DOI: 10.16781/j.CN31-2187/R.20250457

·专题报道·

# 基于 Mamba 与注意力机制的三阴性乳腺癌超声图像分类方法

杨　颖[1]，宋元冰[1]，张一峰[2]，吴　蓉[2]，杜　宇[2]，郑　祎[2*]
1. 上海理工大学健康科学与工程学院，上海 200093
2. 上海交通大学医学院附属第一人民医院超声医学科，上海 200080

［摘要］ **目的**　探讨一种基于 Mamba 与注意力机制的混合神经网络架构（MECSA-Net）在三阴性乳腺癌（TNBC）超声图像分类中的应用效果。**方法**　回顾性收集 1 059 幅乳腺超声图像，其中 TNBC 图像 166 幅，非 TNBC 图像 893 幅。采用图像增强技术缓解类别不平衡问题。提出轻量级混合神经网络架构 MECSA-Net，其特征提取模块为高效混洗感知块（SAEffBlock），由状态空间建模分支（SSM-Branch）与轻量卷积分支（EffConvBranch）组成。在分类器前端引入多尺度空洞融合注意力（MDFA）模块，以提升模型对多尺度结构的感知能力和上下文信息建模能力。**结果**　在 TNBC 分类任务中，MECSA-Net 准确率为 93.9%、精确率为 94.4%、F1 分数为 93.9%、AUC 为 0.976，整体性能优于 ResNet-18、ResNet-50、EfficientNet-B0、ViT-Base 和 MedMamba-T 等主流模型。混淆矩阵分析显示，该模型对 TNBC 与非 TNBC 样本均具备较高的识别准确性和较低的误判率。消融实验进一步验证了 EffConvBranch 与 MDFA 模块在局部纹理建模与多尺度结构判别中的关键作用，显著增强了模型的分类性能与鲁棒性。**结论**　MECSA-Net 在 TNBC 超声图像分类中表现出优异的准确性与鲁棒性，具备良好的临床应用前景，可为 TNBC 术前智能辅助诊断提供技术支持。

［关键词］ 三阴性乳腺癌；超声图像；深度学习；Mamba；状态空间建模；注意力机制
［引用本文］ 杨颖，宋元冰，张一峰，等. 基于Mamba与注意力机制的三阴性乳腺癌超声图像分类方法［J］. 海军军医大学学报，2026，47（1）：37-45. DOI: 10.16781/j.CN31-2187/R.20250457.

## Classification of triple-negative breast cancer ultrasound images based on Mamba and attention mechanisms

YANG Ying[1], SONG Yuanbing[1], ZHANG Yifeng[2], WU Rong[2], DU Yu[2], ZHENG Yi[2*]
1. School of Health Science Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China
2. Department of Ultrasound, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200080, China

［**Abstract**］ **Objective**　To investigate the effectiveness of a Mamba-enhanced convolutional and state-attention network (MECSA-Net) for classifying triple-negative breast cancer (TNBC) in ultrasound images. **Methods**　A total of 1 059 breast ultrasound images were retrospectively collected, including 166 TNBC and 893 non-TNBC images. Data augmentation techniques were applied to mitigate class imbalance. A lightweight hybrid architecture MECSA-Net was proposed, featuring an extraction module named SAEffBlock (shuffle-aware efficient block), which integrated a state-space modeling branch (SSM-Branch) and a lightweight convolutional branch (EffConvBranch). Additionally, a multi-scale dilated fusion attention module (MDFA) was incorporated before the classifier to enhance the model's ability to perceive multi-scale structures and model contextual information. **Results**　MECSA-Net achieved an accuracy of 93.9%, a precision of 94.4%, an F1-score of 93.9%, and an area under curve of 0.976, outperforming mainstream models, including ResNet-18, ResNet-50, EfficientNet-B0, ViT-Base, and MedMamba-T. Confusion matrix analysis demonstrated high classification accuracy and low misclassification rates for both TNBC and non-TNBC samples. Ablation studies confirmed the crucial roles of both the EffConvBranch and MDFA modules in local texture representation and multi-scale structure discrimination, significantly

［收稿日期］ 2025-07-07　［接受日期］ 2025-09-22
［基金项目］ 国家自然科学基金青年科学基金（82302201），上海市科学技术委员会“科技创新行动计划”启明星项目（A 类）（24QA2707500）. Supported by Youth Project of National Natural Science Foundation of China (82302201) and Rising Star Project of “Scientific and Technological Innovation Action Plan” of Science and Technology Commission of Shanghai Municipality (Class A) (24QA2707500).
［作者简介］ 杨　颖，硕士生. E-mail: 1594861374@qq.com
*通信作者（Corresponding author）. E-mail: zhengyichn@sjtu.edu.cn

enhancing classification performance and robustness. **Conclusion**　MECSA-Net exhibits excellent accuracy and robustness in TNBC ultrasound image classification, indicating strong potential for clinical application and providing technical support for the intelligent preoperative diagnosis of TNBC.

［**Key words**］ triple-negative breast cancer; ultrasound images; deep learning; Mamba; state-space modeling; attention mechanisms

［**Citation**］ YANG Y, SONG Y, ZHANG Y, et al. Classification of triple-negative breast cancer ultrasound images based on Mamba and attention mechanisms[J]. Acad J Naval Med Univ, 2026, 47(1): 37-45. DOI: 10.16781/j.CN31-2187/R.20250457.

