how to reduce metal artifacts has become an important problem in CT imaging [4].

Numerous metal artifact reduction (MAR) methods have been proposed in the past decades, while there is no standard solution in clinical practice [3], [5], [6]. Since the metal artifacts are structured and non-local in the reconstructed CT images, the previous metal artifact reduction approaches mainly addressed this problem in the X-ray projections (sinogram). The metal-affected regions in the sinogram domain were corrected by modeling the underlying physical effects of imaging [7]–[10]. For example, Park *et al.* [10] proposed a method to correct beam hardening artifacts caused by the presence of metal in polychromatic X-ray CT. However, with the presence of high-atom number metals, the metal trace regions in sinogram are often severely corrupted and the above methods are limited in achieving satisfactory results [11]. Therefore, the other MAR methods regarded the metal-affected regions as the missing areas and filled them with estimated values [2], [12]. The early Linear interpolation (LI) approach [2] filled the missing regions by the linear interpolation of its neighboring unaffected data for each projection view. As interpolation cannot completely recover the metal trace information, the inconsistency between interpolated values and those unaffected values often results in strong new artifacts in the reconstructed images. To improve the sinogram interpolation quality, recent methods involved the forward projection of a prior image to complete the sinogram [13]–[17]. These methods first estimated prior images with various tissue information from the uncorrected image and then performed forward projection on the prior image to conduct sinogram completion. For example, Meyer *et al.* [15] improved the LI approach by generating a prior image with tissue processing and normalizing the projection with a forward projection of the prior image before interpolation. As the inaccurate prior images would lead to unfaithful structures in the reconstructed images, a key factor for prior-image-based approaches is to generate a good prior image to provide a more accurate surrogate for the missing data in the sinogram. Also, some researchers focused on designing new iterative reconstruction algorithms to reconstruct artifact-free images from the unaffected or corrected projections [18]–[21]. For example, Zhang *et al.* [21] proposed an iterative metal artifact reduction algorithm based on constrained optimization. However, these iterative reconstruction methods often

suffer from heavy computation and require proper hand-crafted regularizations.

With the development of deep learning in medical image reconstruction and analysis [22]–[25], recent progress of MAR has featured neural networks [4], [11], [26]. Park *et al.* [4] employed a U-Net [25] in the sinogram domain to deal with beam-hardening related artifacts in polychromatic CT. Gjesteby *et al.* [26] utilized deep learning to refine the result of NMAR [15] for achieving additional correction in critical image regions. Zhang and Yu [11] proposed to generate a reduced-artifact prior image with CNN to help correct the metal-corrupted regions in the sinogram. Although these methods show reasonable results on MAR, they are limited in handling with the remaining new artifacts in reconstructed CT images.

To improve the quality of the reconstructed CT images, inspired by the success of deep learning in solving ill-posed inverse problems in natural image processing [27]–[29], very recent works formulated MAR as an image restoration problem and reduced the metal artifacts with image-to-image translation networks [26], [30]–[36]. Gjesteby *et al.* [32] employed a deep neural network to reduce the new artifacts after the NMAR method with a perceptual loss. The RL-ARCNN [37] introduced deep residual learning to reduce metal artifacts in cervical CT images and Wang *et al.* [38] proposed to use the conditional generative adversarial network (cGAN) [39] to reduce metal artifacts in CT images. Very recently, Lin *et al.* [35] developed a dual-domain learning method to improve the image-restoration-based MAR results by involving sinogram enhancement as a procedure. These image-restoration-based methods demonstrated good performance on their experimental datasets due to the powerful representation capability of deep neural networks. However, in our experiments, we find that these methods tend to degrade on other site data, as the training samples hardly cover the unseen artifacts patterns. Although DuDoNet [35] introduces the sinogram enhancement procedure to improve the network performance, it still directly adopts the image-domain-refinement output (CNN output) as the final reconstructed image. As there is no geometry (physical) constraints to regularize the neural networks, there would be some tiny anatomical structure changes in the output image (see Fig. 6 for an example), which limits the usage of image domain methods in real clinical scenarios.

In this work, we present a novel image and sinogram domain joint learning framework for generalizable metal artifact reduction. Different from the previous image-restoration-based solutions, we formulate the MAR as the deep-learning-based sinogram completion task and train a deep neural network, *i.e.*, SinoNet, to restore the unreliable projections within the metal trace region. To ease the SinoNet learning and improve the completion quality, we simultaneously train another neural network, *i.e.*, PriorNet, to generate a good prior image with less metal artifact and guide the SinoNet learning with the forward projection of the prior image; see a sinogram completion result in Fig. 1. Moreover, we design a novel residual sinogram learning strategy to fully utilize the prior sinogram guidance to improve the continuity of sinogram completion and thus alleviate the new artifacts in the reconstructed CT images. The final

(a) Original Sinogram (b) Metal Trace (c) Metal-free Sinogram

(d) LI Sinogram (e) Deep Corrected Sinogram (f) Deep Corrected Sinogram with Prior Image
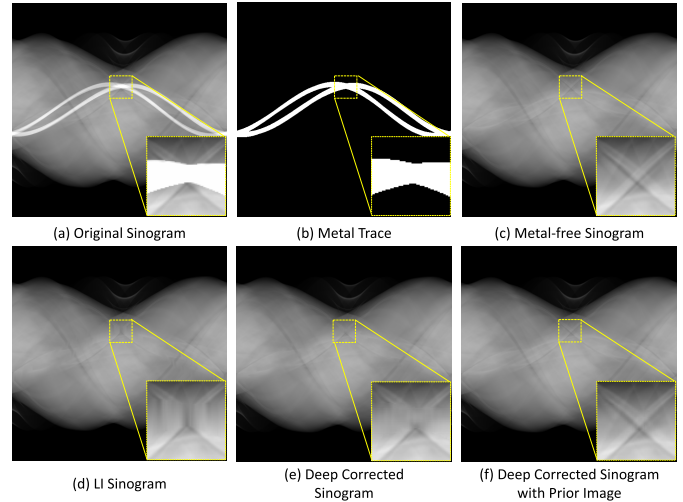
Fig. 1. The qualitative comparison of sinogram completion. An ROI is enlarged and shown with a narrower window to better visualize the difference. The linear interpolation (LI) [2] produces a poor estimation of the missing projections (d), while the deep network can generate a relatively good corrected sinogram (e). With the guidance of prior image, our method predicts more accurate projections (f), which are very close to the metal-free one.

CT image is then reconstructed from the completed sinogram with the conventional FBP algorithm. Compared with the previous prior-image-based MAR approaches, the whole framework is trained in an efficient end-to-end manner so that the prior image generation and deep sinogram completion procedures can be learned in a collaborative manner and benefit from each other. We extensively evaluate our framework on CT images with simulated and real metal artifacts, demonstrating that our method produces superior artifact-reduced results and outperforms other MAR methods.

Our main contributions are summarized as follows.

1) We present a novel image and sinogram domain joint learning framework for metal artifact reduction by simultaneously leveraging the advantages of image domain and sinogram domain-based MAR techniques. The proposed framework achieves superior performance on CT images with simulated and real metal artifacts.
2) We propose to train a deep prior image network to provide a good estimation of missing projections and thus enhance sinogram completion network learning. The two networks are trained in an end-to-end manner and can benefit from each other.
3) We design a novel residual sinogram learning scheme to facilitate sinogram completion. The scheme is able to fully utilize prior image information and alleviate the new artifacts on the reconstructed CT image.

The remainders of this article are organized as follows. We elaborate our framework in Section II. The experiments and results are presented in Section III. We further discuss the key issues of our method in Section IV and draw the conclusions in Section V.

