Directional Adaptive Shuffle-Based Visual State-Space Models for Medical Image Restoration

Simon C. K. Chan, Lulin Shi, Bingxin Huang, Terence T. W. Wong(✉)

Translational and Advanced Bioimaging Laboratory, Department of Chemical and Biological Engineering,
Hong Kong University of Science and Technology, Hong Kong, China
ttwwong@ust.hk

Abstract. Medical image restoration (MedIR) demands precise modeling of anisotropic spatial dependencies, where directional anatomical patterns are frequently degraded by conventional methods. We propose Directional Adaptive Shuffle Mamba (DASMamba), a state-space model architecture that addresses this challenge through two novel components: (1) the Directional Adaptive Shuffle Module (DASM), which captures long-range dependencies via directional adaptive random shuffle and selective scanning, and (2) the Dual-path Feedforward Network (DPFN), enhancing feature representation through multi-scale learning and dynamic channel fusion. By integrating these modules into a hierarchical U-shaped architecture, DASMamba achieves state-of-the-art performance on MRI super-resolution, CT denoising, and PET synthesis tasks while maintaining linear computational complexity. Our framework's ability to preserve diagnostically critical structural details underscores its clinical value. The code is available at https://github.com/cc111mp/DASMamba-MedIR.

Keywords: Medical Image Restoration · State-Space Models · Anisotropic Processing.

1 Introduction

Modern clinical diagnostics rely on high-quality medical imaging, yet routine acquisitions are often compromised by noise (CT), limited resolution (MRI), or artifacts (PET). Medical image restoration (MedIR) enhances low-quality (LQ) inputs into diagnostically useful high-quality (HQ) outputs, particularly critical for MRI super-resolution [19,28], CT denoising [1,14,24], and PET synthesis [23]. A key challenge lies in medical images' inherent anisotropy—directional intensity variations from anatomical structures like vascular networks. Conventional methods often oversmooth these patterns [4,20], risking loss of critical diagnostic details. The challenge of anisotropy is not confined to 3D volumes. Even within 2D slices, crucial medical structures such as vascular networks, muscle fibers, and organ boundaries exhibit strong, direction-specific patterns and textures.

Treating these features isotropically can lead to blurring and the loss of fine details, which are often vital for accurate diagnosis.

Current deep learning approaches have complementary limitations. While CNNs [17] effectively model local textures, their limited receptive fields hinder long-range dependency capture. Transformers [6] address this through global attention but incur quadratic complexity, making them impractical for high-resolution data. SSMs [7] offer linear complexity alternatives but struggle with complex spatial relationships. Additionally, shuffle-based strategies [22] have shown promise in homogenizing spatial distributions, which inspires our approach.

We propose DASMamba, an SSM-based framework combining: DASM preserving anisotropic structures via adaptive shuffling while modeling global contexts through selective scanning, and DPFN enhancing feature learning via parallel multi-scale processing and dynamic channel interactions. Integrated in a hierarchical U-shaped architecture, our framework achieves state-of-the-art performance across MRI/CT/PET tasks with linear complexity, enabling clinical practicality for high-resolution imaging. Our work explicitly addresses this by employing row and column shuffling to present varied spatial arrangements to a four-directional scanning mechanism, ensuring that the model can adapt to dominant structural axes rather than averaging them away.

2 Method

2.1 Vision State-Space Module

State-Space Models (SSMs) are powerful computer vision frameworks for modeling long-range dependencies. The 2D Selective Scan Module (2D-SSM) [12] captures spatial relationships through multi-directional feature scanning using discretized state-space equations along horizontal, vertical, and diagonal axes. Summation-based fusion of directional outputs preserves 2D structural integrity, demonstrating SSMs' effectiveness in vision tasks through systematic integration of cross-dimensional spatial information.

2.2 Network Architecture

The DASMamba network employs a hierarchical U-shaped architecture (Figure 1(a)) to process low-quality input I_{LQ}. It generates shallow features F_S through an initial convolution. The encoder and decoder levels, with block depths of $[2, 2, 4, 4]$, utilize the DASM for anisotropic feature extraction. Three refinement blocks produce the residual image I_R, yielding the final output as $\hat{I}_{HQ} = I_{LQ} + I_R$.

