Proceedings of Machine Learning Research vol 278:1–6, 2025 2025 2nd International Conference on Machine Learning and Intelligent Computing

MUnet-Lite: A Mamba-Based Lightweight Network for Efficient Abdominal Image Segmentation

Ziyang Hou
School of Cyber Science and Engineering, Zhengzhou University, Zhengzhou, Henan, China
Collaborative Innovation Center for Internet Healthcare, Zhengzhou University, Zhengzhou, Henan, China

Zhe Zhao
School of Cyber Science and Engineering, Zhengzhou University, Zhengzhou, Henan, China
Collaborative Innovation Center for Internet Healthcare, Zhengzhou University, Zhengzhou, Henan, China

Yusong Lin*　YSLIN@HA.EDU.CN
School of Cyber Science and Engineering, Zhengzhou University, Zhengzhou, Henan, China
Collaborative Innovation Center for Internet Healthcare, Zhengzhou University, Zhengzhou, Henan, China

Editors: Nianyin Zeng, Ram Bilas Pachori and Dongshu Wang

Abstract

The human abdomen houses multiple vital organs, and medical imaging technology precisely captures pathological features, providing a foundation for clinical diagnosis and treatment. High-precision abdominal image segmentation is crucial for lesion localization, organ measurement, and surgical planning. However, existing methods face challenges in local feature extraction and multi-scale information modeling. To overcome the limitations of Transformer-based approaches, such as insufficient local information perception, large model size, and high computational cost, we propose MUnet-Lite, a lightweight segmentation model. It combines the Mamba method with a U-Net architecture, incorporating a residual spatial modeling unit for enhanced feature extraction and an efficient decoding unit to reduce computation. Experiments on the Synapse dataset show that MUnet-Lite achieves a Dice score of 83.79% and a Hausdorff distance of 16.43mm, with only 26.71M parameters and 925.9 GFLOPs, significantly lowering computational cost while maintaining high segmentation accuracy. This provides a practical solution for real-world applications.

Keywords: Medical image processing, Image segmentation, Abdominal multi-organ.

1. Introduction

Abdominal multi-organ medical image segmentation is crucial in medical diagnosis, disease prevention, and treatment planning. U-Net (Ronneberger et al., 2015), as a classic segmentation model, has demonstrated excellent performance in abdominal segmentation tasks due to its advantage in local feature extraction. However, as it is based on convolutional neural networks (CNNs), it has limitations in capturing global semantic information and long-range dependencies. To enhance global information perception, many studies have introduced methods such as dilated convolutions (Gao, 2023) and self-attention (Zhong et al., 2025) mechanisms. However, these improvements often increase computational complexity and the number of parameters, limiting their application in high-resolution medical images, especially in resource-constrained environments. To address this issue, Gu and Dao (2023) proposed the lightweight model Mamba, based on the State Space Model (SSM), which can effectively model long-range dependencies while reducing computational costs.

Based on this, this paper proposes a lightweight abdominal medical image segmentation model, MUnet-Lite. The model combines the advantages of U-Net and Mamba, designing a Residual

Spatial Modeling Block (RSMB) and an Efficient Decoding Block (EDB), significantly improving segmentation accuracy and computational efficiency. Experimental results show that MUnet-Lite effectively reduces computational costs while ensuring high accuracy, providing a feasible solution for practical applications.

2. Methods

MUnet-Lite enhances long-range dependency modeling and computational efficiency by designing residual space modeling block and efficient decoding block. It significantly reduces computational overhead, making it ideal for resource-constrained medical environments. The overall architecture is as follows (Figure 1):

