—Supplementary Material—
Unsupervised Polychromatic Neural Representation for CT Metal Artifact Reduction

Qing Wu♣ **Lixuan Chen**♣ **Ce Wang**♢ **Hongjiang Wei**♡
S. Kevin Zhou♠ **Jingyi Yu**♣ **Yuyao Zhang**♣,*

♣ShanghaiTech University
♡Shanghai Jiao Tong University
♠University of Science and Technology of China
♢Institute of Computing Technology, Chinese Academy of Sciences
{wuqing, chenlx1, yujingyi, zhangyy8}@shanghaitech.edu.cn
wangce@ict.ac.cn skevinzhou@ustc.edu.cn hongjiang.wei@sjtu.edu.cn
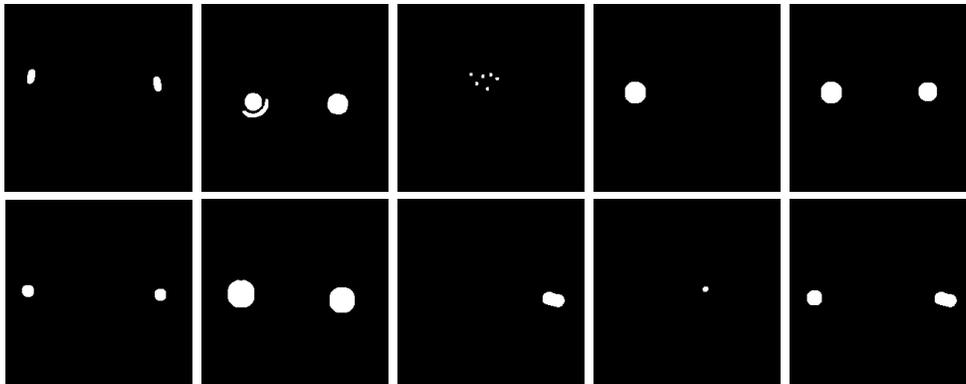
Figure 1: Ten shapes of metals for synthesizing metal-corrupted measurements in the DeepLesion [1] dataset. These metals are supposed as Titanium. The white regions denote metals.

1 Additional Details of Datasets

In this work, we perform experiments on four datasets: DeepLesion [1], XCOM [2], and our Walnut Sample. DeepLesion and XCOM are simulation datasets, while Walnut Sample and Mouse Thigh are real-world datasets. *Note that our method is fully unsupervised, and thus all the data are exclusively used for testing purposes.*

DeepLesion.   To simulate metal-corrupted measurements in the DeepLesion dataset [1], we follow the pipelines outlined in [4, 5, 6]. For our experiments, we extract 200 2D images from the DeepLesion dataset as test GT samples [1]. As depicted in Fig. 1, we leverage ten shapes of metallic implants from [4, 5, 6] and consider them Titanium. To simulate the X-ray source, we employ a polychromatic X-ray with an energy range of [20, 120] KeV and a minimum energy unit of 1 KeV. The number of photons emitted by the X-ray source is set to 2×10^7. The normalized energy spectrum η of the X-ray source is illustrated in Fig. 2 (*Left*). We adopt an equiangular fan-beam CT acquisition geometry,

---

*Corresponding author.
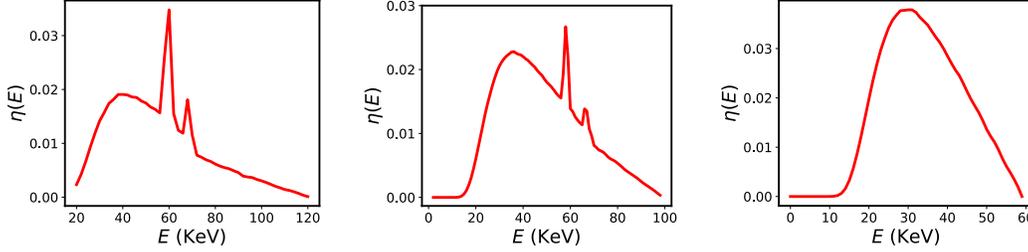
Figure 2: (*Left*) The simulated spectrum within an energy range of [20, 120] for synthesizing metal-corrupted measurements of the DeepLesion [1] and XCOM [2] datasets. The spectrums estimated within the energy ranges of [0, 100] and [0, 60] estimated by the SPEKTR toolkit [3] for the real walnut sample (*Middle*) and mouse thigh (*Right*).