2.3 2D-Directional Adaptive Shuffle Scan Block

To effectively capture the anisotropic nature of medical images, we propose a 2D-Directional Adaptive Shuffle Scan Module(2D-ASM) within our Directional

Adaptive Shuffle Block (DASB). This approach begins with spatially shuffling the input tensor $x \in \mathbb{R}^{B\times C\times H\times W}$ while preserving channel-wise information integrity. The shuffling operation can be formalized as:

$$x, \mathbf{H}, \mathbf{W} = \text{shuffle}(x), \quad x_{\text{restore}} = \text{reverse}(x, \mathbf{H}, \mathbf{W}) \tag{1}$$

where $\mathbf{H}$ and $\mathbf{W}$ store the shuffling indices for exact spatial restoration. Following the shuffle operation, our 2D-ASM employs a guided filtering mechanism that adaptively enhances directional features through a series of operations:

$$f_{avg} = \text{AP}(\text{Conv}(x)), \quad f_{max} = \text{MP}(\text{Conv}(x)) \tag{2}$$

$$f = \text{Conv}([f_{avg}, f_{max}]), \quad f_g = \text{reshape}(f, [B, K, G]), \quad \text{prompt} = f_g \cdot E \tag{3}$$

where AP and MP denote adaptive average pooling and adaptive max pooling respectively, $K = 4$ represents our directional scanning paths, G is the number of feature groups, and $E \in \mathbb{R}^{G\times d_{\text{state}}}$ represents learnable embeddings for feature enhancement. The concatenation of average and max pooling features enables the module to capture both global statistics and salient local features.

The 2D-DAS then processes these enhanced features through multiple directional state-space operations:

$$Y = \{y_k \mid k \in [1, 4]\}, \quad y_k = \text{SSM}_k(x, \text{prompt}) \tag{4}$$

The final output combines filtered and aggregated features through an adaptive weighting mechanism:

$$Y_{\text{filter}} = \sum_{k=1}^{4}(y_k \cdot f_g), \quad Y_{\text{sum}} = \sum_{k=1}^{4} y_k, \quad Y_{\text{out}} = \alpha Y_{\text{filter}} + \beta Y_{\text{sum}} \tag{5}$$

where α and β are learnable parameters that balance the contribution of filtered and aggregated features. Through this adaptive mechanism, our DASB achieves three key advantages: maintaining channel-wise information integrity during spatial reorganization, capturing directional features through multiple selective scanning paths, and adaptively enhancing features based on input characteristics. This framework effectively processes anisotropic medical image features, making it particularly suitable for medical image restoration tasks where directional patterns and structural details are crucial.

2.4 Dual-path Feed-forward Network Module

We propose a DPFN module that incorporates dual-path processing with dynamic feature fusion. The module is designed to enhance feature representation through multi-scale learning and efficient channel interactions. Given an input $X \in \mathbb{R}^{C\times H\times W}$, the module operates through several carefully designed stages: The initial transformation expands the feature space to allow for richer representation learning:

$$X_{exp} = W_{in}(X), \quad X_1, X_2 = \text{split}(X_{exp}) \tag{6}$$

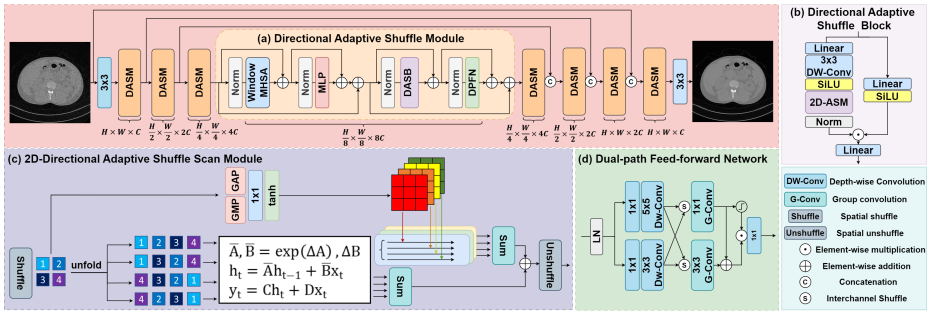

Fig. 1. Overview of the DASMamba architecture. (a) Directional Adaptive Shuffle Module (DASM) showing the complete processing pipeline. (b) Directional Adaptive Shuffle Block (DASB). (c) 2D-Directional Adaptive Shuffle Scan illustrating the state-space formulation and shuffle mechanism (2D-ASM). (d) Dual-path Feed-forward Network (DPFN) with depth-wise and group convolutions.

W_{in} is a 1×1 convolution that expands the channels by a factor of 2β, creating a broader feature space for subsequent multi-scale processing. This bifurcation facilitates parallel feature extraction paths, utilizing depthwise convolutions with varying receptive fields.

$$Y_1 = D_{3\times3}(X_1), \quad Y_2 = D_{5\times5}(X_2) \tag{7}$$

where $D_{k\times k}$ denotes depthwise convolution. The 3×3 kernel captures local structural patterns, while the 5×5 kernel aggregates broader contextual information, enabling complementary feature learning at different scales.

To enable cross-path information exchange and feature refinement, we introduce an interleaved channel fusion mechanism followed by group convolutions and adaptive modulation:

$$F_1 = \mathcal{F}(Y_1^{odd}, Y_2^{even}), \quad F_2 = \mathcal{F}(Y_2^{even}, Y_1^{odd}) \tag{8}$$

$$F_{out} = W_{1\times1}(F_1) + \text{GELU}(W_{3\times3}(F_2)) \tag{9}$$

$$Y = W_{out}(F_{out} \odot \text{GELU}(Y_1)) \tag{10}$$

where $\mathcal{F}$ represents channel-wise concatenation, $W_{1\times1}$, $W_{3\times3}$ are group convolutions (groups=C/β) for channel mixing and spatial modeling, and W_{out} performs channel reduction with GELU-based feature recalibration.

3 Experimental and Results

3.1 Datasets

For experimental evaluation, we conduct studies on three medical imaging tasks: **MRI Super-Resolution:** The experiments use T2-weighted MRI scans from

Table 1. Quantitative comparison of different medical image restoration methods on MRI, CT, and PET modalities. The best results are highlighted in **bold** and the second-best results are underlined. Methods marked with * are re-implemented by us.

Method	Params	FLOPs	MRI Super-Resolution			CT Denoising			PET Synthesis			Average		
	(M)	(G)	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
SwinIR [11]	11.50	187.9	31.5549	0.9334	30.5788	/	/	/	/	/	/	/	/	/
Restormer [26]	26.12	35.20	31.8474	0.9378	29.2095	/	/	/	/	/	/	/	/	/
MambaIR [5]	31.50	34.34	31.7700	0.9369	29.8372	/	/	/	/	/	/	/	/	/
CTformer [21]	1.45	3.44	/	/	/	33.2500	0.9134	8.8974 /	/	/	/	/	/	/
DenoMamba [16]	112.6	110.2	/	/	/	33.5300	0.9149	8.6115 /	/	/	/	/	/	/
SpachTransformer [10]	19	/	/	/	/	/	/	/	37.1371	0.9456	0.0871	/	/	/
ARGAN [13]	31.14	8.48	/	/	/	/	/	/	36.7300	0.9406	0.0902	/	/	/
AMIR [23]	23.54	31.77	31.9923	0.9393	29.2095	33.6738	0.9183	8.4773	37.2121	0.9473	0.0863	34.2927	0.9350	12.5910
Restore-RWKV [25]	27.91	37.45	32.0913	**0.9408**	28.9713	33.7988	0.9188	8.3600	37.3314	0.9474	0.0852	34.4072	0.9357	12.5722
NAFNet* [2]	67.89	15.80	31.9078	0.9376	29.5394	33.7848	0.9200	8.3742	37.2672	0.9453	0.0859	34.3199	0.9343	12.6668
CODEFormer* [29]	12.18	22.44	31.8879	0.9371	29.5584	33.8066	0.9203	8.3532	37.2846	0.9448	0.0856	34.3264	0.9341	12.6657
XFormer* [27]	25.21	35.72	31.9122	0.9373	29.4700	33.7610	0.9200	8.3947	37.2542	0.9450	0.0859	34.3091	0.9341	12.6502
Histoformer* [19]	26.24	36.30	31.9210	0.9378	29.4177	33.8115	0.9206	8.3480	37.3628	0.9472	0.0850	34.3651	0.9352	12.6169
VmambaIR* [18]	22.97	19.21	31.9277	0.9380	29.3439	33.7627	0.9203	8.3943	37.3673	0.9468	0.0847	34.3526	0.9350	12.6076
DASMamba-S	4.76	3.39	31.7919	0.9355	29.8132	33.7948	0.9202	8.3646	37.2895	0.9460	0.0857	34.2921	0.9339	12.7545
DASMamba	27.13	12.96	**32.1546**	0.9406	**28.6585**	**33.8458**	**0.9210**	**8.3192**	**37.3726**	**0.9477**	**0.0845**	**34.4577**	**0.9364**	**12.3541**