乳腺癌是全球女性发病率最高的恶性肿瘤之一，其发病率在过去 10 年持续攀升，严重威胁女性健康[1-2]。三阴性乳腺癌（triple-negative breast cancer，TNBC）是一种缺乏雌激素受体（estrogen receptor，ER）、孕激素受体（progesterone receptor，PR）和人表皮生长因子受体 2（human epidermal growth factor receptor 2，HER-2）的特殊乳腺癌亚型，约占所有乳腺癌的 11%～20%[3]。由于缺乏有效的靶向治疗手段，TNBC 常表现出高转移率和复发率，预后较差，是最难治愈的乳腺癌亚型之一[4-5]。

超声成像因无创、实时、便携及适用于致密型乳腺等优势，在乳腺癌早期筛查和治疗随访中应用广泛[6-7]。然而，TNBC 在超声图像中常呈边界模糊、形态不规则、内部回声不均等特征，人工阅片存在主观性和解释差异，影响诊断一致性及临床决策效率[8-10]。因此，开发高效、稳定且具泛化能力的智能识别模型，辅助 TNBC 与非 TNBC 的影像分型诊断，具有重要的临床价值。

近年来，深度学习方法特别是卷积神经网络（convolutional neural network，CNN）在医学图像分析中取得显著进展。CNN 在提取局部纹理特征、边缘结构等方面具有优势，被广泛应用于乳腺超声图像分类任务[11-14]。但受限于感受野范围，CNN 缺乏对长程依赖关系的建模能力，在处理结构复杂、异质性强的肿瘤区域时表现出一定局限性[15]。视觉 Transformer（vision Transformer，ViT）引入自注意力机制，能够有效实现长程依赖建模，已在自然图像识别中取得突破[16-17]。然而，ViT 结构复杂、计算开销大，对样本规模高度敏感，在医学超声图像小样本、高噪声场景下训练不稳定、易过拟合[18]。状态空间建模（state space modeling，SSM）通过结构化隐状态的动态传播，可高效表达长程依赖，具备线性时间复杂度和并行计算优势[19-20]。Mamba 架构作为 SSM 的代表，通过门控机制与卷积增强策略，在高效建模与长程依赖表达方面表现出色[21]。同时，注意力机制在医学图像分割、目标检测等任务中被广泛应用，有助于提升模型对复杂结构及模糊边界的特征建模能力[22]。为此，我们提出一种基于 Mamba 与注意力机制的混合神经网络架构（Mamba-enhanced convolutional and state-attention network，MECSA-Net）。MECSA-Net 结合 Mamba 架构的 SSM 机制、卷积增强策略与多尺度空洞融合注意力（multi-dilated fusion attention，MDFA）模块，设计双路径特征提取单元——高效混洗感知块（shuffle-aware efficient block，SAEffBlock），在兼顾局部纹理与远程空间依赖建模的同时提升对肿瘤异质性区域的识别能力，为 TNBC 超声图像智能分型诊断提供了新型技术方案。

## 1　资料和方法

1.1　*数据来源与预处理*　本研究数据回顾性采集自上海交通大学医学院附属第一人民医院超声医学科，连续纳入 2022 年 1 月至 2024 年 5 月经病理确诊为 TNBC 以及非 TNBC 的病例。纳入标准：（1）经手术切除病理确诊；（2）在诊断前 4 周内获取目标乳腺病灶的超声图像；（3）无新辅助化疗、放疗或乳腺手术史。排除标准：（1）临床或影像数据不全；（2）术前接受过治疗；（3）图像质量不佳（如存在伪影、病灶显示不全）。常规超声检查遵循美国超声医学会指南规范，使用 Aplio 500（日本佳能医疗系统）和 LOGIQ E9（美国 GE HealthCare 公司）超声设备完成。检查时患者采取适宜体位以确保乳房充分暴露。常规超声检查使用频率范围为 7～12 MHz 的线阵探头，在可能情况下，成像切面均涵盖病灶及周围正常组织。

所有图像均被保存以供后续分析。最终，共收集乳腺超声图像 1 059 幅，其中 TNBC 图像 166 幅，非 TNBC 图像 893 幅。

由 2 位分别拥有 5 年与 3 年乳腺超声诊断经验的医师，在不进行检查操作且不知晓患者病理资料的前提下，对超声图像进行回顾性分析。当出现诊断意见不一致时，则由第 3 位具有 10 年乳腺超声诊断经验的资深医师进行独立审阅，并通过共同讨论达成最终一致意见。具体流程如下：第 3 位医师在不知晓前 2 位医师判断的情况下独立解读图像，随后 3 位医师共同复核影像表现并进行讨论协商以达成一致；若经讨论仍无法形成统一意见，则采纳多数医师的诊断结论。所有病例经病理检查确认，保证标注信息的准确性和权威性。

由于数据集中 TNBC 与非 TNBC 样本比例存在明显不均衡，为缓解类别不平衡对模型训练的影响并提升模型对 TNBC 的识别能力，采用多种数据增强策略对 TNBC 样本进行扩增，包括随机旋转、水平和垂直翻转、亮度调节、饱和度增强和对比度增强等，以丰富 TNBC 样本特征表现形式。数据增强后，TNBC 样本数量由 166 幅扩增至 893 幅，与非 TNBC 样本数量保持一致。

所有图像统一调整分辨率为 224 像素 ×224 像素，以适配模型输入尺寸，并对像素值进行标准化处理以加快模型收敛速度。数据增强完成后，数据集共包含 1 786 幅乳腺超声图像，TNBC 与非 TNBC 图像数量各为 893 幅。采用分层抽样方法按 8 ∶ 1 ∶ 1 比例将数据集划分为训练集、验证集和测试集。其中，训练集包含 TNBC 与非 TNBC 图像各 714 幅，验证集各 89 幅，测试集各 90 幅，确保各数据子集类别分布一致，从而保证模型评估结果的科学性与可比性。

