



MEPNet: A Model-Driven Equivariant Proximal Network for Joint Sparse-View Reconstruction and Metal Artifact Reduction in CT Images

Hong Wang¹(✉), Minghao Zhou^{1,2}, Dong Wei¹, Yuexiang Li¹,
and Yefeng Zheng¹

¹ Tencent Jarvis Lab, Shenzhen, People's Republic of China

{hazelhwang, hippomhzhou, donwei, vicyxli, yefengzheng}@tencent.com

² Xi'an Jiaotong University, Xi'an, Shaan'xi, People's Republic of China

Abstract. Sparse-view computed tomography (CT) has been adopted as an important technique for speeding up data acquisition and decreasing radiation dose. However, due to the lack of sufficient projection data, the reconstructed CT images often present severe artifacts, which will be further amplified when patients carry metallic implants. For this joint sparse-view reconstruction and metal artifact reduction task, most of the existing methods are generally confronted with two main limitations: 1) They are almost built based on common network modules without fully embedding the physical imaging geometry constraint of this specific task into the dual-domain learning; 2) Some important prior knowledge is not deeply explored and sufficiently utilized. Against these issues, we specifically construct a dual-domain reconstruction model and propose a model-driven equivariant proximal network, called MEPNet. The main characteristics of MEPNet are: 1) It is optimization-inspired and has a clear working mechanism; 2) The involved proximal operator is modeled via a rotation equivariant convolutional neural network, which finely represents the inherent rotational prior underlying the CT scanning that the same organ can be imaged at different angles. Extensive experiments conducted on several datasets comprehensively substantiate that compared with the conventional convolution-based proximal network, such a rotation equivariance mechanism enables our proposed method to achieve better reconstruction performance with fewer network parameters. We will release the code at <https://github.com/hongwang01/MEPNet>.

Keywords: Sparse-view reconstruction · Metal artifact reduction · Rotation equivariance · Proximal network · Generalization capability

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43999-5_11.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14229, pp. 109–120, 2023.
https://doi.org/10.1007/978-3-031-43999-5_11

1 Introduction

Computed tomography (CT) has been widely adopted in clinical applications. To reduce the radiation dose and shorten scanning time, sparse-view CT has drawn much attention in the community [10, 34]. However, sparse data sampling inevitably degenerates the quality of CT images and leads to adverse artifacts. In addition, when patients carry metallic implants, such as hip prostheses and spinal implants [11, 13, 24], the artifacts will be further aggravated due to beam hardening and photon starvation. For the joint sparse-view reconstruction and metal artifact reduction task (SVMAR), how to design an effective method for artifact removal and detail recovery is worthy of in-depth exploration.

For the sparse-view (SV) reconstruction, the existing deep-learning (DL)-based methods can be roughly divided into three categories based on the information domain exploited, *e.g.*, sinogram domain, image domain, and dual domains. Specifically, for the sinogram-domain methods, sparse-view sinograms are firstly repaired based on deep networks, such as U-Net [10] and dense spatial-channel attention network [37], and then artifact-reduced CT images are reconstructed via the filtered-back-projection (FBP) process. For the image-domain methods, researchers have proposed to learn the clean CT images from degraded ones via various structures [18, 34, 35]. Alternatively, both sinogram and CT images are jointly exploited for the dual reconstruction [4, 21, 32, 36].

For the metal artifact reduction (MAR) task, similarly, the current DL-based approaches can also be categorized into three types. To be specific, sinogram-domain methods aim to correct the sinogram for the subsequent CT image reconstruction [6, 33]. Image-domain-based works have proposed different frameworks, such as simple residual network [8] and an interpretable structure [22, 23, 26], to learn artifact-reduced images from metal-affected ones. The dual-domain methods [12, 24, 25, 36] focus on the mutual learning between sinogram and CT image.

Albeit achieving promising performance, these aforementioned methods are sub-optimal for the SVMAR task. The main reasons are: 1) Most of them do not consider the joint influence of sparse data sampling and MAR, and do not fully embed the physical imaging constraint between the sinogram domain and CT image domain under the SVMAR scenario; 2) Although a few works focus on the joint SVMAR task, such as [36], the network structure is empirically built based on off-the-shelf modules, *e.g.*, U-Net and gated recurrent units, and it does not fully investigate and embed some important prior information underlying the CT imaging procedure. However, for such a highly ill-posed restoration problem, the introduction of the proper prior is important and valuable for constraining the network learning and helping it evolve in a right direction [24].

