



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

SRMA-Mamba: Spatial Reverse Mamba Attention Network for Pathological Liver Segmentation in MRI Volumes

Jun Zeng¹, Quoc-Huy Trinh³, Deepak Ranjan Nayak⁴, Nikhil Kumar Tomar⁵,
Ulas Bagci⁵, and Debesh Jha² *

¹ Chongqing University of Posts and Telecommunications, China

² Biomedical Perception & Intelligence Lab, University of South Dakota, USA

³ Aalto University, Finland

⁴ Malaviya National Institute of Technology Jaipur, India

⁵ Machine & Hybrid Intelligence Lab, Northwestern University, USA
debesh.jha@usd.edu

Abstract. Liver cirrhosis plays a critical role in the prognosis of chronic liver disease. Early detection and timely intervention are essential for reducing mortality rates. However, the intricate anatomical architecture and diverse pathological changes of liver tissue complicate the accurate detection and characterization of pathological liver structures in clinical settings. Existing methods underutilize spatial anatomical details in volumetric MRI data, thereby hindering their clinical effectiveness and explainability. To address this challenge, we introduce a novel Mamba-based network, SRMA-Mamba, designed to model the spatial relationships within complex anatomical structures of MRI volumes. By integrating the Spatial Anatomy-Based Mamba module (SABMamba), SRMA-Mamba performs selective Mamba scans within pathological liver tissues and combines anatomical information from the sagittal, coronal, and axial planes to construct a global spatial context representation, enabling efficient volumetric segmentation of pathological liver structures. Furthermore, we introduce the Spatial Reverse Mamba Attention module (SRMA), designed to progressively refine boundary details in the segmentation map, utilizing both the coarse segmentation map and hierarchical encoding features. Extensive experiments demonstrate that SRMA-Mamba surpasses state-of-the-art methods, delivering exceptional performance in 3D pathological liver segmentation. The source code is available at <https://github.com/JunZengz/SRMA-Mamba>.

Keywords: Volumetric segmentation · Liver cirrhosis · Vision Mamba · State-space models · Reverse attention.

1 Introduction

Cirrhosis constitutes a significant global health burden, ranking as the 11th leading cause of mortality worldwide in 2019 [6]. As the end-stage manifestation

* Corresponding author.

of progressive liver fibrosis, its precise assessment is paramount for determining clinical prognosis and intervention strategies. While liver biopsy remains the diagnostic gold standard [1], its clinical utility is frequently hampered by its invasive nature and inherent risks, such as post-procedural hemorrhage. Consequently, Magnetic Resonance Imaging (MRI) has emerged as a cornerstone of non-invasive diagnostics [14], leveraging superior soft-tissue contrast to characterize cirrhotic architecture. However, manual delineation of pathological liver structures from high-dimensional MRI volumes is labor-intensive and susceptible to significant inter-observer variability, underscoring an urgent need for high-precision automated segmentation algorithms.

Recently, state-space models (SSMs) [3, 13, 2] have demonstrated exceptional performance in capturing long-range dependencies. Mamba [2] employs a selection mechanism and a hardware-aware algorithm to model global dependencies, achieving state-of-the-art performance in linear-time sequence modeling with linear complexity. To extend Mamba into the vision domain, VMamba [9] performs four distinct traversal path scans to achieve a 2D global receptive field. Based on this foundation, Mamba-based methods [12, 15–17] have emerged in the field of medical image segmentation. SegMamba [16] takes the whole volumetric data as a linear sequence and explores the forward, reverse, and inter-slice features interactions. However, such linearization inherently disrupts the intrinsic 3D spatial topology. For pathological liver segmentation, where cirrhotic structures exhibit highly irregular and anisotropic morphologies, discarding the spatial context can lead to suboptimal boundary definition and fragmented results. In this paper, we propose SRMA-Mamba to model the spatial relationships within the complex anatomical structures of MRI volumes for efficient volumetric pathological liver segmentation. The main contributions of the work are as follows:

- We present a novel spatial reverse mamba attention network, SRMA-Mamba, for efficient volumetric pathological liver segmentation in MRI volumes.
- We propose a novel Spatial Anatomy-Based Mamba (SABMamba) module to perform scans within multiple anatomical perspectives, including sagittal, coronal, and axial planes. Additionally, we introduce the Spatial Reverse Mamba Attention (SRMA) module, which progressively refines the segmentation map using hierarchical features and a coarse segmentation map.
- We introduce the Anatomy-Based Selective Scan (ABSS) module to model 3D visual feature representations for pathological liver tissues, enabling direct processing of 3D volumetric data with low computational complexity.
- We conduct extensive experiments on the CirrMRI600+ T1W and T2W datasets and demonstrate that SRMA-Mamba achieves state-of-the-art performance in 3D pathological liver segmentation, surpassing existing volumetric segmentation methods.

