Supplementary Materials for Prompting multi-scale network with provably Lipschitz continuity for universal CT reconstruction

I. More visualization results

The quantitative comparisons of SVCT reconstruction methods across various view numbers are shown in Fig. 1. It can be observed that PRIOR-S exhibits superior reconstruction performance across all test views compared to the encoder-decoder architecture RED-CNN, the dual-domain architecture CNNMAR, and the multi-scale architecture MSANet. Compared to the best-performing MSANet among benchmark methods, the average performances of PRIOR-S are improved by 1.25 dB in PSNR, 0.0110 in SSIM, and reduced by 0.0017 in RMSE, respectively. This demonstrates the effectiveness of our proposed sparse representation model-driven network that takes full advantage of the multi-scale information. Fig. 2 presents the visualization of degraded features after adding prompt features in TGM for the LACT reconstruction task using t-SNE. Specifically, Fig. 2 (a) shows that features are indistinct without using prompts. In contrast, as illustrated in Fig. 2 (b), EP enables clear and linear classification. This indicates that our explicit prompt embedding module can achieve desirable view-distinction, thereby guiding reconstructions under multiple view numbers.

To further evaluate the effectiveness of PRIOR, we supplement the comparison results of different algorithms in the SVCT reconstruction task and LACT reconstruction task, as shown in Fig. 3 and Fig. 4. The reconstruction results of various SVCT reconstruction methods from 120 views on the AAPM dataset are presented in Fig. 3. The FBP method yields images of inferior quality, rendering crucial details and structures indistinct. While RED-CNN and FBPConvNet mitigate some artifacts, they lead to a loss of vital details. Moreover, CNNMAR and MSANet fall short of accurately recovering intricate details. Besides, results of DuDoTrans and FreeSeed exhibit blurred edges. Remarkably, PRIOR generates visual results with the most texture details and the least noise. Fig. 4 displays the reconstruction results of different LACT reconstruction methods from 150 views on the AAPM dataset. As we can observe, PRIOR consistently outperforms other methods. It yields enhanced organ boundary recovery, detailed structural information recovery, and boundary artifact elimination.

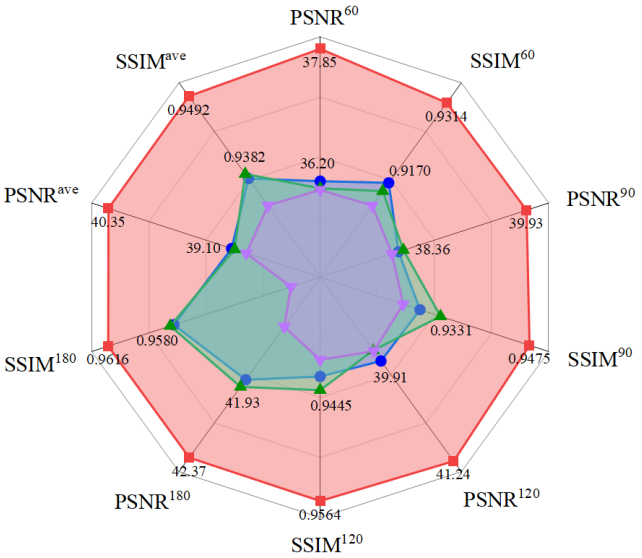


Fig. 1. Score comparisons with benchmark methods under various SVCT view numbers. We compare PRIOR-S (red) with the top 3 (blue, green and purple) benchmark SVCT reconstruction methods, i.e., RED-CNN [1], CNNMAR [2] and MSANet [3]. 60, 90, 120, and 180 are the view numbers we choose and their average scores are calculated for comparisons.

