

Towards Globally Predictable k -Space Interpolation: A White-box Transformer Approach

Chen Luo^{1,†}, Qiyu Jin^{1,†}, Taofeng Xie², Xuemei Wang³, Huayu Wang⁴, Congcong Liu⁴, Liming Tang⁵, Guoqing Chen¹, Zhuo-Xu Cui^{4,6,*}, and Dong Liang^{4,6,*}

¹ School of Mathematical Sciences, Inner Mongolia University, China

² College of Computer and Information, Inner Mongolia Medical University, China

³ Inner Mongolia Medical University Affiliated Hospital, China

⁴ Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

⁵ School of Mathematics and Statistics, Hubei Minzu University, China

⁶ Key Laboratory of Biomedical Imaging Science and System, Chinese Academy of Sciences, China

zx.cui@siat.ac.cn, dong.liang@siat.ac.cn

Abstract. Interpolating missing data in k -space is essential for accelerating imaging. However, existing methods, including convolutional neural network-based deep learning, primarily exploit local predictability while overlooking the inherent global dependencies in k -space. Recently, Transformers have demonstrated remarkable success in natural language processing and image analysis due to their ability to capture long-range dependencies. This inspires the use of Transformers for k -space interpolation to better exploit its global structure. However, their lack of interpretability raises concerns regarding the reliability of interpolated data. To address this limitation, we propose GPI-WT, a white-box Transformer framework based on *Globally Predictable Interpolation (GPI)* for k -space. Specifically, we formulate GPI from the perspective of annihilation as a novel k -space structured low-rank (SLR) model. The global annihilation filters in the SLR model are treated as learnable parameters, and the subgradients of the SLR model naturally induce a learnable attention mechanism. By unfolding the subgradient-based optimization algorithm of SLR into a cascaded network, we construct the first white-box Transformer specifically designed for accelerated MRI. Experimental results demonstrate that the proposed method significantly outperforms state-of-the-art approaches in k -space interpolation accuracy while providing superior interpretability.

Keywords: White-box Transformer · k -Space Interpolation · MRI.

[†] These authors contributed equally to this work.

^{*} Corresponding authors.

1 Introduction

Magnetic resonance imaging (MRI) is a cornerstone clinical imaging modality, widely recognized for its non-invasive nature, absence of ionizing radiation, and exceptional soft tissue contrast. However, the inherently prolonged acquisition time imposes significant practical limitations, often requiring the acquisition of only partial k -space data [11, 19]. Consequently, advanced reconstruction algorithms are employed to recover high-quality images from these incomplete measurements [16, 1]. A central challenge in MRI reconstruction is the accurate prediction of missing k -space data.

The efficacy of k -space interpolation methods fundamentally relies on the assumption that missing k -space data can be reliably predicted [6]. Traditional approaches, such as those based on image sparsity [19, 20, 28], phase smoothness [12, 26, 15, 8, 9], and coil sensitivity smoothness [24, 23], have enabled the estimation of missing k -space data from local neighboring information [27, 6, 21]. However, these assumptions often do not hold rigorously in practical scenarios. When these assumptions are violated, the predictability of k -space data transitions from a localized interpolation problem to a more complex global dependency, leading to significant errors and degraded image quality.

Recent advancements have explored deep learning (DL)-based methodologies for k -space interpolation [10, 4, 34, 2, 5]. Notably, studies such as [22, 13, 18] have replaced the linear interpolation in traditional methods with nonlinear convolutional neural networks (CNNs), leveraging large-scale datasets to enhance interpolation accuracy. Despite their effectiveness in capturing local structures, CNNs inherently possess a limited receptive field, making them insufficient for modeling long-range dependencies in k -space. This limitation highlights the need for alternative approaches capable of capturing global relationships within k -space.

Transformer-based models [29] have shown remarkable success in capturing long-range dependencies through self-attention mechanisms and fully connected layers, making them highly promising for k -space interpolation. However, existing Transformer-based approaches [17, 34, 31, 7] typically function as black-box models, lacking theoretical interpretability, which raises concerns regarding their reliability in k -space interpolation tasks. A significant advancement by Yu et al. [33] introduced a white-box Transformer for image classification, optimizing sparse rate reduction and providing a theoretical foundation for multi-head self-attention in terms of encoding efficiency. This work also links self-attention mechanisms to iterative optimization algorithms. Although the concept of sparse rate reduction is not directly applicable to k -space interpolation, this insight inspires the development of a novel, interpretable white-box Transformer tailored to the global interpolation properties inherent in k -space.