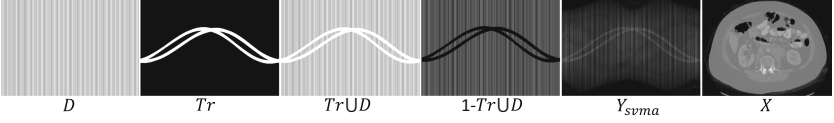


Fig. 1. Illustration of the elements in the model Eq. (2) for easy understanding.

To alleviate these issues, in this paper, we propose a model-driven equivariant proximal network, called MEPNet, which is naturally constructed based on the CT imaging geometry constraint for this specific SVMAR task, and takes into account the inherent prior structure underlying the CT scanning procedure. Concretely, we first propose a dual-domain reconstruction model and then correspondingly construct an unrolling network framework based on a derived optimization algorithm. Furthermore, motivated by the fact that the same organ can be imaged at different angles making the reconstruction task equivariant to rotation [2], we carefully formulate the proximal operator of the built unrolling neural network as a rotation-equivariant convolutional neural network (CNN). Compared with the standard-CNN-based proximal network with only translation-equivariance property [3], our proposed method effectively encodes more prior knowledge, *e.g.*, rotation equivariance, possessed by this specific task. With such more accurate regularization, our proposed MEPNet can achieve higher fidelity of anatomical structures and has better generalization capability with fewer network parameters. This is finely verified by comprehensive experiments on several datasets of different body sites. To the best of our knowledge, we should be the first to study rotation equivariance in the context of SVMAR and validate its utility, which is expected to make insightful impacts on the community.

2 Preliminary Knowledge About Equivariance

Equivariance of a mapping w.r.t. a certain transformation indicates that executing the transformation on the input produces a corresponding transformation on the output [3, 28]. Mathematically, given a group of transformations G , a mapping Φ from the input feature space to the output feature space is said to be group equivariant about G if

$$\Phi(T_g(f)) = T'_g(\Phi(f)), \forall g \in G, \quad (1)$$

where f is any input feature map in the input feature space; T_g and T'_g represent the actions of g on the input and output, respectively.

The prior work [3] has shown that adopting group equivariant CNNs to encode symmetries into networks would bring data efficiency and it can constrain the network learning for better generalization. For example, compared with the fully-connected layer, the translational equivariance property enforces weight sharing for the conventional CNN, which makes CNN use fewer parameters to preserve the

representation capacity and then obtain better generalization ability. Recently, different types of equivariant CNNs have been designed to preserve more symmetries beyond current CNNs, such as rotation symmetry [1, 2, 27] and scale symmetry [7, 20]. However, most of these methods do not consider specific designs for the SVMAR reconstruction. In this paper, we aim to build a physics-driven network for the SVMAR task where rotation equivariance is encoded.

3 Dual-Domain Reconstruction Model for SVMAR

In this section, for the SVMAR task, we derive the corresponding dual domain reconstruction model and give an iterative algorithm for solving it.

Dual-Domain Reconstruction Model. Given the captured sparse-view metal-affected sinogram $Y_{svma} \in \mathbb{R}^{N_b \times N_p}$, where N_b and N_p are the number of detector bins and projection views, respectively, to guarantee the data consistency between the reconstructed clean CT image $X \in \mathbb{R}^{H \times W}$ and the observed sinogram Y_{svma} , we can formulate the corresponding optimization model as [36]:

$$\min_X \|(1 - Tr \cup D) \odot (\mathcal{P}X - Y_{svma})\|_F^2 + \mu R(X), \quad (2)$$

where $D \in \mathbb{R}^{N_b \times N_p}$ is the binary sparse downsampling matrix with 1 indicating the missing region; $Tr \in \mathbb{R}^{N_b \times N_p}$ is the binary metal trace with 1 indicating the metal-affected region; \mathcal{P} is forward projection; $R(\cdot)$ is a regularization function for capturing the prior of X ; \cup is the union set; \odot is the point-wise multiplication; H and W are the height and width of CT images, respectively; μ is a trade-off parameter. One can refer to Fig. 1 for easy understanding.

