GF-BrainSR: Gated Frequency-Aware Selective State Space Model-Based Brain MRI Image Super-resolution

Boren Tong, Linxuan Fan, Panhui Yang*, and Juntao Jiang*

[1] Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, Zhejiang 315211, China
236002394@nbu.edu.cn
[2] Data Science Institute, Vanderbilt University, Nashville, TN 37240, USA
linxuan.fan@vanderbilt.edu
[3] Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, Zhejiang 315211, China
yangpanhui@nbu.edu.cn
[4] College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang 310000, China
juntaojiang@zju.edu.cn
*Corresponding authors.

Abstract. Deep learning-based super-resolution (SR) mitigates resolution bottlenecks in brain MRI, yet existing CNNs and Transformers struggle to balance global context, computational efficiency, and fidelity of high-frequency details. We propose GF-BrainSR, an efficient linear-complexity State Space Model (SSM) framework. GF-BrainSR introduces a DWT-driven Frequency Adaptive Gate that explicitly decouples and enhances structural boundaries, while a depthwise spatial gate robustly suppresses noise in homogeneous soft tissues. Additionally, an Implicit Coordinate Upsampling module fuses continuous spatial coordinates with local features, thereby eliminating discrete interpolation artifacts and reconstructing morphologically continuous tissue interfaces. Experiments demonstrate that GF-BrainSR achieves competitive performance, delivering superior anatomical clarity and spatial resolution with linear computational complexity. The source code can be accessed at https://github.com/qianchang680/GFBrainSR.

Keywords: Magnetic Resonance Imaging · Brain Imaging · Super-Resolution · State Space Models · Frequency-Aware Network · Implicit Coordinate Upsampling

1 Introduction

High-quality brain MRI is critical for neurological diagnosis and automated neuroimage analysis. However, prohibitive hardware costs and clinical constraints, such as accelerated scanning to minimize patient motion, often yield scans

with insufficient spatial resolution, Rician noise, and truncation artifacts. Deep learning-based super-resolution (SR) offers a cost-effective software alternative to enhance degraded images without costly hardware upgrades, democratizing high-quality imaging across under-resourced clinical settings [1,2,3].

While general SR methods have advanced significantly, they face distinct challenges in medical imaging. CNN-based methods [4,5,6,7] mainly capture local features and may struggle with long-range anatomical dependencies, while Transformer-based methods [8,9,10,11,12,13] rely on explicit global attention that becomes increasingly demanding at high resolutions. Diffusion model- based methods [14,15], meanwhile, introduce substantial inference latency due to iterative denoising. Recently, State Space Models (SSMs) like Mamba [16] and recurrent networks like RWKV [17] have emerged as an alternative for global representation learning and have been increasingly explored in image super-resolution [18,19,20,21,22,23,24]. Their ability to propagate information over long spatial ranges provides an effective alternative to local convolution and explicit attention. However, existing SSM or RWKV-based SR methods largely overlook the distinctive anatomical characteristics of brain MRI, which combines extensive homogeneous soft-tissue regions with intricate high-frequency structures such as cortical folds and pathological boundaries. Without explicit anatomical priors, generic SSMs may propagate noise in homogeneous regions while oversmoothing clinically important details, motivating anatomy-aware SSMs for brain MRI super-resolution.

