

A Mamba-Spectral Hybrid U-Net for Breast Tumor Segmentation

Chi-An Chen

Department of Computer Science and
Information Engineering
National Ilan University
Yilan County, Taiwan

Email: b1043003@ems.niu.edu.tw

Tzu-Hung Chien

Taipei American School
Taipei City, Taiwan

Email: 27henryel@students.tas.tw

Liang-Ying Ke

Department of Engineering Science
National Cheng Kung University
Tainan City, Taiwan

Email: n98131041@gs.ncku.edu.tw

Abstract—動態對比增強磁振造影(Dynamic Contrast-Enhanced Magnetic Resonance Imaging, DCE-MRI)在乳房腫瘤檢測中具有高度敏感性，然而，由於腫瘤形狀與大小變化多端，以及背景組織的不均勻性，使得準確分割仍面臨重大挑戰。有鑑於此，本研究提出一種結合狀態空間模型(State Space Model, SSM)以及頻率域(spectral domain)特徵表示的U型卷積網路(U-Net convolutional network, U-Net)架構(Mamba-Spectral Hybrid U-Net, MSU-Net)，用於乳房DCE-MRI影像中的腫瘤分割任務。MSU-Net藉由引入Mamba概念與頻率域特徵處理，有效強化影像中模糊病灶的辨識能力，進而提升特徵的表達與提取效率。實驗結果顯示，本研究所提出之MSU-Net模型在BreastDM公開資料集上達到82.35%的Dice係數(Dice-Sorenson Coefficient, Dice)，並在準確率(Positive Predictive Value, PPV)與召回率(Recall)兩項指標上分別取得89.73%與88.25%的優異表現，證實其在乳房DCE-MRI影像分割任務中的準確性與穩定性，為臨床乳癌早期檢測與輔助診斷提供一項具潛力的有效工具。

Keywords—Computer Vision, Image Segmentation, State Space Models, U-net, Spectral Domain, Breast Tumor.

I. INTRODUCTION

根據國際癌症研究機構(International Agency for Research on Cancer, IARC)於2022年的統計資料，在所有癌症(cancer)類型中，乳癌(breast cancer)是全球第二常見的癌症，僅次於肺癌(lung cancer)。在死亡率方面，乳癌則排名第四，每年約造成670,000人死亡[1]。儘管如此，自1989年以來，乳癌的死亡率已穩定下降，截至2022年，總體下降幅度達44%[2]。這一改善主要歸因於早期篩檢、公眾對乳癌認識的提升，以及治療技術的進步。隨著乳癌篩檢、分期與治療在臨床上的重要性日益提升，選擇合適且準確的檢測方式變得尤為關鍵。在眾多檢測技術中，例如乳房攝影(mammography)與超音波(ultrasound)，動態對比增強磁振造影(Dynamic Contrast-Enhanced Magnetic Resonance Imaging, DCE-MRI)因其在檢測軟組織病變方面具備高敏感度(sensitivity)，已成為乳癌診斷中不可或缺的重要工具。與傳統X光不同，DCE-MRI不受乳腺組織密度影響，且無輻射風險[3]。此外，DCE-MRI能夠清晰顯示腫瘤的新生血管形成(angiogenesis)，進而捕捉腫瘤的細節特徵，為病灶評估與治療反應監測提供了關鍵依據[4]。

在乳癌的診斷與治療過程中，準確的標註腫瘤區域是一項至關重要的任務。其中最直接的方法是由放射科醫師手動標註腫瘤區域，然而這樣不僅耗時，亦容易出錯。隨著技術的進步與深度學習(Deep Learning, DL)的發展，目前已經有許多研究文獻將DL和電腦視覺(Computer Vision, CV)技術應用於乳房腫瘤影像分割當中，並提出

多種影像分割模型。特別是以卷積神經網路(Convolutional Neural Networks, CNN)為核心的模型，取得了顯著成果。例如：U型卷積網路(U-Net convolutional network, U-Net)，透過編碼器-解碼器(encoder-decoder)架構，能有效提取與重建影像特徵，進而精確描繪腫瘤邊界。然而，CNN由於其局部感受野(local receptive field)的限制，對於長期依賴性(long-range dependencies)的能力不足，可能導致在結構複雜或邊界模糊的腫瘤影像中的辨識效果下降[5]。為解決此問題，近年來基於Transformer架構的模型逐漸受到重視。Transformer擅長捕捉全局上下文資訊，並在大規模預訓練後展現出優異的泛化能力與遷移學習(transfer learning)效果，可在資料量有限的情況下轉移至特定醫療任務中並取得良好表現[6]。然而，由於其使用自注意力機制(self-attention)的計算複雜度為 $O(n^2)$ ，且參數量龐大，限制了在臨床端等資源受限環境中的實際應用。針對上述限制，近期的狀態空間模型(State Space Model, SSM)提供了兼顧效能與效率的新方向。Mamba架構[7]作為代表，在自然語言處理與高解析度影像處理任務中展現優異表現。與Transformer相比，Mamba在處理長序列資料時具有線性擴展能力，能大幅降低運算成本。