1.2　模型整体架构　本研究提出了一种适用于 TNBC 与非 TNBC 超声图像分类任务的混合神经网络架构——MECSA-Net（图 1）。该架构受到 Medmamba 模型[23]的启发，采用编码器－分类器结构，编码器部分由 4 个特征提取阶段（stage 1～4）堆叠而成，用于逐步提取多尺度、多层次的图像语义特征。Mamba 架构引入的 SSM 机制有效弥补了传统 CNN 的局限性，在处理复杂肿瘤区域时具有优势。

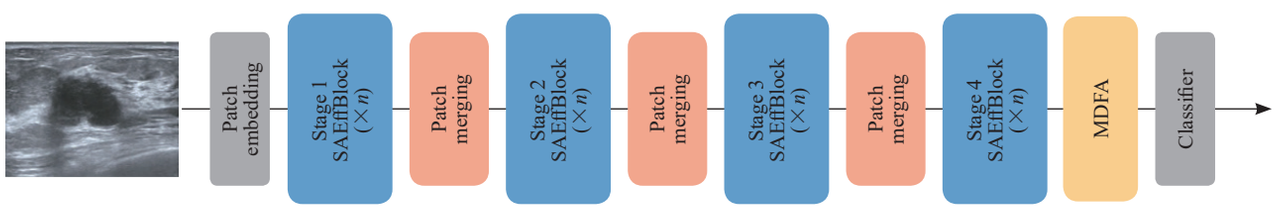


**图 1　MECSA-Net 的总体架构**

**Fig 1　Overall architecture of MECSA-Net**

Patch embedding denotes the process of splitting and embedding image patches, while patch merging refers to the hierarchical merging of image blocks. MECSA-Net: Mamba-enhanced convolutional and state-attention network; SAEffBlock: Shuffle-aware efficient block; MDFA: Multi-dilated fusion attention.

在输入端，MECSA-Net 通过 patch embedding 模块对原始图像进行切分与特征投影，保持空间结构的同时降低计算维度。随后输入 4 个特征提取阶段，每个阶段由多个自定义模块 SAEffBlock 组成，集成了轻量卷积与空间注意力机制，用于联合建模局部与全局特征关系。第 1～3 阶段输出后，均通过 patch merging 模块进行空间下采样与通道扩展，从而构建层次化特征表示。

在编码器末端，MECSA-Net 引入 MDFA，进一步整合深层特征的空间与上下文信息，增强模型对细粒度区域差异的判别能力。最终，融合特征通过自适应平均池化与全连接层输出分类预测结果。该架构在保持低计算复杂度的同时，兼顾全局建模能力与局部细节提取性能，具备良好的可推广性与应用潜力。

1.3　SAEffBlock 模块　SAEffBlock 是 MECSA-Net 中用于高效特征提取的核心模块，采用双分支设计，结合了轻量卷积分支（efficient convolution branch，EffConvBranch）与状态空间建模分支（state space model branch，SSM-Branch），旨在联合建

模乳腺超声图像中的局部纹理和长程依赖信息（图 2）。灵感来源于 Mamba 架构[21]，SAEffBlock 通过引入 SSM 机制与卷积增强策略，提升了对复杂特征区域的感知能力。

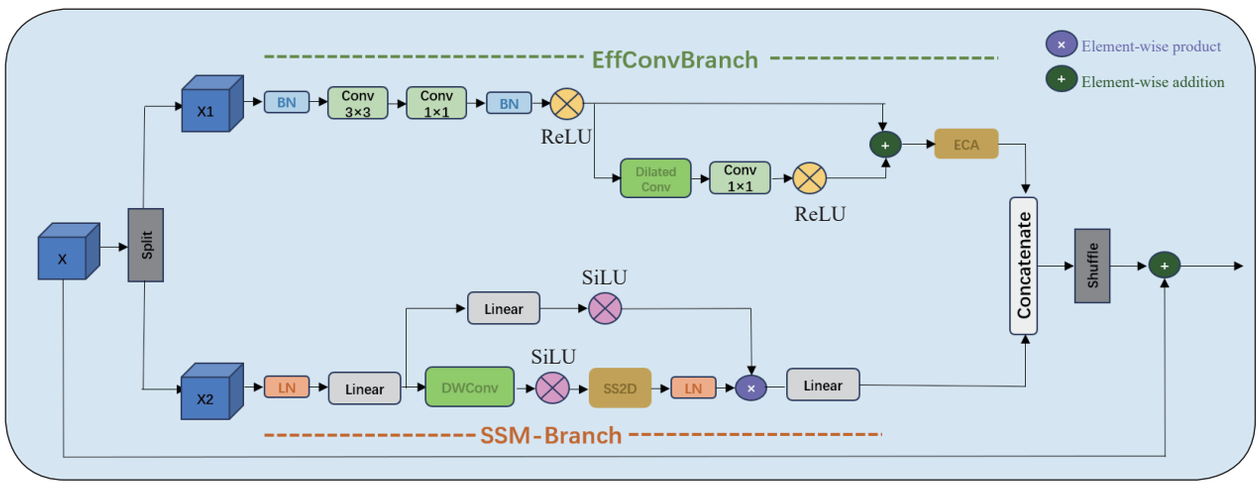


**图 2　SAEffBlock 模块结构示意图**

**Fig 2　Schematic diagram of SAEffBlock module structure**

SAEffBlock: Shuffle-aware efficient block; EffConvBranch: Efficient convolution branch; BN: Batch normalization; Conv: Convolution; LN: Layer normalization; ReLU: Rectified linear unit; ECA: Efficient channel attention; DWConv: Depth-wise convolution; SiLU: Sigmoid linear unit; SS2D: 2D selective scan; SSM: State space modeling.