the IXI dataset [9]. We select the central 100 slices (256×256) from each 3D volume to rich content can be shown in those 100 images. Low-resolution images are simulated by downsampling the k-space data by a factor of $4\times$. We split the dataset into 405, 59, and 114 cases for training, validation, and testing.

CT Denoising: Our evaluation uses the 2016 NIH AAPM-Mayo Clinic Low-Dose CT Challenge dataset [15] with paired normal and quarter-dose scans at 512×512 resolution. The dataset includes 10 patients, with 8 used for training, 1 for validation, and 1 for testing.

PET Synthesis: We evaluate a PET imaging dataset [8] comprising 159 scans ($192 \times 192 \times 400$, voxel size: $3.15\text{mm} \times 3.15\text{mm} \times 1.87\text{mm}$). The low-dose images are obtained by reducing the original dose by a factor of 12. The dataset is divided into 120/10/29 volumes for training, validation, and testing, excluding air-only slices.

3.2 Loss Function

DASMamba optimizes both spatial and frequency domain dependencies through a dual-domain loss function:

$$\mathcal{L}_{\text{total}} = \frac{1}{S_r}\|I_{HQ} - I_{LQ}\|_1 + \lambda \cdot \frac{1}{S_r}\|\mathcal{F}(I_{HQ}) - \mathcal{F}(I_{LQ})\|_1, \tag{11}$$

where I_{HQ} and I_{LQ} denote high-quality ground truth and low-quality inputs respectively, S_r is the scale factor, $\mathcal{F}(\cdot)$ represents the Fourier transform, and λ balances the spatial and spectral components. The FFT loss guides the recovery of anisotropic patterns and ensures global frequency coherence through learnable channel-wise frequency embeddings, preventing blurring artifacts while maintaining critical spectral features.

3.3 Implementation

The proposed network uses a base channel dimension of 48 and a feedforward expansion ratio of 2 for DPFN modules. It is trained on 128×128 image patches

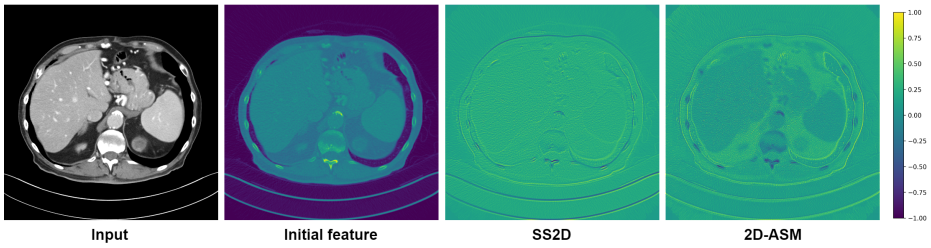

Fig. 2. Visualization of feature maps showing 2D-ASM effectiveness on a low-dose CT image: original image, initial features, SS2D-generated maps, and 2D-ASM-enhanced maps.