Table 1: Detailed parameters of the acquisition geometry for the four datasets.

Parameters	Simulation Datasets		Real Datasets	
	DeepLesion [1]	XCOM [2]	Walnut	Mouse Thigh
Type of geometry	2D fan-beam	2D fan-beam	2D fan-beam	3D cone-beam
Source Voltage (kV)	120	120	100	60
Source Current (uA)	-	-	200	200
Exposure Time (ms)	-	-	276	730
Image size	256×256	512×512	650×650	200×200×150
Voxel size (mm)	1×1	0.8×0.8	0.05×0.05	0.06×0.06×0.06
Angle range (°)	[0, 360)	[0, 360)	[0, 360)	[0, 360)
The number of the angles	360	984	720	900
Angular spacing (°)	0.1	0.057	-	-
Detector spacing (mm)	-	-	0.069	0.069
Distance from source to center (mm)	362	743	92.602	92.602
Distance from center to detector (mm)	362	743	65.946	65.946

and the detailed parameters are provided in Table 1. Additionally, we incorporate Poisson noise and consider the partial volume effect in the sinogram domain during the simulation process.

XCOM. For the XCOM dataset [2], we use two samples provided by Zhang *et al*. [5]. These two cases are simulated using two 2D clean CT images sourced from the XCOM [2] database. Zhang *et al*. [5] consider a polychromatic X-ray source with an energy range of [20, 120] KeV and a minimum energy unit of 1 KeV. The corresponding normalized energy spectrum η is depicted in Fig. 2 (*Left*). In oder to generate metal-corrupted sinograms, an equiangular fan-beam CT acquisition geometry is employed, and the geometry parameters are specified in Table 1. Similar to the DeepLesion [1] dataset, Zhang *et al*. [5] also simulate Poisson noise and consider the partial volume effect in the sinogram domain during the simulation process.

Walnut Sample. To assess the performance of our proposed method on real CT data, we employ a commercial Bruker SKYSCAN 1276 micro-CT scanner to scan a walnut sample that contains a metal paper clip. Detailed parameters of the acquisition geometry can be found in Table 1. To estimate the X-ray spectrum of the micro-CT scanner, we leverage the SPEKTR toolkit [3]. The estimated spectrum is illustrated in Fig. 2 (*Middle*).

Mouse Thigh. We also scan a mouse thigh containing a metal intramedullary needle on the same micro-CT scanner. This sample is 3D cone-beam data. Detailed parameters of the acquisition geometry are shown in Table 1. We leverage the SPEKTR toolkit [3] to estimate the X-ray spectrum of the micro-CT scanner. The estimated spectrum is illustrated in Fig. 2 (*Right*).

2 Additional Details of Baselines

In our experiments, we compare our proposed method against eight baseline MAR approaches. *Notably, for the five DL-based methods, we evaluate their performance using the pre-trained models provided by the respective authors.*

FBP. The FBP [7] is a conventional approach used for reconstructing CT images. It involves re-projecting the acquired sinogram data onto the image domain using the corresponding projection angles and geometric parameters to obtain an approximate estimate of the unknown image. In our experiments, we use the in-build function `ifanbem` in MATLAB (`https://ww2.mathworks.cn/help/images/ref/ifanbeam.html?requestedDomain=cn`).

LI. A simple strategy to mitigate metal artifacts in CT imaging involves the direct linear interpolation of the sinogram [8] to fill in the regions affected by metal. This approach does not require any network training but may result in imperfect sinogram completion, which in turn could introduce undesired artifacts in the reconstructed image. In our experiments, we use the implementation provided by Zhang *et al.* [5] (`https://github.com/yanbozhang007/CNN-MAR/blob/master/cnnmar`).