输入特征首先沿通道维度均匀分配，分别送入 EffConvBranch 和 SSM-Branch 并行建模。SSM-Branch 以二维选择性扫描（2D selective scan，SS2D）机制为核心，通过深度卷积和层归一化（layer normalization，LayerNorm）标准化提取空间上下文信息[24]。EffConvBranch 则采用 3×3 分组卷积、1×1 点卷积、批归一化（batch normalization，BatchNorm）与线性整流单元（rectified linear unit，ReLU）等结构高效提取局部纹理特征，并通过空洞卷积扩展感受野。两分支的输出在通道维度拼接后，通过 1×1 卷积进行特征融合，接着经 BatchNorm 与 ReLU 统一特征空间，并通过 channel shuffle 增强跨通道特征交互。

最后，模块通过残差连接将融合特征与原始输入相加，确保深层特征的稳定传递与表达能力，并通过 dropout 提升模型的泛化性能。整体而言，SAEffBlock 在保证计算效率和结构对称性的基础上，结合高效通道注意力（efficient channel attention，ECA）机制[25]、空洞增强路径和 SSM，充分兼顾局部细节感知与全局依赖建模能力，适用于乳腺超声图像中边界模糊与结构异质性显著的病灶区域。

1.4　MDFA 模块　为了增强模型对多尺度结构的建模能力，并提升对乳腺肿瘤区域空间与通道特征的响应能力，本研究在分类器前端引入了 MDFA 模块（图 3）。该模块结合多种感受野的空洞卷积结构与通道－空间注意力机制，实现局部与全局信息的深度融合，提升了对复杂肿瘤区域的感知与判别能力。

MDFA 模块由 5 个并行分支组成，用于提取不同尺度的语义特征。第 1 分支使用 1×1 标准卷积保持细粒度特征的完整性；第 2～4 分支采用不同膨胀率（6、12、18）的 3×3 空洞卷积，扩大感受野以增强对中大尺度病灶的感知；第 5 分支通过全局平均池化和 1×1 卷积提取全局上下文信息。所有分支输出在通道维度拼接，形成融合后的多尺度特征张量。

MDFA 模型进一步引入联合注意力机制，对关键特征进行加权表达。通道注意力模块通过全局平均池化与 1×1 卷积生成通道权重，空间注意力模块通过 1×1 卷积与 Sigmoid 激活生成空间注意图。两者的输出经逐元素最大融合后，与原始融合特征进行逐元素乘法，增强对关键区域的响应能力，最后通过 1×1 卷积进行通道压缩，并辅以 BatchNorm 与 ReLU 激活，提升训练稳定性与特征一致性。

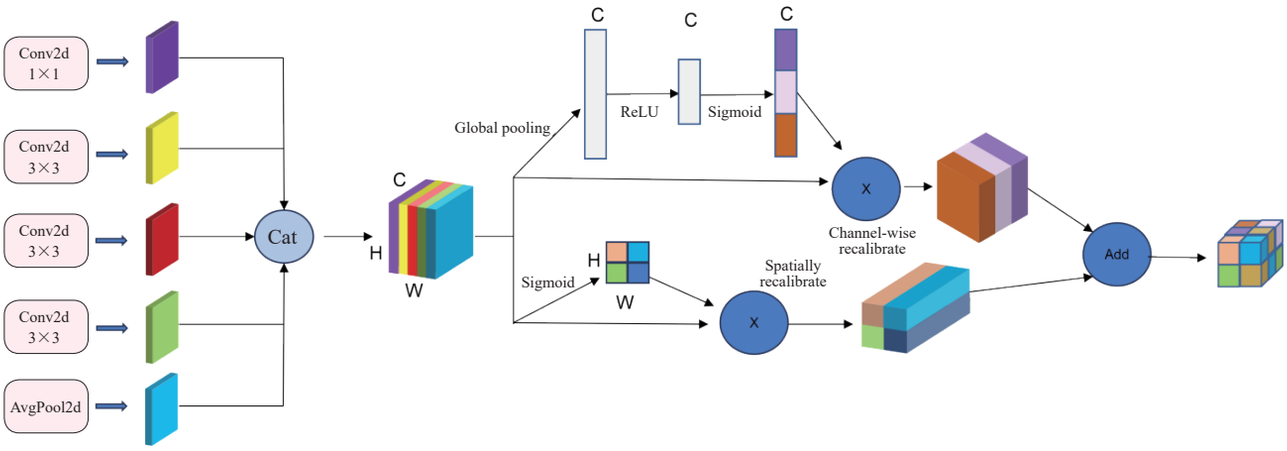


**图3　MDFA模块结构示意图**

**Fig 3　Schematic diagram of MDFA module structure**

MDFA: Multi-dilated fusion attention; Conv2d: 2D convolution; AvgPool2d: 2D average pooling; Cat: Concatenation; C: Channel; H: Height; W: Width; ReLU: Rectified linear unit; Add: Addition.