with a batch size of 4, using a cosine annealing learning rate schedule that decays from 2×10^{-4} to 1×10^{-6}. The model is trained for 300,000 iterations for CT denoising, PET synthesis, and MRI reconstruction tasks. All experiments are implemented in PyTorch and executed on an NVIDIA RTX 3090 GPU.

3.4 Comparative Experiments

DASMamba leverages directional adaptive shuffle scanning and a dual-path feed-forward network to effectively recover fine anatomical details from low-resolution MRI images. As shown in Table 1, our method achieves the highest PSNR (32.1546 dB) coupled with excellent RMSE performance, surpassing methods such as Restore-RWKV that rely on recurrent attention. Figure 3 further illustrates that DASMamba preserves contrast and subtle structural patterns—particularly fine line details—by mitigating blurring artifacts and enhancing local content recovery.

MRI Super-Resolution Results. Our DASMamba achieves superior performance in MRI super-resolution across major evaluation metrics, particularly in PSNR and RMSE, while maintaining competitive SSIM scores. Figure 3 demonstrates that our method better preserves contrast and structural patterns in brain MRI images, particularly in fine line details. While Restore-RWKV shows strong results with its recurrent attention mechanism, our approach further improves reconstruction quality through effective selective scanning.

CT Denoising Results. DASMamba attains state-of-the-art CT denoising performance, obtaining a PSNR of 33.85 dB. The architecture's selective scanning mechanism, which effectively captures anisotropic spatial dependencies, provides substantial improvements over conventional methods. In addition, our lightweight variant DASMamba-S delivers competitive results (33.79 dB PSNR) while dramatically reducing the parameter count (by 23.7×) and FLOPs (by 32.5×) compared to more complex approaches like DenoMamba. These results underscore the benefit of our adaptive shuffle strategy in efficiently restoring low-dose images while preserving critical anatomical structures.

PET Synthesis Results. Figure 3 presents the qualitative evaluation of different restoration approaches. The experimental results demonstrate that tra-

Table 2. Ablation experiments for the DASM module components.

Component	PSNR ↑	SSIM ↑	RMSE ↓
Baseline (W-MSA)	33.8064	0.9201	8.3549
+ SS2D	33.8124	0.9205	8.3474
+ 2D-ASM	33.8223	0.9205	8.3392
+ Shuffle + SS2D	33.8172	0.9204	8.3431
+ Shuffle + 2D-ASM	33.8317	0.9206	8.3305
+ Shuffle + DGFF	33.8317	0.9205	8.3308
+ Shuffle + MSFN	33.8324	0.9205	8.3284
2D-ASM + Shuffle + DPFN	**33.8352**	**0.9208**	**8.3267**

Table 3. Ablation studies on the group size for 2D-ASM.

Group	PSNR ↑
2	33.8331
4	33.8335
8	**33.8352**
12	33.8304

ditional methods have limitations in distinguishing signal from noise, resulting in detail degradation during the restoration process. Although existing approaches achieve broad spatial coverage, they show reduced effectiveness in handling oriented structural elements. DASMamba demonstrates superior capability in preserving anatomical boundaries and textural details. Quantitative metrics confirm this performance advantage, even when compared to recent approaches like VmambaIR and Restore-RWKV. DASMamba's consistent effectiveness across all metrics validates its strength in PET image processing.