II. METHODOLOGY

A. Overview

Fig. 2 depicts the overview of our proposed image and sinogram domain joint learning framework for metal artifact
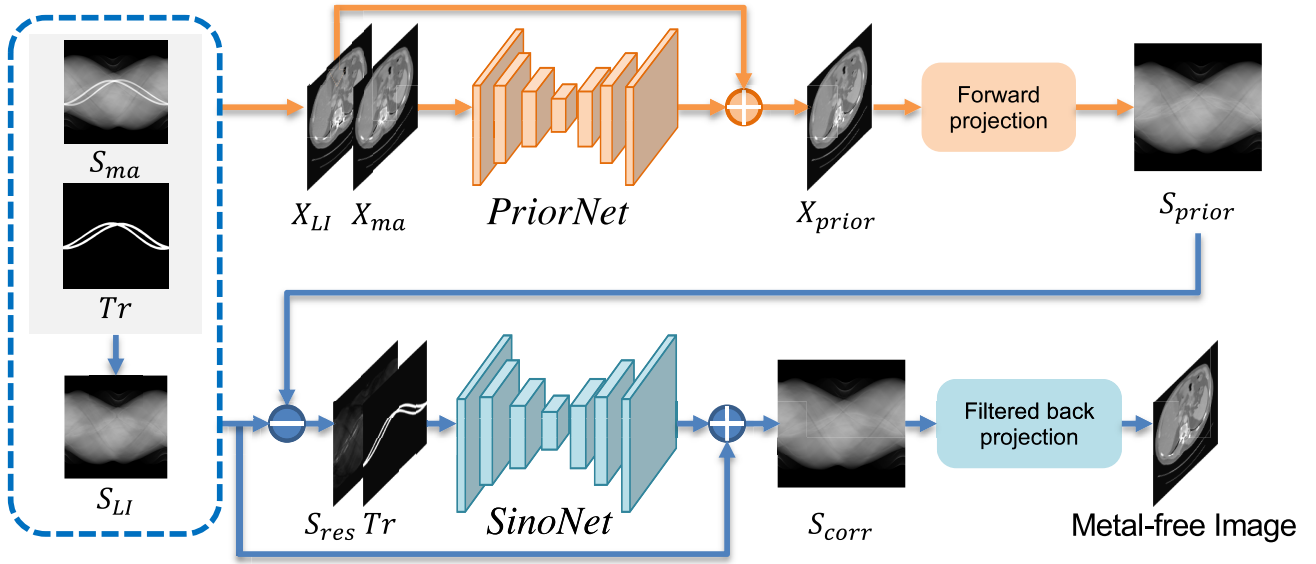
Fig. 2.   Schematic diagram of our proposed image and sinogram domain joint learning framework for metal artifact reduction. Given the metal-affected sinogram S_{ma} and metal trace mask Tr, we use linear interpolation to acquire LI corrected sinogram S_{LI}. We jointly train a prior image generation network, *i.e.*, PriorNet, to generate a good prior image X_{prior} and a singram completion network, *i.e.*, SinoNet, to restore the metal-affected sinogram with the guidance of the prior sinogram S_{prior}, which is the forward projection of the prior image X_{prior}. The S_{res} is the residual sinogram map between S_{LI} and S_{prior}. The final metal-free image is reconstructed from the corrected sinogram S_{corr} with the FBP algorithm.

reduction in CT images. The whole framework integrates the image domain learning (prior image generation) and sinogram domain learning (sinogram completion). Given the original metal-corrupted sinogram $S_{ma} \in \mathbb{R}^{H \times W}$ and the metal trace mask $Tr \in \{0, 1\}^{H \times W}$, we first apply the linear interpolation [2] to produce an initial estimation for the projections within the metal trace region and acquire the LI corrected sinogram S_{LI} for the following procedures. To ease the sinogram completion procedure, we train an image domain network, *i.e.*, PriorNet, to produce a good prior image X_{prior} with less metal artifact and acquire the prior sinogram S_{prior} with the forward projection of the generated prior image X_{prior} to guide the sinogram domain learning. We simultaneously train another deep neural network, *i.e.*, SinoNet, to restore the metal-affected projections to acquire the corrected sinogram S_{corr} by taking the LI corrected sinogram S_{LI}, the prior sinogram S_{prior}, and the metal trace mask Tr as input. Particularly, we design a novel residual learning strategy and make the SinoNet refine the residual sinogram map between S_{LI} and S_{prior}. The final metal-free CT image is then reconstructed from the corrected sinogram S_{corr} with the conventional FBP algorithm. The whole framework is trained in an end-to-end manner so that the prior image generation and sinogram completion procedures can benefit from each other.

B. Deep Prior Image Generation

In this step, we propose to generate a prior image with a deep neural network to facilitate the sinogram completion procedure, as the metal-free prior image would provide a good estimation for the missing projections in the original sinogram. A straightforward solution for this procedure is to take the original CT image with metal artifacts as input and train a neural network to generate the prior image with less

metal artifact. However, when the metal objects are relatively large, the metal artifacts in the original CT image would be very strong and it is difficult for the neural network to reduce the metal artifacts. Therefore, besides the original CT image, we also involve the LI corrected image into the prior image generation procedure and employ a neural network, *i.e.*, PriorNet, to refine the LI corrected image by residual learning. Specifically, we first reconstruct the original metal-corrupted CT image X_{ma} and LI corrected image X_{LI} from the original metal-affected sinogram S_{ma} and the linear interpolated sinogram S_{LI}, respectively. Then the artifact-reduced prior image is represented as

$$X_{prior} = X_{LI} + f_P([X_{ma}, X_{LI}]), \qquad (1)$$

where f_P denotes the prior image generation network and $[a, b]$ represents the concatenation operation of image a and b.

The PriorNet is based on the U-Net [25] architecture, but we halve the channel number to reduce the total number of parameters. To optimize the network, we employ the L1 loss to minimize the difference between the network output and the ground truth CT image X_{gt} without metal artifacts

$$\mathcal{L}_{prior} = ||X_{prior} - X_{gt}||_1. \qquad (2)$$

We further acquire the prior sinogram S_{prior} by performing forward projection on the generated prior image X_{prior}

$$S_{prior} = \mathcal{P}(X_{prior}), \qquad (3)$$

where $\mathcal{P}$ denotes the forward projection operator. The prior sinogram S_{prior} is then used to guide the network to complete the missing projections in the sinogram domain.

C. Deep Sinogram Completion

With the guidance of prior sinogram S_{prior}, we train another neural network, *i.e.*, SinoNet, to restore the projections within the metal trace region Tr in the sinogram domain. Specifically, the SinoNet takes the LI corrected sinogram S_{LI}, the prior sinogram S_{prior}, and the metal trace Tr as input, and outputs the missing projections in metal trace region Tr by utilizing the contextual information of the sinogram. To improve the continuity of the completed projections at the boundary of the metal trace region, we design a *residual sinogram learning* strategy and make the SinoNet refine the residual sinogram between S_{LI} and S_{prior}. Particularly, we calculate the residual sinogram map S_{res}, which can be treated as a smooth transition between the prior sinogram S_{prior} and the LI corrected sinogram S_{LI}, and then we employ the SinoNet to refine the *residual projections* within the metal trace region Tr. The corrected sinogram S'_{corr} can be written as:

$$S'_{corr} = f_S([S_{prior} - S_{LI}, Tr]) + S_{LI}, \qquad (4)$$

where f_S represents the sinogram completion network. As the network estimates the residual values instead of the absolute projection values, it can alleviate the discontinuity at the boundary of the metal trace [17]. Considering that the metals only affect projection data in metal trace region, we further composite the output of SinoNet and S_{LI} with respect to Tr to get the final corrected result:

$$\begin{aligned} S_{corr} &= S'_{corr} \odot Tr + S_{LI} \odot (1 - Tr) \\ &= f_S([S_{prior} - S_{LI}, Tr]) \odot Tr + S_{LI}, \qquad (5) \end{aligned}$$

where $\odot$ denotes element-wise multiplication.