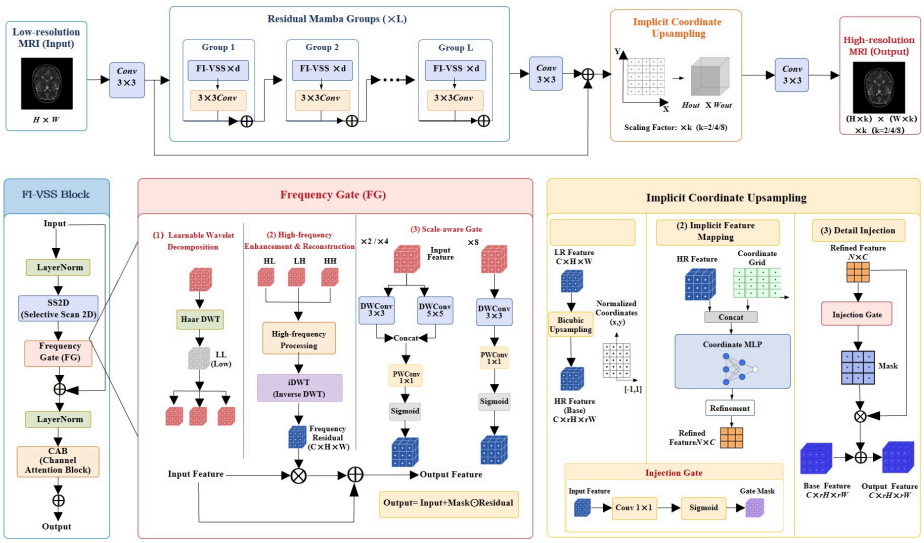

Fig. 1: Architecture of the proposed GF-BrainSR. Each FI-VSS block integrates fixed orthogonal Haar wavelet decomposition, multi-scale frequency gating, and selective state-space scanning for efficient global-local feature refinement.

To address these challenges, we propose an efficient linear-complexity SSM framework tailored to cerebral MRI. Our main contributions are:

- A selective-scanning SR architecture that achieves a strict balance between global anatomical context modeling and linear computational efficiency.
- A DWT-driven frequency-adaptive gate that couples orthogonal frequency decoupling with depthwise spatial masking, selectively enhancing cortical boundaries while suppressing noise in homogeneous tissues.
- An implicit coordinate upsampling module that fuses features with continuous 2D coordinates, eliminating checkerboard artifacts to reconstruct morphologically continuous tissue interfaces.

2 Methodology: GF-BrainSR

Fig. 1 illustrates the proposed GF-BrainSR. A 3×3 convolution first extracts shallow features F_0 from the low-resolution MRI I_{in}. These are hierarchically refined by N cascaded Residual Mamba Groups, comprising Frequency-Integrated Visual State Space (FI-VSS) and Channel Attention Blocks, to jointly model long-range anatomical dependencies and localized high-frequency structures. To preserve low-frequency structural consistency, the output deep feature F_{deep} is residually fused with F_0. Finally, an Implicit Coordinate Upsampling module (H_{UP}) reconstructs the high-resolution image via $I_{\text{out}} = H_{\text{UP}}(F_{\text{deep}} + F_0)$. Unlike conventional discrete operators, this continuous coordinate-guided reconstruction effectively alleviates interpolation discontinuities and sharpens anatomical boundaries.

2.1 Frequency-Integrated Visual State Space Block

MRI acquisition artifacts (e.g., Rician noise) are highly coupled with high-frequency anatomical details. Consequently, standard SR models often blindly amplify noise in homogeneous tissues like white matter and cerebrospinal fluid (CSF). To mitigate this, we propose a frequency-aware gate that decouples structural boundaries from noise, selectively enhancing valid details prior to state-space propagation.

Frequency Gate via DWT Given an input feature map $X \in \mathbb{R}^{C \times H \times W}$, we first apply a 2D discrete wavelet transform (DWT) to decompose it into one low-frequency approximation subband X_{LL} and three high-frequency detail subbands $\{X_{HL}, X_{LH}, X_{HH}\}$:

$$X_{LL}, \{X_{HL}, X_{LH}, X_{HH}\} = \text{DWT}(X). \tag{1}$$

The high-frequency components, which encapsulate critical directional edge information, are concatenated along the channel dimension to form:

$$X_{\text{high}} = \text{Concat}(X_{HL}, X_{LH}, X_{HH}). \tag{2}$$

The processed high-frequency representation is formulated as:

$$\hat{X}_{\text{high}} = W_{1\times1} * \delta\left(\text{DWConv}_{3\times3}(X_{\text{high}})\right), \tag{3}$$