Overall Performance. Across all three medical imaging tasks, DASMamba not only achieves state-of-the-art quantitative performance but also offers significant computational advantages. By combining state-space modeling with directional adaptive shuffling and dual-path fusion, our framework effectively reconstructs directional patterns and clinically significant high-frequency details with linear complexity. These advantages are reflected in the comprehensive evaluation metrics of Table 1, where DASMamba outperforms established CNN, Transformer, and SSM counterparts. Overall, our results demonstrate that the proposed method delivers both enhanced restoration quality and computational efficiency, underscoring its potential for practical clinical applications. The overall model complexity remains linear, as the 2D-ASM relies on a fixed-window multi-head self-attention (MHSA), and both the SSM blocks and the DPFN operate with linear complexity.

3.5 Ablation Studies

To evaluate the contributions of key components in DASMamba, we conduct comprehensive ablation experiments on CT image denoising. In these studies, we employ the window-based multi-head self-attention (W-MSA) block as a baseline.

Effect of Selective Scanning and Shuffling. Table 2 shows that while the baseline SS2D module successfully extracts features, incorporating the 2D-Directional Adaptive Shuffle Scan module (2D-ASM) further improves the capture of local patterns. As illustrated in Figure 2, the 2D-ASM highlights degraded regions and preserves anatomical details, with enhanced edge information and structural patterns apparent in the progression of features. The most significant

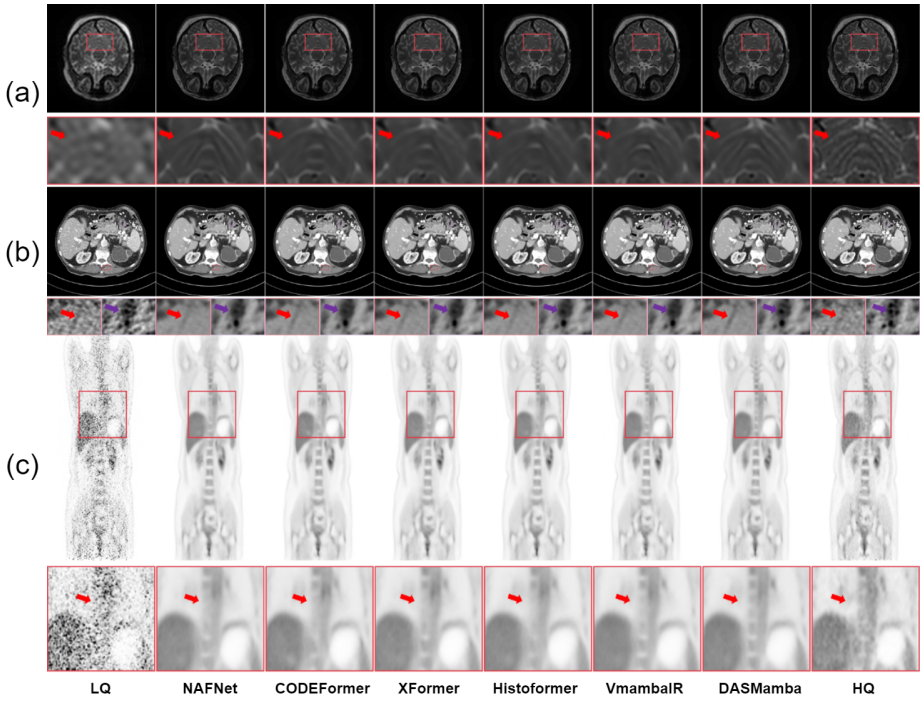

Fig. 3. Visual comparison of different methods: (a) CT image denoising, (b) PET image synthesis, and (c) MRI image super-resolution.

gains are achieved when shuffling is combined with 2D-ASM; here, shuffling introduces diverse spatial arrangements, while adaptive scanning efficiently detects directional patterns in the rearranged feature maps.

Ablation on Feed-Forward Network Design. Table 2 also reveals limitations in existing feed-forward network designs. The Mixed-scale FFN (MSFN) [3] processes 3×3 and 5×5 features sequentially, which restricts cross-scale interaction. In contrast, the Dual-scale Gated FFN (DGFF) [19] relies on pixel shuffling, a process that can distort fine details. Our proposed Dual-path Feedforward Network (DPFN) addresses these issues by processing multi-scale features in parallel, dynamically recombining channels across scales and adaptively fusing them via grouped and pointwise convolutions. The final feature calibration, enhanced by GELU activation, refines the representation without compromising spatial integrity, leading to superior performance compared to DGFF and MSFN.