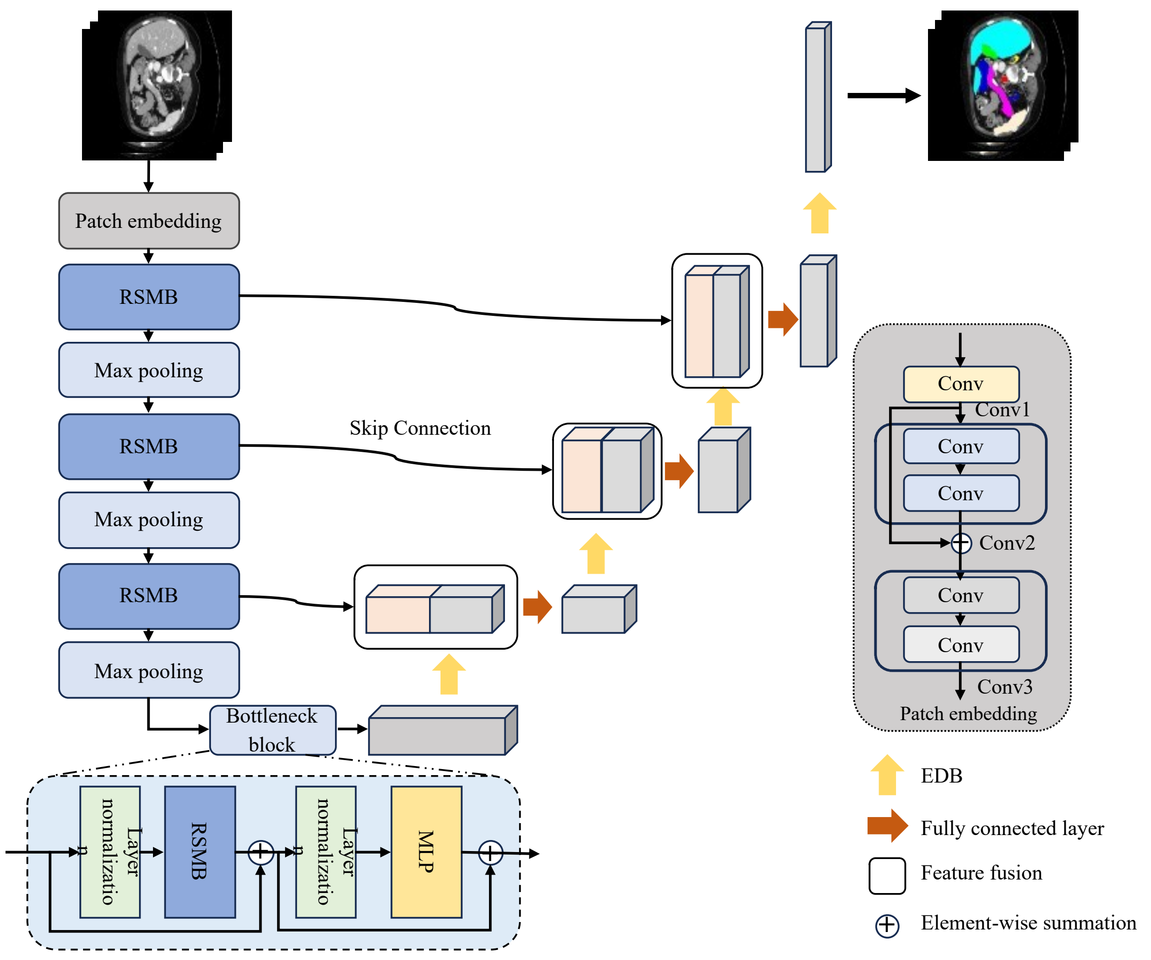

Figure 1: The overall architecture of MUnet-Lite.

2.1. Residual Spatial Modeling Block

MUnet-Lite designs an Enhanced State Space Block (ESSB) for long-range spatial modeling (Figure 2), where the included SSM unit follows the method in VMamba (Liu et al., 2024).

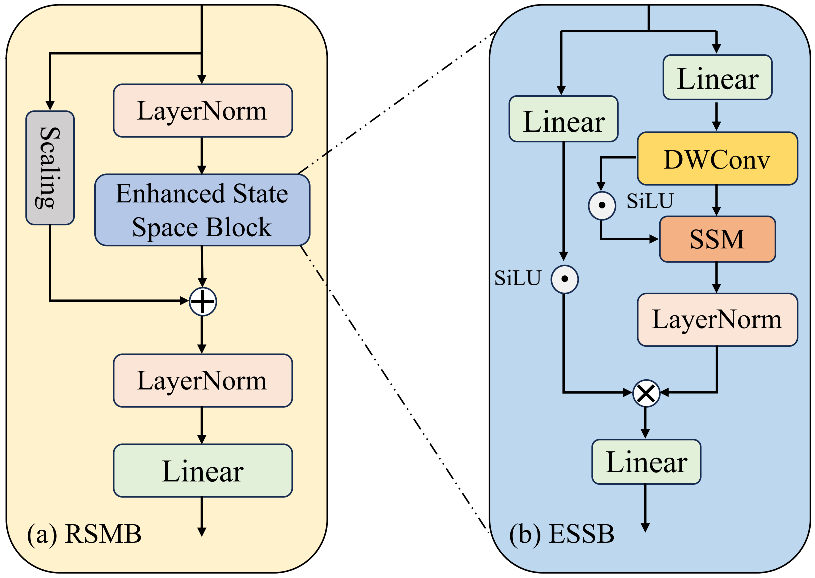

Figure 2: Residual Spatial Modeling Block (RSMB) structure.