In this paper, we introduce a novel k -space structured low-rank (SLR) model, formulated from the perspective of globally predictable interpolation (GPI) in k -space through annihilation. Within this framework, the annihilation filters are defined as learnable parameters, and the subgradients of the SLR model naturally induce a learnable attention mechanism. By unfolding the subgradient-based optimization process for SLR-based k -space interpolation into a cascaded

network, we develop a white-box Transformer specifically designed for k -space interpolation.

The main contributions of this paper are as follows: (1) *Interpretable white-box Transformer for MRI reconstruction*: We introduce the first theoretically grounded white-box Transformer model specifically designed for MRI reconstruction. (2) *Effective modeling of global dependencies*: Unlike conventional CNN-based k -space interpolation methods, the proposed Transformer model effectively captures long-range dependencies within k -space, enabling more accurate interpolation of missing data. (3) *Comprehensive experimental validation*: The proposed approach is trained and evaluated on MRI datasets, and benchmarked against state-of-the-art methods. Experimental results demonstrate that the proposed GPI-WT consistently outperforms competing approaches across various undersampling patterns, achieving superior performance in both qualitative and quantitative assessments.

2 Methods

2.1 Global White-box Transformer Reconstruction Model

For parallel MRI, the forward model of k -space measurements can be mathematically represented as follows:

$$\mathbf{y} = M_\Omega \mathbf{k} + \mathbf{n}, \quad (1)$$

where $\mathbf{k} = [\mathbf{k}_1, \dots, \mathbf{k}_{N_c}]$ and $\mathbf{y} \in \mathbb{C}^{N_1 \times N_2 \times N_c}$ denote the multi-coil fully sampled k -space data and the undersampled k -space data, respectively, with $N_c \geq 1$ coils. \mathbf{n} represents the noise. The sampling pattern $M_\Omega \in \mathbb{C}^{N_1 \times N_2 \times N_c}$ is set to 1 at the sampled positions and 0 otherwise. We assume that the missing k -space data in \mathbf{y} can be reliably predicted, which forms the basis for accurately reconstructing \mathbf{k} from \mathbf{y} .

The SLR model, conceptualized from the perspective of globally predictable interpolation in k -space through annihilation, can be formulated as a summation $\sum_{h=1}^H \|\mathcal{H}(\mathbf{k}, d) \mathbf{s}_h\|_F^2$, where \mathcal{H} represents the Hankelization operation, d is the window size used for sliding over the data, and \mathbf{s}_h are the annihilation filters [32]. In this framework, we consider the globally predictable and interpolated annihilation dependencies in k -space, thus d is chosen to match the size of the k -space data. A well-established result is that $\|\mathcal{H}(\mathbf{k}, d) \mathbf{s}_h\|_F^2 = \|\mathbf{Q}_h \mathbf{k}\|_F^2$, where \mathbf{Q}_h denotes the Hankel transformation of the annihilating filter \mathbf{s}_h [22]. By leveraging the equivalence between the Frobenius norm and the trace of a matrix $\|\mathbf{X}\|_F^2 = \text{Tr}(\mathbf{X}^* \mathbf{X})$, the SLR constraint $\|\mathbf{Q}_h \mathbf{k}\|_F^2$ can be reformulated as $\text{Tr}[(\mathbf{Q}_h \mathbf{k})^* (\mathbf{Q}_h \mathbf{k})]$.

To further refine the SLR model, we introduce a penalty function ρ_γ applied to $\mathbf{Q} \mathbf{k}$, extending the SLR constraint to:

$$R(\mathbf{k}; \mathbf{Q}_{[H]}) := \sum_{h=1}^H \text{Tr}[\rho_\gamma((\mathbf{Q}_h \mathbf{k})^* (\mathbf{Q}_h \mathbf{k}))] \quad (2)$$

where $\rho_\gamma(\mathbf{X}) = \mathbf{U}\text{Diag}(\rho_\gamma(\sigma_1), \dots, \rho_\gamma(\sigma_r))\mathbf{V}^*$ and σ_i denotes the singular values of \mathbf{X} . In this work, we adopt a sparse prompting function

$$\rho_\gamma(\mathbf{X}) := \mathbf{U}\text{diag}(\ln(\mathbf{1} + \gamma\sigma_1), \dots, \ln(\mathbf{1} + \gamma\sigma_r))\mathbf{V}^*.$$