NMAR. The NMAR [9] introduces a generalized normalization technique that extends previously developed interpolation-based MAR techniques (*e.g.*, LI [8]), which also does not require any network training. Specifically, it normalizes the projections before interpolation based on forward projections of a prior image obtained through multi-threshold segmentation of the initial image. In our experiments, we use the implementation provided by Zhang *et al.* [5] (`https://github.com/yanbozhang007/CNN-MAR/tree/master/cnnmar`).

CNN-MAR. Zhang *et al.* [5] proposed a CNN-based MAR framework, which uses a CNN to estimate a prior image and subsequently apply a sinogram correction. However, despite the strong representation ability of CNNs, these approaches are still susceptible to secondary artifacts resulting from inconsistent sinograms. In our experiments, we use its official implementation and pre-trained model (`https://github.com/yanbozhang007/CNN-MAR/tree/master/cnnmar`).

DICDNet. Wang *et al.* [10] propose a deep interpretable convolutional dictionary network (DICDNet) for the MAR task, which explicitly formulates the prior structures underlying metal artifacts in CT images as a convolutional dictionary model. In our experiments, we use its official implementation and pre-trained model (`https://github.com/hongwang01/DICDNet/tree/main`).

ACDNet. Similarly to DICDNet [10], the adaptive convolutional dictionary network (ACDNet) [11] explicitly encodes the prior observations underlying the MAR task into an adaptive convolutional dictionary network. In our experiments, we use its official implementation and pre-trained model (`https://github.com/hongwang01/ACDNet`).

ADN. The ADN [4] is an unsupervised learning approach for the MAR problem. Specifically, it takes unpaired metal-corrupted and clean CT images as inputs to learn the transformation between these two distributions. In our experiments, we use its official implementation and pre-trained model (`https://github.com/liaohaofu/adn/tree/master`).

Score-MAR. Song *et al.* [12] demonstrated that unconditional diffusion models can be adapted to various inverse problems, such as the MAR. Specifically, it learns the prior distribution of metal-free CT images with a generative model in order to infer the lost sinogram in the metal-affected regions. In our experiments, we use its official implementation and pre-trained model (`https://github.com/yang-song/score_inverse_problems`).

3 Additional Visual Results

Figs. 3, 4, 5, 6, and 7 demonstrate some additional visual results. We observe that the proposed Polyner generally obtains the best MAR results.

4 Broader Impacts

Our Polyner is expected to have significant broader impacts in the field of medical imaging. By effectively reducing metal artifacts in CT scans, our research has the potential to improve diagnostic accuracy, leading to more precise diagnoses and enhanced patient care. However, it is important to address potential limitations and concerns, such as the possibility of introducing false positive or negative results. Thorough evaluation and validation of our method is crucial to ensure its reliability and minimize any adverse effects. Overall, our work contributes to advancing medical imaging technology and has the potential to positively impact healthcare.

References

[1] Ke Yan, Xiaosong Wang, Le Lu, Ling Zhang, Adam P Harrison, Mohammadhadi Bagheri, and Ronald M Summers. 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*, pages 9261–9270, 2018.

[2] Martin J Berger. Xcom: Photon cross sections database, version 1.4. *http://physics. nist. gov/xcom*, 2009.

[3] J Punnoose, J Xu, A Sisniega, W Zbijewski, and JH Siewerdsen. Technical note: spektr 3.0—a computational tool for x-ray spectrum. *Medical Physics*, 43(8):4711–4717, 2016.

[4] H. Liao, W. Lin, S. K. Zhou, and J. Luo. Adn: Artifact disentanglement network for unsupervised metal artifact reduction. *IEEE Transactions on Medical Imaging*, 2019.

[5] Yanbo Zhang and Hengyong Yu. Convolutional neural network based metal artifact reduction in x-ray computed tomography. *IEEE transactions on medical imaging*, 37(6):1370–1381, 2018.

[6] Wei-An Lin, Haofu Liao, Cheng Peng, Xiaohang Sun, Jingdan Zhang, Jiebo Luo, Rama Chellappa, and Shaohua Kevin Zhou. Dudonet: Dual domain network for ct metal artifact reduction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10512–10521, 2019.