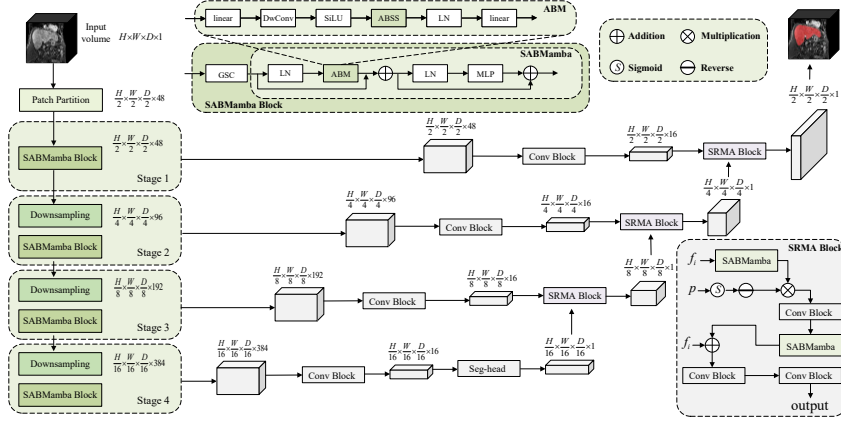


Fig. 1. Overview of the proposed *SRMA-Mamba* architecture.

2 Methodology

2.1 Overall Architecture of SRMA-Mamba

SRMA-Mamba is an innovative Mamba-based network specifically designed for volumetric segmentation of pathological liver in MRI volumes. The network leverages the Mamba-based encoder for hierarchical feature extraction and integrates a spatial reverse mamba attention module to progressively refine the segmentation map with features from each encoding stage. As illustrated in Figure 1, the proposed network employs an encoder-decoder architecture. The encoder extracts different levels of features from the input cirrhotic MRI volume and the decoder incrementally refines the segmentation maps, utilizing the coarse segmentation map alongside the features obtained from the encoding stages.

2.2 Spatial Anatomy-Based Mamba (SABMamba)

In clinical practice, radiologists typically review MRI volumes from multiple anatomical planes to develop a comprehensive spatial awareness of liver pathology. Inspired by this, we introduce the Spatial Anatomy-Based Mamba module (SABMamba), designed to model complex anatomical relationships between spatial voxels by integrating information from three orthogonal planes.

As shown in Figure 1, SABMamba consists of consecutive residual branches. The first branch includes a Layer Normalization followed by an ABM block, while the second branch contains a Layer Normalization and a multi-layer perceptron layer. This process can be formulated as follows:

$$x_1 = x + ABM(LN(x)), \quad (1)$$

$$x_2 = x_1 + MLP(LN(x_1)), \quad (2)$$

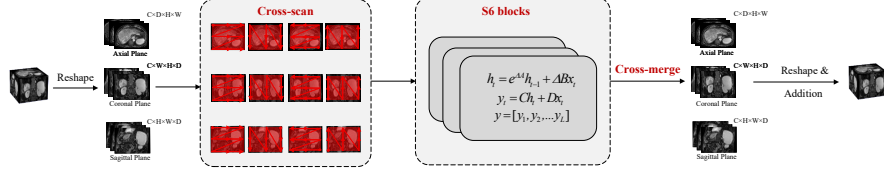


Fig. 2. Details of the *Anatomy-Based Selective Scan (ABSS)* module

where MLP stands for a multi-layer perceptron, and LN refers to Layer Normalization.

In the ABM block, the input features first pass through a linear layer and a depthwise separable convolution. The resulting features are then passed through the SiLU activation function before being fed into an ABSS module. Finally, the features undergo Layer Normalization, and a linear layer is applied to produce the final output of the ABM block.