In traditional methods, \mathbf{Q}_h is obtained by applying Hankelization to low-dimensional filters \mathbf{s}_h , leveraging the local predictability of k -space. This approach involves a small number of parameters, which can typically be estimated from calibration data. However, in our method, \mathbf{s}_h is treated as a global annihilation filter with dimensions matching those of the k -space data \mathbf{k} , resulting in a substantially larger parameter space. This increased complexity makes it difficult to accurately estimate \mathbf{s}_h using calibration data alone. To address this, we relax the Hankel structural constraint on \mathbf{Q}_h and define it as a set of learnable parameters, allowing it to be directly learned from large datasets.

On the other hand, the subgradient of the specially designed SLR model (2) can be approximated as the following structure:

$$\begin{aligned} \nabla_{\mathbf{k}} R(\mathbf{k}; \mathbf{Q}_{[H]}) &= \sum_{h=1}^H \nabla_{\mathbf{k}} \text{Tr} [\ln (\mathbf{I} + \gamma (\mathbf{Q}_h \mathbf{k})^* (\mathbf{Q}_h \mathbf{k}))] \\ &= \sum_{h=1}^H \nabla_{\mathbf{k}} \ln \text{Det} [(\mathbf{I} + \gamma (\mathbf{Q}_h \mathbf{k})^* (\mathbf{Q}_h \mathbf{k}))] \\ &= \gamma \sum_{h=1}^H \mathbf{Q}_h^* \mathbf{Q}_h \mathbf{k} (\mathbf{I} + \gamma (\mathbf{Q}_h \mathbf{k})^* (\mathbf{Q}_h \mathbf{k}))^{-1} \\ &\approx \gamma \mathbf{k} - \gamma^2 \sum_{h=1}^H \mathbf{Q}_h^* \mathbf{Q}_h \mathbf{k} \text{softmax} ((\mathbf{Q}_h \mathbf{k})^* \mathbf{Q}_h \mathbf{k}), \end{aligned} \tag{3}$$

where the last approximate equality follows from reference [33]. Based on the above analysis, the subgradient of (2) can fundamentally be interpreted as a multi-head subspace self-attention (MSSA):

$$\text{MSSA}(\mathbf{k} | \mathbf{Q}_{[H]}) := \gamma^2 [\mathbf{Q}_1^*, \dots, \mathbf{Q}_H^*] \begin{bmatrix} \text{SSA}(\mathbf{k} | \mathbf{Q}_1) \\ \vdots \\ \text{SSA}(\mathbf{k} | \mathbf{Q}_H) \end{bmatrix}, \tag{4}$$

where subspace self-attention (SSA) is defined as

$$\text{SSA}(\mathbf{k} | \mathbf{Q}_h) := \mathbf{Q}_h \mathbf{k} \text{softmax} ((\mathbf{Q}_h \mathbf{k})^* \mathbf{Q}_h \mathbf{k}). \tag{5}$$

Unlike conventional Transformers, the MSSA mechanism utilizes a single matrix to derive the Query, Key, and Value representations in the attention mechanism of a white-box Transformer, where $Q = K = V = \mathbf{Q}\mathbf{k}$. Therefore, we derive an interpretable white-box attention mechanism through the SLR model (2).

Based on the white-box attention mechanism constructed above, by unfolding the subgradient-based optimization algorithm for the (2)-regularized k -space

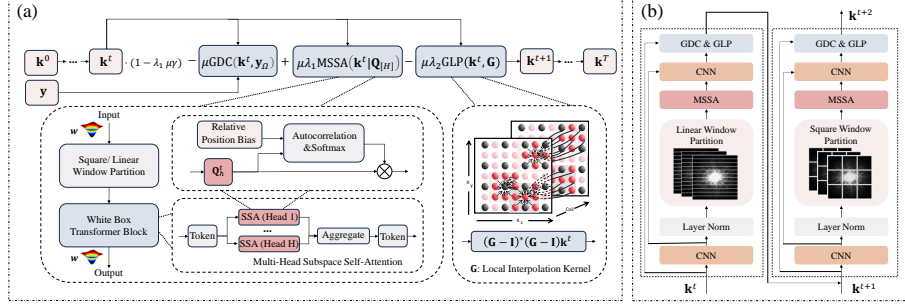


Fig. 1. (a) Schematic of the our proposed unfolded network. (b) Diagram of window partitions. Smoothing filter w is applied to facilitate the computation in k -space.