[7] Avinash C Kak and Malcolm Slaney. *Principles of computerized tomographic imaging*. SIAM, 2001.

[8] Willi A Kalender, Robert Hebel, and Johannes Ebersberger. Reduction of ct artifacts caused by metallic implants. *Radiology*, 164(2):576–577, 1987.

[9] Esther Meyer, Rainer Raupach, Michael Lell, Bernhard Schmidt, and Marc Kachelrieß. Normalized metal artifact reduction (nmar) in computed tomography. *Medical physics*, 37(10):5482–5493, 2010.

[10] Hong Wang, Yuexiang Li, Nanjun He, Kai Ma, Deyu Meng, and Yefeng Zheng. Dicdnet: Deep interpretable convolutional dictionary network for metal artifact reduction in ct images. *IEEE Transactions on Medical Imaging*, 41(4):869–880, 2021.

[11] Hong Wang, Yuexiang Li, Deyu Meng, and Yefeng Zheng. Adaptive convolutional dictionary network for ct metal artifact reduction. In *The 31st International Joint Conference on Artificial Intelligence*. IEEE, 2022.

[12] Yang Song, Liyue Shen, Lei Xing, and Stefano Ermon. Solving inverse problems in medical imaging with score-based generative models. In *International Conference on Learning Representations*, 2022.