To jointly reconstruct sinogram and CT image, we introduce the dual regularizers $R_1(\cdot)$ and $R_2(\cdot)$, and further derive Eq. (2) as:

$$\min_{S, X} \|\mathcal{P}X - S\|_F^2 + \lambda \|(1 - Tr \cup D) \odot (S - Y_{svma})\|_F^2 + \mu_1 R_1(S) + \mu_2 R_2(X), \quad (3)$$

where S is the to-be-estimated clean sinogram; λ is a weight factor. Following [24], we rewrite S as $\bar{Y} \odot \bar{S}$ for stable learning, where \bar{Y} and \bar{S} are the normalization coefficient implemented via the forward projection of a prior image, and the normalized sinogram, respectively. Then we can get the final dual-domain reconstruction model for this specific SVMAR task as:

$$\min_{\bar{S}, X} \|\mathcal{P}X - \bar{Y} \odot \bar{S}\|_F^2 + \lambda \|(1 - Tr \cup D) \odot (\bar{Y} \odot \bar{S} - Y_{svma})\|_F^2 + \mu_1 R_1(\bar{S}) + \mu_2 R_2(X). \quad (4)$$

As observed, given Y_{svma} , we need to jointly estimate \bar{S} and X . For $R_1(\cdot)$ and $R_2(\cdot)$, the design details are presented below.

Iterative Optimization Algorithm. To solve the model (4), we utilize the classical proximal gradient technique [15] to alternatively update the variables \bar{S} and X . At iterative stage k , we can get the corresponding iterative rules:

$$\begin{aligned} \bar{S}_k &= \text{prox}_{\mu_1 \eta_1}(\bar{S}_{k-1} - \eta_1(\bar{Y} \odot (\bar{Y} \odot \bar{S}_{k-1} - \mathcal{P}X_{k-1}) + \lambda(1 - Tr \cup D) \odot \bar{Y} \odot (\bar{Y} \odot \bar{S}_{k-1} - Y_{svma}))), \\ X_k &= \text{prox}_{\mu_2 \eta_2}(X_{k-1} - \eta_2 \mathcal{P}^T(\mathcal{P}X_{k-1} - \bar{Y} \odot \bar{S}_k)), \end{aligned} \quad (5)$$

where η_i is stepsize; $\text{prox}_{\mu_i \eta_i}(\cdot)$ is proximal operator, which relies on the regularization term $R_i(\cdot)$. For any variable, its iterative rule in Eq. (5) consists of two steps: an explicit gradient step to ensure data consistency and an implicit proximal computation $\text{prox}_{\mu_i \eta_i}(\cdot)$ which enforces the prior $R_i(\cdot)$ on the to-be-estimated variable. Traditionally, the prior form $R_i(\cdot)$ is empirically designed, *e.g.*, l_1 penalty, which may not always hold in real complicated scenarios. Due to the high representation capability, CNN has been adopted to adaptively learn the proximal step in a data-driven manner for various tasks [5, 14, 30]. Motivated by their successes, in the next section, we will deeply explore the prior of this specific SVMAR task and carefully construct the network for $\text{prox}_{\mu_i \eta_i}(\cdot)$.

4 Equivariant Proximal Network for SVMAR

By unfolding the iterative rules (5) for K iterations, we can easily build the unrolling neural network. Specifically, at iteration k , the network structure is sequentially composed of:

$$\begin{aligned} \bar{S}_k &= \text{proxNet}_{\theta_{\bar{s}}^{(k)}}(\bar{S}_{k-1} - \eta_1(\bar{Y} \odot (\bar{Y} \odot \bar{S}_{k-1} - \mathcal{P}X_{k-1}) + \lambda(1 - \text{Tr} \cup D) \odot \bar{Y} \odot (\bar{Y} \odot \bar{S}_{k-1} - Y_{svma}))), \\ X_k &= \text{proxNet}_{\theta_x^{(k)}}(X_{k-1} - \eta_2 \mathcal{P}^T(\mathcal{P}X_{k-1} - \bar{Y} \odot \bar{S}_k)), \end{aligned} \quad (6)$$

where $\text{proxNet}_{\theta_{\bar{s}}^{(k)}}$ and $\text{proxNet}_{\theta_x^{(k)}}$ are proximal networks with parameters $\theta_{\bar{s}}^{(k)}$ and $\theta_x^{(k)}$ to execute the proximal operators $\text{prox}_{\mu_1 \eta_1}$ and $\text{prox}_{\mu_2 \eta_2}$, respectively.

To build $\text{proxNet}_{\theta_{\bar{s}}^{(k)}}$, we follow [24] and choose a standard-CNN-based structure with four [*Conv+BN+ReLU+Conv+BN+Skip Connection*] residual blocks, which do not change image sizes. While for $\text{proxNet}_{\theta_x^{(k)}}$, we carefully investigate that during the CT scanning, the same body organ can be imaged at different rotation angles. However, the conventional CNN for modeling $\text{proxNet}_{\theta_x^{(k)}}$ in [24] has only the translation equivariance property and it cannot preserve such an intrinsic rotation equivariance structure [3]. Against this issue, we propose to replace the standard CNN in [24] with a rotation equivariant CNN. Then we can embed more useful prior, such as rotation equivariance, to constrain the network, which would further boost the quality of reconstructed CT images (refer to Sect. 5.2).