interpolation model, we derive a cascaded white-box Transformer network specifically tailored for k -space interpolation. In particular, the (2)-regularized k -space interpolation model can be formulated as follows:

$$\min_{\mathbf{k}} \frac{1}{2} \|\mathbf{M}_\Omega \mathbf{k} - \mathbf{y}\|_2^2 + \lambda_1 R(\mathbf{k}; \mathbf{Q}_{[H]}) + \lambda_2 \|(\mathbf{G} - \mathbf{I})\mathbf{k}\|_2^2, \quad (6)$$

It is worth noting that, considering that the global predictability enforced by (2) may weaken local dependencies, we have retained a local predictability term $\|(\mathbf{G} - \mathbf{I})\mathbf{k}\|_2^2$ in the above model. Here, \mathbf{G} represents the local linear interpolation kernel estimated by SPIRiT [21]. The parameters λ_1 and λ_2 are regularization coefficients. The gradient descent method for (6) yields a cascaded white-box Transformer, as detailed below:

$$\mathbf{k}^{t+1} = (1 - \lambda_1 \mu \gamma) \mathbf{k}^t - \mu \text{GDC}(\mathbf{k}^t, \mathbf{y}) + \mu \lambda_1 \text{MSSA}(\mathbf{k}^t | \mathbf{Q}_{[H]}) - \mu \lambda_2 \text{GLP}(\mathbf{k}^t, \mathbf{G}), \quad (7)$$

where μ denotes the step size, GDC is the gradient of the data consistency term, i.e., $\text{GDC}(\mathbf{k}, \mathbf{y}) = \mathbf{M}_\Omega^* \mathbf{M}_\Omega \mathbf{k} - \mathbf{M}_\Omega^* \mathbf{y}$, and GLP is the gradient of the local predictability term, i.e., $\text{GLP}(\mathbf{k}, \mathbf{G}) = (\mathbf{G} - \mathbf{I})^* (\mathbf{G} - \mathbf{I}) \mathbf{k}$. The specific structural framework of this approach is illustrated in Fig. 1(a).

2.2 Network Architecture

Square and linear window partiton In visual Transformer-based methods [3, 33], images are divided into patches, leading to quadratic computational growth as image size increases. To address this, we adopt Swin Transformer’s [17] window partition, treating each pixel as a patch within fixed-size non-overlapping windows. This allows our MSSA to maintain linear computational scaling.

Unlike the image domain, where neighbouring pixels are highly correlated, k -space exhibits long-range dependencies between symmetric signals. Standard self-attention with square windows limits cross-window interactions, restricting

distant information capture. To overcome this limitation, we propose a novel linear window strategy that preserves efficiency while enhancing attention to distant signals by using non-overlapping lines as windows. Linear and square partitions alternate in successive GPI-WT iterations, improving the network’s nonlinear representation and reconstruction performance. Further details are illustrated in Fig. 1(b).

Relative position bias As the Transformer architecture is inherently order-agnostic and relies on self-attention mechanisms, it is essential to explicitly incorporate positional information through positional embeddings. In our approach, we use learnable relative position bias [17] when computing self-attention, which not only reduces the number of trainable parameters but also enhances the performance of our model in square and linear window partitions. Relative position bias \mathbf{B} is incorporated into each head of MSSA, thus the SSA module can ultimately be modified to:

$$\text{SSA}(\mathbf{k}|\mathbf{Q}_h) = \mathbf{Q}_h \mathbf{k} \text{softmax}((\mathbf{Q}_h \mathbf{k})^* \mathbf{Q}_h \mathbf{k} + \mathbf{B}). \quad (8)$$

3 Data and Experiments

The raw knee data was acquired from a 3T Siemens scanner. Data acquisition used a 15 channel knee coil array and conventional Cartesian 2D TSE protocol. We randomly selected 31 individuals (840 slices in total) as training data and 3 individuals (96 slices in total) as test data. The image size is cropped to 320×300 .