To estimate the residual projections within the metal trace region, the network should be better aware of the metal trace information. However, the metal mask or metal trace regions are usually small and occupy a small portion of the whole sinogram, directly concatenating the residual sinogram map $S_{prior} - S_{LI}$ and Tr as network input would weaken the metal trace information due to the down-sampling operations of the network. Therefore, we employ the mask pyramid U-Net [40] to retain the metal trace information into each layer explicitly so that the network is able to extract more discriminative feature for restoring the missing information at metal trace region.

To optimize the SinoNet, we adopt the L1 loss to minimize the differences between the corrected sinogram and the ground truth sinogram S_{gt} without metals. However, as the composited sinogram S_{corr} has the identical values with the ground truth sinogram outside the metal trace region, directly minimizing the difference between S_{corr} and S_{gt} would provide supervision for the network output only within the metal trace region. As we mentioned above, the metal region occupies only a small portion of the whole sinogram. To improve the training efficiency, we also encourage the pre-composited sinogram S'_{corr} to be close to the ground truth sinogram S_{gt} so that the loss function can also provide supervision for those network outputs outside of the metal trace region. The total objective of the deep sinogram completion can be represented as

$$\mathcal{L}_{sino} = ||S_{gt} - S_{corr}||_1 + \beta ||S_{gt} - S'_{corr}||_1, \qquad (6)$$

where β is a hyper-parameter to control the trade-off between two difference items. we find that it is not sensitive to the network performance and we empirically set it as 0.1 in our experiments.

D. Overall Objective Function and Technical Details

The above L1 loss for SinoNet optimization only penalizes single projection value inconsistency in the sinogram domain, without considering the geometry-consistency of the completed values or penalizing the new artifacts in the reconstructed CT images. Therefore, we further design a filtered back-prorogation (FBP) loss to alleviate the new artifacts in the reconstructed CT image

$$\mathcal{L}_{FBP} = ||(\mathcal{P}^{-1}(S_{corr}) - X_{gt}) \odot (1 - M)||_1, \qquad (7)$$

where $\mathcal{P}^{-1}$ represents the FBP operator and M is the metal mask. Here we adopt the masked L1 loss to penalize the intensity difference only in the non-metal regions, as it is difficult to accurately reconstruct the original image at the metal position. Note that the FBP operation $\mathcal{P}^{-1}$ is *differentiable*, so that the gradient of $\mathcal{L}_{FBP}$ is able to back-propagate to SinoNet, encouraging it to generate geometry-consistent completion results. We jointly train the PriorNet and SinoNet in an end-to-end manner and the total objective function is

$$\mathcal{L}_{total} = \mathcal{L}_{prior} + \alpha_1 \mathcal{L}_{sino} + \alpha_2 \mathcal{L}_{FBP}, \qquad (8)$$

where α_1 and α_2 are hyperparameters to balance the weight of different loss items. We empirically set them as 1.0 in our experiments.

Our whole framework takes original metal-affected sinogram and metal trace as input. In the training phase, we use the simulated data to train the whole framework so that we can acquire the metal trace mask Tr for the simulated training data by performing the forward projection on the simulated metal mask M. In the testing phase, given the metal-affected sinogram S_{ma}, we can segment the metal mask M from the reconstructed metal-corrupted CT images X_{ma} with simple thresholding method or other advanced metal segmentation algorithms, and then conduct the similar forward projection to get the metal trace mask Tr.

III. EXPERIMENTS

A. Dataset and Simulation

We evaluated our method with simulated metal artifacts on CT images and real CT images with metal artifacts. For the simulation data, we randomly selected a subset of CT images from the recently released DeepLesion dataset [41] to synthesize metal artifacts. For the simulated metal masks, we employed the previous metal mask collection in [11], which contains 100 manually segmented metal implants with different shapes and sizes. Specifically, we randomly chose 1000 CT images and 90 metal masks to synthesize the training data. The remaining 10 metal masks were paired with an additional 200 CT images from 12 patients to generate 2000 combinations for network evaluation.

We followed the procedure in [11], [35] to simulate the metal-corrupted sinograms and CT images by inserting

metallic implants into clean CT images, where beam hardening and Poisson noise are simulated. We employed a polychromatic X-ray source and assumed the incident X-ray has 2×10^7 photons. The partial volume effect was also considered during the simulation. A fan-beam geometry was adopted and we uniformly sampled 640 projection views between 0-360 degrees. Before the simulation, the CT images were resized to 416×416, resulting in the sinogram with the size of 641×640.

B. Implementation Details

The framework was implemented in Python based on PyTorch [42] deep learning library. We trained the PriorNet and SinoNet in an end-to-end manner with differential forward projection (FP) and filtered backprojection (FBP) operations provided in ODL library.[1] In the network training, all the images had a size of 416×416 and the sinograms were with a size of 641×640. The Adam optimizer [43] was used to optimize the whole framework with the parameters $(\beta 1, \beta 2) = (0.5, 0.999)$. We totally trained 400 epochs with a mini-batch size of 8 on one Nvidia 1080Ti GPU and the learning rate was set as $1e^{-4}$. In each training iteration, we randomly chose one CT image with synthesized metal artifacts from the pool of 90 different metal mask pairs and the different CT images were formed as one mini-batch data to be fed into the network for calculating the total objective function.

C. Experimental Results on DeepLesion Data

1) Quantitative Comparisons With State-of-the-Art Methods:
We compare our method with conventional interpolation-based methods: linear interpolation (LI) [2] and normalized metal artifact reduction (NMAR) [15], which are widely employed approaches in MAR. Also, we compare our method with the recent deep-learning-based methods CNNMAR [11], cGANMAR [38], and the state-of-the-art method DuDoNet [35]. The CNNMAR approach also adopts a CNN to generate a reduced-artifact prior image and then uses traditional interpolation to correct the metal-corrupted regions in the sinogram. Both cGANMAR and DuDoNet employ the image domain network to generate the final results, where cGANMAR directly uses an image-to-image translation network to reduce artifacts on original metal images while DuDoNet further incorporates sinogram enhancement to ease image domain learning. For CNNMAR, we used the public released code and model. We re-implemented DuDoNet [35] and cGANMAR [38], since there is no public implementation.

Table I shows the quantitative comparison results of our method and other methods on the DeepLesion dataset. It is observed that the prior-image-based interpolation method NMAR outperforms LI approach on both root mean square error (RMSE) and structured similarity index (SSIM) metrics, as the prior image information improves the accuracy of interpolation for missing projection values. The deep-learning-based methods CNNMAR and cGANMAR achieves much lower RMSE and higher SSIM values than conventional MAR methods, showing the advantage of data-driven deep neural

[1] https://github.com/odlgroup/odl

TABLE I
QUANTITATIVE COMPARISON OF DIFFERENT METHODS ON DEEPLESION DATASET

Method	RMSE (HU)	SSIM
LI [2]	50.31±19.41	0.9455±0.0315
NMAR [15]	47.03±20.67	0.9594±0.0299
CNNMAR [11]	43.27±14.44	0.9706±0.0159
cGANMAR [38]	39.01±12.66	0.9754±0.0055
DuDoNet [35]	38.00±13.31	0.9766±0.0072
Ours	**31.15±5.81**	**0.9784±0.0048**

networks for metal artifact reduction. The DuDoNet achieves better RMSE and SSIM performance when compared with cGANMAR, as it integrates sinogram enhancement to reduce artifacts before conducting the image refinement procedure. Compared with DuDoNet, our method further reduces RMSE with 6.85 HU and achieves slightly better SSIM values. Overall, our framework attains the best performance among different methods in terms of RMSE and SSIM, showing the effectiveness of our method for metal artifact reduction.