1.5　*模型训练与参数设置*　模型训练基于PyTorch框架，在NVIDIA GeForce RTX 2080 Ti显卡平台上完成。训练总轮数设为120，批次大小为32。优化器选用Adam，初始学习率为$1\times10^{-4}$，损失函数为交叉熵损失（cross-entropy loss），以优化模型在二分类任务中的判别性能。为提高训练稳定性与模型泛化能力，训练过程中引入学习率衰减与早停机制，并在验证集上动态监控性能指标，以选择最佳模型参数并防止过拟合。

1.6　*模型性能评估*　为全面评估所提出模型的性能，本研究选取多种具有代表性的深度学习模型进行对比，包括经典卷积神经网络ResNet-18与ResNet-50，高效结构代表EfficientNet-B0，Transformer架构的ViT-Base，以及基于SSM的MedMamba-T。这些模型在医学图像分类领域具有广泛应用或良好性能，涵盖不同的建模策略和结构特征，可从多个维度验证所提MECSA-Net模型在TNBC超声图像分类任务中的有效性与优势。测试集共包含180例样本，其中TNBC与非TNBC各90例。所有定量性能指标均基于该测试集计算。计算各模型的准确率、精确率、召回率和F1分数；绘制ROC曲线评估模型在不同判别阈值下的分类性能；同时生成混淆矩阵分析MECSA-Net的分类结果。

为评估MECSA-Net中关键模块对整体性能的具体贡献，设计3组消融对比实验，分别考察EffConvBranch、MDFA模块的独立作用及其协同效果。实验设置包括同时移除EffConvBranch与MDFA模块的基础模型C1、去除EffConvBranch的模型C2、去除MDFA模块的模型C3，以及完整结构的MECSA-Net模型C4。在相同训练策略与测试集条件下评估各模型的作用。

## 2　结　果

2.1　*整体性能对比分析*　由表1可见，MECSA-Net在准确率、精确率和F1分数3项指标上均表现最优，显示出优越的分类性能及良好的精度与召回平衡。ResNet-50虽接在准确率和F1分数上接近MECSA-Net，但精确率与召回率均低于MECSA-Net，分类均衡性和稳定性不及MECSA-Net。ResNet-18的召回率虽最高，但精确率偏低，导致F1分数下降，提示其在提高灵敏性的同时存在明显误报风险。EfficientNet-B0整体性能低于MECSA-Net，精确率偏低影响了综合表现。MedMamba-T精确率与召回率差异较大，提示存在误报风险，分类稳定性不及MECSA-Net。ViT-Base各项指标均处于最低水平，整体分类能力较差，提示其在小样本医学图像任务中训练不稳定，泛化能力不足。

2.2　*ROC曲线与AUC评估*　ROC曲线分析（图4）显示，MECSA-Net的ROC曲线整体最贴近左上角，AUC最高（0.976），显示出最优的分类性能、稳定性与泛化能力。MedMamba-T和EfficientNet-B0的ROC曲线也表现出较好的判别能力，AUC分别为0.966和0.971，整体走势紧随MECSA-Net，但在低假阳性率区段略逊于MECSA-Net，提示其在特

定阈值下的检出能力稍弱。ResNet-50 和 ResNet-18 的 ROC 曲线在低假阳性率区段的提升相对平缓，AUC 分别为 0.970 和 0.961，说明在高判别阈值条件下识别效率尚有提升空间。ViT-Base 的 ROC 曲线波动较大且 AUC 最低（0.950），整体走势不够平稳，反映其性能易受训练扰动影响，稳定性较差。

**表 1　不同模型在 TNBC 与非 TNBC 分类任务中的性能比较**

**Tab 1　Performance comparison of different models in TNBC and non-TNBC classification tasks**

| Model | Accuracy/% (*n*/*N*) | Precision/% (*n*/*N*) | Recall/% (*n*/*N*) | F1-score/% |
|---|---|---|---|---|
| ResNet-18 | 89.4 (161/180) | 84.5 (87/103) | 96.7 (87/90) | 90.2 |
| ResNet-50 | 92.2 (166/180) | 93.2 (82/88) | 91.1 (82/90) | 92.1 |
| EfficientNet-B0 | 88.9 (160/180) | 85.0 (85/100) | 94.4 (85/90) | 89.5 |
| ViT-Base | 66.7 (120/180) | 70.3 (52/74) | 57.8 (52/90) | 63.4 |
| MedMamba-T | 88.9 (160/180) | 84.3 (86/102) | 95.6 (86/90) | 89.6 |
| MECSA-Net | 93.9 (169/180) | 94.4 (84/89) | 93.3 (84/90) | 93.9 |

TNBC: Triple-negative breast cancer; MECSA-Net: Mamba-enhanced convolutional and state-attention network.

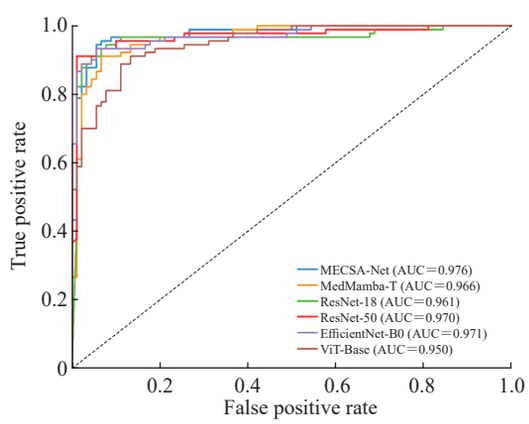


**图 4　MECSA-Net 与对比模型在 TNBC 分类任务中的 ROC 曲线分析**

**Fig 4　ROC curve analysis of MECSA-Net and comparison model in TNBC classification task**

MECSA-Net: Mamba-enhanced convolutional and state-attention network; TNBC: Triple-negative breast cancer; ROC: Receiver operating characteristic; AUC: Area under curve.