The model unfolded 10 iterations empirically with 4×4 square window and 6 head in self-attention. We used ADAM optimizer [14] starting with learning rate $lr = 0.0001$, which decayed exponentially at a rate of 0.99. The parameters λ_1 , λ_2 , μ , and γ are learnable parameters. Experiments were on Nvidia Tesla A6000 GPU. Three metrics were used to evaluate the results quantitatively, including normalized mean square error (NMSE), the peak signal-to-noise ratio (PSNR), and the structural similarity index (SSIM) [30].

4 Results and Discussion

We compared our proposed approach with state-of-the-art k -space MR reconstruction methods, including SPIRiT [21], a classical training-free k -space interpolation method; KNet [10], a DL method based on U-Net [25]; Swin Transformer [17]; and DSLR [22], an iterative regularization method that learns SLR constraints using CNNs. All these methods were trained exclusively in k -space. To further demonstrate the superiority of our GPI-WT in k -space, we also compared it with a CNN-based version of GPI-WT called GPI-CNN.

Comparative experiments were conducted on knee MRI data using random and uniform 2D sampling trajectories with acceleration factors (AF) of AF = 4, 6

Table 1. Quantitative comparison for various methods on knee dataset with different mask patterns M_Ω and AF. \mathcal{R} : random mask. \mathcal{U} : uniform mask. The optimal values are denoted in **bold** and the proposed method is highlighted with gray background.

Method	M_Ω	NMSE (%) \downarrow		PSNR \uparrow		SSIM (%) \uparrow	
		AF=4	AF=6	AF=4	AF=6	AF=4	AF=6
SPiRiT	\mathcal{R}	1.59 \pm 1.20	1.71 \pm 1.00	29.63 \pm 3.15	28.95 \pm 2.50	73.33 \pm 10.28	73.01 \pm 8.50
KNet		0.84 \pm 0.30	1.38 \pm 0.68	31.73 \pm 1.44	29.75 \pm 1.78	85.28 \pm 3.08	80.91 \pm 3.95
Swin		0.74 \pm 0.38	1.17 \pm 0.57	32.46 \pm 1.84	30.40 \pm 1.70	85.86 \pm 4.10	81.29 \pm 4.23
DSLr		0.76 \pm 0.28	1.23 \pm 0.50	32.19 \pm 1.64	30.13 \pm 1.77	85.84 \pm 3.28	81.02 \pm 4.02
GPI-CNN		0.57 \pm 0.29	0.96 \pm 0.38	33.48 \pm 1.69	31.18 \pm 1.67	87.42 \pm 2.87	83.41 \pm 3.27
GPI-WT		0.48\pm0.18	0.81\pm0.33	34.13\pm1.64	31.92\pm1.65	88.94\pm3.24	84.93\pm3.84
SPiRiT	\mathcal{U}	1.53 \pm 1.00	2.23 \pm 1.02	29.65 \pm 3.02	27.58 \pm 2.08	74.88 \pm 9.39	71.99 \pm 7.28
KNet		1.13 \pm 0.55	2.32 \pm 1.04	30.60 \pm 1.70	27.52 \pm 1.83	84.35 \pm 3.14	76.88 \pm 4.03
Swin		0.87 \pm 0.44	2.08 \pm 1.20	31.72 \pm 1.85	28.06 \pm 1.91	85.65 \pm 4.16	77.75 \pm 4.55
DSLr		1.09 \pm 0.52	2.17 \pm 0.87	30.74 \pm 1.78	27.72 \pm 1.68	84.54 \pm 3.25	76.70 \pm 4.25
GPI-CNN		0.69 \pm 0.46	1.36 \pm 0.59	32.79 \pm 1.86	29.77 \pm 1.73	87.31 \pm 3.05	80.98 \pm 3.24
GPI-WT		0.55\pm0.25	1.10\pm0.41	33.64\pm1.78	30.59\pm1.39	88.79\pm3.30	82.97\pm3.48

Table 2. Ablation experiment on knee dataset with mask pattern M_Ω and AF = 4. \mathcal{R} : random mask. The optimal values are denoted in **bold** and the proposed method is highlighted with gray background..