Specifically, from Eq. (1), for a rotation group G and any input feature map f , we expect to find a properly parameterized convolutional filter ψ which is group equivariant about G , satisfying

$$[T_\theta[f]] \star \psi = T_\theta[f \star \psi] = f \star \pi_\theta[\psi], \forall \theta \in G, \quad (7)$$

where π_θ is a rotation operator. Due to its solid theoretical foundation, the Fourier-series-expansion-based method [28] is adopted to parameterize ψ as:

$$\psi(x) = \sum_{m=0}^{p-1} \sum_{n=0}^{p-1} a_{mn} \varphi_{mn}^c(x) + b_{mn} \varphi_{mn}^s(x), \quad (8)$$

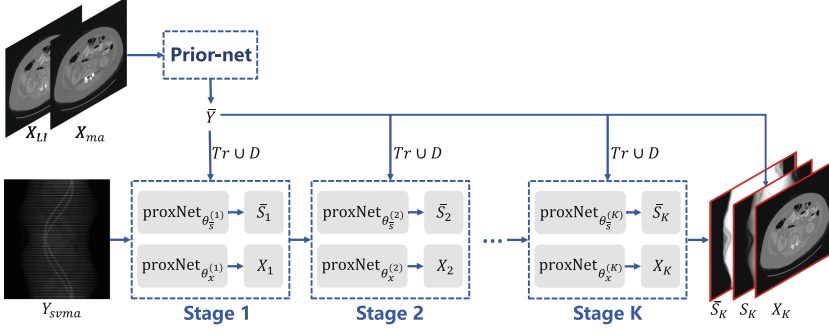


Fig. 2. The framework of the proposed MEPNet where “prior-net” is designed in [24].

where $x = [x_i, x_j]^T$ is 2D spatial coordinates; a_{mn} and b_{mn} are learnable expansion coefficients; $\varphi_{mn}^c(x)$ and $\varphi_{mn}^s(x)$ are 2D fixed basis functions as designed in [28]; p is chosen to be 5 in experiments. The action π_θ on ψ in Eq. (7) can be achieved by coordinate transformation as:

$$\pi_\theta[\psi](x) = \psi(U_\theta^{-1}x), \text{ where } U_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \forall \theta \in G. \quad (9)$$

Based on the parameterized filter in Eq. (8), we follow [28] to implement the rotation-equivariant convolution for the discrete domain. Compared with other types, *e.g.*, harmonics and partial-differential-operator-like bases [19, 27], the basis in Eq. (8) has higher representation accuracy, especially when being rotated.

By implementing $\text{proxNet}_{\theta_s^{(k)}}$ and $\text{proxNet}_{\theta_x^{(k)}}$ in Eq. (6) with the standard CNN and the rotation-equivariant CNN with the $p8$ group,¹ respectively, we can then construct the model-driven equivariant proximal network, called MEPNet, as shown in Fig. 2. The expansion coefficients, $\{\theta_s^{(k)}\}_{k=1}^K$, θ_{prior} for learning \bar{Y} [24], η_1 , η_2 , and λ , are all flexibly learned from training data end-to-end.

Remark: Our MEPNet is indeed inspired by InDuDoNet [24]. However, MEPNet contains novel and challenging designs: 1) It is specifically constructed based on the physical imaging procedure for the SVMAR task, leading to a clear working mechanism; 2) It embeds more prior knowledge, *e.g.*, rotation equivariance, via advanced filter parametrization method, which promotes better reconstruction; 3) It is desirable that the usage of more transformation symmetries would further decrease the number of model parameters and improve the generalization. These advantages are validated in the experiments below.

¹ Considering performance and efficiency, we follow [28] and chose $p8$ group for discretized equivariance convolution on CT images. The parameterized filters for eight different rotation orientations share a set of expansion coefficients, largely reducing the network parameters (validated in Sect. 5.2).