[13] Lee A Feldkamp, Lloyd C Davis, and James W Kress. Practical cone-beam algorithm. *Josa a*, 1(6):612–619, 1984.
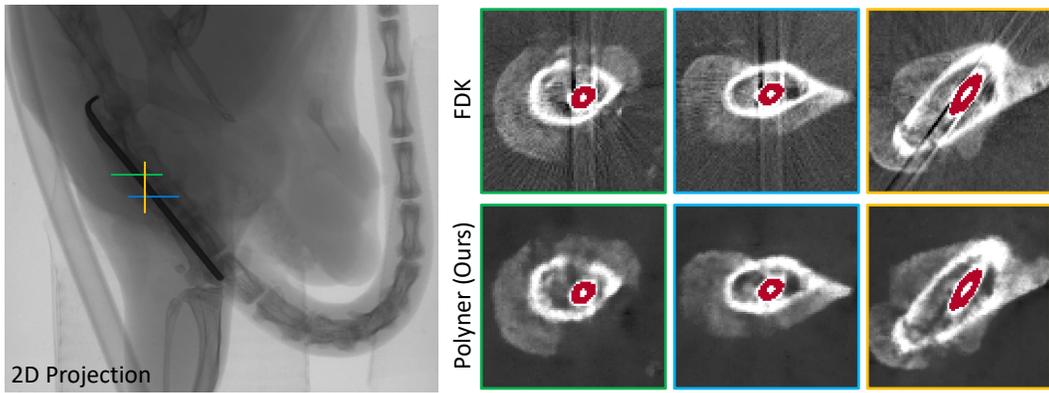
Figure 3: (*Left*) A sample among 2D projections of a mouse thigh containing a metal intramedullary needle scanned by the micro-CT scanner. (*Right*) Qualitative results of FDK [13] and our Polyner on the sample. Note that the acquisition geometry is the 3D cone beam. The reconstructed