2.3 Anatomy-Based Selective Scan (ABSS)

The overall structure of Anatomy-Based Selective Scan (ABSS) is illustrated in Figure 2. ABSS is employed to learn the spatial feature representation of cirrhotic liver tissue. It comprises three key components: scan expanding operation, scan merging operation, and the S6 module [2]. The selective mechanism within the S6 module effectively filters out irrelevant information while preserving relevant features.

Specifically, the input MRI volume is initially rearranged into three anatomical planes.

$$\begin{cases} f_h = \text{Reshape}(x) \in \mathbb{R}^{C \times H \times L_1}, \\ f_w = \text{Reshape}(x) \in \mathbb{R}^{C \times W \times L_2}, \\ f_d = \text{Reshape}(x) \in \mathbb{R}^{C \times D \times L_3}. \end{cases} \quad (3)$$

where $L_1 = W \times D$, $L_2 = H \times D$, $L_3 = H \times W$.

The scan expanding operation then unfolds the data from these planes into sequences along four distinct directions. Next, the S6 module selectively processes these sequences to extract features from the anatomical planes. After that, the scan merging operation integrates these features and restores them to their original resolution. This process can be summarized as follows:

$$Y = \text{CrossMerge}(\text{S6}(\text{CrossScan}(f))) \quad (4)$$

where $f \in \{f_h, f_w, f_d\}$, and CrossScan denotes the scan expanding operation, which rearranges the input features along multiple plane-wise scanning paths, and CrossMerge denotes the scan merging operation, which integrates the processed sequences from the S6 blocks into the final output feature map [9].

2.4 Spatial Reverse Mamba Attention (SRMA)

As depicted in Figure 1, the architecture of the Spatial Reverse Mamba Attention (SRMA) consists of SABMamba and convolutional blocks. SRMA takes the coarse segmentation map and feature map from the encoding stage as input, progressively refining edge details and enhancing feature representation. Specifically, the feature map from the current level and the segmentation map from the previous level are first fed into the SRMA module. Then the segmentation map undergoes a Sigmoid function and a reverse attention operation to generate the reverse attention map weights A_i , which can be formulated as:

$$A_i = 1 - \text{Sigmoid}(S_{i+1}), \quad (5)$$

where $i \in \{1, 2, 3\}$, S_{i+1} denotes the coarse segmentation map. Subsequently, element-wise multiplication is applied between the attention weights and the enhanced feature map. The resulting features F are further processed by a SABMamba module, which is summarized as follows:

$$F = \delta(\text{Conv}(A_i \cdot \delta(f_i))) \quad (6)$$

where $\delta(\cdot)$ denotes a SABMamba module. Finally, a residual connection is applied, followed by two convolutional layers to generate the output.

3 Experiments

3.1 Dataset

CirrMRI600+ [7] is a single-center, multi-vendor, multi-sequence dataset designed to enhance cirrhotic liver research in MRI volumes. The dataset includes 628 abdominal MRI scans, consisting of 310 T1-weighted (T1W) images and 318 T2-weighted (T2W) images. For the T1W dataset, 248 images were split for training, 31 for validation, and 31 for testing. Similarly, for the T2W dataset, 256 images were allocated for training, 31 for validation, and 31 for testing.

3.2 Implementation details

We implemented our model using the PyTorch framework on a single Tesla V100 GPU with 32 GB of memory. During the training phase, each volume is randomly cropped to a size of $224 \times 224 \times 64$. The Adam optimizer is employed to optimize the model parameters with the initial learning rate of $1e-4$. In our experiments, all models are trained for 500 epochs with an early stopping patience of 50. The batch size was set to 2. The data augmentation techniques, including spatial cropping, flipping, and intensity scaling, were employed to mitigate overfitting. We use Dice loss and cross-entropy loss to supervise the network training. We use the mean Dice similarity coefficient (mDSC), mean intersection over union (mIoU), recall, precision, Hausdorff distance (HD95), and average symmetric surface distance (ASSD) to evaluate the performance of our network.

Table 1. Model performance on the CirrMRI600+ T1W Liver Cirrhosis MRI dataset [7].