2.3　混淆矩阵分析　图 5 展示了 MECSA-Net 在 TNBC 与非 TNBC 分类任务中的混淆矩阵结果。TNBC 类别共 90 例样本，模型正确识别 84 例，误判 6 例为非 TNBC；非 TNBC 类别同为 90 例样本，模型正确识别 85 例，误判 5 例为 TNBC。结果表明，模型在两类样本上的识别能力均较高，分类结果未见明显偏倚，整体误判率较低，进一步说明 MECSA-Net 在不同类别间具备良好的区分能力和判别稳定性。

2.4　消融实验分析　由表 2 可见，完整结构 MECSA-Net 模型 C4 在准确率、精确率、召回率、F1 分数与 AUC 等各项指标中均表现最优，验证了 EffConvBranch 与 MDFA 协同设计在提升模型特征建模能力和分类性能方面的重要作用。与模型 C4 相比，模型 C2 在去除 EffConvBranch 后各项性能均有下降，尤其是召回率和 F1 分数，提示该模块在捕捉局部纹理特征、增强模型灵敏性方面发挥了重要作用。模型 C3 在去除 MDFA 后，召回率和 F1 分数明显下降，尽管精确率仍保持较高水平，但整体分类性能受到影响，说明该模块在提取多尺度上下文信息、提升对结构复杂病灶判别能力方面具有积极贡献。模型 C1 同时缺失 EffConvBranch 与 MDFA，整体性能最弱，各项指标均为最低水平，进一步验证了 EffConvBranch 与 MDFA 在 TNBC 分类任务中具有互补性和必要性。上述结果表明，MECSA-Net 的模块化设计具有良好的协同增强作用，可有效提升模型在 TNBC 分类任务中的精度与鲁棒性。

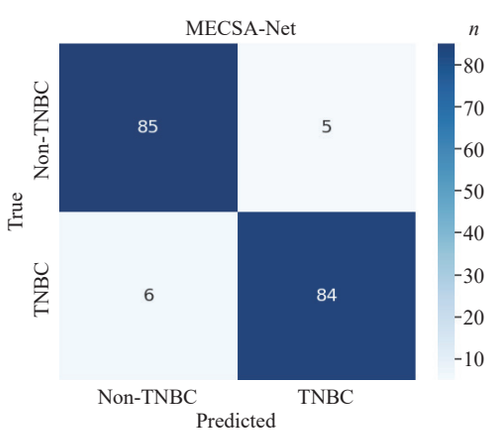


**图 5　MECSA-Net 在 TNBC 与非 TNBC 分类任务中的混淆矩阵**

**Fig 5　Confusion matrix of MECSA-Net in TNBC versus non-TNBC classification tasks**

MECSA-Net: Mamba-enhanced convolutional and state-attention network; TNBC: Triple-negative breast cancer.

表2　MECSA-Net模块消融实验结果

Tab 2　Ablation experimental results of MECSA-Net module

| Model | Accuracy/% (*n*/*N*) | Precision/% (*n*/*N*) | Recall/% (*n*/*N*) | F1-score/% | AUC |
|---|---|---|---|---|---|
| C1 | 86.1 (155/180) | 84.2 (80/95) | 88.9 (80/90) | 86.5 | 0.918 |
| C2 | 88.3 (159/180) | 89.7 (78/87) | 86.7 (78/90) | 88.1 | 0.947 |
| C3 | 87.8 (158/180) | 89.5 (77/86) | 85.6 (77/90) | 87.5 | 0.953 |
| C4 | 93.9 (169/180) | 94.4 (84/89) | 93.3 (84/90) | 93.9 | 0.976 |

C1: The baseline model with both EffConvBranch and MDFA removed; C2: The model with EffConvBranch removed; C3: The model with MDFA removed; C4: The complete MECSA-Net model. MECSA-Net: Mamba-enhanced convolutional and state-attention network; EffConvBranch: Efficient convolution branch; MDFA: Multi-dilated fusion attention; AUC: Area under curve.

## 3　讨　论

针对TNBC在超声图像中存在的结构复杂、边界模糊等挑战，本研究提出了一种基于Mamba与注意力机制的混合神经网络结构——MECSA-Net。该模型通过模块化设计，联合增强局部纹理建模能力与全局依赖表达能力，具有卓越的分类性能和稳定性，这在真实乳腺超声图像数据集上得到了验证。

MECSA-Net的性能提升主要得益于结构中引入的2个关键模块：SAEffBlock与MDFA模块。SAEffBlock采用双分支设计，其中EffConvBranch通过轻量化卷积、空洞增强卷积与通道注意力机制有效提升局部纹理的感知能力，同时扩展感受野；SSM-Branch引入了SS2D，通过多方向交叉扫描与SSM提高长程依赖与全局上下文感知能力。双分支通过特征融合与通道重组实现信息交互与表达互补。

为提升模型对病灶尺度变化与边界模糊区域的识别能力，MDFA模块被引入特征提取末端，进一步增强了对复杂结构区域的感知能力。通过多路空洞卷积提取多尺度语义信息，并结合通道与空间注意机制进行加权整合，提升了对复杂病灶区域的特征建模能力。消融实验表明，EffConvBranch在细粒度纹理建模中显著提升了局部特征的捕捉能力，MDFA模块则在多尺度特征融合与复杂区域的判别中发挥了重要作用。两者协同作用极大地增强了模型的整体性能与鲁棒性。