Method	M_Ω	NMSE (%) \downarrow	PSNR \uparrow	SSIM (%) \uparrow
GPI-WT w/o LW w/o GLP	\mathcal{R}	0.69 \pm 0.32	32.70 \pm 1.70	86.83 \pm 3.35
GPI-WT w/o GLP		0.60 \pm 0.30	33.34 \pm 1.77	87.73 \pm 3.19
GPI-BT		0.56 \pm 0.24	33.60 \pm 1.78	88.61 \pm 3.22
GPI-WT		0.48\pm0.18	34.13\pm1.64	88.94\pm3.24

and an autocalibration signal (ACS) region of ACS = 24. Table 1 summarizes the quantitative results, including average evaluation metrics. Fig. 2 illustrates reconstructed MRI slices at various acceleration rates. Both qualitative and quantitative results demonstrate that our k -space GPI-WT outperforms other approaches. The zoomed-in regions in Fig. 2 and corresponding error maps clearly demonstrate that, compared to CNN-based unrolled methods for k -space interpolation, our proposed GPI-WT significantly reduces aliasing artifacts and delivers more stable reconstructions, underscoring its enhanced capability to exploit global structural information.

We conducted ablation experiments using random undersampling knee data at AF = 4 to evaluate the performance of key components in GPI-WT. Specifically, we trained and tested GPI-WT without linear window partitioning (LW) and GLP, denoted as GPI-WT w/o LW w/o GLP, which uses only square windows. To validate the effectiveness of LW, we compared it with GPI-WT without GLP but with alternating linear and square windows, denoted as GPI-WT w/o GLP. Additionally, we evaluated GPI-WT which have GLP and alternating windows against a black-box Transformer variant, denoted as GPI-BT, to assess the contributions of GLP and the white-box design. Qualitative and quantitative results are presented in Table 2 and Figure 3. Results show LW combined with square windows better capture long-range dependencies, while GLP boosts lo-

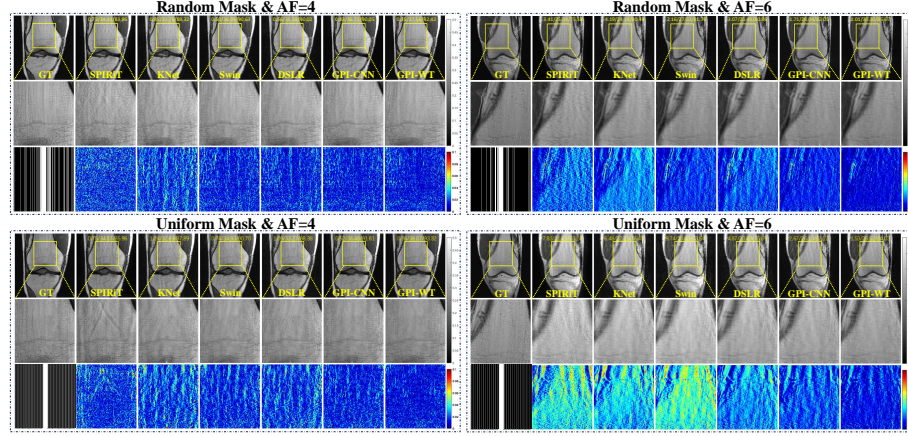


Fig. 2. Reconstruction results under random and uniform masks and $AF = 4, 6$ with 24 ACS lines. Values of $NMSE(\%)/PSNR/SSIM(\%)$ are given. The second and third rows illustrate the enlarged and error views. Grayscale bars for reconstructed images and color bars for error maps are on figures' right.

cal consistency. GPT-WT surpasses GPT-BT, proving its interpretable design's efficacy, and reduces aliasing best (Fig. 3).

5 Conclusion

This paper introduced GPI-WT, the first white-box Transformer for k -space interpolation, built upon the concept of GPI. Specifically, GPI was formulated from the perspective of annihilation as a novel k -space SLR model. Within this framework, the global annihilation filters were treated as learnable parameters, and the subgradients of the SLR model naturally induced a learnable attention mechanism. By unfolding the subgradient-based optimization algorithm of the SLR model into a cascaded network, the first white-box Transformer network tailored for MRI reconstruction was developed. Experimental results demonstrated that the proposed method achieved superior k -space interpolation accuracy compared to state-of-the-art approaches while offering enhanced interpretability. Currently, we focus solely on k -space feature learning. Future work will extend the white-box Transformer to the image domain.

