A Frequency-Aware Hybrid U-Mamba for Breast Tumor Segmentation

Chi-An Chen
Department of Computer Science and Information Engineering
National Ilan University
Yilan, Taiwan
Email: b1043003@ems.niu.edu.tw

Sara Katharina Alessia Bosshart
Leysin American School
Leysin, Switzerland
Email: bosshas116@las.ch

Liang-Ying Ke
Department of Engineering Science
National Cheng Kung University
Tainan, Taiwan
Email: n98131041@gs.ncku.edu.tw

Chih-Hsien Hsia, *Senior Member*, *IEEE*
Department of Computer Science and Information Engineering
National Ilan University
Yilan, Taiwan
Email: hsiach@niu.edu.tw

***Abstract*— Recently, dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) has become an important tool for breast tumor examination. However, challenges such as varying tumor sizes and shapes, as well as blurred boundaries in breast tumor images, hinder deep learning models from accurately segmenting breast tumors. Given these issues, this work proposes a frequency-aware hybrid U-Mamba (FAHU-Mamba) model architecture, which integrates a state space model (SSM) and a frequency-aware module (FAM) for breast tumor segmentation. In particular, the model utilizes FAM to learn the frequency-domain features from breast tumor images, enhancing feature representation for blurred tumor regions and thereby improving segmentation performance. According to the experimental results, FAHU-Mamba model achieved overall performance scores of 82.35%, 89.73%, and 88.25% in terms of dice-sørensen coefficient (DSC), positive predictive value (PPV), and recall metrics, respectively, on the BreastDM database. As results, the proposed model not only delivers robust segmentation performance but also holds promise to support early diagnosis and treatment planning in clinical practice.**

Keywords— Computer vision, Image segmentation, State space models, Breast tumor.

I. INTRODUCTION

During breast cancer diagnosis and treatment, accurately identifying the location and extent of the tumor is critical for assisting doctors in treatment planning. However, challenges such as varying tumor sizes and shapes, as well as blurred boundaries in breast tumor images, hinder the precise identification of tumor extent.

In recent years, with the rapid advancement of computer vision and deep learning [1], numerous studies have introduced these technologies into breast tumor image segmentation and proposed various DL models. Saifullah *et al.* [2] proposed a U-Net architecture based on convolutional neural networks (CNNs) for breast cancer segmentation. However, since CNN-based models primarily capture local receptive fields, U-Net struggles to model long-range dependencies (LRD). This limitation hinders the model's ability to effectively capture complex image structures, thereby degrading segmentation performance. To overcome this limitation, Sabry *et al.* [3] proposed a DL model based on vision transformer (ViT) architecture for breast cancer segmentation. The ViT-based model not only captures global features from images to better represent complex image structures but also addresses the shortcomings of modeling LRD in CNN-based architectures. However, ViT-based models generally lack inductive bias and involve higher computational complexity, making them difficult to train with limited data and thereby limiting their applicability in clinical practice.

Given these challenges, this work proposes a frequency-aware hybrid U-Mamba (FAHU-Mamba) model, which integrates state space models (SSMs) [4] and frequency-aware modules (FAMs) [5] for breast tumor segmentation. Specifically, SSM architecture not only reduces the computational complexity of DL models but also introduces inductive bias through depth-wise convolution (DWConv), thereby enhancing training stability and improving model generalizability in clinical practice. In addition, FAM is designed to learn frequency-domain features from breast tumor images, enhancing the model's perceptual ability at tumor boundaries and further improving the overall segmentation performance of FAHU-Mamba model.

II. PROPOSED METHOD

A. Overall Architecture

To accurately segment breast tumor images, this work proposes a FAHU-Mamba model, which incorporates SSMs and FAMs to emphasize tumor boundaries and enhance feature extraction for modeling LRD, as shown in Fig. 1. First, FAHU-Mamba model utilizes patch embedding to transform the input image into embedding features, which are then used for subsequent feature extraction. The model then employs encoding blocks to progressively reduce spatial resolution while capturing multi-scale semantic features. Next, FAHU-Mamba model upsamples the features through decoding blocks to restore their spatial resolution, thereby producing the final breast tumor segmentation results. Both the encoding and decoding blocks utilize stacked frequency-aware selective scan blocks (FA-SSBs) to extract global features, thereby facilitating a comprehensive understanding of tumor structures and enhancing feature representation capabilities.

B. Frequency-Aware Selective Scan Block

To effectively improve the model's perceptual ability at tumor boundaries, this work proposes the FA-SSB incorporating a FAM to enhance feature representation of high-frequency components, thereby improving segmentation performance. The architectures of FA-SSB and FAM are illustrated in Fig. 2(a) and (b), respectively. First, FA-SSB utilizes a normalization layer to stabilize the distribution of the input feature maps, and projects them into a high-dimensional space using a linear layer, thus enhancing feature representation capability. By leveraging DWConv and the SS2D module, the block captures fine-grained texture and structural features of breast tumors, significantly improving the model's ability to distinguish tumor regions from surrounding tissue. Next, the block applies another normalization layer to stabilize the feature distribution and adaptively reweights the features using feature maps transformed by another linear layer, thereby suppressing redundant spatial information. Moreover, FA-SSB transforms the feature maps from the spatial domain into the frequency domain using FAM, and adopts learnable parameters to reweight different frequency components, enabling the model to capture critical frequency-specific features. Finally, the block facilitates the fusion of spatial- and frequency-domain features, further enhancing the feature representation capability.

