Rethinking Diffusion Model for Multi-Contrast MRI Super-Resolution Supplementary Material

Guangyuan Li, Chen Rao, Juncheng Mo, Zhanjie Zhang, Wei Xing*, Lei Zhao*
College of Computer Science and Technology, Zhejiang University, China
{cslgy, raochen, csmjc, cszzj, wxing, cszhl}@zju.edu.cn

1. Prior Extraction

The architecture of the PE is shown in Figure 1. As can be seen, PE mainly consists of 9 residual blocks and 2 linear layers. Specifically, we first concatenate the target LR image $I_{LR} \in \mathbb{R}^{H\times W\times 2}$ and the target HR image $I_{HR}^{P} \in \mathbb{R}^{H\times W\times 2}$ after the PixelUnshuffle operation along the channel dimension to obtain $X \in \mathbb{R}^{H\times W\times 4}$. Then, input X into the PE to generate the prior knowledge $Z \in \mathbb{R}^{4\hat{C}}$. Note that when employed as the condition extraction (CE) module, only the target LR image I_{LR} is utilized as input.

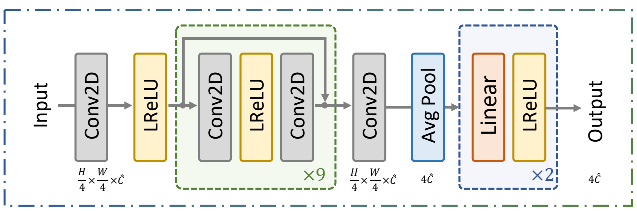

Figure 1. The architecture of the prior extraction module.

2. Frequency-Domain Data Consistency Loss

Unlike natural images, MR images undergo data acquisition in the frequency domain. Therefore, in MR image SR reconstruction, frequency domain loss is crucial for maintaining data consistency by using sampled values to replace specific *k*-space positions. Specifically, Fourier transforms are performed on both I_{SR} and I_{HR} to obtain *k*-space data K_{SR} and K_{HR}, respectively. Then, a sampling mask M is employed to evaluate *k*-space sampling. If the coefficients in K_{SR} have been sampled, they are replaced with the corresponding coefficients in K_{HR}; otherwise, they stay unchanged, as follows:

$$K_{DC}[a,b] = \begin{cases} K_{SR}[a,b] & \text{if } (a,b) \notin M \\ \frac{K_{SR}[a,b]+nK_{HR}[a,b]}{1+n} & \text{if } (a,b) \in M \end{cases}, \quad (1)$$

where $n \geq 0$ is noise (n is set to infinity), [a, b] is a matrix indexing operation, and K_{DC} is the *k*-space data after fidelity. The sampling mask M employs a commonly used *k*-space downsampling mask [5, 9]. Then, the mean squared error is used to constrain K_{DC} and K_{HR}:

$$\mathcal{L}_{dc} = \left\| K_{DC} - K_{HR} \right\|_2 . \quad (2)$$

Method	Window Size	PSNR	SSIM	FLOPs
SwinIR [7]	8×8	30.13	0.8107	32.99G
SwinIR [7]	16×16	30.54	0.8183	39.79G
PLWformer	16×16	30.52	0.8180	**29.45G**
PLWformer	32×32	**30.78**	**0.8242**	36.25G

Table 1. Performance comparison between SwinIR [7] and PLWformer at different window sizes. The best result is marked in **bold**.

3. Window Size Analyses

We further show the comparison of window size and computational complexity (*e.g.*, FLOPs) in Table 1. For a fair comparison, we conduct single-contrast SR reconstruction employing SwinIR [7] and PLWformer, with the optimization function utilizing the L1 loss. Note that in this case, the PLWformer does not utilize the prior knowledge represented by Z. The performance metrics, including PSNR, SSIM, and FLOPs, are evaluated with the image size set to 64×64, the upsampling factor at 4, and using the FastMRI dataset. For SwinIR, expanding the window size can improve the performance of the network, but it also increases computational complexity. In contrast, PLWformer employs permutation operations to transfer some spatial information to the channel dimension. Therefore, even with an expanded window size, the computational complexity does not increase significantly. When the window size is 16×16, the FLOPs required for PLWformer are smaller than those for SwinIR.