where $\delta(\cdot)$ denotes the GELU activation function, $\text{DWConv}_{3\times3}$ represents the 3×3 depthwise convolution operator for spatial grouping, and $W_{1\times1}$ denotes pointwise frequency fusion. The enhanced residual component is then reconstructed through inverse discrete wavelet transform (IDWT):

$$X_{\text{res}} = \text{IDWT}(\hat{X}_{\text{high}}). \tag{4}$$

To suppress noise in homogeneous tissues and adapt to varying degradation levels, we introduce a **scale-aware depthwise gating module**. Under severe degradation ($\times8$), spatial dimensions are highly compressed; thus, compact 3×3 kernels provide an adequate relative receptive field while preventing the oversmoothing of fragile high-frequency signals. Conversely, milder degradations ($\times2$ and $\times4$) preserve higher resolution and richer structural information. For these scales, we fuse both 3×3 and 5×5 convolutions to simultaneously capture fine local details and broader spatial context, effectively balancing detail retention and structural continuity. The operation $\mathcal{G}(X)$ is formalized as:

$$\mathcal{G}(X) = \begin{cases} \text{Concat}\left(\text{DWConv}_{3\times3}(X), \text{DWConv}_{5\times5}(X)\right), & \text{for } \times2, \times4 \\ \text{DWConv}_{3\times3}(X), & \text{for } \times8 \end{cases} \tag{5}$$

The gating mask M is then estimated through a lightweight bottleneck network:

$$M = \sigma\left(W_{1\times1}^{up} * \text{LReLU}\left(W_{1\times1}^{down} * \mathcal{G}(X)\right)\right), \tag{6}$$

where $\sigma(\cdot)$ denotes the Sigmoid activation function and LReLU represents Leaky ReLU.

The final gated frequency enhancement is:

$$X_{\text{out}} = X + M \odot X_{\text{res}}, \tag{7}$$

where $\odot$ denotes element-wise multiplication.

Frequency-Integrated 2D Selective Scan Standard state-space models exhibit strong long-range dependency modeling capability but tend to suppress subtle high-frequency structural variations. To alleviate this limitation, we inject frequency-enhanced priors before selective state propagation:

$$Z_{\text{freq}} = \text{Gate}_{\text{freq}}\left(\text{SiLU}\left(\text{Conv}_{3\times3}(Z)\right)\right). \tag{8}$$

Unlike post-hoc frequency refinement strategies, the proposed frequency gate is inserted before selective scanning, enabling the state-space model to encode enhanced anatomical boundaries during long-range dependency propagation. To capture global non-causal anatomical dependencies, the enhanced feature Z_{freq} is

processed through a 4-direction (horizontal, vertical, and their reverses) selective state-space model (SSM). The multi-directional outputs are then reshaped and aggregated to form the global representation:

$$Y = \sum_{k=1}^{4} \text{Unflatten}^{(k)} \left(\text{SSM}^{(k)} \left(\text{Flatten}^{(k)}(Z_{\text{freq}}) \right) \right), \tag{9}$$

where $k \in \{1, 2, 3, 4\}$ denotes the four scanning directions. Outputs from all directions are reshaped back to the spatial domain and aggregated to form the global contextual representation:

$$Y = \sum_{k=1}^{4} \text{Unflatten}^{(k)} \left(y^{(k)} \right). \tag{10}$$

2.2 Implicit Coordinate Upsampling

Inspired by implicit image representation methods such as LIIF [25], we introduce an implicit coordinate-based upsampling module to enable continuous-scale reconstruction. Unlike conventional discrete operators, such as transposed convolution and pixel shuffle, which are restricted to integer scale factors, our approach conditions feature reconstruction on continuous spatial coordinates, facilitating flexible resolution adaptation for multi-center MR images. Given a low-resolution feature map F_{LR} and scale factor s, we first obtain a bicubic-interpolated base feature:

$$F_{\text{up}} = \text{Bicubic}(F_{\text{LR}}, s) \in \mathbb{R}^{C \times sH \times sW}. \tag{11}$$