在TNBC与非TNBC分类任务中，MECSA-Net在准确率、F1分数与AUC等多个指标上表现优异，显示了其出色的分类精度与泛化能力。基于混淆矩阵分析，误判样本主要集中于边界明显模糊或内部回声异质性显著的TNBC病例，这与临床超声阅片中医师判读存在一定主观差异一致，也提示在处理边界不清和局部纹理信息不足的病灶时，模型仍有进一步优化空间。相较于传统卷积模型（如ResNet-18、EfficientNet-B0），MECSA-Net在边界模糊与异质性较强的样本中具有更强的稳定性，有效缓解了局部感受野受限的问题；与ViT-Base相比，MECSA-Net在小样本高噪声场景中表现出更高的训练稳定性与抗过拟合能力，体现了融合轻量卷积与SSM机制的优势。

近年来，多中心超声影像组学及深度学习研究显示，通过跨机构异构数据训练的模型在TNBC分类上可获得更高泛化能力[26-28]；同时，多模态超声（灰阶、彩色多普勒、弹性成像）融合策略被证实能显著提高乳腺癌分型任务的准确性[29-31]，为未来拓展模型临床应用提供了方向。

在实际应用中，MECSA-Net在TNBC智能识别任务中表现出较高的准确性与稳定性，尤其在类别不均衡、小样本与噪声干扰等典型临床环境下具有良好的适应性。该模型结构轻量、参数量适中，具有较高的推理效率与部署灵活性，适合嵌入式智能诊断系统，具备较强的工程可行性与临床转化潜力。结合临床意义，MECSA-Net可辅助超声医师进行快速初筛，尤其在低年资医师和高工作负荷场景中减少漏诊。

尽管本研究取得了积极结果，但仍存在局限性。首先，数据来源于单中心，可能存在样本偏倚，未来可扩展至多中心并进行广泛的模型泛化验证。其次，当前模型基于灰阶超声图像进行建模，尚未整合彩色超声、弹性成像等多模态信息，后续可通过多源图像融合进一步丰富特征表达维度。此外，本研究聚焦于静态二维图像建模，尚未涉及动态图像序列与三维成像，未来可结合时间序列与体积建模技术进一步提升模型的时空感知能力。

综上所述，MECSA-Net在结构设计、性能表现与临床可行性方面展现出优势，为乳腺癌复杂亚型的智能识别提供了有力支持。结合多模态、多中心及动态图像研究，该模型有望在乳腺癌智能筛查与辅助诊疗中得到广泛应用，为临床决策提供技术支撑，同时为未来多源影像融合与临床人工智能辅助系统提供参考。

## ［参 考 文 献］


［1］ KIM J, HARPER A, MCCORMACK V, et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries[J]. Nat Med, 2025, 31(4): 1154-1162. DOI: 10.1038/s41591-025-03502-3.

［2］ 李小钰，黄青，吴雨濛，等. 2022年全球癌症统计报告的窄谱总结和展望[J]. 肿瘤防治研究，2024，51(5)：307-312. DOI: 10.3971/j.issn.1000-8578.2024.24.0437.

［3］ GLUZ O, LIEDTKE C, GOTTSCHALK N, et al. Triple-negative breast cancer: current status and future directions[J]. Ann Oncol, 2009, 20(12): 1913-1927. DOI: 10.1093/annonc/mdp492.

［4］ LI Y, ZHANG H, MERKHER Y, et al. Recent advances in therapeutic strategies for triple-negative breast cancer[J]. J Hematol Oncol, 2022, 15(1): 121. DOI: 10.1186/s13045-022-01341-0.

［5］ XU F, XU K, FAN L, et al. Estrogen receptor beta suppresses the androgen receptor oncogenic effects in triple-negative breast cancer[J]. Chin Med J (Engl), 2024, 137(3): 338-349. DOI: 10.1097/CM9.0000000000002930.

［6］ MA D, WANG C, LI J, et al. Analysis of the diagnostic efficacy of ultrasound, MRI, and combined examination in benign and malignant breast tumors[J]. Front Oncol, 2025, 15: 1494862. DOI: 10.3389/fonc.2025.1494862.

［7］ LEE Y W, WANG M Y, CHEN H Y, et al. Automatic slice selection and diagnosis of breast ultrasound image using deep learning[J]. Biomed Signal Process Contr, 2024, 97: 106688. DOI: 10.1016/j.bspc.2024.106688.

［8］ LIU F, WANG Z, WAN Q, et al. Study on the correlation between ultrasonic features and subtypes of non-special invasive breast cancer[J/OL]. SSRN 4882361 (2024-07-11)[2025-07-01]. *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4882361*.

［9］ ZHANG Q, XIAO J, ZHENG B. Image segmentation of triple-negative breast cancer by incorporating multiscale and parallel attention mechanisms[J]. Sci Program, 2023, 2023: 6629189. DOI: 10.1155/2023/6629189.

［10］ 兰梓涵，彭玉兰. 基于乳腺超声图像的人工智能诊断[J]. 中南大学学报(医学版)，2022，47(8)：1009-1015. DOI: 10.11817/j.issn.1672-7347.2022.220110.