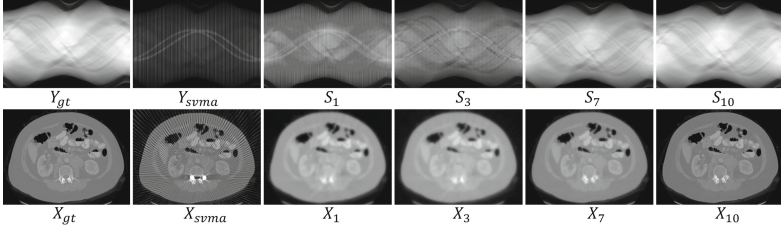


Fig. 3. The sinogram S_k and CT image X_k reconstructed by our MEPNet ($K = 10$).

5 Experiments

5.1 Details Description

Datasets and Metrics. Consistent with [24], we synthesize the training set by randomly selecting 1000 clean CT images from the public DeepLesion [29] and collecting 90 metals with various sizes and shapes from [33]. Specifically, following the CT imaging procedure with fan-beam geometry in [24, 31, 36], all the CT images are resized as 416×416 pixels where pixel spacing is used for normalization, and 640 fully-sampled projection views are uniformly spaced in 360° . To synthesize sparse-view metal-affected sinogram Y_{svma} , similar to [36], we uniformly sample 80, 160, and 320 projection views to mimic 8, 4, and 2-fold radiation dose reduction. By executing the FBP process on Y_{svma} , we can obtain the degraded CT image X_{svma} .

The proposed method is tested on three datasets including DeepLesion-test (2000 pairs), Pancreas-test (50 pairs), and CLINIC-test (3397 pairs). Specifically, DeepLesion-test is generated by pairing another 200 clean CT images from DeepLesion [29] with 10 extra testing metals from [33]. Pancreas-test is formed by randomly choosing 5 patients with 50 slices from Pancreas CT [17] and pairing each slice with one randomly-selected testing metal. CLINIC-test is synthesized by pairing 10 volumes with 3397 slices randomly chosen from CLINIC [13] with one testing metal slice-by-slice. The 10 testing metals have different sizes as [35] in pixels. For evaluation on different sizes of metals as listed in Table 2 below, we merge the adjacent two sizes into one group. Following [12, 24], we adopt peak signal-to-noise ratio (PSNR) and structured similarity index (SSIM) for quantitative analysis.

Implementation Details. Our MEPNet is trained end-to-end with a batch size of 1 for 100 epochs based on PyTorch [16] on an NVIDIA Tesla V100-SMX2 GPU card. An Adam optimizer with parameters $(\beta_1, \beta_2) = (0.5, 0.999)$ is exploited. The initial learning rate is 2×10^{-4} and it is decayed by 0.5 every 40 epochs. For a fair comparison, we adopt the same loss function as [24] and also select the total number of iterations K as 10.

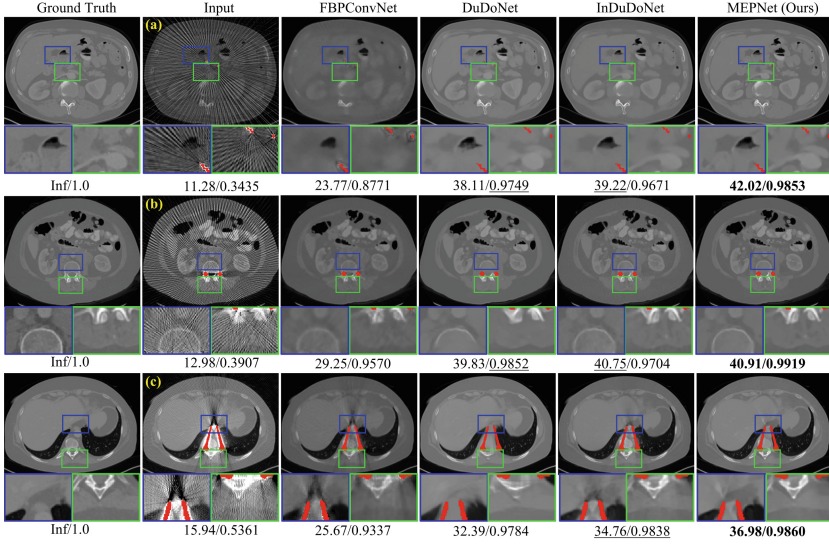


Fig. 4. DeepLesion-test: Artifact-reduced images (the corresponding PSNRs/SSIMs are shown below) of the comparing methods under different sparse-view under-sampling rates (a) $\times 8$, (b) $\times 4$, (c) $\times 2$, and various sizes of metals marked by red pixels. (Color figure online)

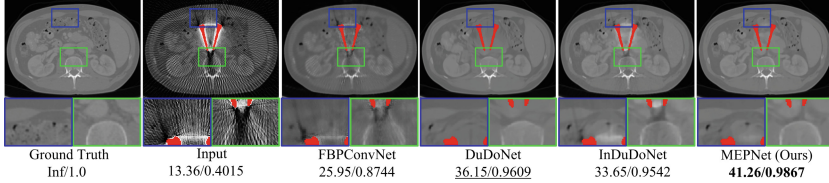


Fig. 5. Cross-domain: Reconstruction results with the corresponding PSNR/SSIM of different methods on Pancreas-test under $\times 4$ under-sampling rate.