A normalized coordinate grid $V_{\text{coord}} \in \mathbb{R}^{2 \times sH \times sW}$ is then constructed with coordinates linearly distributed in $[-1, 1]$. The coordinate grid is concatenated with F_{up} and projected into coordinate-conditioned features through a lightweight fusion network ϕ_{coord}:

$$F_{\text{imp}} = \phi_{\text{coord}} \left(\text{Concat}(F_{\text{up}}, V_{\text{coord}}) \right). \tag{12}$$

A refinement network ψ_{refine} predicts high-frequency residuals, while a sigmoid gate adaptively controls their contribution:

$$F_{\text{HR}} = F_{\text{up}} + \sigma \left(W_{\text{gate}} * F_{\text{imp}} \right) \odot \psi_{\text{refine}}(F_{\text{imp}}). \tag{13}$$

This design enhances anatomical boundaries while suppressing artifacts in homogeneous tissue regions.

3 Experiments

3.1 Datasets

IXI Dataset. The IXI dataset [26] is a publicly available brain imaging dataset comprising multi-modal MRI scans (T1, T2, and PD sequences) of healthy adults

from diverse scanning sources. In this study, we use only the T2-weighted images, as T2 provides rich soft-tissue contrast and is well suited for evaluating the reconstruction of fine anatomical structures. Restricting the experiments to a single modality also avoids confounding effects from cross-modal information and enables a focused assessment of the proposed super-resolution framework. Following preprocessing and spatial sampling, we obtain 4,524 slices for training, 1,508 for validation, and 1,536 for testing. The partitioning is performed at the patient level, ensuring that no subjects are shared across the three sets and thus preventing patient-level data leakage.

3.2 Implementation Details

Experiments are implemented in PyTorch and conducted on NVIDIA RTX 4090 GPUs. High-resolution training patches are cropped to 192×192 and augmented with random flips and rotations. Our model is trained for 250K iterations using L_1 loss (batch size 4) and the Adam optimizer ($\beta_1 = 0.9, \beta_2 = 0.99$). The initial learning rate is 1×10^{-4}, halved at 125K, 200K, 225K, and 237.5K iterations. Low-resolution inputs are generated from high-resolution images through MRI-specific k-space truncation, which simulates resolution degradation by removing high-frequency components. The same degradation process is consistently applied to both training and testing datasets. Notably, for fair comparison, baselines including Restore-RWKV [24] and SwinIR [9] are trained for 750K iterations with a 2×10^{-4} initial learning rate to ensure full convergence. Performance is evaluated every 5K iterations using PSNR, and the best model is selected.

3.3 Results and Analysis

Table 1 compares the quantitative performance on the IXI dataset. The proposed GF-BrainSR consistently outperforms representative CNN, Transformer, Mamba, and RWKV baselines across all scales. Notably, under severe degradation, it surpasses the runner-up MambaIR by 0.30 dB and 0.22 dB in PSNR at $\times 4$ and $\times 8$, respectively. These improvements demonstrate that our scale-adaptive frequency gating and implicit upsampling effectively capture unique anatomical boundaries while maintaining linear-complexity global context modeling. Qualitative comparisons are visualized in Fig. 2. Admittedly, these performance gains come at the cost of substantially increased computational complexity and parameter count.

3.4 Ablation Study

Frequency-Aware Encoder Mechanisms. Table 2 compares encoder variants for $\times 8$ SR with our implicit decoder. The baseline VSS block achieves 26.44 dB, while naive DWT integration slightly degrades performance (26.43 dB). Pseudo-IDWT improves PSNR to 26.58 dB but remains affected by aliasing, whereas removing the spatial gate reduces PSNR to 26.57 dB, highlighting the importance

Table 1: Quantitative results with different methods and upsampling factors on the IXI dataset. Best and second-best results are highlighted in **bold** and underlined, respectively.