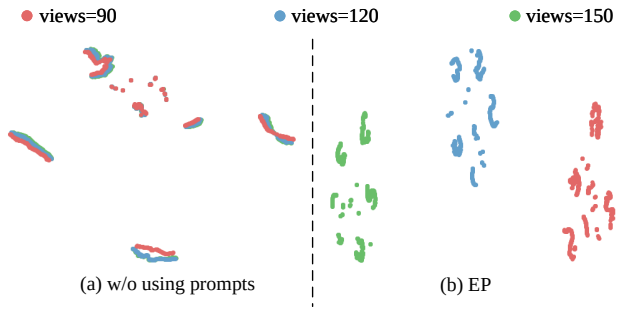


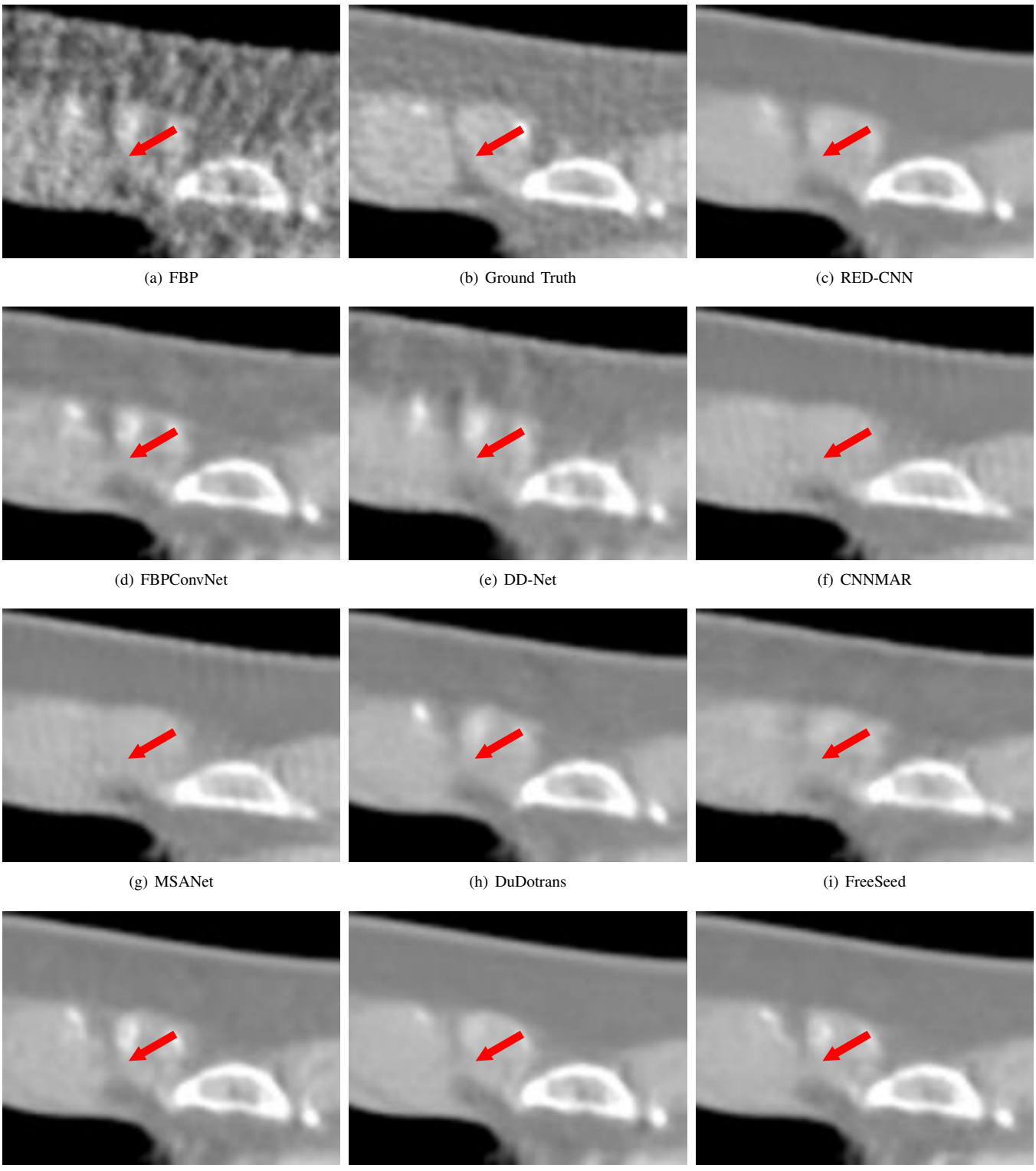
Fig. 2. The t-SNE plots of degraded features after adding prompt features in TGM for the LACT reconstruction task. Our explicit prompt strategy results in the clear distinction of features.

References

[1] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.

[2] M. D. Ketcha, M. Marrama, A. Souza, A. Uneri, P. Wu, X. Zhang, P. A. Helm, and J. H. Siewerdsen, "Sinogram + image domain neural network approach for metal artifact reduction in low-dose cone-beam computed tomography," *J. Med. Imaging*, vol. 8, no. 5, pp. 052 103–052 103, 2021.

[3] Y. Gou, P. Hu, J. Lv, J. T. Zhou, and X. Peng, "Multi-scale adaptive network for single image denoising," in *Advances in Neural Information Processing Systems*, vol. 35. Curran Associates, Inc., 2022, pp. 14 099–14 112.