Impact of Group Size in 2D-ASM. Experiments varying the group size within the 2D-ASM design (see Table 3) demonstrate that a group size of 8 achieves optimal performance. This setting effectively captures multi-directional features through the four complementary scanning paths and maintains computational efficiency. In contrast, smaller group sizes limit directional coverage,

while larger group sizes increase complexity without substantial gains in feature representation.

4 Conclusion

In this paper, we present DASMamba, a novel state-space model framework that combines adaptive spatial processing with a hierarchical U-shaped architecture to achieve state-of-the-art medical image restoration across MRI, CT, and PET, effectively preserving anisotropic structural details with linear computational complexity.

Acknowledgment

This work was supported by a startup grant at The Hong Kong University of Science and Technology (R9421).

Disclosure of Interest

T. T. W. W. has a financial interest in PhoMedics Limited, which, however, did not support this work. The authors declare no other conflicts of interest.

References

1. Chen, H., Zhang, Y., Kalra, M.K., Lin, F., Chen, Y., Liao, P., Zhou, J., Wang, G.: Low-dose ct with a residual encoder-decoder convolutional neural network. IEEE Transactions on Medical Imaging **36**(12), 2524–2535 (Dec 2017). https://doi.org/10.1109/tmi.2017.2715284, http://dx.doi.org/10.1109/TMI.2017.2715284
2. Chen, L., Chu, X., Zhang, X., Sun, J.: Simple baselines for image restoration (2022), https://arxiv.org/abs/2204.04676
3. Chen, X., Li, H., Li, M., Pan, J.: Learning a sparse transformer network for effective image deraining (2023), https://arxiv.org/abs/2303.11950
4. Dong, Z., He, Y., Qi, X., Chen, Y., Shu, H., Coatrieux, J.L., Yang, G., Li, S.: Mnet: Rethinking 2d/3d networks for anisotropic medical image segmentation (2022), https://arxiv.org/abs/2205.04846
5. Guo, H., Li, J., Dai, T., Ouyang, Z., Ren, X., Xia, S.T.: Mambair: A simple baseline for image restoration with state-space model (2024), https://arxiv.org/abs/2402.15648
6. Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., et al.: A survey on vision transformer. IEEE transactions on pattern analysis and machine intelligence **45**(1), 87–110 (2022)
7. Hatamizadeh, A., Kautz, J.: Mambavision: A hybrid mamba-transformer vision backbone. arXiv preprint arXiv:2407.08083 (2024)
8. Hudson, H., Larkin, R.: Accelerated image reconstruction using ordered subsets of projection data. IEEE Transactions on Medical Imaging **13**(4), 601–609 (1994). https://doi.org/10.1109/42.363108