5.2 Performance Evaluation

Working Mechanism. Figure 3 presents the sinogram S_k and CT image X_k reconstructed by MEPNet at different stages. We can easily observe that S_k and X_k are indeed alternatively optimized in information restoration and artifact reduction, approaching the ground truth Y_{gt} and X_{gt} , respectively. This finely shows a clear working mechanism of our proposed MEPNet, which evolves in the right direction specified by Eq. (5).

Visual Comparison. Figure 4 shows the reconstructed results of different methods, including FBPCConvNet [9], DuDoNet [12], InDuDoNet [24], and the proposed MEPNet, on three degraded images from DeepLesion-test with different

sparse-view under-sampling rates and various sizes of metallic implants.² As seen, compared with these baselines, our proposed MEPNet can consistently produce cleaner outputs with stronger artifact removal and higher structural fidelity, especially around the metals, thus leading to higher PSNR/SSIM values.

Figure 5 presents the cross-domain results on Pancreas-test with $\times 4$ under-sampling rate where DL-based methods are trained on synthesized DeepLesion. As seen, DuDoNet produces over-smoothed output due to the lack of physical geometry constraint on the final result. In contrast, MEPNet achieves more efficient artifact suppression and sharper detail preservation. Such favorable generalization ability is mainly brought by the dual-domain joint regularization and the fine utilization of rotation symmetry via the equivariant network, which can reduce the model parameters from 5,095,703 (InDuDoNet) to 4,723,309 (MEPNet). Besides, as observed from the bone marked by the green box, MEPNet alleviates the rotational-structure distortion generally existing in other baselines. This finely validates the effectiveness of embedding rotation equivariance.

Quantitative Evaluation. Table 1 lists the average PSNR/SSIM on three testing sets. It is easily concluded that with the increase of under-sampling rates, all these comparison methods present an obvious performance drop. Nevertheless, our MEPNet still maintains higher PSNR/SSIM scores on different testing sets, showing good superiority in generalization capability. Table 2 reports the results on the DeepLesion-test with different sizes of metals under the $\times 4$ sparse-view under-sampling rate. We can observe that MEPNet almost outperforms others, especially for the large metal setting, showing good generality.³

Table 1. Average PSNR (dB) and SSIM of different methods on three testing sets.

Methods	DeepLesion-test			Pancreas-test			CLINIC-test		
	$\times 8$	$\times 4$	$\times 2$	$\times 8$	$\times 4$	$\times 2$	$\times 8$	$\times 4$	$\times 2$
Input	12.65	13.63	16.55	12.37	13.29	16.04	13.95	14.99	18.04
	0.3249	0.3953	0.5767	0.3298	0.3978	0.5645	0.3990	0.4604	0.6085
FBPConvNet [9]	25.91	27.38	29.11	24.24	25.44	26.85	27.92	29.62	31.56
	0.8467	0.8851	0.9418	0.8261	0.8731	0.9317	0.8381	0.8766	0.9362
DuDoNet [12]	34.33	36.83	38.18	30.54	35.14	36.97	34.47	37.34	38.81
	0.9479	0.9634	0.9685	0.9050	0.9527	0.9653	0.9157	0.9493	0.9598
InDuDoNet [24]	37.50	40.24	40.71	36.86	38.17	38.22	38.39	39.67	40.86
	0.9664	0.9793	0.9890	0.9664	0.9734	0.9857	0.9572	0.9621	0.9811
MEPNet (Ours)	38.48	41.43	42.66	36.76	40.69	41.17	39.04	41.58	42.30
	0.9767	0.9889	0.9910	0.9726	0.9872	0.9896	0.9654	0.9820	0.9857

Table 2. Average PSNR (dB)/SSIM of the comparing methods on DeepLesion-test with the $\times 4$ under-sampling rate and different sizes of metallic implants.