Acknowledgments. This study was funded by Shenzhen Science and Technology Program under grant no. JCYJ20240813155840052; the National Key R&D Program of China (2022YFA1004203, 2021YFF0501503), the National Natural Science Foundation of China (62125111, 62331028, 62476268, 62206273, 12061052, 62271474), Natural Science Fund of Inner Mongolia Autonomous Region (2024LHMS01006).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

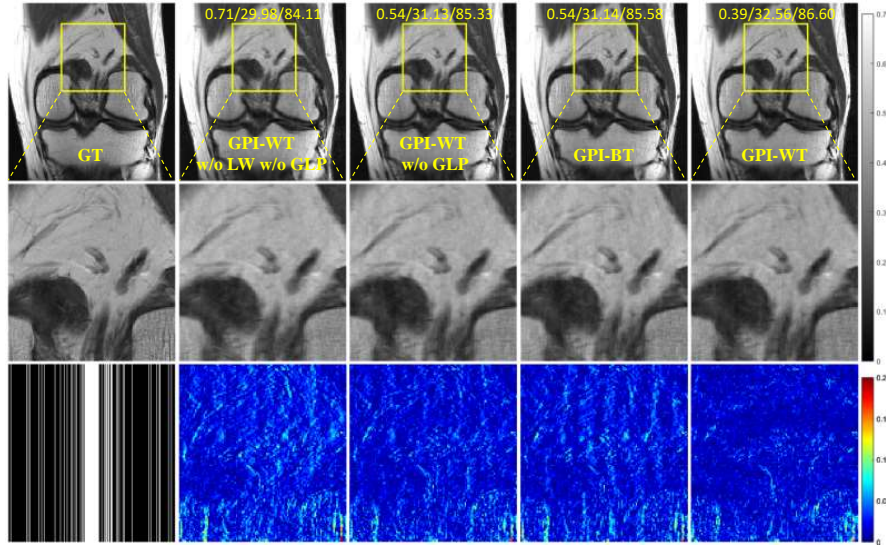


Fig. 3. Ablation results under random mask at $AF = 4$ with 24 ACS lines. Values of NMSE(%) / PSNR / SSIM(%) are given. The second and third rows illustrate the enlarged and error views. Grayscale bars for reconstructed images and color bars for error maps are on figure's right.

References

1. Aggarwal, H.K., Mani, M.P., Jacob, M.: Modl: Model-based deep learning architecture for inverse problems. *IEEE Transactions on Medical Imaging* **38**(2), 394–405 (2019)
2. Cui, Z.X., Jia, S., Cao, C., Zhu, Q., Liu, C., Qiu, Z., Liu, Y., Cheng, J., Wang, H., Zhu, Y., Liang, D.: K-unn: k-space interpolation with untrained neural network. *Medical Image Analysis* **88**, 102877 (2023)
3. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An image is worth 16x16 words: Transformers for image recognition at scale. In: *International Conference on Learning Representations* (2021)
4. Du, T., Zhang, H., Li, Y., Pickup, S., Rosen, M., Zhou, R., Song, H.K., Fan, Y.: Adaptive convolutional neural networks for accelerating magnetic resonance imaging via k-space data interpolation. *Medical Image Analysis* **72**, 102098 (2021)
5. Eo, T., Jun, Y., Kim, T., Jang, J., Lee, H.J., Hwang, D.: Kiki-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. *Magnetic Resonance in Medicine* **80**(5), 2188–2201 (2018)
6. Griswold, M.A., Jakob, P.M., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A.: Generalized autocalibrating partially parallel acquisitions (grappa). *Magnetic Resonance in Medicine* **47**(6), 1202–1210 (2002)
7. Guo, P., Mei, Y., Zhou, J., Jiang, S., Patel, V.M.: Reconformer: Accelerated mri reconstruction using recurrent transformer. *IEEE Transactions on Medical Imaging* **43**(1), 582–593 (2024)