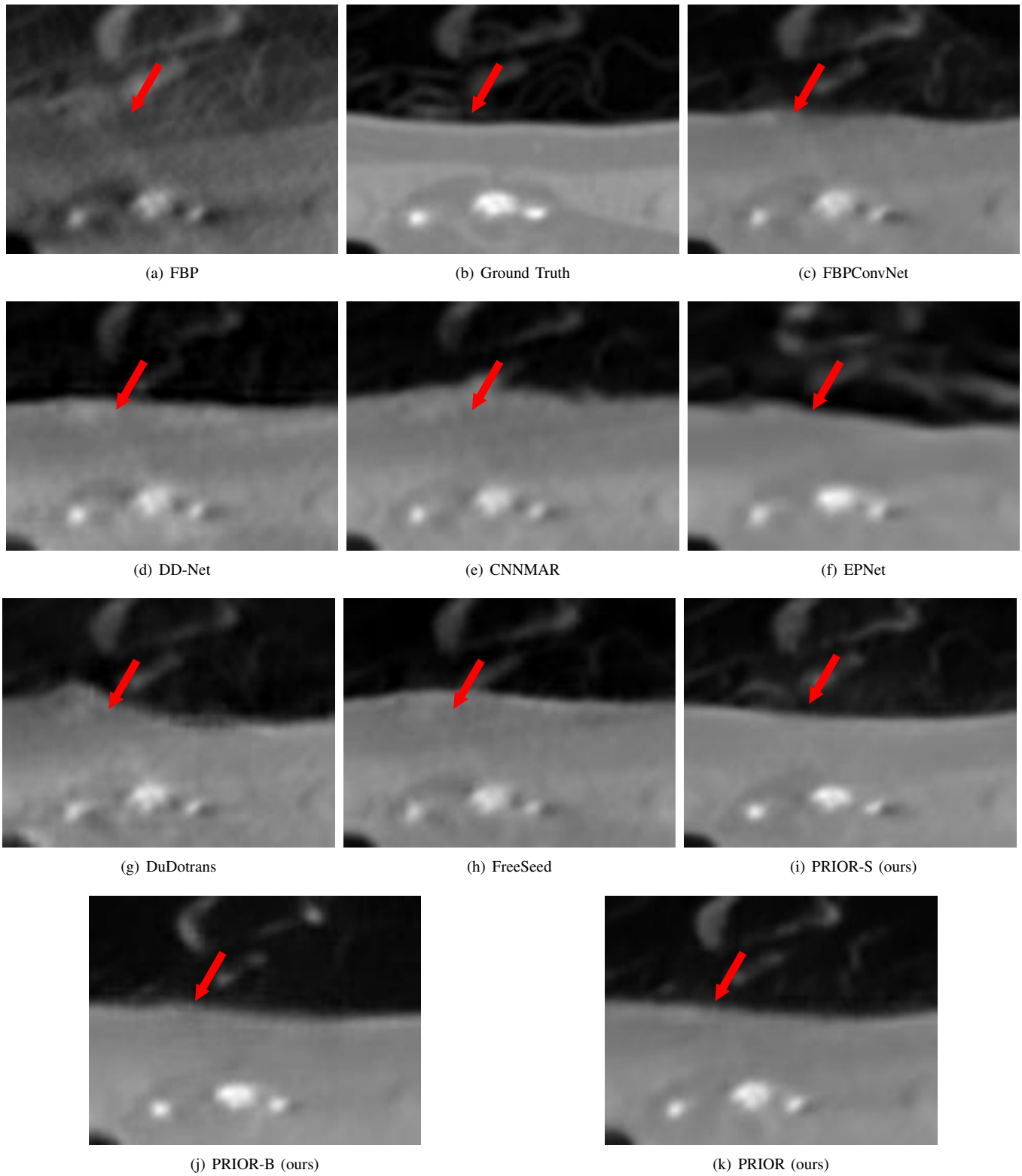
(a) FBP (b) Ground Truth (c) RED-CNN

(d) FBPConvNet (e) DD-Net (f) CNNMAR

(g) MSANet (h) DuDotrans (i) FreeSeed

(j) PRIOR-S (ours) (k) PRIOR-B (ours) (l) PRIOR (ours)

Fig. 3. SVCT reconstruction results recovered from 120 views using different methods. Regions of interest are zoomed in for better viewing. The display window of reconstructed results is [-600, 600] HU for a better observation of small details.



(a) FBP (b) Ground Truth (c) FBPConvNet

(d) DD-Net (e) CNNMAR (f) EPNet

(g) DuDotrans (h) FreeSeed (i) PRIOR-S (ours)

(j) PRIOR-B (ours) (k) PRIOR (ours)

Fig. 4. LACT reconstruction results recovered from 150 views using different methods. Regions of interest are zoomed in for better viewing. The display window of reconstructed results is [-1000, 1000] HU for a better observation of small details.