Methods	Large Metal	→	Small Metal	Average
Input	13.68/0.3438	13.63/0.3736	13.61/0.4046	13.61/0.4304
	13.60/0.4240			13.63/0.3953
FBPConvNet [9]	26.15/0.7865	26.96/0.8689	27.77/0.9154	27.98/0.9216
	28.03/0.9331			27.38/0.8851
DuDoNet [12]	31.73/0.9519	33.89/0.9599	37.81/0.9667	40.19/0.9688
	40.54/0.9696			36.83/0.9634
InDuDoNet [24]	33.78/0.9540	38.15/0.9746	41.96/0.9873	43.48/0.9898
	43.83/0.9910			40.24/0.9793
MEPNet (Ours)	37.51/0.9797	39.45/0.9879	42.78/0.9920	43.92/0.9924
	43.51/0.9924			41.31/0.9889

² Here InDuDoNet is a particularly strong baseline and it is exactly an ablation study, which is the degenerated form of MEPNet with removing group equivariance.

³ More experimental results are included in *supplementary material*.

6 Conclusion and Future Work

In this paper, for the SVMAR task, we have constructed a dual-domain reconstruction model and built an unrolling model-driven equivariant network, called MEPNet, with a clear working mechanism and strong generalization ability. These merits have been substantiated by extensive experiments. Our proposed method can be easily extended to more applications, including limited-angle and low-dose reconstruction tasks. A potential limitation is that consistent with [24, 36], the data pairs are generated based on the commonly-adopted protocol, which would lead to a domain gap between simulation settings and clinical scenarios. In the future, we will try to collect clinical data captured in the sparse-view metal-inserted scanning configuration to evaluate our method.

Acknowledgements. This work was supported by the National Key R&D Program of China under Grant 2020AAA0109500/2020AAA0109501.

References

1. Celledoni, E., Ehrhardt, M.J., Etmann, C., Owren, B., Schönlieb, C.B., Sherry, F.: Equivariant neural networks for inverse problems. *Inverse Prob.* **37**(8), 085006 (2021)
2. Chen, D., Tachella, J., Davies, M.E.: Equivariant imaging: learning beyond the range space. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4379–4388 (2021)
3. Cohen, T., Welling, M.: Group equivariant convolutional networks. In: *International Conference on Machine Learning*, pp. 2990–2999 (2016)
4. Ding, Q., Ji, H., Gao, H., Zhang, X.: Learnable multi-scale Fourier interpolation for sparse view CT image reconstruction. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 286–295 (2021)
5. Fu, J., Wang, H., Xie, Q., Zhao, Q., Meng, D., Xu, Z.: KXNet: a model-driven deep neural network for blind super-resolution. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) *Computer Vision, ECCV 2022*. LNCS, vol. 13679, pp. 235–253. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19800-7_14
6. Ghani, M.U., Karl, W.C.: Fast enhanced CT metal artifact reduction using data domain deep learning. *IEEE Trans. Comput. Imag.* **6**, 181–193 (2019)
7. Gunel, B., et al.: Scale-equivariant unrolled neural networks for data-efficient accelerated MRI reconstruction. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 737–747 (2022)
8. Huang, X., Wang, J., Tang, F., Zhong, T., Zhang, Y.: Metal artifact reduction on cervical CT images by deep residual learning. *Biomed. Eng. Online* **17**(1), 1–15 (2018)
9. Jin, K.H., McCann, M.T., Froustey, E., Unser, M.: Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**(9), 4509–4522 (2017)
10. Lee, H., Lee, J., Kim, H., Cho, B., Cho, S.: Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. *IEEE Trans. Radiat. Plasma Med. Sci.* **3**(2), 109–119 (2018)