Method	2×		4×		8×	
	PSNR↑	**SSIM↑**	**PSNR↑**	**SSIM↑**	**PSNR↑**	**SSIM↑**
SRResNet [5]	31.78	0.9467	27.73	0.8598	24.23	0.7389
EDSR [4]	34.38	0.9605	29.22	0.8841	24.51	0.7499
RCAN [6]	36.98	0.9702	28.98	0.8802	24.30	0.7463
ECBSR [7]	31.24	0.9374	25.32	0.7953	22.29	0.6796
SwinIR [9]	38.69	0.9782	30.79	0.9075	25.92	0.7891
FreMamba [27]	34.25	0.9582	26.49	0.8492	23.67	0.7375
IRSRMamba [20]	39.17	0.9797	31.27	0.9149	26.30	0.8045
MambaIR [18]	39.36	0.9805	31.25	0.9139	26.42	0.8068
Restore-RWKV [24]	38.13	0.9782	30.08	0.9117	25.23	0.7973
MambaIRV2 [19]	38.97	0.9790	30.78	0.9085	25.80	0.7866
GF-BrainSR (ours)	**39.46**	**0.9809**	**31.55**	**0.9201**	**26.64**	**0.8163**

Table 2: Quantitative ablation study on the evolution of the frequency-aware encoder for the ×8 super-resolution task.

Encoder Architecture	**PSNR↑**	**SSIM↑**
VSS Block (Baseline)	26.44	0.8094
Parallel DWT VSS Block	26.43	0.8081
Pseudo-IDWT VSS Block	26.58	0.8150
FI-VSS Block w/o gate	26.57	**0.8164**
FI-VSS Block (Ours)	**26.64**	0.8163

Table 3: Comparison of frequency decomposition methods for the ×8 super-resolution task.

Decomposition	**PSNR↑**	**SSIM↑**
FFT-based	25.91	0.7953
DWT-based (Ours)	**26.48**	**0.8131**

of localized frequency modulation. Our complete FI-VSS Block achieves the best performance (26.64 dB) by combining orthogonal IDWT with spatial gating for effective boundary enhancement and noise suppression. We further replace DWT with FFT under the same architecture. As shown in Table 3, the FFT variant drops PSNR/SSIM from 26.48/0.8131 to 25.91/0.7953, demonstrating that DWT's localized multi-scale representation is better suited to capturing anatomical high-frequency structures in brain MRI.

Scale-aware Depthwise Gating. Table 4 demonstrates a clear scale-dependent effect of kernel size in the frequency gating module. The joint $3 \times 3 + 5 \times 5$ kernels perform best at ×2 and ×4 degradation (39.46 and 31.55 dB), while the compact 3×3 kernel is optimal at ×8 (26.64 dB). This suggests that mild-to-moderate degradation benefits from multi-scale context, whereas severe degradation favors localized modeling to preserve fragile high-frequency anatomical details, supporting our scale-aware gating design.

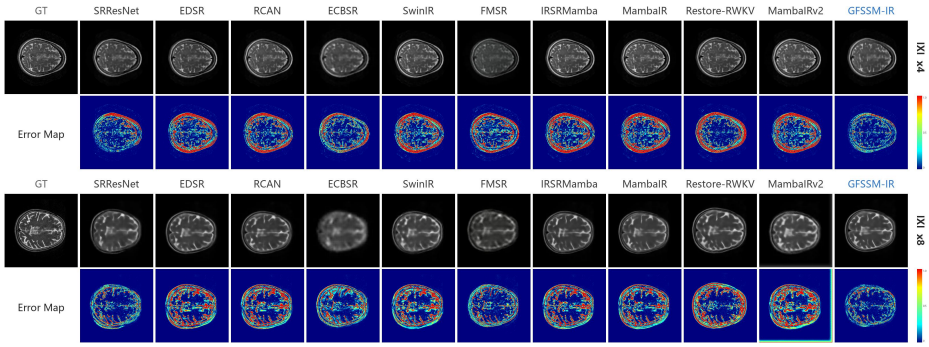

Fig. 2: Qualitative comparison of $\times 4$ and $\times 8$ MRI super-resolution on the IXI dataset. Error maps visualize per-pixel absolute residuals between SR predictions and ground truth.

Table 4: Ablation study of different kernel sizes on the IXI dataset for depthwise gating in the frequency gate of the FI-VSS block. Best results are highlighted in **bold**.