2) Qualitative Analysis:
Fig. 3 and Fig. 4 shows the visual comparisons of our method and other methods on DeepLesion simulation data. We show the refer metal-free CT images, simulated metal artifact images (Metal Image), and metal artifact reduction results of different methods. The simulated metal masks are colored in red for better visualization. It is observed that severe streaking artifacts are in the original metal images and a severe dark strip exists between two metal implants (see Fig. 3(B1)). Generally, the deep-learning-based methods CNNMAR, cGANMAR and DuDoNet can reduce more artifacts than conventional LI and NMAR approaches. When the metal implants are small (Fig. 3), the DuDoNet and cGANMAR can achieve better visual results than CNNMAR, while there are still some mild artifacts in the DuDoNet and cGANMAR results; see the dashed blue ovals in Fig. 3(F3&G3). Compared with these methods, our method effectively reduces most artifacts and retains the fine details of the structures. Fig. 4 shows the results when the metal implants are large. It is observed that the conventional interpolation methods LI and NMAR, and image domain method cGANMAR and DuDoNet cannot preserve the details of the original structures and there are some new secondary-artifacts in the cGANMAR results; see the green arrows in Fig. 4(F1&F2). Compared to CNNMAR, our method preserves better structure details, showing the effectiveness of the deep signogram completion mechanism. We also calculate the ROI RMSE and ROI SSIM for the red box patches in Fig. 5 to quantitatively compare different methods. As shown in Fig. 5 (A2-H2&A4-H4), our method achieves the lowest RMSE and highest SSIM values among different methods.

D. Generalization to Different Site Data

The selected CT images from the DeepLesion dataset are samples of abdomen and thorax. To show the feasibility of our method applied to different site data, we directly evaluate the model trained with DeepLesion data on the head CT images collected from the online website with simulated metal artifacts. Fig. 5 shows the visual metal artifact reduction
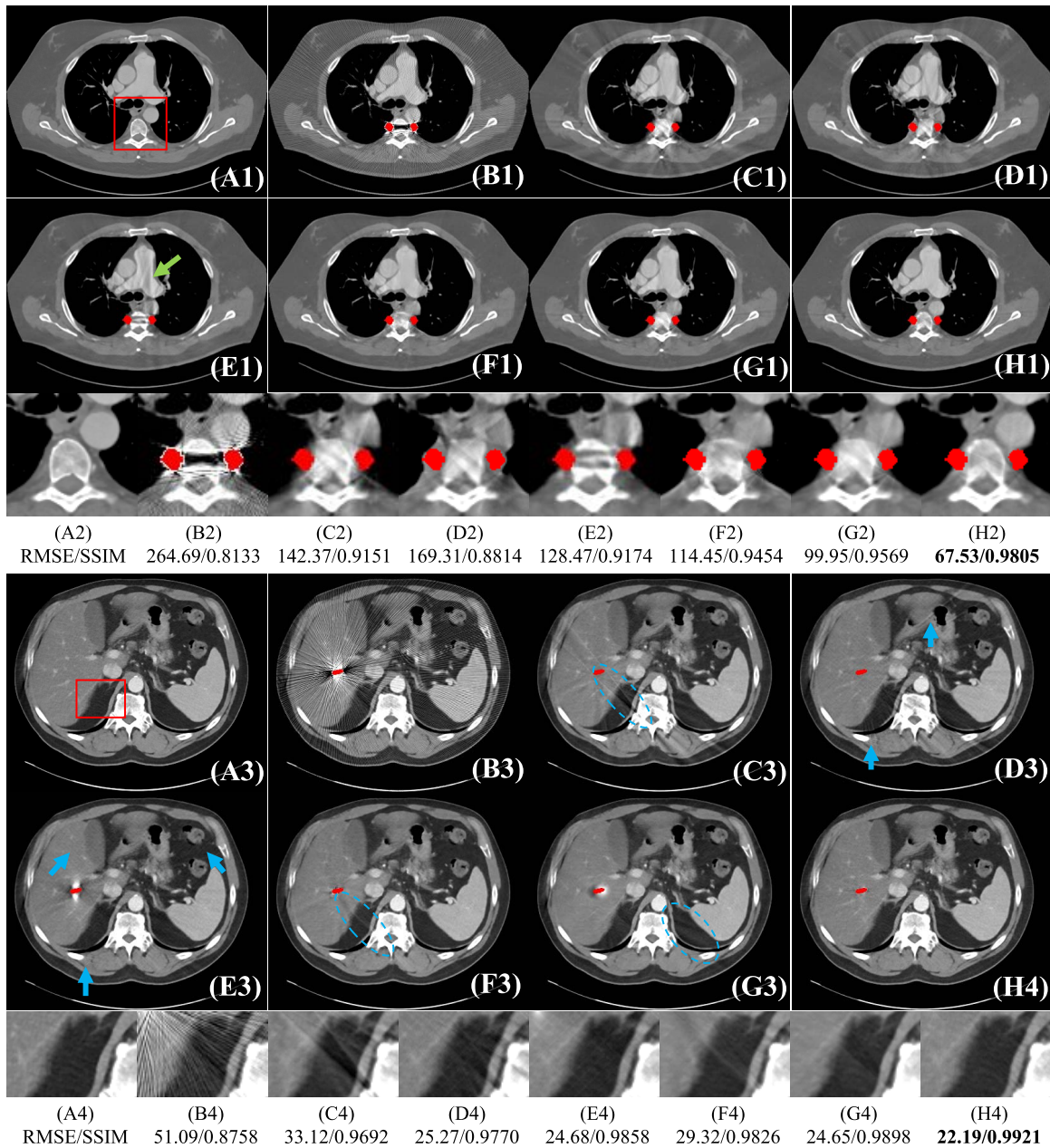
Fig. 3. Visual comparison with different methods on DeepLesion dataset. The simulated metal masks are colored in red for better visualization. The (A1-A4) are refer images. We show the MAR results of LI [2] (C1-C4), NMAR [15] (D1-D4), CNNMAR [11] (E1-E4), cGANMAR [38] (F1-F4), DuDoNet [35] (G1-G4), and our method (H1-H4). The display window of the first and second samples are [−480 560] and [−175, 275] HU, respectively. We also use ROI RMSE and ROI SSIM to show quantitative results for a better comparison.

results of our method on the head CT images with simulated dental fillings. We also show the conventional LI and NMAR results and the deep-learning-based CNNMAR and DuDoNet results. It is observed that the LI and NMAR introduce several secondary-artifacts and could change the anatomical structures of the tooth; see blue arrows in Fig. 5(d). The DuDoNet can further reduce the artifacts, while there are still several shading artifacts in the output images; see green arrows in Fig. 5(f). Although without training with head CT images, our method effectively reduces artifacts, indicating that the proposed method has the potential to handle different site data. Notably, the MAR result of our method is even comparable with CNNMAR, which is also trained with head CT images.

