

# Supplementary Materials

Xiao-Xin Li<sup>1,2</sup>, Fang-Zheng Zhu<sup>1</sup>, Junwei Yang<sup>5</sup>, Yong Chen<sup>6</sup>, and Dinggang Shen<sup>2,3,4,\*</sup>

<sup>1</sup> College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China

<sup>2</sup> School of Biomedical Engineering & State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai 201210, China

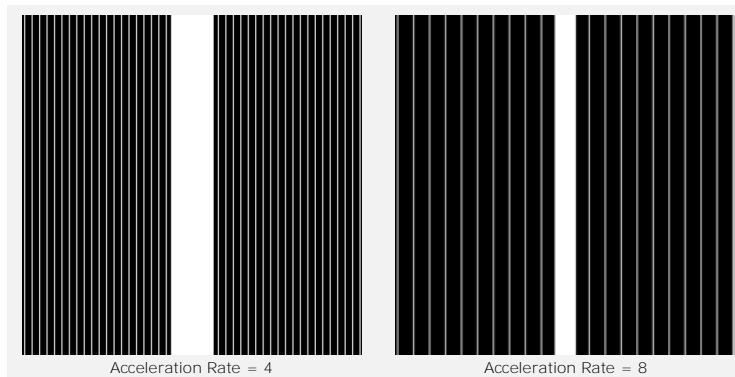
<sup>3</sup> Shanghai United Imaging Intelligence Co., Ltd., Shanghai 200232, China

<sup>4</sup> Shanghai Clinical Research and Trial Center, Shanghai 201210, China

<sup>5</sup> Department of Computer Science and Technology, University of Cambridge, Cambridge, UK

<sup>6</sup> Department of Radiology, Case Western Reserve University, Cleveland, Ohio, USA

## 1 Implementation Details



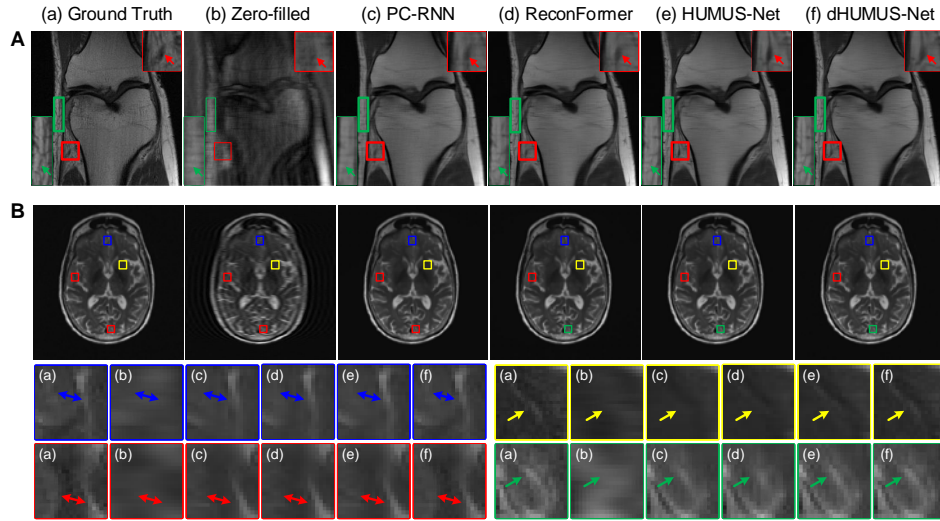
**Fig. 1.** The equispaced undersampling masks we used.

We adopt the equispaced undersampling masks [1] for  $4\times$  and  $8\times$  accelerations, as shown in Fig. 1. The equispaced undersampling masks continuously sample partial low-frequency lines from the center of  $k$ -space and uniformly sample high-frequency lines from the remaining  $k$ -space.

We use the lookahead version of Adam optimizer [2] with a momentum of 0.9 and a weight decay of  $10^{-4}$  and perform 60 epochs in the training stage. The learning rate is initially set to  $10^{-4}$  and dropped by a factor of 10 at epoch 40. Mini-batch-based training was adopted. We set the number of cascades  $T = 8$ , the size of local windows to be 8, and MLP ratio to be 2. All RSTBs consist of 2 STLs with embedding dimension of 66. We implement the proposed model using PyTorch on NVIDIA RTX5000 24GB GPUs.

Performance comparison tables report metrics' mean and standard deviation across test subjects. We evaluate the statistical significance of performance differences using Wilcoxon sign-rank tests.

## 2 Supplemental Experiments



**Fig. 2.** Visual comparison of the reconstructed images for AR=8. A: PD of knee, B: T2 of brain.

## References

1. Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C.L., Yakubova, N., Knoll, F., Johnson, P.: End-to-end variational networks for accelerated MRI reconstruction. In: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention. pp. 64–73. Springer International Publishing, Cham (2020)
2. Zhang, M., Lucas, J., Ba, J., Hinton, G.E.: Lookahead optimizer: k steps forward, 1 step back. *Advances in neural information processing systems* **32** (2019)