11. Liao, H., Lin, W.A., Zhou, S.K., Luo, J.: ADN: artifact disentanglement network for unsupervised metal artifact reduction. *IEEE Trans. Med. Imaging* **39**(3), 634–643 (2019)
12. Lin, W.A., et al.: DuDoNet: dual domain network for CT metal artifact reduction. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10512–10521 (2019)
13. Liu, P., et al.: Deep learning to segment pelvic bones: large-scale CT datasets and baseline models. *arXiv preprint [arXiv:2012.08721](https://arxiv.org/abs/2012.08721)* (2020)
14. Liu, X., Xie, Q., Zhao, Q., Wang, H., Meng, D.: Low-light image enhancement by retinex-based algorithm unrolling and adjustment. *IEEE Trans. Neural Netw. Learn. Syst.* (2023)
15. Parikh, N., Boyd, S., et al.: Proximal algorithms. *Found. Trends Optim.* **1**(3), 127–239 (2014)
16. Paszke, A., et al.: Automatic differentiation in PyTorch (2017)
17. Roth, H.R., et al.: DeepOrgan: multi-level deep convolutional networks for automated pancreas segmentation. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 556–564 (2015)
18. Shen, L., Pauly, J., Xing, L.: NeRP: implicit neural representation learning with prior embedding for sparsely sampled image reconstruction. *IEEE Trans. Neural Netw. Learn. Syst.* (2022)
19. Shen, Z., He, L., Lin, Z., Ma, J.: PDO-eConvs: partial differential operator based equivariant convolutions. In: *International Conference on Machine Learning*, pp. 8697–8706 (2020)
20. Sosnovik, I., Szmaja, M., Smeulders, A.: Scale-equivariant steerable networks. *arXiv preprint [arXiv:1910.11093](https://arxiv.org/abs/1910.11093)* (2019)
21. Wang, C., Shang, K., Zhang, H., Li, Q., Hui, Y., Zhou, S.K.: DuDoTrans: dual-domain transformer provides more attention for sinogram restoration in sparse-view CT reconstruction. *arXiv preprint [arXiv:2111.10790](https://arxiv.org/abs/2111.10790)* (2021)
22. Wang, H., Li, Y., He, N., Ma, K., Meng, D., Zheng, Y.: DICDNet: deep interpretable convolutional dictionary network for metal artifact reduction in CT images. *IEEE Trans. Med. Imaging* **41**(4), 869–880 (2021)
23. Wang, H., Li, Y., Meng, D., Zheng, Y.: Adaptive convolutional dictionary network for CT metal artifact reduction. *arXiv preprint [arXiv:2205.07471](https://arxiv.org/abs/2205.07471)* (2022)
24. Wang, H., et al.: InDuDoNet: an interpretable dual domain network for CT metal artifact reduction. In: *Medical Image Computing and Computer Assisted Intervention*, pp. 107–118 (2021)
25. Wang, H., Li, Y., Zhang, H., Meng, D., Zheng, Y.: InDuDoNet+: a deep unfolding dual domain network for metal artifact reduction in CT images. *Med. Image Anal.* **85**, 102729 (2022)
26. Wang, H., Xie, Q., Li, Y., Huang, Y., Meng, D., Zheng, Y.: Orientation-shared convolution representation for CT metal artifact learning. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 665–675 (2022)
27. Weiler, M., Hamprecht, F.A., Storath, M.: Learning steerable filters for rotation equivariant CNNs. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 849–858 (2018)
28. Xie, Q., Zhao, Q., Xu, Z., Meng, D.: Fourier series expansion based filter parametrization for equivariant convolutions. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 4537–4551 (2022)

29. Yan, K., 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: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9261–9270 (2018)
30. Yang, Y., Sun, J., Li, H., Xu, Z.: ADMM-Net: a deep learning approach for compressive sensing MRI. arXiv preprint [arXiv:1705.06869](https://arxiv.org/abs/1705.06869) (2017)
31. Yu, L., Zhang, Z., Li, X., Xing, L.: Deep sinogram completion with image prior for metal artifact reduction in CT images. *IEEE Trans. Med. Imaging* **40**(1), 228–238 (2020)
32. Zhang, H., Liu, B., Yu, H., Dong, B.: MetaInv-Net: meta inversion network for sparse view CT image reconstruction. *IEEE Trans. Med. Imaging* **40**(2), 621–634 (2020)
33. Zhang, Y., Yu, H.: Convolutional neural network based metal artifact reduction in X-ray computed tomography. *IEEE Trans. Med. Imaging* **37**(6), 1370–1381 (2018)
34. Zhang, Z., Liang, X., Dong, X., Xie, Y., Cao, G.: A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution. *IEEE Trans. Med. Imaging* **37**(6), 1407–1417 (2018)
35. Zhang, Z., Yu, L., Liang, X., Zhao, W., Xing, L.: TransCT: dual-path transformer for low dose computed tomography. In: Medical Image Computing and Computer Assisted Intervention, pp. 55–64 (2021)
36. Zhou, B., Chen, X., Zhou, S.K., Duncan, J.S., Liu, C.: DuDoDR-Net: dual-domain data consistent recurrent network for simultaneous sparse view and metal artifact reduction in computed tomography. *Med. Image Anal.* **75**, 102289 (2022)
37. Zhou, B., Zhou, S.K., Duncan, J.S., Liu, C.: Limited view tomographic reconstruction using a cascaded residual dense spatial-channel attention network with projection data fidelity layer. *IEEE Trans. Med. Imaging* **40**(7), 1792–1804 (2021)