images have a size of 200×200×150. Our Polyner takes about 32 minutes on a single NVIDIA RTX TITAN GPU. The red regions denote the metal needle tubing. This data collection is approved ethically.
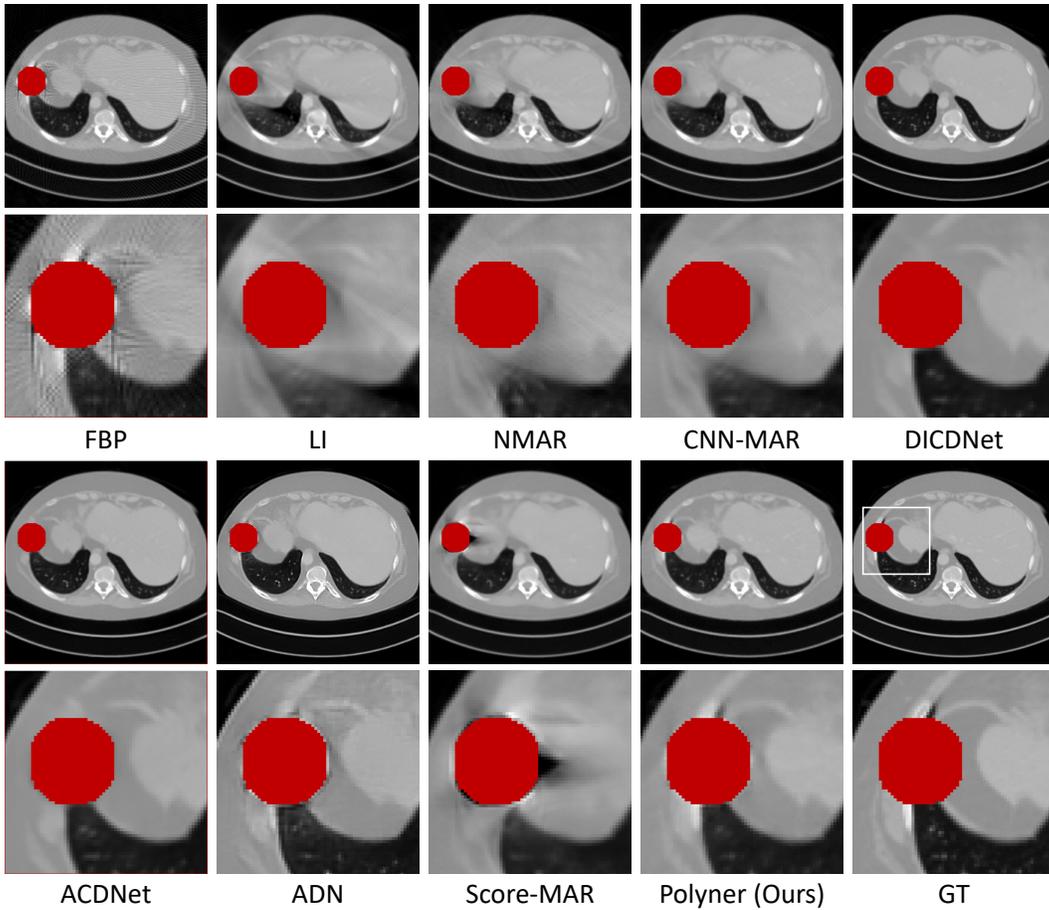
Figure 4: Qualitative results of the compared methods on a sample (#74) of DeepLesion [1] dataset. The white regions denote metals.
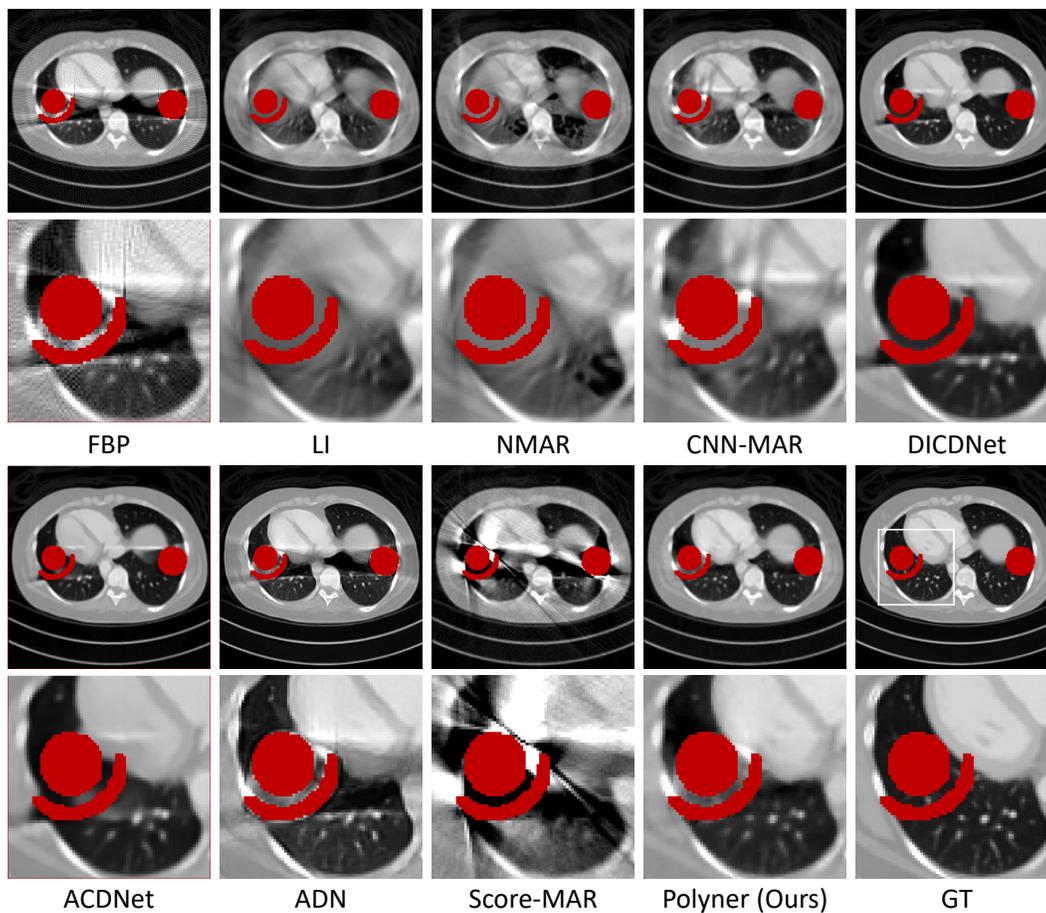