［11］ 唐蕴芯，廖梅，张艳玲，等. 基于乳腺超声视频流和自监督对比学习的肿瘤良恶性分类系统[J]. 南京大学学报(自然科学)，2024，60(1)：26-37. DOI: 10.13232/j.cnki.jnju.2024.01.004.

［12］ BOUKAACHE A, NASSER EDİNNE B, BOUDJEHEM D. Breast cancer image classification using convolutional neural networks (CNN) models[J]. Int J Inform Appl Math, 2024, 6(2): 20-34. DOI: 10.53508/ijiam.1407152.

［13］ HE Q, YANG Q, XIE M. HCTNet: a hybrid CNN-Transformer network for breast ultrasound image segmentation[J]. Comput Biol Med, 2023, 155: 106629. DOI: 10.1016/j.compbiomed.2023.106629.

［14］ JIN Z, ZHANG Q, ZHAO J. CNN-based breast cancer ultrasound medical image recognition[C]// 2024 6th International Conference on Industrial Artificial Intelligence (IAI). August 21-24, 2024, Shenyang, China. IEEE, 2024: 1-5. DOI: 10.1109/IAI63275.2024.10730068.

［15］ 杨杰，蒋严宣，熊欣燕. 结合Transformer和SimAM轻量化路面损伤检测算法[J]. 铁道科学与工程学报，2024，21(9)：3911-3920. DOI: 10.19713/j.cnki.43-1423/u.T20232012.

［16］ SONG B, KC D R, YANG R Y, et al. Classification of mobile-based oral cancer images using the vision transformer and the swin transformer[J]. Cancers (Basel), 2024, 16(5): 987. DOI: 10.3390/cancers16050987.

［17］ SARKER P K, ZHAO Q. Enhanced visible-infrared person re-identification based on cross-attention multiscale residual vision transformer[J]. Pattern Recognit, 2024, 149: 110288. DOI: 10.1016/j.patcog.2024.110288.

［18］ AKKAYA I B, KATHIRESAN S S, ARANI E, et al. Enhancing performance of vision transformers on small datasets through local inductive bias incorporation[J]. Pattern Recognit, 2024, 153: 110510. DOI: 10.1016/j.patcog.2024.110510.

［19］ 李雪，李栋，房建东，等. 基于变化引导和双向Mamba网络的遥感影像变化检测方法[J]. 光学学报，2025，45(5)：218-229. DOI: 10.3788/AOS241826.

［20］ RUAN J, LI J, XIANG S. Vm-unet: vision Mamba UNet for medical image segmentation[J/OL]. arXiv: 2402.02491v2 (2024-11-08)[2025-07-01]. *https://doi.org/10.48550/arXiv.2402.02491*.

［21］ GU A, DAO T. Mamba: linear-time sequence modeling with selective state spaces[J/OL]. arXiv: 2312.00752v1 (2024-05-31)[2025-07-01]. *https://doi.org/10.48550/arXiv.2312.00752*.

［22］ TAO Y, TANG J, ZHAO X, et al. Multi-scale network with attention mechanism for underwater image

enhancement[J]. Neurocomputing, 2024, 595: 127926. DOI: 10.1016/j.neucom.2024.127926.

[23] YUE Y, LI Z. Medmamba: vision Mamba for medical image classification[J/OL]. arXiv: 2403.03849v5 (2024-09-29)[2025-07-01]. *https://doi.org/10.48550/arXiv.2403.03849*.

[24] LIU Y, TIAN Y, ZHAO Y, et al. Vmamba: visual state space model[J/OL]. arXiv: 2401.10166v4 (2024-12-29)[2025-07-01]. *https://doi.org/10.48550/arXiv.2401.10166*.

[25] WANG Q, WU B, ZHU P, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020. Seattle, WA, USA. IEEE, 2020: 11531-11539. DOI: 10.1109/cvpr42600.2020.01155.

[26] LI H, CHENG T. Multicenter and multimodal ultrasound-based radiomics and transformer-driven end-to-end deep learning for breast cancer molecular subtype classification[J]. J Radiat Res Appl Sci, 2025, 18(3): 101656. DOI: 10.1016/j.jrras.2025.101656.

[27] WANG T, ZHU Q, YU T, et al. Multimodal deep learning-based classification of breast non-mass lesions using gray scale and color Doppler ultrasound[J]. Diagnostics (Basel), 2025, 15(23): 2967. DOI: 10.3390/diagnostics15232967.

[28] HUANG Z, ZHANG X, JU Y, et al. Explainable breast cancer molecular expression prediction using multi-task deep-learning based on 3D whole breast ultrasound[J]. Insights Imaging, 2024, 15(1): 227. DOI: 10.1186/s13244-024-01810-9.

[29] SAINI M, PARVAR T A, GRAHAM C, et al. Deep learning-powered multi-parametric ultrasound for classifying metastatic versus reactive axillary lymph nodes[J]. Breast Cancer Res, 2025, 27(1): 185. DOI: 10.1186/s13058-025-02138-y.

[30] XIONG L, TANG X, JIANG X, et al. Automatic segmentation-based multi-modal radiomics analysis of US and MRI for predicting disease-free survival of breast cancer: a multicenter study[J]. Breast Cancer Res, 2024, 26(1): 157. DOI: 10.1186/s13058-024-01909-3.

[31] LI H, ZHAO J, JIANG Z. Deep learning-based computer-aided detection of ultrasound in breast cancer diagnosis: a systematic review and meta-analysis[J]. Clin Radiol, 2024, 79(11): e1403-e1413. DOI: 10.1016/j.crad.2024.08.002.

［本文编辑］　孙　岩