Furthermore, we observe that when the window size is the same, the performance of PLWformer is slightly lower than SwinIR, due to the reduction of tokens K and V in

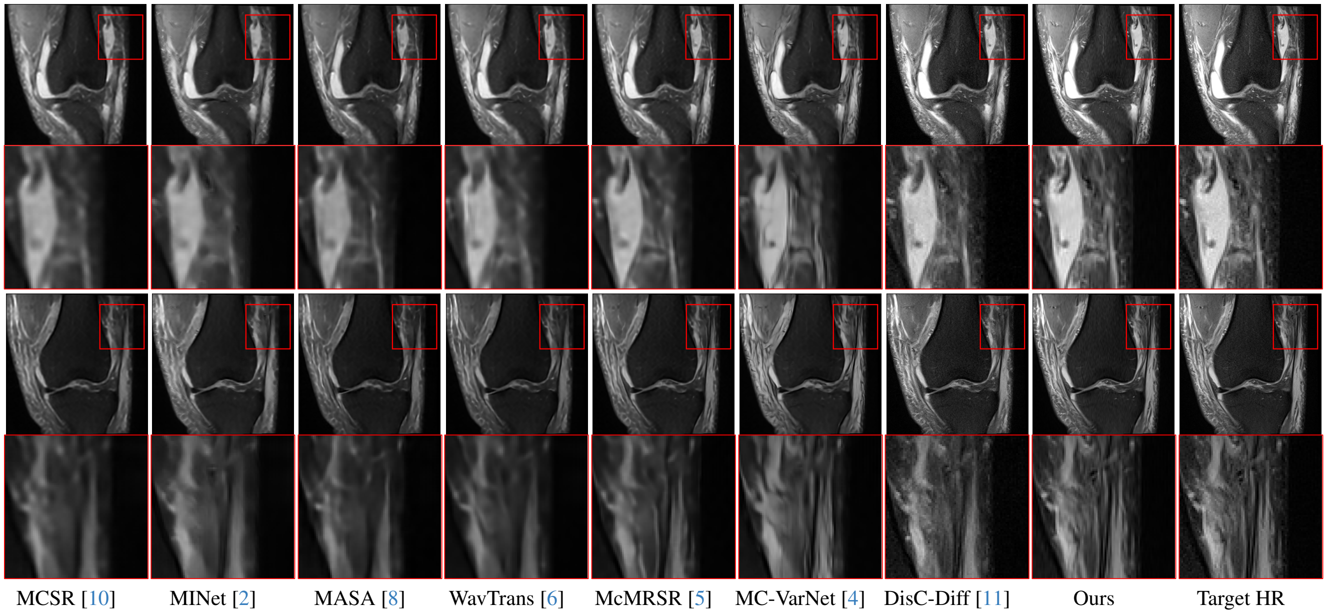

Figure 2. Qualitative visual comparison of various methods on the FastMRI dataset (4×).