Figure 5: Qualitative results of the compared methods on a sample (#162) of DeepLesion [1] dataset. The white regions denote metals.
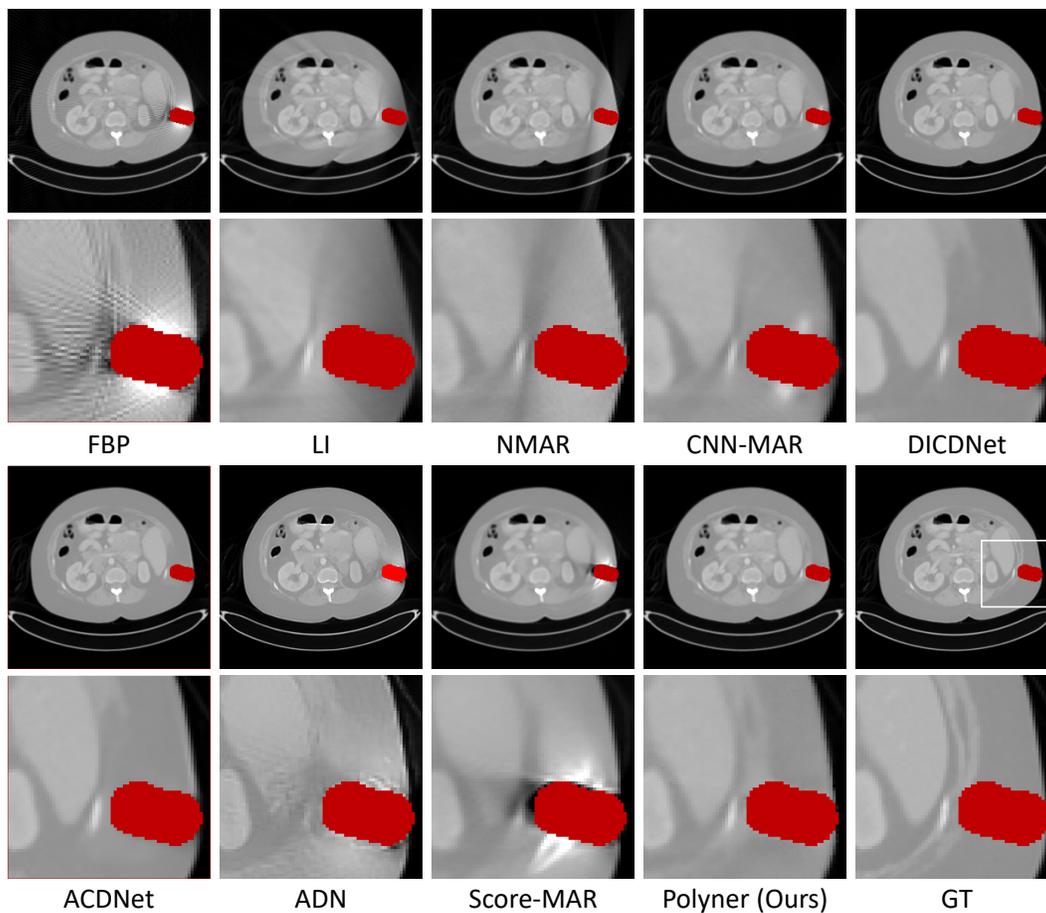
Figure 6: Qualitative results of the compared methods on a sample (#158) of DeepLesion [1] dataset. The white regions denote metals.
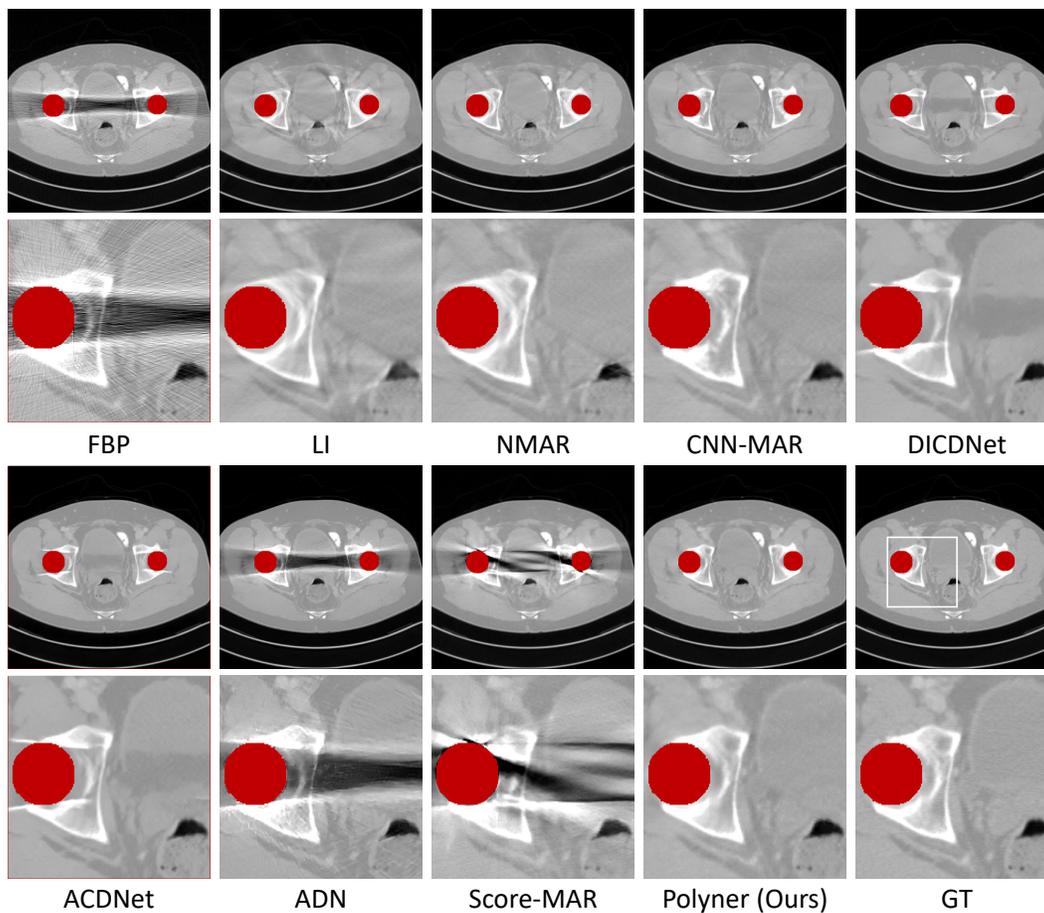
Figure 7: Qualitative results of the compared methods on a sample (#1) of XCOM [2] dataset. The white regions denote metals.