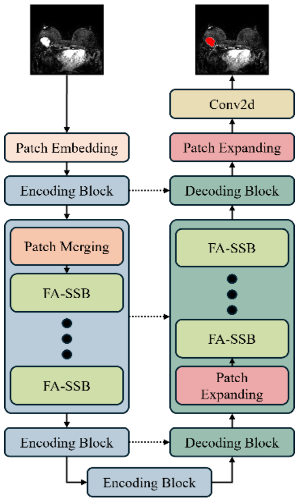
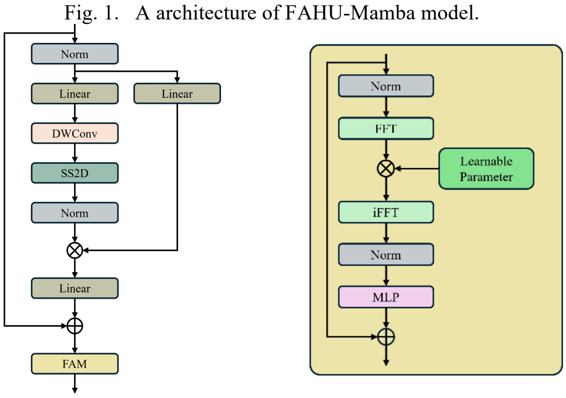


Fig. 1. A architecture of FAHU-Mamba model.



(a) (b)

Fig. 2. (a) An architecture of FA-SSB; (b) An architecture of FAM.

III. Experimental Results And Analysis

A. *Breast Tumor Image Database*

To evaluate the segmentation performance of breast tumor segmentation models, this work adopts the BreastDM database [6] for model training and testing. The BreastDM database comprises 29,274 DCE-MRI images with a resolution of 369×369, collected from 232 subjects aged between 22 and 79 years. Among these images, 7,905 are labeled as benign tumor images and 21,369 as malignant tumor images. For effective evaluation, the database is divided into 80% training and 20% testing sets.

In the hyperparameter settings, the image size, batch size, and number of epochs are set to 320×384, 6, and 100, respectively, for training. Moreover, the Adam optimizer is employed within the PyTorch DL framework, using a learning rate of 0.0001 and a weight decay of 0.00001 to optimize the model parameters.

B. *Experimental Results and Analysis*

To assess the overall performance of the segmentation model, this work employs dice-sørensen coefficient (DSC), positive predictive value (PPV), and recall as evaluation metrics to enable a fair comparison with previous studies. According to the experimental results on the BreastDM public database, FAHU-Mamba model achieves a DSC of 82.35%, a PPV of 89.73%, and a recall of 88.25%, as shown in Table I. Compared with previous studies, the proposed model effectively captures tumor boundaries through FA-SSBs and accurately identifies the extent of breast tumors, thereby enhancing segmentation performance.

TABLE I. Performance Comparison with Previous Studies on The BreastDM Database.

Methods	DSC (%)	PPV (%)	Recall (%)
PSPNet [6]	72.7	72.9	72.5
UNeXt [7]	70.1	86.3	59.0
This Work	**82.35**	**89.73**	**88.25**

IV. Conclusion

To address the challenge of blurred tumor boundaries, this work proposes a FAHU-Mamba model, which integrates SSMs and FAMs for breast tumor segmentation. The model utilizes FA-SSB to learn different frequency components from feature maps and to capture the structural features of breast tumors, thereby enhancing feature representation capability. According to the experimental results on the BreastDM database, FAHU-Mamba model achieved 82.35% in DSC, 89.73% in PPV, and 88.25% in recall, respectively. Compared with previous studies, the proposed model demonstrates the ability to capture complex structures and blurred boundaries of breast tumor features, highlighting its potential for clinical practice.

References

[1] C.-H. Hsia and C.-H. Liu, "New hierarchical finger-vein feature extraction method for iVehicles," *IEEE Sensors Journal*, vol. 22, no. 13, pp. 13612-13621, 2022.

[2] S. Saifullah and R. Dreżewski, "Enhancing breast cancer diagnosis: a CNN-based approach for medical image segmentation and classification," *International Conference on Computational Science*, pp. 155-162, 2024.

[3] H. Wang, L. Wei, B. Liu, J. Li, J. Li, J. Fang, and C. Mooney, "Transformer-based explainable model for breast cancer lesion segmentation," *Applied Sciences*, vol. 15, no. 3 pp. 1295, 2025.

[4] A. Gu and T. Dao, "Mamba: Linear-time sequence modeling with selective state spaces," *arXiv preprint arXiv:2312.00752*, 2023.

[5] A. Chang, J. Zeng, R. Huang, and D. Ni, "EM-Net: efficient channel and frequency learning with mamba for 3d medical image segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 266-275, 2024.

[6] X. Zhao, Y. Liao, J. Xie, X. He, S. Zhang, G. Wang, J. Fang, H. Lu, and J. Yu, "BreastDM: a DCE-MRI dataset for breast tumor image segmentation and classification," *Computers in Biology and Medicine*, vol. 164, pp. 107255, 2023.

[7] J. M. J. Valanarasu and V. M. Patel, "UNeXt: MLP-based rapid medical image segmentation network," *Medical Image Computing and Computer Assisted Intervention*, pp. 23-33, 2022.