9. Imperial College London, Guy's Hospital, Institute of Psychiatry: IXI Dataset - Information eXtraction from Images. https://brain-development.org/ixi-dataset/ (nd), ePSRC Project GR/S21533/02
10. Jang, S.I., Pan, T., Li, Y., Heidari, P., Chen, J., Li, Q., Gong, K.: Spach transformer: Spatial and channel-wise transformer based on local and global self-attentions for pet image denoising (2023), submitted to IEEE Transactions on Medical Imaging
11. Liang, J., Cao, J., Sun, G., Zhang, K., Gool, L.V., Timofte, R.: Swinir: Image restoration using swin transformer (2021), https://arxiv.org/abs/2108.10257
12. Liu, Y., Tian, Y., Zhao, Y., Yu, H., Xie, L., Wang, Y., Ye, Q., Liu, Y.: Vmamba: Visual state space model (2024), https://arxiv.org/abs/2401.10166
13. Luo, Y., Zhou, L., Zhan, B., Fei, Y., Zhou, J., Wang, Y., Shen, D.: Adaptive rectification based adversarial network with spectrum constraint for high-quality PET image synthesis (2022). https://doi.org/10.1016/j.media.2021.102335, https://www.sciencedirect.com/science/article/pii/S1361841521003807
14. Luthra, A., Sulakhe, H., Mittal, T., Iyer, A., Yadav, S.: Eformer: Edge enhancement based transformer for medical image denoising (2021), https://arxiv.org/abs/2109.08044
15. McCollough, C.H., Bartley, A.C., Carter, R.E., Chen, B., Drees, T.A., Edwards, P., Holmes, D.R., Huang, A.E., Khan, F., Leng, S., McMillan, K.L., Michalak, G.J., Nunez, K.M., Yu, L., Fletcher, J.G.: Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 Low Dose CT Grand Challenge. Medical Physics **44**(10), e339–e352 (oct 2017). https://doi.org/10.1002/mp.12345
16. "Ozt"urk, Ş., Duran, O.C., Çukur, T.: Denomamba: A fused state-space model for low-dose ct denoising (2024), https://arxiv.org/abs/2409.13094
17. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation (2015), https://arxiv.org/abs/1505.04597
18. Shi, Y., Xia, B., Jin, X., Wang, X., Zhao, T., Xia, X., Xiao, X., Yang, W.: Vmambair: Visual state space model for image restoration. arXiv preprint arXiv:2403.11423 (2024)
19. Sun, S., Ren, W., Gao, X., Wang, R., Cao, X.: Restoring images in adverse weather conditions via histogram transformer. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T., Varol, G. (eds.) Computer Vision – ECCV 2024. pp. 111–129. Springer Nature Switzerland, Cham (2025)
20. Tebini, S., Seddik, H., Ben Braiek, E.: Medical image enhancement based on new anisotropic diffusion function. In: 2017 14th International Multi-Conference on Systems, Signals & Devices (SSD). pp. 456–460 (2017). https://doi.org/10.1109/SSD.2017.8166957
21. Wang, D., Fan, F., Wu, Z., Liu, R., Wang, F., Yu, H.: Ctformer: Convolution-free token2token dilated vision transformer for low-dose ct denoising (2023), http://dx.doi.org/10.1088/1361-6560/acc000
22. Xiao, J., Fu, X., Zhu, Y., Li, D., Huang, J., Zhu, K., Zha, Z.J.: Homoformer: Homogenized transformer for image shadow removal. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 25617–25626 (June 2024)
23. Yang, Z., Chen, H., Qian, Z., Yi, Y., Zhang, H., Zhao, D., Wei, B., Xu, Y.: All-in-one medical image restoration via task-adaptive routing (2024), https://arxiv.org/abs/2405.19769

24. Yang, Z., Chen, H., Qian, Z., Zhou, Y., Zhang, H., Zhao, D., Wei, B., Xu, Y.: Region attention transformer for medical image restoration. In: Proceedings of the 27th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (2024), https://github.com/Yaziwel/Region-Attention-Transformer-for-Medical-Image-Restoration.git, arXiv:2407.09268
25. Yang, Z., Zhang, H., Zhao, D., Wei, B., Xu, Y.: Restore-rwkv: Efficient and effective medical image restoration with rwkv (2024), https://arxiv.org/abs/2407.11087
26. Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.H.: Restormer: Efficient transformer for high-resolution image restoration (2022), https://arxiv.org/abs/2111.09881
27. Zhang, J., Zhang, Y., Gu, J., Dong, J., Kong, L., Yang, X.: Xformer: Hybrid x-shaped transformer for image denoising (2024), https://arxiv.org/abs/2303.06440
28. Zhang, Y., Li, K., Li, K., Fu, Y.: Mr image super-resolution with squeeze and excitation reasoning attention network. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 13420–13429 (2021). https://doi.org/10.1109/CVPR46437.2021.01322
29. Zhao, H., Gou, Y., Li, B., Peng, D., Lv, J., Peng, X.: Comprehensive and delicate: An efficient transformer for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14122–14132. IEEE, New York, NY, USA (Jun 2023)