8. Haldar, J.P.: Low-rank modeling of local k -space neighborhoods (loraks) for constrained mri. *IEEE transactions on medical imaging* **33**(3), 668–681 (2013)
9. Haldar, J.P., Zhuo, J.: P-loraks: Low-rank modeling of local k -space neighborhoods with parallel imaging data. *Magnetic Resonance in Medicine* **75**(4), 1499–1514 (2016)
10. Han, Y., Sunwoo, L., Ye, J.C.: k -space deep learning for accelerated mri. *IEEE Transactions on Medical Imaging* **39**(2), 377–386 (2020)
11. Hutchinson, M., Raff, U.: Fast mri data acquisition using multiple detectors. *Magnetic Resonance in Medicine* **6**(1), 87–91 (1988)
12. Jacob, M., Mani, M.P., Ye, J.C.: Structured low-rank algorithms: Theory, magnetic resonance applications, and links to machine learning. *IEEE Signal Processing Magazine* **37**(1), 54–68 (2020)
13. Kim, T.H., Garg, P., Haldar, J.P.: Loraki: Autocalibrated recurrent neural networks for autoregressive mri reconstruction in k -space (2019), <https://arxiv.org/abs/1904.09390>
14. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: *Proceedings of the International Conference on Learning Representations* (2014)
15. Lee, D., Jin, K.H., Kim, E.Y., Park, S.H., Ye, J.C.: Acceleration of mr parameter mapping using annihilating filter-based low rank hankel matrix (aloha). *Magnetic Resonance in Medicine* **76**(6), 1848–1864 (2016)
16. Liang, Z.P., Boada, F., Constable, R., Haacke, E., Lauterbur, P., Smith, M.: Constrained reconstruction methods in mr imaging. *Rev Magn Reson Med* **4**(2), 67–185 (1992)
17. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. pp. 9992–10002 (2021)
18. Luo, C., Wang, H., Xie, T., Jin, Q., Chen, G., Cui, Z.X., Liang, D.: Matrix completion-informed deep unfolded equilibrium models for self-supervised k -space interpolation in mri (2023), <https://doi.org/10.48550/arXiv.2309.13571>
19. Lustig, M., Donoho, D., Pauly, J.M.: Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine* **58**(6), 1182–1195 (2007)
20. Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M.: Compressed sensing mri. *IEEE Signal Processing Magazine* **25**(2), 72–82 (2008)
21. Lustig, M., Pauly, J.M.: Spirit: Iterative self-consistent parallel imaging reconstruction from arbitrary k -space. *Magnetic Resonance in Medicine* **64**(2), 457–471 (2010)
22. Pramanik, A., Aggarwal, H.K., Jacob, M.: Deep generalization of structured low-rank algorithms (deep-slr). *IEEE Transactions on Medical Imaging* **39**(12), 4186–4197 (2020)
23. Pruessmann, K.P., Weiger, M., Börnert, P., Boesiger, P.: Advances in sensitivity encoding with arbitrary k -space trajectories. *Magnetic Resonance in Medicine* **46**(4), 638–651 (2001)
24. Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P.: Sense: Sensitivity encoding for fast mri. *Magnetic Resonance in Medicine* **42**(5), 952–962 (1999)
25. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. pp. 234–241 (2015)
26. Shin, P.J., Larson, P.E.Z., Ohliger, M.A., Elad, M., Pauly, J.M., Vigneron, D.B., Lustig, M.: Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion. *Magnetic Resonance in Medicine* **72**(4), 959–970 (2014)

27. Smith, M.R., Nichols, S.T., Henkelman, R.M., Wood, M.L.: Application of autoregressive moving average parametric modeling in magnetic resonance image reconstruction. *IEEE Transactions on Medical Imaging* **5**(3), 132–139 (1986)
28. Uecker, M., Lai, P., Murphy, M.J., Virtue, P., Elad, M., Pauly, J.M., Vasanawala, S.S., Lustig, M.: Espirit—an eigenvalue approach to autocalibrating parallel mri: Where sense meets grappa. *Magnetic Resonance in Medicine* **71**(3), 990–1001 (2014)
29. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. p. 6000–6010. NIPS’17 (2017)
30. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**(4), 600–612 (2004)
31. Wu, Z., Liao, W., Yan, C., Zhao, M., Liu, G., Ma, N., Li, X.: Deep learning based mri reconstruction with transformer. *Computer Methods and Programs in Biomedicine* **233**, 107452 (2023)
32. Ye, J.C., Han, Y., Cha, E.: Deep convolutional framelets: A general deep learning framework for inverse problems. *SIAM Journal on Imaging Sciences* **11**(2), 991–1048 (2018)
33. Yu, Y., Buchanan, S., Pai, D., Chu, T., Wu, Z., Tong, S., Haeffele, B.D., Ma, Y.: White-box transformers via sparse rate reduction. In: *Proceedings of the 37th International Conference on Neural Information Processing Systems*. NIPS ’23 (2023)
34. Zhao, Z., Zhang, T., Xie, W., Wang, Y., Zhang, Y.: K-space transformer for undersampled mri reconstruction. In: *British Machine Vision Conference (BMVC)* (2022)