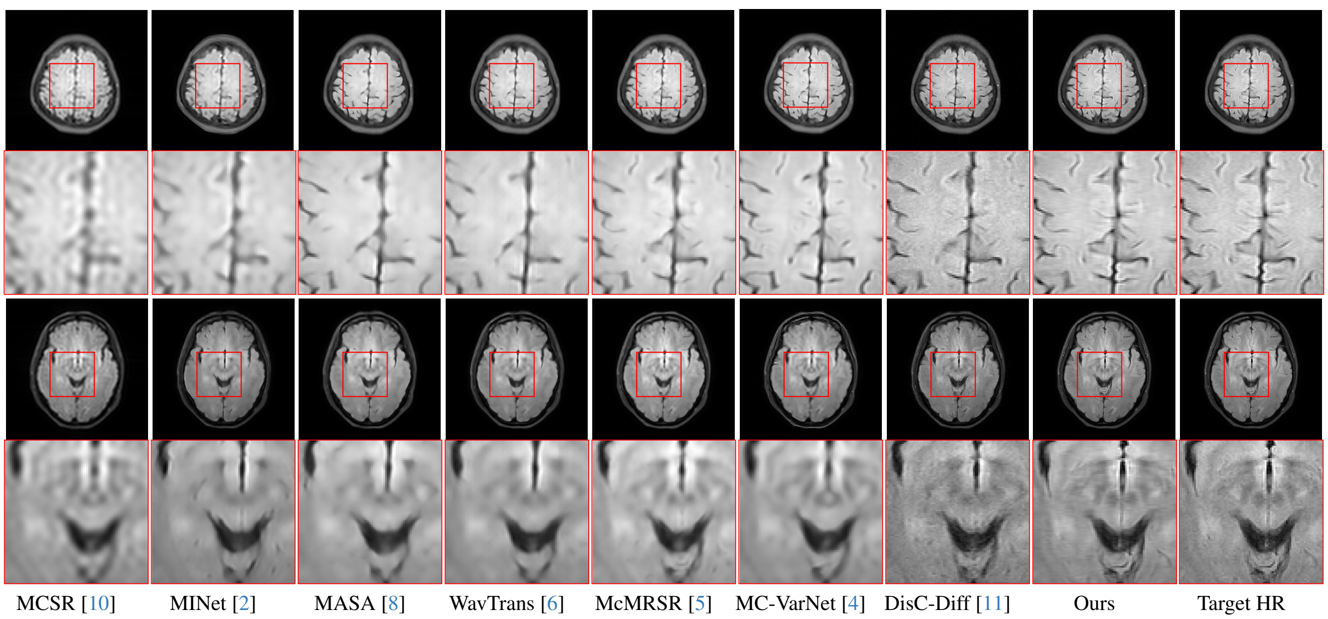

Figure 3. Qualitative visual comparison of various methods on the clinical brain dataset (4×).

PLWformer, which leads to the loss of a small portion of the structural information of the image. However, PLWformer effectively reduces the computational burden, so a slight performance reduction is acceptable.

4. More Visual Comparisons

In this section, we present more visual qualitative comparisons. Figures 2, 3, 4, and 5 show the reconstruction results of each method in FastMRI, clinical brain, clinical tumor, and clinical pelvic, respectively. As can be seen, although DisC-Diff can reconstruct MR images with high-frequency information, it fails to preserve the structure and content of the original Target HR image effectively, resulting in image distortion. In contrast, our proposed DiffMSR can restore high-frequency information while preserving the structure of the original HR image, indicating the effectiveness of the joint use of DM and PLWformer.

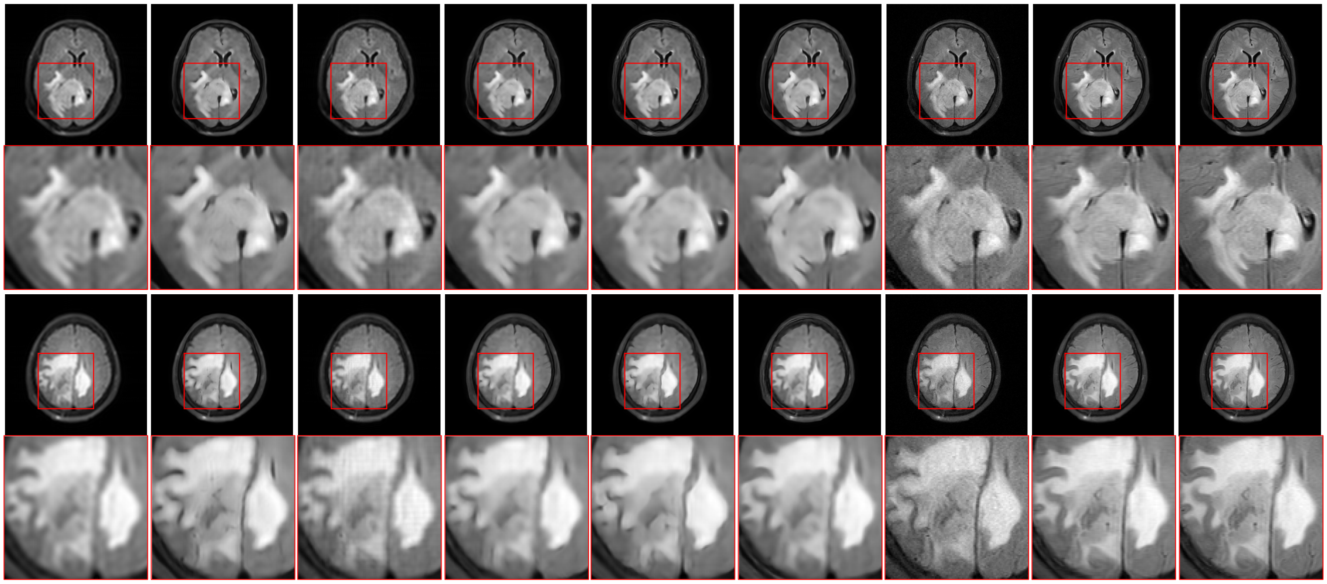

Figure 4. Qualitative visual comparison of various methods on the clinical tumor dataset (4×).