E. Experiments on CT Images With Real Metal Artifacts

1) Results on CT Images With Real Metal Artifacts: Since the original sinogram data with metal artifacts in the real clinical scenario are difficult to access, we follow the previous work [35] to evaluate our method on clinical CT images with metal artifacts. Specifically, we collected some clinical CT images with metal artifacts and segmented the metal mask from the clinical CT images with a simple thresholding method (*i.e.*, 2000 HU in our experiments). The forward projection of the metal masks was conducted to generate the metal projection and the pixels with the projection value greater than zero were regarded as the metal trace region Tr. We also performed the forward projection on the clinical CT image
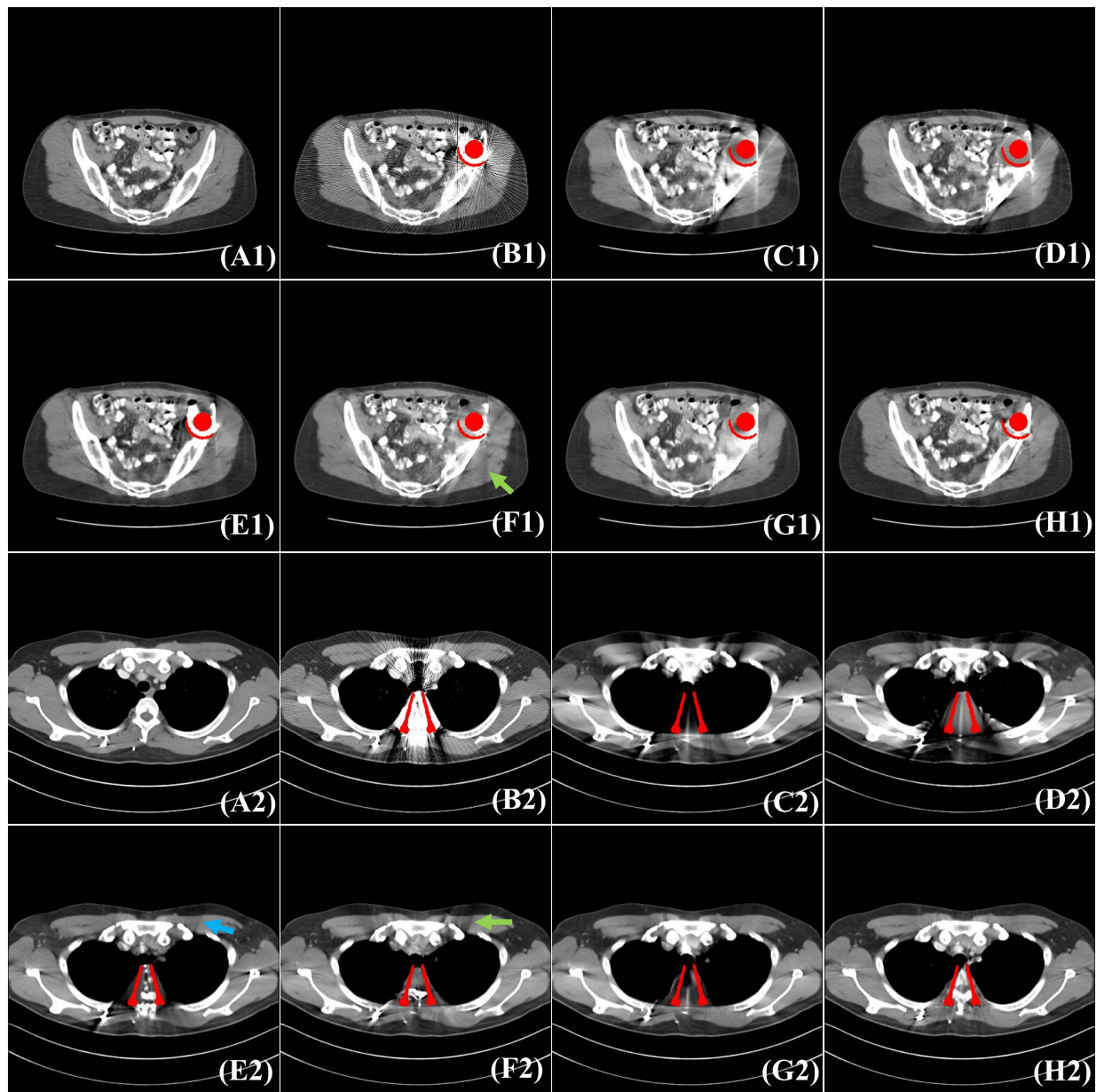
Fig. 4. Visual comparison with different methods on DeepLesion dataset. The simulated metal masks are colored in red for better visualization. The (A1-A2) are refer images. We show the MAR results of LI [2] (C1-C2), NMAR [15] (D1-D2), CNNMAR [11] (E1-E2), cGANMAR [38] (F1-F2), DuDoNet [35] (G1-G2), and our method (H1-H2). The display window is [−175 275] HU.

with the same imaging geometry as the above simulation procedure to acquire the metal-corrupted sinogram S_{ma}. The LI corrected sinogram S_{LI} was then generated from S_{ma} and Tr with linear interpolation. Finally, we fed S_{ma}, S_{LI} and Tr into our framework to get the meta artifact reduction images. Fig. 6 presents the visual results of different methods. Our method effectively reduces metal artifacts compared with the original metal image. From the yellow zoomed patches, it is observed that the other methods change some tiny anatomical structures of the original image, while our method can preserve the fine-grained anatomical structures.

2) The Influence of Metal Mask Segmentation: An accurate metal trace mask (or metal mask equally) is vital for the good performance of metal artifacts reduction in our framework.

In practice, we can manually segment the metals or use some automatic metal segmentation methods (*e.g.*, thresholding method) to segment the metals. To investigate the influence of metal mask segmentation on the final MAR results, we take different metal masks to acquire the metal traces and then adopt our method to conduct MAR with taking these different metal traces as input. Fig. 7 shows the MAR results of our method by taking the original thresholding-based metal mask (a), the dilation metal masks (b & c), and erosion metal masks (e & f) as input. In general, our method can achieve reasonable MAR results under slightly metal segmentation errors. The over-segmented metal masks would lead to a relatively large metal trace and our method needs to complete more projection values. As the network has been trained to

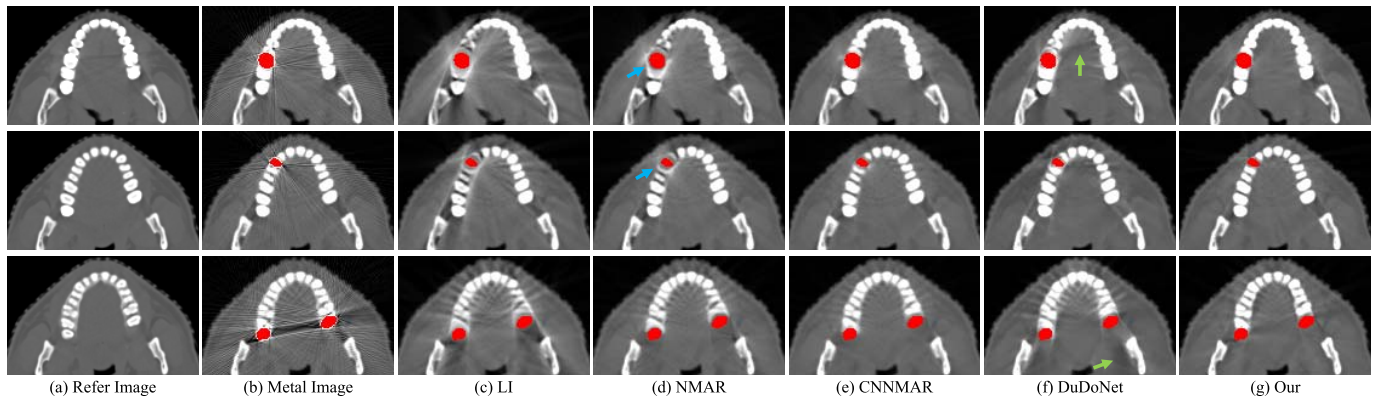(a) Refer Image   (b) Metal Image   (c) LI   (d) NMAR   (e) CNNMAR   (f) DuDoNet   (g) Our

Fig. 5. Visual results on head CT images with different numbers of simulated dental fillings. The simulated dental fillings are colored in red for better visualization. The display window is [−1000 1600] HU.