此外，在影像分割中，另一個常見的問題是全局(global)特徵與局部(local)特徵之間的取捨。與分類任務不同，準確的分割同時需要模型同時具有全局特徵(low-frequency)以及局部紋理特徵(high-frequency)的理解[8]。由於頻率域(Spectral Domain, SD)能夠完整地表示局部與全局訊號[8]。因此Chang et al.[9]提出透過快速傅立葉轉換(Fast Fourier Transform, FFT)構建一個可訓練的頻率濾波器，以自動學習並強調與分割任務相關的不同頻率訊號，進而實現兩者之間的平衡。

有鑑於上述發展與挑戰，本研究提出一種結合Mamba與SD概念的U-Net架構(Mamba-Spectral Hybrid U-Net, MSU-Net)，應用於乳房腫瘤DCE-MRI影像之分割任務。該模型在設計上兼顧運算效率與分割準確性，並針對醫療影像中病灶邊界模糊等特性，提供有效的解決方案。

II. PROPOSED METHOD

A. Overall Architecture

為有效提升模型對乳房腫瘤DCE-MRI影像之特徵表徵能力，本研究提出一種整合頻域狀態空間模塊(State Space Spectral Block, SSS Block)於U-Net架構中的新型分割模型，稱為MSU-Net，以增強模型對長距離依賴(long-range dependencies)之能力，並強化全域與局部特徵的提取效能。如圖1所示，整體模型沿用編碼器-解碼

器架構設計，首先將輸入影像透過圖塊嵌入層(patch embedding)轉換為嵌入特徵後，進入編碼階段，並經由圖塊合併層(patch merging)進行空間維度壓縮與特徵融合。編碼器中堆疊多層 SSS Block，以提升深層語義特徵的表徵能力。接著解碼器部分結合圖塊擴展層(patch expanding)模組實現逐層上採樣與尺寸恢復。解碼器亦配置多層 SSS Block，以進一步細化語義資訊並強化分割精度。最後，透過卷積層輸出最終之乳房腫瘤分割結果。本模型兼具頻域與時域特徵的處理能力，不僅提升對腫瘤形態與邊界的辨識能力，亦展現優異的分割性能與泛化能力。