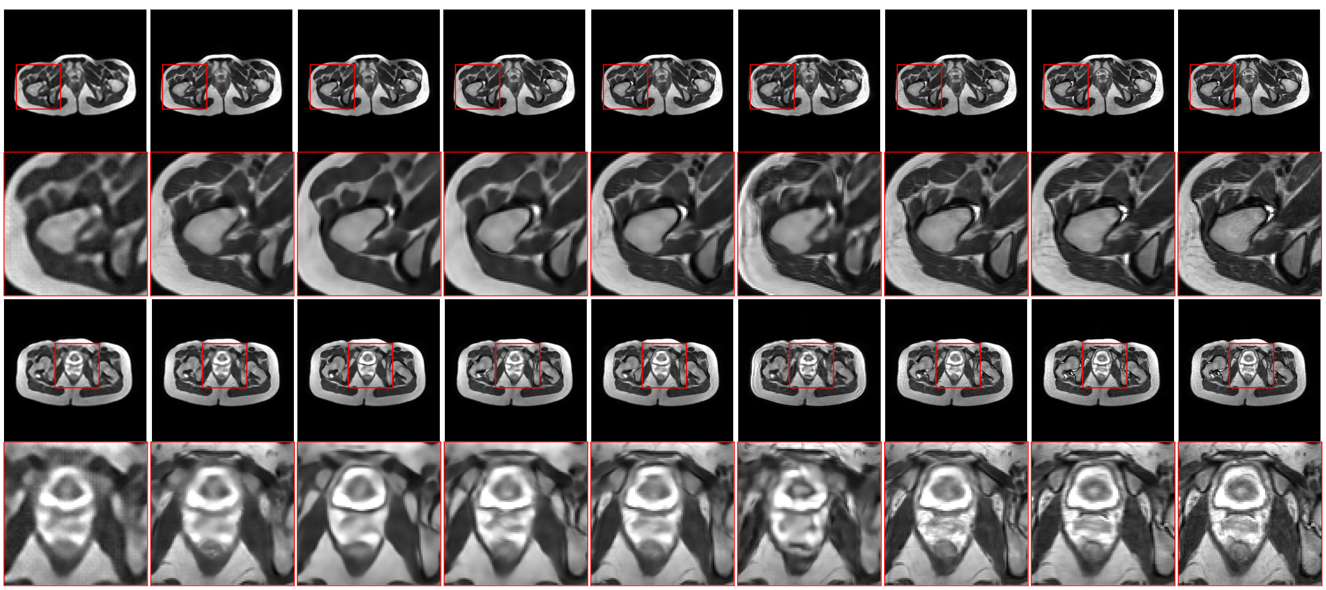

Figure 5. Qualitative visual comparison of various methods on the clinical pelvic dataset (4×).

Besides, our proposed method can be modified to the Single-Contrast Super-Resolution (SCSR) method by removing the cross-attention Transformer layer. The visual qualitative results are shown in Figure 6. As can be seen, our method outperforms other DM-based SCSR methods in the SCSR task.

References

[1] Hyungjin Chung and Jong Chul Ye. Score-based diffusion models for accelerated mri. *Medical image analysis*, 80: 102479, 2022. 4

[2] Chun-Mei Feng, H. Fu, Shuhao Yuan, and Yong Xu. Multi-contrast mri super-resolution via a multi-stage integration network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2021. 2, 3