(A) Metal Image   (B) LI   (C) NMAR   (D) CNNMAR   (E) DuDoNet   (F) Ours
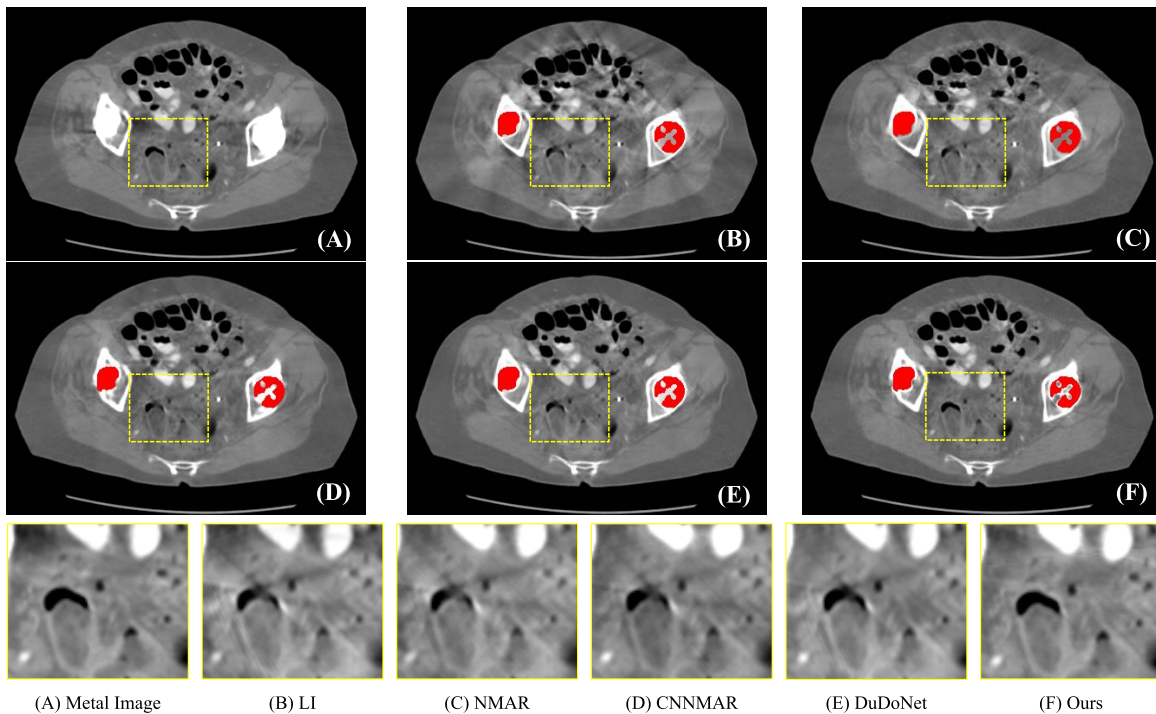
Fig. 6. Visual comparison on CT images with real metal artifacts. The segmented metals are colored in red for better visualization. Our method effectively reduces metal artifacts and preserves the fine-grained anatomical structures. The display window of whole image is [−480 560] HU and the display window of cropped patches is [−400 300] HU.

completed different metal traces, it can estimate the missing projection values without leading to strong new artifacts in the reconstructed CT images, although there are additional shading artifacts in the CT images. As for the under-segmented case, the corresponding metal trace is narrower than the original metal trace. The sinogram completion network would only complete some projection values of the original metal trace region and reuse some unreliable projection data. Therefore, our method can only reduce some metal artifacts and there still remains some residual streaking artifacts in the reconstructed CT image.

F. Analytical Studies

1) Effectiveness of Prior Image Generation: In our framework, we train the PriorNet to generate a good prior image

to ease sinogram completion. To show the effectiveness of this procedure, we directly train a neural network to complete the sinogram without taking the prior sinogram as input. The quantitative results of this method on DeepLesion simulated dataset are shown in Table II. It is observed that this method (Deep sinogram completion) achieves much higher RMSE and lower SSIM values than our method, verifying the effectiveness of prior image generation procedure. In Fig. 8, we show some generated prior images and final metal artifact reduction images on DeepLesion and head CT data. We can see that the prior images (Fig. 8(b)) have less artifacts than original metal image (Fig. 8(a)) and our final results (Fig. 8(c)) further reduce artifacts compared with the prior images. We also train our framework by taking only original metal image as PriorNet input. As shown in Table II, this method

(a) Our Result   (b) Dilation Mask 1   (c) Dilation Mask 2

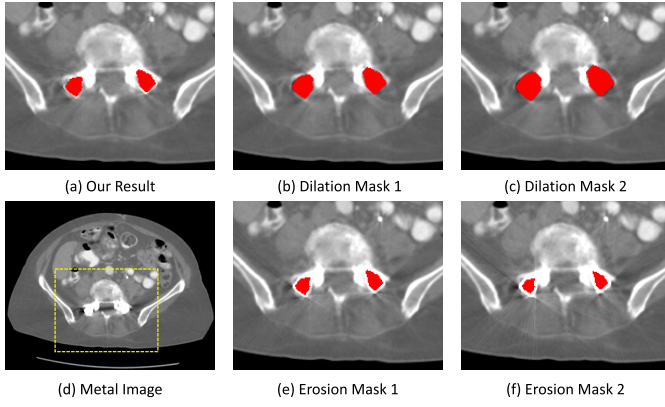(d) Metal Image   (e) Erosion Mask 1   (f) Erosion Mask 2

Fig. 7. The results of our method with different segmented metal masks as input. We show the MAR results with thresholding-based metal segmentation (a); we also dilate the metal mask to get over-segmented masks (b & c) and erode the metal mask to get under-segmented masks (e & f). The display window is [−480 560] HU.

TABLE II
QUANTITATIVE ANALYSIS OF OUR METHODS ON SIMULATED DEEPLESION DATASET

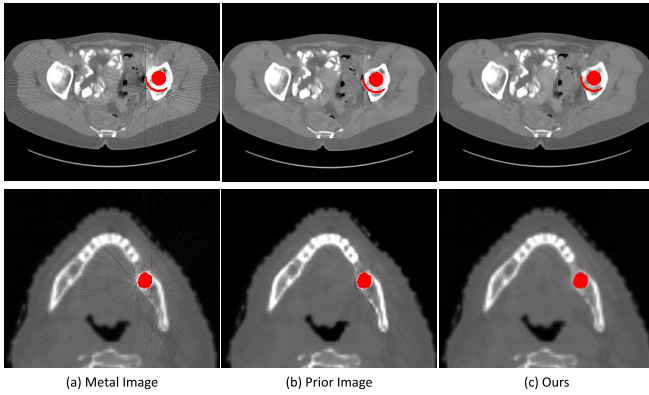| Method	RMSE (HU)	SSIM
Deep sinogram completion	43.65±17.61	0.9720±0.0082
With tissue processing	35.64±7.91	0.9768±0.0047
w/o residual sinogram learning	31.86±4.69	0.9781±0.0048
Only metal image	31.60±4.95	0.9778±0.0048
Ours	**31.15±5.81**	**0.9784±0.0048**

(a) Metal Image   (b) Prior Image   (c) Ours

Fig. 8. Illustration of the generated prior image. The display window of first row and seconed row are [−480 560] HU and [−1000 1600] HU, respectively.

(Only metal image) generates slightly worse results than our method on the simulated dataset, indicating that incorporating LI corrected image as input can facilitate the prior image generation.