Method	2×		4×		8×	
	PSNR↑	**SSIM↑**	**PSNR↑**	**SSIM↑**	**PSNR↑**	**SSIM↑**
3×3	39.33	0.9803	31.37	0.9167	**26.64**	**0.8163**
5×5	39.38	0.9805	31.47	0.9170	26.48	0.8131
3×3 + 5×5	**39.46**	**0.9809**	**31.55**	**0.9201**	26.56	0.8150

Decoder Architecture. Table 5 validates our continuous spatial decoder. The PixelShuffle baseline yields the lowest performance (26.38 dB), as its discrete nature introduces severe aliasing artifacts and struggles to reconstruct smooth anatomical boundaries at large scales. Replacing it with coordinate-based implicit upsampling improves PSNR to 26.59 dB, demonstrating that continuous spatial modeling effectively mitigates discrete interpolation bottlenecks. Finally, incorporating our spatial gating mechanism achieves the best result (26.64 dB). This gate dynamically regulates high-frequency detail injection, successfully preventing spurious textures in homogeneous soft-tissue regions.

4 Conclusion

We present GF-BrainSR, an efficient linear-complexity super-resolution framework tailored for brain MRI. By coupling a DWT-driven frequency gate with selective state-space models, it successfully decouples high-frequency anatomical boundaries from acquisition noise. Furthermore, our gated implicit upsampling eliminates discrete interpolation artifacts, reconstructing morphologically continuous tissue interfaces. Experiments confirm that GF-BrainSR consistently outperforms state-of-the-art baselines, delivering superior anatomical clarity for

Table 5: Ablation study on different upsampling decoder architectures for the $\times 8$ super-resolution task on the IXI dataset. "Implicit" denotes continuous coordinate upsampling, and "Gate" refers to the spatial gating mechanism for detail refinement. Best results are highlighted in **bold**.

Decoder Architecture	**PSNR↑**	**SSIM↑**
Standard PixelShuffle	26.38	0.8095
Implicit Upsample w/o Gate	26.59	0.8156
Implicit Upsample w/ Gate (Ours)	**26.64**	**0.8163**

clinical diagnosis. Future work will extend the framework to a 3D state-space architecture to exploit full spatial continuity. We will also validate its generalization on multi-center pathological datasets and explore its utility in enhancing downstream clinical tasks, such as automated brain MRI image classification and segmentation.

Disclosure of Interests

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

References

1. B Sahu and G Madani. Imaging inequality: exploring the differences in radiology between high-and low-income countries, 2024.
2. TA Awolowo, GM OLUMIDE, OO OLAPOSI, OO ADENIYI, AO AYOBAMI, DO ALFRED, AC IYABO, and A BENJAMIN. A critical appraisal of the unmet needs for cross-sectional imaging techniques in nigeria. *International Journal of Radiology and Diagnostic Imaging*, 7:19–24, 2024.
3. Chun Wai Chin, Haniza Yazid, and Hoi Leong Lee. Challenges, advances, and evaluation metrics in medical image enhancement: A systematic literature review. *arXiv preprint arXiv:2510.13638*, 2025.
4. Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 136–144, 2017.
5. Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.
6. Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)*, pages 286–301, 2018.