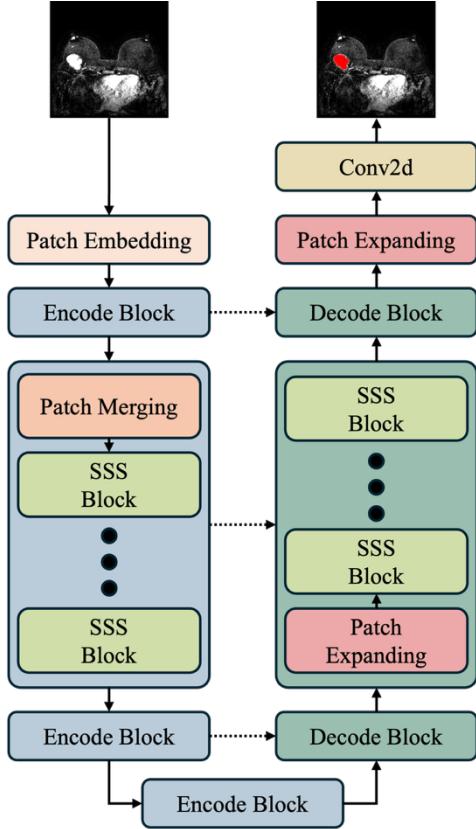


Fig. 1. 本研究提出 MSU-Net 之架構圖

B. SSS Block

為有效增強模型對乳房腫瘤 DCE-MRI 影像中全域與區域關鍵特徵的提取能力，本研究設計並提出一種結合 Mamba 結構的選擇性二維掃描(2D Selective Scan, SS2D)與 FFT 之特徵擷取模組，稱為 SSS Block，其架構如圖 2 所示。整體流程首先透過一層標準化(normalization)處理輸入特徵，以穩定特徵分佈並提升後續模組的學習表現。接著，透過線性層將特徵映射至更高維度，進一步擴展其表達能力。為強化局部資訊擷取，模型引入深度可分離卷積(depth-wise convolution)，以低計算成本有效提取局部空間特徵；隨後結合 SS2D 模組進一步捕捉空間中的結構性訊息。經再次標準化處理以穩定模組間特徵後，透過線性層完成特徵變換並作為殘差連接的橋樑，以利與頻域模組進行融合。最後，模型導入頻域模組(spectral layer)，透過 FFT 將特徵轉換至頻域，整合全局與區域特徵。該模組於 FFT 前後皆加入標準化層，並搭配殘差連接以提升特徵的整合與傳遞能力。

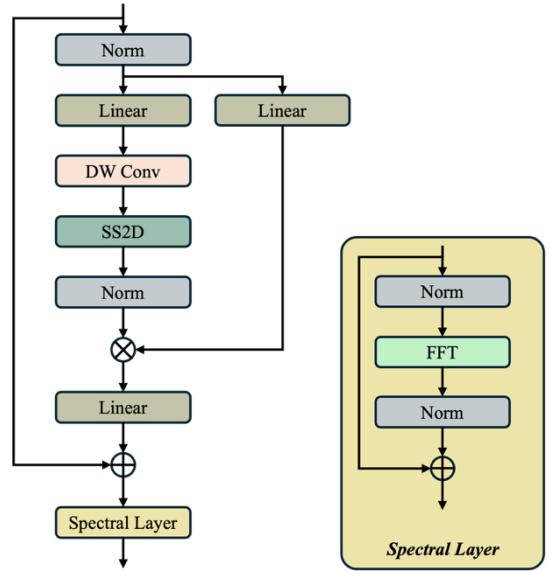


Fig. 2. SSS Block 之架構圖

透過上述設計，使得 MSU-Net 整體架構能有效捕捉不同層次語義特徵與空間資訊。有效的提升乳房腫瘤 DCE-MRI 影像分割任務中的準確性。

III. EXPERIMENTAL RESULTS AND ANALYSIS

本研究在此章節會詳細說明本研究之實驗設計與相關細節。首先，實驗所使用之資料庫將於第 III-A 節進行介紹；接著，模型訓練與實作環境的設定將於第 III-B 節說明；最後，各項實驗結果與分析則整理於第 III-C 節中進行呈現與討論。

A. 乳房腫瘤影像資料集

為評估本研究所提出之乳房腫瘤分割模型的辨識能力與泛化能力，本研究採用公開可用的 BreastDM [10] 資料集作為訓練與測試之基礎。BreastDM 資料集對年齡介於 22 至 79 歲之間的 232 名受試者收集術前 DCE-MRI 資料。其影像解析度為 369×369 。為有效評估所提模型於該資料集上的腫瘤分割效能，實驗中將資料集依照 80% 與 20% 的比例劃分為訓練集與測試集。

B. 實驗環境設置

本研究將影像尺寸(image size)設為 320×384 、批次量(batch size)設為 6、並以 100 個訓練週期(epochs)進行模型訓練。為對模型權重進行有效更新，本研究會採用 Adam 優化器(optimizer)，並且將其初始學習率(learning rate)和權重衰減(weight decay)分別設置為 0.0001 和 0.00001。此外，本研究所提出的模型架構使用 nnUNet 進行開發，並且採用 PyTorch DL 框架來進行實現，並且於 Intel(R) Core(TM) i7-12700 CPU 和 Nvidia RTX 4080 圖形處理器(graphics processing unit, GPU)的硬體環境上進行訓練。

C. 實驗結果與分析

為評估本研究所提出之影像分割模型的整體性能，本研究採用 BreastDM 公開資料集來對模型進行訓練與測試。為有效評估模型的分割能力，本研究選用四種常見的醫學影像分割效能評估指標，包括 Dice (Dice-Sørensen

coefficient)、IoU (Intersection over Union)、PPV (Positive Predictive Value)以及召回率 (Recall)。其中，Dice 指標用以衡量模型預測結果與實際標註區域的重疊程度，是最常用的分割準確率評估方式；IoU 則計算預測區域與實際區域交集與聯集的比值，對邊界誤差較為敏感；PPV 反映模型在預測為正類樣本時的準確性；而召回率則用於評估模型在所有實際正樣本中能成功預測出的比例。透過上述多指標的綜合分析，可全面掌握模型的分割效能。