[3] Alper Güngör, Salman UH Dar, Şaban Öztürk, Yilmaz Ko-

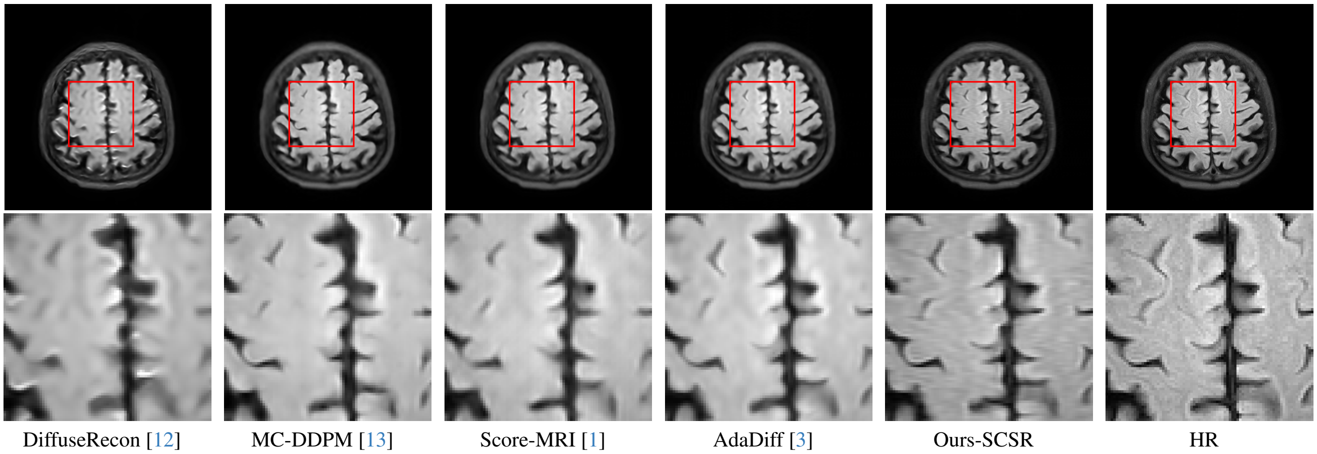

Figure 6. Qualitative results of our method versus other DM-based methods under single-contrast MRI super-resolution ($4\times$).

rkmaz, Hasan A Bedel, Gokberk Elmas, Muzaffer Ozbey, and Tolga Çukur. Adaptive diffusion priors for accelerated mri reconstruction. *Medical Image Analysis*, page 102872, 2023. 4

[4] Pengcheng Lei, Faming Fang, Guixu Zhang, and Tieyong Zeng. Decomposition-based variational network for multi-contrast mri super-resolution and reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 21296–21306, 2023. 2, 3

[5] Guangyuan Li, Jun Lv, Yapeng Tian, Qingyu Dou, Cheng-Yan Wang, Chenliang Xu, and Jing Qin. Transformer-empowered multi-scale contextual matching and aggregation for multi-contrast mri super-resolution. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20604–20613, 2022. 1, 2, 3

[6] Guangyuan Li, Jun Lyu, Chengyan Wang, Qi Dou, and Jin Qin. Wavtrans: Synergizing wavelet and cross-attention transformer for multi-contrast mri super-resolution. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2022. 2, 3

[7] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1833–1844, 2021. 1

[8] Liying Lu, Wenbo Li, Xin Tao, Jiangbo Lu, and Jiaya Jia. Masa-sr: Matching acceleration and spatial adaptation for reference-based image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6368–6377, 2021. 2, 3

[9] Jun Lyu, Guangyuan Li, Chengyan Wang, Qing Cai, Qi Dou, David Zhang, and Jing Qin. Multicontrast mri super-resolution via transformer-empowered multiscale contextual matching and aggregation. *IEEE Transactions on Neural Networks and Learning Systems*, 2023. 1

[10] Qing Lyu, Hongming Shan, Cole R. Steber, Corbin A. Helis, Chris Whitlow, Michael Chan, and Ge Wang. Multi-contrast super-resolution mri through a progressive network. *IEEE Transactions on Medical Imaging*, 39:2738–2749, 2019. 2, 3

[11] Ye Mao, Lan Jiang, Xi Chen, and Chao Li. Disc-diff: Disentangled conditional diffusion model for multi-contrast mri super-resolution. Springer, 2023. 2, 3

[12] Cheng Peng, Pengfei Guo, S Kevin Zhou, Vishal M Patel, and Rama Chellappa. Towards performant and reliable undersampled mr reconstruction via diffusion model sampling. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 623–633. Springer, 2022. 4

[13] Yutong Xie and Quanzheng Li. Measurement-conditioned denoising diffusion probabilistic model for under-sampled medical image reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 655–664. Springer, 2022. 4