Method	mIoU	mDSC	Recall	Precision	HD95	ASSD
SegResNet	81.14	88.65	94.66	84.85	34.63	7.55
UX-Net	79.12	87.43	91.21	84.77	47.52	7.19
MedNeXt	82.49	89.97	93.62	87.53	29.41	4.75
SwinUNETR	80.61	88.71	92.73	86.13	36.79	4.66
SwinUNETRv2	79.68	87.30	89.99	85.87	33.27	6.07
SegMamba	85.70	91.80	95.56	88.98	26.97	4.69
SRMA-Mamba	87.59	92.95	97.13	89.77	17.52	3.05

Table 2. Model performance on the CirrMRI600+ T2W Liver Cirrhosis MRI dataset [7].

Method	mIoU	mDSC	Recall	Precision	HD95	ASSD
SegResNet	76.74	84.86	86.55	86.93	28.47	4.60
UX-Net	72.35	82.01	84.33	84.80	43.87	5.86
MedNeXt	69.03	79.87	83.32	81.28	44.83	6.33
SwinUNETR	74.90	83.91	87.01	84.83	34.61	4.84
SwinUNETRv2	75.21	84.89	85.51	86.96	29.37	4.26
SegMamba	76.99	85.36	87.41	87.00	30.75	4.52
SRMA-Mamba	78.24	86.25	87.71	87.52	26.04	4.57

3.3 Quantitative Comparison

To evaluate both the efficiency and effectiveness of the proposed SRMA-Mamba, we conducted a comparative analysis against state-of-the-art volumetric segmentation methods, including those based on CNNs (SegResNet [10], UX-Net [8], MedNeXt [11]), Transformers (SwinUNETR [4], SwinUNETRv2 [5]), and Mamba (SegMamba [16]). As demonstrated in Table 1 and Table 2, we evaluated the performance of SRMA-Mamba on the CirrMRI600+ T1W and T2W datasets, respectively. To ensure a fair comparison, all methods were trained and tested using the same codebase, dataset partitioning, and hyperparameter tuning procedures. On the T1W dataset, SRMA-Mamba exhibited outstanding performance, surpassing all other leading methods across key metrics, including mDSC, HD95, and ASSD. Specifically, our model achieves mDSC and mIoU scores of 92.95% and 87.59% on the T1W dataset, outperforming SegMamba by 1.15% and 1.89%, respectively. SRMA-Mamba also achieves impressive results with HD95 and ASSD values of 17.52 and 3.05, respectively. On the T2W dataset, SRMA-Mamba outperformed SegMamba with an impressive mDSC score of 86.25% and an IoU score of 78.24%, demonstrating its superior ability to delineate pathological liver structures. Additionally, compared to SegMamba, SRMA-Mamba demonstrates lower computational complexity, as shown in Table 3, while delivering superior segmentation results.

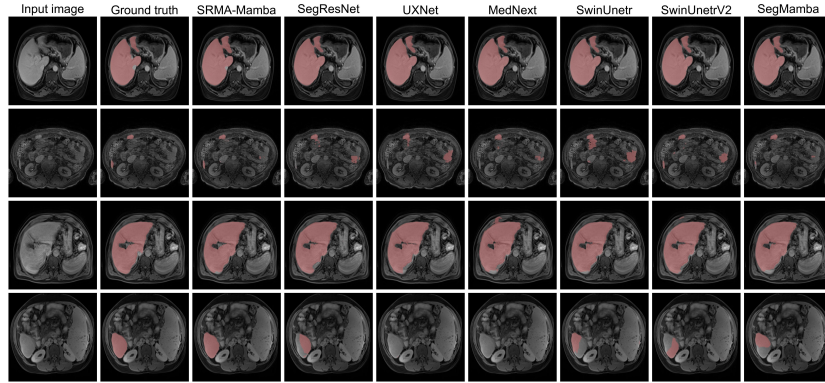


Fig. 3. Qualitative results of different methods on the CirrMRI600+ T1W dataset.

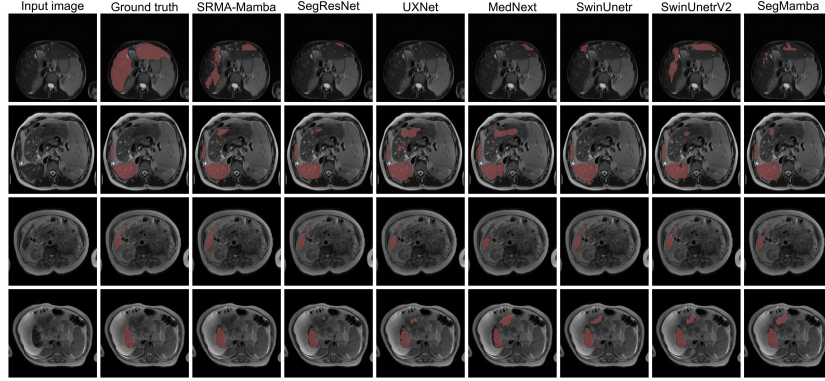


Fig. 4. Qualitative results of different methods on the CirrMRI600+ T2W dataset.