根據實驗結果，本研究所提出的 MSU-Net 模型在 BreastDM 公開資料集上的四項指標分別達到：Dice 82.35%、IoU 74.96%、PPV 89.73%，以及召回率 88.25%，如表 1 所示。與既有影像分割方法相比，MSU-Net 在該資料集上展現出更為優異的表現。較高的 Dice 值顯示該模型能夠有效掌握腫瘤病灶的輪廓與空間位置；而 PPV 與召回率的優異結果則代表模型兼具高準確性與敏感度，能在減少誤判的同時，成功標記出多數病灶區域。此結果證明，MSU-Net 具備處理乳房 DCE-MRI 影像中腫瘤分割任務的強大特徵擷取與辨識能力。

TABLE I. 與過往的影像分割研究在 BREASTDM 資料集上的比較

Methods	Dice (%)	IoU (%)	PPV (%)	Recall (%)
Random forest [10]	66.2	74.5	75.3	59.1
FCN-50 [10]	71.1	77.7	77.0	66.0
FCN-101 [10]	71.4	79.0	76.2	67.2
DeepLabV3-50 [10]	72.9	79.3	80.3	66.8
DeepLabV3-101 [10]	73.4	78.7	86.5	63.8
U-net [10]	73.7	80.3	83.6	65.9
Unet-VGG16 [10]	70.4	77.3	79.6	63.1
UNeXt [10]	70.1	78.5	86.3	59.0
PSPNet [10]	72.7	79.0	72.9	72.5
U-net-finetune [11]	80.1	--	--	--
This Work	82.35	74.96	89.73	88.25

IV. CONCLUSION

為解決過往影像分割研究中常見的運算複雜度高、以及醫療影像中病灶邊界模糊導致難以正確標記病灶區域等問題，本文提出一種新穎的 U-Net 變體架構 MSU-Net。該模型將 Mamba 與 SD 概念融合於 U-Net 結構設計中，以增強模型對複雜結構資訊的學習能力。

根據實驗結果顯示，MSU-Net 在 BreastDM 公開資料集上展現出穩定且優異的影像分割效能。與既有的影像分割方法相比，本研究所提出的模型更能因應醫學影像中複雜且具高變異性的特徵分佈，展現出良好的實務應用潛力與發展價值。

REFERENCES

- [1] F. Bray, et al., “Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA: a cancer journal for clinicians*, pp. 229-263, 2024.
- [2] [Online] <https://www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html>
- [3] S. Bellhouse, et al., “Breast cancer risk assessment and primary prevention advice in primary care: a systematic review of provider attitudes and routine behaviours,” *Cancers*, pp. 4150, 2021.
- [4] M.F. Vocaturo, et al., “Unsupervised Learning Techniques for Breast Lesion Segmentation on MRI Images: Are We Ready for Automation?,” *Applied Sciences*, pp. 2401, 2025.
- [5] F. Shamshad et al., “Transformers in medical imaging: A survey,” *Medical Image Analysis*, vol. 88, pp. 102802, 2023.
- [6] A. Dosovitskiy et al., “An image is worth 16x16 words: Transformers for image recognition at scale,” *arXiv preprint*, arXiv:2010.11929, 2020.
- [7] A. Gu and T. Dao, “Mamba: Linear-time sequence modeling with selective state spaces,” *arXiv preprint*, arXiv:2312.00752, 2023.
- [8] W. Wang, et al., “FreMIM: fourier transform meets masked image modeling for medical image segmentation,” *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 7860-7870 ,2024.
- [9] A. Chang, et al., “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.
- [10] X. Zhao, Y. Liao, J. Xie, X. He, S. Zhang, G. Wang, et al., “BreastDM: A DCE-MRI dataset for breast tumor image segmentation and classification”, *Comput. Biol. Med.*, vol. 164, Sep. 2023.
- [11] R. Bousmaha, I. Benguedda and S. M. Benslimane, “Transfer Learning for Breast Tumor Segmentation in DCE MRI,” *International Conference on Embedded & Distributed Systems (EDiS)*, pp. 55-60 , 2024.
- [12] F. Isensee, et al. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation,” *Nature methods*, 18.2, pp. 203-211, 2021.
- [13] J. Liu et al., “Swin-UMamba†: Adapting Mamba-based vision foundation models for medical image segmentation,” *IEEE Transactions on Medical Imaging*, pp. 1-1, 2024.