7. Xindong Zhang, Hui Zeng, and Lei Zhang. Edge-oriented convolution block for real-time super resolution on mobile devices. In *Proceedings of the 29th ACM international conference on multimedia*, pages 4034–4043, 2021.
8. Fuzhi Yang, Huan Yang, Jianlong Fu, Hongtao Lu, and Baining Guo. Learning texture transformer network for image super-resolution. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5790–5799. IEEE, 2020.
9. 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.
10. Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5728–5739, 2022.
11. Zhisheng Lu, Juncheng Li, Hong Liu, Chaoyan Huang, Linlin Zhang, and Tieyong Zeng. Transformer for single image super-resolution. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 456–465. IEEE, 2022.
12. Marcos V Conde, Ui-Jin Choi, Maxime Burchi, and Radu Timofte. Swin2sr: Swinv2 transformer for compressed image super-resolution and restoration. In *European conference on computer vision*, pages 669–687. Springer, 2022.
13. Yufei Wang, Wenhan Yang, Xinyuan Chen, Yaohui Wang, Lanqing Guo, Lap-Pui Chau, Ziwei Liu, Yu Qiao, Alex C Kot, and Bihan Wen. Sinsr: diffusion-based image super-resolution in a single step. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 25796–25805. IEEE, 2024.
14. Sicheng Gao, Xuhui Liu, Bohan Zeng, Sheng Xu, Yanjing Li, Xiaoyan Luo, Jianzhuang Liu, Xiantong Zhen, and Baochang Zhang. Implicit diffusion models for continuous super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10021–10030, 2023.
15. Zongsheng Yue, Jianyi Wang, and Chen Change Loy. Resshift: Efficient diffusion model for image super-resolution by residual shifting. *Advances in neural information processing systems*, 36:13294–13307, 2023.
16. Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*, 2023.
17. Bo Peng, Eric Alcaide, Quentin Anthony, Alon Albalak, Samuel Arcadinho, Stella Biderman, Huanqi Cao, Xin Cheng, Michael Chung, Leon Derczynski, et al. Rwkv: Reinventing rnns for the transformer era. In *Findings of the association for computational linguistics: EMNLP 2023*, pages 14048–14077, 2023.
18. Hang Guo, Jinmin Li, Tao Dai, Zhihao Ouyang, Xudong Ren, and Shu-Tao Xia. Mambair: A simple baseline for image restoration with state-space model. In *European conference on computer vision*, pages 222–241. Springer, 2024.
19. Hang Guo, Yong Guo, Yaohua Zha, Yulun Zhang, Wenbo Li, Tao Dai, Shu-Tao Xia, and Yawei Li. Mambairv2: Attentive state space restoration. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 28124–28133, 2025.
20. Yongsong Huang, Tomo Miyazaki, Xiaofeng Liu, and Shinichiro Omachi. Irsrmamba: Infrared image super-resolution via mamba-based wavelet transform feature modulation model. *IEEE Transactions on Geoscience and Remote Sensing*, 2025.
21. Wenfeng Huang, Xiangyun Liao, Wei Cao, Wenjing Jia, and Weixin Si. Versatile and efficient medical image super-resolution via frequency-gated mamba. In *2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 01–08. IEEE, 2025.

22. Kui Jiang, Mengru Yang, Yi Xiao, Jianbo Wu, Guangcheng Wang, Xiaocheng Feng, and Junjun Jiang. Rep-mamba: Re-parameterization in vision mamba for lightweight remote sensing image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 2025.
23. Yuzhen Du, Teng Hu, Jiangning Zhang, Ran Yi Chengming Xu, Xiaobin Hu, Kai Wu, Donghao Luo, Yabiao Wang, and Lizhuang Ma. Exploring real&synthetic dataset and linear attention in image restoration. *arXiv preprint arXiv:2412.03814*, 2024.
24. Zhiwen Yang, Jiayin Li, Hui Zhang, Dan Zhao, Bingzheng Wei, and Yan Xu. Restore-rwkv: Efficient and effective medical image restoration with rwkv. *IEEE Journal of Biomedical and Health Informatics*, 30(1):513–526, 2025.
25. Hanting Chen, Yunhe Wang, Tianyu Guo, Chang Xu, Yiping Deng, Zhenhua Liu, Siwei Ma, Chunjing Xu, Chao Xu, and Wen Gao. Pre-trained image processing transformer. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12299–12310, 2021.
26. IXI Dataset. Ixi dataset. https://brain-development.org/ixi-dataset/, 2026.
27. Yi Xiao, Qiangqiang Yuan, Kui Jiang, Yuzeng Chen, Qiang Zhang, and Chia-Wen Lin. Frequency-assisted mamba for remote sensing image super-resolution. *IEEE Transactions on Multimedia*, 27:1783–1796, 2024.