3.4 Qualitative Comparison

Figure 3 and Figure 4 present a qualitative comparison of state-of-the-art methods on the CirrMRI600+ T1W and T2W datasets, highlighting the superior capability of SRMA-Mamba in delineating pathological liver structures. The proposed SABMamba effectively models spatial voxel dependencies by integrating feature information across multiple anatomical planes. Additionally, SRMA-Mamba achieves impressive segmentation performance by optimizing boundary details at multiple decoding stages. On the CirrMRI600+ T1W dataset, SRMA-Mamba demonstrates superior performance in segmenting both large and small pathological regions. Despite the challenging cases in the CirrMRI600+ T2W dataset, which occasionally result in false positives and missed detections of cirrhotic tissues, SRMA-Mamba consistently outperforms most state-of-the-art models, showcasing exceptional segmentation capabilities.

Table 3. Computational complexity of various methods.

Method	Parameters	GFLOPs
SegResNet	1.17M	55.95
UX-Net	53.00M	2292.44
MedNeXt	5.54M	210.91
SwinUNETR	15.51M	303.33
SwinUNETRv2	18.15M	326.79
SegMamba	64.24M	2379.03
SRMA-Mamba	17.22M	149.14

Table 4. Ablation study on the CirrMRI600+ T1W dataset [7]. ✓ denotes the module is included, ✗ denotes the module is removed, corresponding to the “w/o” notation in the text.

Exp.	SABMamba	SRMA	mIoU	mDSC
#1	✗	✗	46.66	61.65
#2	✗	✓	68.61	80.62
#3	✓	✗	83.82	90.63
#4	✓	✓	87.59	92.95

Table 5. Performance comparison of different anatomical plane configurations on the CirrMRI600+ T1W dataset [7].

Anatomical planes	mIoU	mDSC	HD95
Axial	86.61	92.48	18.92
Sagittal	86.69	92.50	19.56
Coronal	82.78	90.22	27.38
Three Planes (Ours)	87.59	92.95	17.52

3.5 Ablation Study

To evaluate the effectiveness of the proposed modules, we perform comprehensive ablation experiments on the CirrMRI600+ T1W dataset. Table 4 presents the ablation results of SABMamba and SRMA. “w/o SABMamba” replaces the SABMamba encoder with the UNETR encoder, while “w/o SRMA” replaces the SRMA decoder with a standard 3D convolutional decoder. Replacing the SABMamba encoder with the UNETR encoder decreases the mDSC from 92.95% to 80.62%, demonstrating the effectiveness of SABMamba in modeling spatial anatomical information. In addition, removing SRMA reduces the mDSC from 92.95% to 90.63%, indicating that SRMA provides complementary improvements by enhancing feature aggregation and boundary delineation. The results in Table 5 further validate the effectiveness of the proposed ABSS module. By integrating information from axial, sagittal, and coronal planes, the three-plane configuration

consistently outperforms all single-plane variants, demonstrating the benefit of multi-plane anatomical representation learning.

4 Conclusion

This paper presents a novel spatial reverse mamba attention network, named SRMA-Mamba, for volumetric segmentation of pathological liver in MRI volumes. We introduce the SABMamba module, which performs planar scanning across the three anatomical planes and combines feature information from these planes to construct a spatial representation of liver cirrhosis regions. Additionally, SRMA-Mamba progressively refines boundary precision through multiple Spatial Reverse Mamba Attention modules, thereby enhancing the overall segmentation performance. We also introduce an Anatomy-Based Selective Scan module that directly processes voxel data to generate 3D feature representations, preserving spatial anatomical information. Extensive experiments demonstrate that SRMA-Mamba achieves state-of-the-art performance on both T1W and T2W modalities of the liver cirrhosis benchmark dataset, surpassing existing methods and highlighting its effectiveness and stability in the volumetric segmentation of pathological liver.