ESSB adopts a dual-branch structure: the first branch expands the channel dimension to $\lambda \times C$ through a linear layer, followed by depthwise separable convolution (DWConv), SiLU activation, SSM, and LayerNorm; the second branch also expands to $\lambda \times C$ and applies SiLU activation. Finally, the features from both branches are fused through element-wise multiplication and projected back to the original channel dimension. To reduce computational cost, MUnet-Lite further introduces a Residual Spatial Modeling Block (RSMB), which integrates residual connections and dynamic scaling factors to enhance long-range dependency modeling while maintaining computational efficiency. RSMB first normalizes the input using LayerNorm, then models long-range dependencies through ESSB, and introduces a scaling factor [89] to adaptively adjust input features for optimizing the fusion of global and local features. The specific formulation is as follows:

$$
\begin{aligned}
Y_1 &= \text{LayerNorm}\left(\text{SSM}\left(\text{SiLU}\left(\text{DWConv}\left(\text{Linear}\left(Y_{\text{in}}\right)\right)\right)\right)\right) \\
Y_2 &= \text{SiLU}\left(\text{Linear}\left(Y_{\text{in}}\right)\right) \\
Y_{\text{out}} &= \text{Linear}\left(Y_1 \otimes Y_2\right)
\end{aligned}
\tag{1}
$$

2.2. Efficient Decoding Block

This study introduces an Efficient Decoding Block (EDB) to enhance decoder performance and efficiency. EDB replaces traditional convolutions with depthwise separable convolutions and integrates upsampling, feature enhancement, and channel adjustment to reduce computational complexity while enhancing feature representation. This module significantly improves decoder efficiency while maintaining segmentation accuracy (Figure 3).

EDB first upsamples feature maps to double their size, restoring spatial details. Then, depthwise separable convolutions preserve spatial structure while reducing computation. Batch normalization (BN) and ReLU activation accelerate convergence and enhance nonlinearity. Finally, convolution adjusts channel numbers to optimize feature expression and interaction. With low computational cost, EDB enables efficient feature extraction and multi-stage fusion, ideal for resource-limited abdominal image segmentation. Its mathematical definition is as follows:

$$x_{\text{out}} = \text{Conv}_{1\times 1}(\text{ReLU}(\text{BatchNorm}(\text{DepthwiseConv}(\text{Upsample}(x + y))))) \tag{2}$$

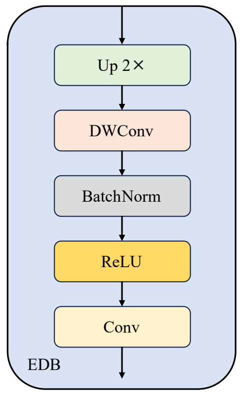

Figure 3: Efficient Decoding Block (EDB) structure.

3. Experiments

3.1. Dataset

Synapse: This is a publicly available dataset for medical image segmentation, primarily used to train and evaluate the performance of machine learning models on abdominal medical image segmentation tasks.

3.2. Experimental Results

To comprehensively validate the segmentation performance of MUnet-Lite, this study compares it with mainstream medical image segmentation methods, including CNN-based models UNet and Att-UNet, Transformer-based models TransUNet and Swin-Unet, as well as Mamba-based models VM-UNet and Swin-UMamba. By comparing the performance of different models, the effectiveness of MUnet-Lite in abdominal medical image segmentation tasks is evaluated.

Table 1: Comparison of multi-organ segmentation performance among different models.

Methods	DSC↑(%)	HD↓(mm)
U-Net	76.85	39.70
Att-UNet	77.77	36.02
TransUNet	77.48	31.69
Swin U-Net	79.13	21.55
VM-UNet	81.08	19.21
Swin-UMamba	82.26	19.51
MUnet-Lite (Ours)	**83.79**	**16.43**

Table 1 presents the performance of different segmentation models on the Synapse dataset. MUnet-Lite demonstrates exceptional performance in abdominal medical image segmentation, achieving the best results in Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD) with 83.79% and 16.43mm, respectively. Compared to the traditional CNN-based Att-UNet, MUnet-Lite improves DSC by 6.02% and reduces HD95 by 19.59mm. When compared with the Transformer-based Swin-Unet, MUnet-Lite shows improvements of 4.66% in DSC and 5.12mm in HD95. Even when compared with the latest Mamba-based Swin-UMamba method, MUnet-Lite still improves DSC by 1.53% and HD95 by 3.08mm, demonstrating strong competitiveness. Figure 1 shows the segmentation performance of different methods. The results confirm that MUnet-Lite outperforms other methods in the Synapse abdominal multi-organ segmentation task, highlighting its potential in the field of medical image segmentation.