2) Effectiveness of Residual Sinogram Learning: We show the qualitative MAR results of our method with and without residual sinogram learning strategy in Fig. 9. The first row is the results on the simulated metal artifact image and the second row shows the results on the real clinical CT image with metal artifacts. In the experiment without residual sinogram learning, the SinoNet directly takes the prior sinogram and metal trace mask as input and outputs the refined projections within the metal trace region. From the visual comparison in Fig. 9, we can observe that our framework further reduces

(a) Metal Image   (b) w/o Residual Sinogram Learning   (c) Our Results
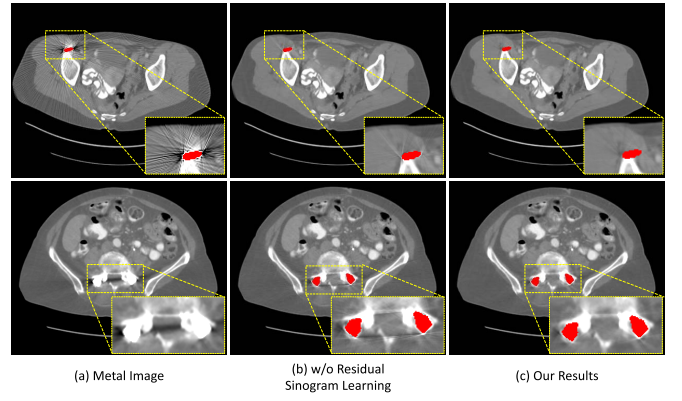
Fig. 9. The results of our method with (b) and without (c) residual sinogram learning strategy. The first row is simulated metal image and the second image is real clinical CT image. The display window is [−480 560] HU.

metal artifacts on both simulated and real samples by adopting the residual sinogram learning strategy. We also present the quantitative results of "w/o residual sinogram learning" in Table II. It is observed that we achieve higher RMSE and lower SSIM values with residual sinogram learning.

3) Compared With Tissue Processing: The tissue processing step is often used to acquire the prior image in previous MAR methods. We conduct another experiment to investigate the effect of this strategy, where we employ the tissue processing and metal trace replacement steps [17], [23] to process the generated prior image X_{prior} of our method and then acquire the final metal-free image with FBP reconstruction. The quantitative result of this method (With tissue processing) is shown in Table II. It is observed that *With tissue processing* achieves satisfying results on the simulated DeepLesion dataset, while its performance is still inferior to our end-to-end deep sinogram completion strategy, indicating that the deep sinogram network could automatically learn how to reduce the mild artifacts in the prior image.

IV. DISCUSSION

MAR is a long-standing problem in CT imaging. In this work, we aim to design a data-driven framework to address this problem by utilizing a large amount of training data. The previous deep-learning-based methods usually formulate the MAR as an image restoration problem. Whereas we borrow the spirit of conventional MAR approaches and formulate the MAR as a deep sinogram completion problem, aiming to improve the generalization and robustness of the framework. Since directly regressing the accurate missing projection data is difficult, we propose to incorporate the deep prior image generation procedure and adopt a residual sinogram completion strategy. This manner can improve the continuity of the projection values at the boundary of metal traces and alleviate the new artifacts, which are the common drawbacks of sinogram completion based MAR methods. In such way, our framework could better utilize the advantages of deep learning techniques while alleviating the risk of overfitting to certain training data.

We solve MAR in both sinogram and image domains, which share the same strategy with DuDoNet [35] and

DuDoNet++ [36]. However, our framework differs from them in a few important aspects. The DuDoNet and DuDoNet++ directly adopt the image-domain-refinement output as the final MAR image, whereas the final MAR image of our method is directly FBP-reconstructed from the completed sinogram. As there is no geometry (physical) constraints to regularize the neural networks, there would be some tiny anatomical structure changes in the CNN-output images. Our FBP-reconstructed image could preserve the anatomical structure of the original image and avoid the resolution loss, as we only modify the metal trace region values in the sinogram; see comparisons in Figs. 3&4. More importantly, we design a novel residual learning strategy for the sinogram enhancement network to refine the residual projections within the metal trace region, and both quantitative and qualitative results show the effectiveness of such residual learning strategy.

It is clinically impractical to acquire metal-free and metal-inserted CT data for network training, We thus simulate metal artifacts from clinical metal-free CT images to acquire synthesized training pairs. In this case, the quality of simulated data would largely influence network performance. Currently, we simulate the metal artifacts without carefully designing the simulated metal masks. In the future, we will investigate how to create a good simulated dataset to further improve the network performance on real clinical CT images.

The previous prior-image-based MAR methods would utilize tissue processing to post-process the generated prior image. Whereas in our method, we directly employ the CNN output as the prior image to guide the sinogram completion network. In this case, we can jointly train the PriorNet and SinoNet, and the prior image generation and sinogram completion procedures can benefit from each other. Although some mild artifacts would remain in the generated prior image, the sinogram completion network would automatically learn how to complete the sinogram from the prior sinogram, so that our final output can remove these mild artifacts.

We have trained and evaluated our method on simulated datasets, but as shown in the experiments, our framework has strong potential to be applied in CT images with real metal artifacts. Since there is no public real projection data and we need to cooperate with CT device venders to acquire such real projection data, in the current study, we use forward projection to simulate the projection data. This is a limitation of our current work and we will evaluate the effectiveness of our method on real project data in the future. When applying the framework into real clinical data, one important practice issue is how to acquire the accurate metal trace and metal masks. Although our framework is relatively robust to the metal mask segmentation, an accurate metal segmentation would further ensure the stability of the MAR results. Deep learning has achieved promising results in various medical image segmentation problems [24], [25], [44]. Incorporating deep learning-based metal segmentation or advanced metal identification algorithm [3] into our framework would further improve the robustness of our method. Recently, some works studied how to locate the shape and location of metal objects directly from the metal-corrupted sinogram [3], [45]. These binary reconstruction works can also be integrated into our framework for better metal artifact reduction. Moreover, it is more interesting to investigate how to simultaneously conduct metal mask identification and metal artifact reduction in a collaborative manner.

V. CONCLUSION

We present a generalizable image and sinogram domain joint learning framework for metal artifact reduction in CT imaging, which integrates the merits of deep learning and conventional MAR methods. Our framework follows the prior-image-based sinogram completion strategy and we employ two networks to conduct prior image generalization and sinogram completion. The whole framework is trained in an end-to-end manner so that the two networks can benefit from each other in network learning. Our framework is trained with the simulated metal artifacts data, while the experimental results show the strong potential of our method to handle CT images with real artifacts. The future works include investigating how to simultaneously conduct metal mask identification and metal artifact reduction, as well as how to perform the procedure in an unsupervised manner.

REFERENCES

[1] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "Metal streak artifacts in X-ray computed tomography: A simulation study," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec., IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, Nov. 1998, pp. 1860–1865.

[2] W. A. Kalender, R. Hebel, and J. Ebersberger, "Reduction of CT artifacts caused by metallic implants.," *Radiology*, vol. 164, no. 2, pp. 576–577, Aug. 1987.

[3] B. Meng, J. Wang, and L. Xing, "Sinogram preprocessing and binary reconstruction for determination of the shape and location of metal objects in computed tomography (CT)," *Med. Phys.*, vol. 37, no. 11, pp. 5867–5875, Oct. 2010.

[4] H. S. Park, S. M. Lee, H. P. Kim, J. K. Seo, and Y. E. Chung, "CT sinogram-consistency learning for metal-induced beam hardening correction," *Med. Phys.*, vol. 45, no. 12, pp. 5376–5384, Dec. 2018.

[5] J. Y. Huang *et al.*, "An evaluation of three commercially available metal artifact reduction methods for CT imaging," *Phys. Med. Biol.*, vol. 60, no. 3, p. 1047, 2015.