Acknowledgments D. Jha is supported in part by the U.S. Department of Education (P116Z240151 to the University of South Dakota and the South Dakota School of Mines & Technology). The views expressed are those of the author(s) and do not necessarily represent the official views of the U.S. Department of Education. D. R. Nayak is supported by the Anusandhan National Research Foundation (ANRF), Government of India (grant number ANRF/ARGM/2025/002890/TS and CRG/2023/007397).

Disclosure of Interests The authors have no competing interests to declare.

References

1. Chowdhury, A.B., Mehta, K.J.: Liver biopsy for assessment of chronic liver diseases: a synopsis. *Clinical and experimental medicine* **23**(2), 273–285 (2023)
2. Gu, A., Dao, T.: Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752* (2023)
3. Gu, A., Goel, K., Ré, C.: Efficiently modeling long sequences with structured state spaces. *arXiv preprint arXiv:2111.00396* (2021)
4. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: *International MICCAI brainlesion workshop*. pp. 272–284. Springer (2021)
5. He, Y., Nath, V., Yang, D., Tang, Y., Myronenko, A., Xu, D.: Swinunetr-v2: Stronger swin transformers with stagewise convolutions for 3d medical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 416–426. Springer (2023)
6. Huang, D.Q., Terrault, N.A., Tacke, F., Glud, L.L., Arrese, M., Bugianesi, E., Loomba, R.: Global epidemiology of cirrhosis—etiology, trends and predictions. *Nature reviews Gastroenterology & hepatology* **20**(6), 388–398 (2023)

7. Jha, D., Susladkar, O.K., Gorade, V., Keles, E., Antalek, M., Seyithanoglu, D., Cebeci, T., Aktas, H.E., Kartal, G.D., Kaymakoglu, S., et al.: Large scale mri collection and segmentation of cirrhotic liver. *Scientific Data* **12**(1), 896 (2025)
8. Lee, H.H., Bao, S., Huo, Y., Landman, B.A.: 3d ux-net: A large kernel volumetric convnet modernizing hierarchical transformer for medical image segmentation. *arXiv preprint arXiv:2209.15076* (2022)
9. Liu, Y., Tian, Y., Zhao, Y., Yu, H., Xie, L., Wang, Y., Ye, Q., Jiao, J., Liu, Y.: Vmamba: Visual state space model. *Advances in neural information processing systems* **37**, 103031–103063 (2024)
10. Myronenko, A.: 3d mri brain tumor segmentation using autoencoder regularization. In: *International MICCAI brainlesion workshop*. pp. 311–320. Springer (2018)
11. Roy, S., Koehler, G., Ulrich, C., Baumgartner, M., Petersen, J., Isensee, F., Jaeger, P.F., Maier-Hein, K.H.: Mednext: transformer-driven scaling of convnets for medical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 405–415. Springer (2023)
12. Ruan, J., Li, J., Xiang, S.: Vm-unet: Vision mamba unet for medical image segmentation. *arXiv preprint arXiv:2402.02491* (2024)
13. Smith, J.T., Warrington, A., Linderman, S.W.: Simplified state space layers for sequence modeling. *arXiv preprint arXiv:2208.04933* (2022)
14. Tonan, T., Fujimoto, K., Qayyum, A.: Chronic hepatitis and cirrhosis on mr imaging. *Magnetic Resonance Imaging Clinics* **18**(3), 383–402 (2010)
15. Wang, Z., Zheng, J.Q., Zhang, Y., Cui, G., Li, L.: Mamba-unet: Unet-like pure visual mamba for medical image segmentation. *arXiv preprint arXiv:2402.05079* (2024)
16. Xing, Z., Ye, T., Yang, Y., Liu, G., Zhu, L.: Segmamba: Long-range sequential modeling mamba for 3d medical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 578–588. Springer (2024)
17. Zeng, J., Jha, D., Aktas, E., Keles, E., Medetalibeyoglu, A., Antalek, M., Lewandowski, R., Ladner, D.P., Borhani, A.A., Durak, G., et al.: A reverse mamba attention network for pathological liver segmentation. In: *2025 5th International Conference on Electrical, Computer, Communications and Mechatronics Engineering*. pp. 1–8. IEEE (2025)