Table 2: Comparison of DSC values for multi-organ segmentation across models.

Methods	Aorta	Gallbladder	Kidney (L)	Kidney (R)	Liver	Pancreas	Spleen	Stomach
U-Net	89.07	69.72	77.77	68.60	93.43	53.98	86.67	75.58
Att-UNet	**89.55**	68.88	77.98	71.11	93.57	58.04	87.30	75.75
TransUNet	87.23	63.13	81.87	77.02	94.08	55.86	85.08	75.62
Swin U-Net	85.47	66.53	83.28	79.61	94.29	56.58	90.66	76.60
VM-UNet	86.40	69.41	**86.16**	**82.76**	94.17	58.80	89.51	81.40
Swin-UMamba	86.32	70.77	83.66	81.60	95.23	69.36	89.95	81.14
MUnet-Lite	89.54	**72.63**	85.25	82.43	**95.82**	**71.27**	**90.53**	**82.91**

Table 2 shows the performance of different methods in the abdominal multi-organ segmentation task. The proposed method performs particularly well in the segmentation of the gallbladder and pancreas, with DSC values reaching 72.63% and 71.27%, respectively. For large organ segmentation, MUnet-Lite achieves a DSC score of 82.91% for the stomach, demonstrating good adaptability. This advantage stems from MUnet-Lite's strong feature extraction capability, which can capture long-range dependencies and integrate local contextual information. By introducing the residual spatial modeling unit, MUnet-Lite enhances its ability to model organs with complex shapes, improving segmentation accuracy, robustness, and generalization ability.

Experimental results demonstrate the remarkable lightweight advantage of MUnet-Lite on the Synapse dataset. With only 26.71M parameters, it reduces 35.48M and 31.76M compared to Swin UNETR and U-Mamba, respectively, significantly lowering computational demands and enhancing deployment flexibility. Its complexity is merely 925.9 GFLOPs, outperforming other models and confirming its efficiency.Specific results are shown in the Table 3.

Table 3: Comparison of Segmentation Models' Parameters and GFLOPs.

Methods	Type	Parameters (M)	GFLOPs
3DUX-Net	CNN	53.01	2805.8
Swin U-Net	Transformer	52.91	3620.6
Swin UNETR	Transformer	62.19	1505.3
U-Mamba	Mamba	58.47	5620.2
Swin-UMamba	Mamba	32.86	4893.7
MUnet-Lite (Ours)	**Mamba**	**26.71**	**925.9**

4. Conclusion

This paper proposes MUnet-Lite, a lightweight abdominal medical image segmentation model. By incorporating residual spatial modeling and efficient decoding units, it optimizes computational complexity while maintaining high segmentation accuracy, ensuring a balance between efficiency and performance for resource-limited medical settings.

Acknowledgments

This paper is supported by the National Natural Science Foundation of China under Grant No. 82441022, and Collaborative Innovation Major Project of Zhengzhou under Grant No.20XTZX05015.

References

Roland Gao. Rethinking dilated convolution for real-time semantic segmentation. pages 4675–4684, 2023. doi: 10.1109/CVPRW59228.2023.00493.

Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. *ArXiv*, abs/2312.00752, 2023.

Yue Liu, Yunjie Tian, Yuzhong Zhao, Hongtian Yu, Lingxi Xie, Yaowei Wang, Qixiang Ye, and Yunfan Liu. Vmamba: Visual state space model. *ArXiv*, abs/2401.10166, 2024.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. pages 234–241, 2015.

Jiahui Zhong, Wenhong Tian, Yuanlun Xie, Zhijia Liu, Jie Ou, Taoran Tian, and Lei Zhang. Pmfsnet: Polarized multi-scale feature self-attention network for lightweight medical image segmentation. *Computer Methods and Programs in Biomedicine*, 261:108611, 2025. doi: https://doi.org/10.1016/j.cmpb.2025.108611.