[6] L. Gjesteby *et al.*, "Metal artifact reduction in CT: Where are we after four decades?" *IEEE Access*, vol. 4, pp. 5826–5849, 2016.

[7] J. Hsieh, R. C. Molthen, C. A. Dawson, and R. H. Johnson, "An iterative approach to the beam hardening correction in cone beam CT," *Med. Phys.*, vol. 27, no. 1, pp. 23–29, Jan. 2000.

[8] M. Kachelrieß, O. Watzke, and W. A. Kalender, "Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT," *Med. Phys.*, vol. 28, no. 4, pp. 475–490, Apr. 2001.

[9] E. Meyer, C. Maas, M. Baer, R. Raupach, B. Schmidt, and M. Kachelries, "Empirical scatter correction (ESC): A new CT scatter correction method and its application to metal artifact reduction," in *Proc. IEEE Nucl. Sci. Symposuim Med. Imag. Conf.*, Oct. 2010, pp. 2036–2041.

[10] H. S. Park, D. Hwang, and J. K. Seo, "Metal artifact reduction for polychromatic X-ray CT based on a beam-hardening corrector," *IEEE Trans. Med. Imag.*, vol. 35, no. 2, pp. 480–487, Feb. 2016.

[11] Y. Zhang and H. Yu, "Convolutional neural network based metal artifact reduction in X-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1370–1381, Jun. 2018.

[12] A. Mehranian, M. R. Ay, A. Rahmim, and H. Zaidi, "X-Ray CT metal artifact reduction using wavelet domain L_0 sparse regularization," *IEEE Trans. Med. Imag.*, vol. 32, no. 9, pp. 1707–1722, Sep. 2013.

[13] J. Müller and T. M. Buzug, "Spurious structures created by interpolation-based ct metal artifact reduction," *Proc. SPIE*, vol. 7258, Mar. 2009, Art. no. 72581Y.

[14] D. Prell, Y. Kyriakou, M. Beister, and W. A. Kalender, "A novel forward projection-based metal artifact reduction method for flat-detector computed tomography," *Phys. Med. Biol.*, vol. 54, no. 21, p. 6575, 2009.

[15] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrieß, "Normalized metal artifact reduction (NMAR) in computed tomography," *Med. Phys.*, vol. 37, no. 10, pp. 5482–5493, Sep. 2010.

[16] J. Wang, S. Wang, Y. Chen, J. Wu, J.-L. Coatrieux, and L. Luo, "Metal artifact reduction in CT using fusion based prior image," *Med. Phys.*, vol. 40, no. 8, Jul. 2013, Art. no. 081903.

[17] Y. Zhang, H. Yan, X. Jia, J. Yang, S. B. Jiang, and X. Mou, "A hybrid metal artifact reduction algorithm for X-ray CT," *Med. Phys.*, vol. 40, no. 4, Mar. 2013, Art. no. 041910.

[18] G. Wang, D. L. Snyder, J. A. O'Sullivan, and M. W. Vannier, "Iterative deblurring for CT metal artifact reduction," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 657–664, Oct. 1996.

[19] G. Wang, M. W. Vannier, and P.-C. Cheng, "Iterative X-ray cone-beam tomography for metal artifact reduction and local region reconstruction," *Microsc. Microanal.*, vol. 5, no. 1, pp. 58–65, Jan. 1999.

[20] C. Lemmens, D. Faul, and J. Nuyts, "Suppression of metal artifacts in CT using a reconstruction procedure that combines MAP and projection completion," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 250–260, Feb. 2009.

[21] X. Zhang, J. Wang, and L. Xing, "Metal artifact reduction in X-ray computed tomography (CT) by constrained optimization," *Med. Phys.*, vol. 38, no. 2, pp. 701–711, Jan. 2011.

[22] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1289–1296, Jun. 2018.

[23] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, "A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1407–1417, Jun. 2018.

[24] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[25] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent*. Cham, Switzerland: Springer, 2015, pp. 234–241.

[26] L. Gjesteby, Q. Yang, Y. Xi, Y. Zhou, J. Zhang, and G. Wang, "Deep learning methods to guide ct image reconstruction and reduce metal artifacts," *Proc. SPIE*, vol. 10132, Mar. 2017, Art. no. 101322W.

[27] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.

[28] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9446–9454.

[29] J. Lehtinen *et al.*, "Noise2noise: Learning image restoration without clean data," in *Proc. ICML*, 2018, pp. 2965–2974.

[30] H. S. Park, S. M. Lee, H. P. Kim, and J. K. Seo, "Machine-learning-based nonlinear decomposition of CT images for metal artifact reduction," 2017, *arXiv:1708.00244*. [Online]. Available: http://arxiv.org/abs/1708.00244

[31] L. Gjesteby *et al.*, "Reducing metal streak artifacts in ct images via deep learning: Pilot results," in *Proc. 14th Int. Meeting Fully Three-Dimensional Image Reconstruction Radiol. Nucl. Med.*, 2017, pp. 611–614.

[32] L. Gjesteby *et al.*, "Deep neural network for ct metal artifact reduction with a perceptual loss function," in *Proc. 5th Int. Conf. Image Formation X-Ray Comput. Tomogr.*, vol. 1, 2018.

[33] K. Liang, L. Zhang, H. Yang, Y. Yang, Z. Chen, and Y. Xing, "Metal artifact reduction for practical dental computed tomography by improving interpolation-based reconstruction with deep learning," *Med. Phys.*, vol. 46, no. 12, pp. e823–e834, Dec. 2019.

[34] H. Liao, W.-A. Lin, S. K. Zhou, and J. Luo, "ADN: Artifact disentanglement network for unsupervised metal artifact reduction," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 634–643, Mar. 2020.

[35] W.-A. Lin *et al.*, "DuDoNet: Dual domain network for CT metal artifact reduction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10512–10521.

[36] Y. Lyu, W.-A. Lin, H. Liao, J. Lu, and S. K. Zhou, "Encoding metal mask projection for metal artifact reduction in computed tomography," 2020, *arXiv:2001.00340*. [Online]. Available: http://arxiv.org/abs/2001.00340

[37] X. Huang, J. Wang, F. Tang, T. Zhong, and Y. Zhang, "Metal artifact reduction on cervical CT images by deep residual learning," *Biomed. Eng. OnLine*, vol. 17, no. 1, p. 175, Dec. 2018.

[38] J. Wang, Y. Zhao, J. H. Noble, and B. M. Dawant, "Conditional generative adversarial networks for metal artifact reduction in ct images of the ear," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent*. Cham, Switzerland: Springer, 2018, pp. 3–11.

[39] 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. (CVPR)*, Jul. 2017, pp. 1125–1134.

[40] H. Liao *et al.*, "Generative mask pyramid network for ct/cbct metal artifact reduction with joint projection-sinogram correction," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent*. Cham, Switzerland: Springer, 2019, pp. 77–85.

[41] K. Yan *et al.*, "Deep lesion graphs in the wild: Relationship learning and organization of significant radiology image findings in a diverse large-scale lesion database," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9261–9270.

[42] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Proc. NIPS*, 2019, pp. 8026–8037.

[43] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980

[44] L. Yu, X. Yang, H. Chen, J. Qin, and P. A. Heng, "Volumetric convnets with mixed residual connections for automated prostate segmentation from 3D mr images," in *Proc. 21st AAAI Conf. Artif. Intell.*, 2017, pp. 36–72.

[45] J. Wang and L. Xing, "A binary image reconstruction technique for accurate determination of the shape and location of metal objects in X-ray computed tomography," *J. X-Ray Sci. Technol.*, vol. 18, no. 4, pp. 403–414, 2010.