

空间 - 频率域分层特征增强的乳腺超声图像分割

蔡嘉丽, 李露希, 钟森洁, 李春艳*

重庆科技大学数理科学学院, 重庆

收稿日期: 2026年9月16日; 录用日期: 2026年10月7日; 发布日期: 2026年10月20日

摘要

乳腺超声图像分割是计算机辅助诊断系统的关键环节, 但斑点噪声干扰与病灶边界模糊给精确分割带来了严峻挑战。现有方法多采用统一策略处理跳跃连接中的多层级特征, 忽视了浅层特征噪声敏感与深层特征定位模糊的本质差异。针对上述问题, 提出一种空间 - 频率域分层特征增强的轻量化分割网络SFH-Net (Spatial-Frequency Hierarchical Network)。首先, 设计交替双路径编码器, 通过标准卷积与深度可分离卷积的交替组合并引入空洞卷积扩展感受野, 在保证特征提取能力的同时有效控制参数规模; 其次, 提出分层跳跃连接策略, 对浅层特征引入自适应频域滤波模块以抑制斑点噪声, 对深层特征采用坐标注意力机制增强空间定位能力。在BUSI和UDIAT两个公开数据集上的实验表明, SFH-Net以仅13.12 M的参数量分别取得了78.73%和85.15%的Dice系数, 所提出方法能够在保持较低模型复杂度的同时实现有效的病灶分割。

关键词

乳腺超声, 图像分割, 傅里叶变换, 坐标注意力, 深度可分离卷积

Spatial-Frequency Domain Hierarchical Feature Enhancement for Breast Ultrasound Image Segmentation

Jiali Cai, Luxi Li, Senjie Zhong, Chunyan Li*

School of Mathematical and Physical Sciences, Chongqing University of Science and Technology, Chongqing

Received: September 16, 2026; accepted: October 7, 2026; published: October 20, 2026

*通讯作者。

文章引用: 蔡嘉丽, 李露希, 钟森洁, 李春艳. 空间-频率域分层特征增强的乳腺超声图像分割[J]. 图像与信号处理, 2026, 15(4): 505-516. DOI: 10.12677/jisp.2026.154045

Abstract

Breast ultrasound image segmentation is a crucial component of computer-aided diagnosis systems, yet speckle noise interference and blurred lesion boundaries pose severe challenges to accurate segmentation. Existing methods mostly adopt a unified strategy to process multi-level features in skip connections, ignoring the essential differences between shallow features that are sensitive to noise and deep features with imprecise localization, and also underutilize frequency domain information. To address these issues, we propose a lightweight segmentation network named SFH-Net with spatial-frequency domain hierarchical feature enhancement. First, an alternating dual-path encoder is designed: the alternating combination of standard convolution and depthwise separable convolution, combined with the introduction of atrous convolution to expand the receptive field, effectively controls the parameter scale while ensuring the feature extraction capability. Second, a hierarchical skip connection strategy is proposed: an adaptive frequency domain filtering module is introduced for shallow features to suppress speckle noise, and the coordinate attention mechanism is adopted for deep features to enhance spatial localization ability. Experiments on two public datasets (BUSI and UDIAT) demonstrate that SFH-Net achieves Dice coefficients of 78.73% and 85.15% with only 13.12 million parameters, the proposed method can achieve effective lesion segmentation while maintaining low model complexity.

Keywords

Breast Ultrasound, Image Segmentation, Fourier Transform, Coordinate Attention, Depthwise Separable Convolution

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

乳腺癌是全球女性最常见的恶性肿瘤,根据 GLOBOCAN 2022 统计数据,2022 年全球女性乳腺癌新发病例约 230 万例,占女性新发癌症病例总数的 23.8% [1]。乳腺癌的早期发现与准确评估对于后续诊疗具有重要意义。超声成像凭借无电离辐射、实时动态及设备便携等优势,已成为主流筛查手段[2]。然而,超声图像普遍存在斑点噪声严重、肿瘤与腺体对比度低、病灶形态各异等问题,对分割算法的鲁棒性提出了严峻考验[3]。

早期的传统图像处理方法依赖人工特征,面对复杂超声图像泛化能力有限。而深度学习技术的兴起为医学图像分割带来了革命性变化。自 Ronneberger 等[4]提出 U-Net 以来,其编码器-解码器与跳跃连接范式已成为该领域的基准。此后,众多改进模型相继涌现,例如,通过密集嵌套或全尺度跳跃连接缓解语义鸿沟的 UNet++ [5]、UNet3+ [6],引入注意力门控机制聚焦病灶区域的 Attention U-Net [7],以及利用残差学习和空洞卷积强化多尺度特征捕获的 ResUNet [8]、CE-Net [9]。

Transformer 在计算机视觉领域的成功,推动其应用于医学图像分割: TransUNet [10]结合 Transformer 编码器与 U-Net 解码器建模全局依赖; Swin-UNet [11]采用 Swin Transformer 构建纯 Transformer 架构,性能优异; UCTransNet [12]、CMUNet [13]也通过不同模块优化模型。但 Transformer 类方法计算开销和参数量较高,限制了其在资源受限临床场景的部署。

现有分割方法虽持续优化性能,但应对乳腺超声图像仍存在不足。首先,跳跃连接采用统一处理,未兼

顾浅层特征高分辨率易受噪声影响、深层特征语义丰富但空间细节模糊的差异，简单融合无法实现最优信息互补；其次，仅在空间域做特征提取增强，未充分发挥频域分析的噪声抑制优势，可学习频域处理融入深度分割网络的研究仍较有限[14]；此外，多数高性能模型以增加参数数量和计算复杂度换取精度，难以满足临床部署的效率要求。

针对上述挑战，本文中提出了一种基于空间-频率域分层特征增强的轻量化分割网络(SFH-Net)。该网络的核心在于根据编码器不同层级特征的语义属性采取差异化处理策略，实现空频域信息的协同增强。本文的主要贡献如下：

(1) 设计了交替双路径模块 ABP。该结构将标准卷积与深度可分离卷积以交替方式组合，充分发挥前者特征表达能力强与后者参数效率高的互补优势。同时，引入空洞卷积分支扩展网络感受野，增强对不同尺度病灶的捕获能力。上述设计使得编码器在保证多尺度特征提取能力的同时，显著降低了整体参数量，让模型更适合资源受限的临床部署场景。

(2) 提出了分层跳跃连接策略 HSC，针对不同语义层级的特征采取差异化增强方案。对于空间分辨率高但噪声敏感的浅层特征，设计了高效的自适应频域滤波模块。该模块采用低分辨率可学习权重插值策略，生成与特征频谱尺寸匹配的频域调制权重，在控制参数规模的同时，对低频结构信息进行自适应调制，并保留包含边缘线索的高频细节分量，从而增强病灶相关结构特征的表达能力，减少有效边界信息的损失。对于语义信息丰富的深层特征，采用坐标注意力机制，沿水平和垂直方向分别编码位置信息，增强网络对病灶空间位置的精确感知能力。上述策略充分尊重了编码器各层级特征的内在属性，实现了去噪与定位的协同优化。

(3) 构建兼顾精度与效率的新型分割网络 SFH-Net。该网络以仅 13.12 M 的参数量在 BUSI 和 UDIAT 两个公开数据集上分别取得了 78.73%和 85.15%的 Dice 系数，实现了分割精度与查全率的双重提升。

2. 本文方法

2.1. 网络整体架构

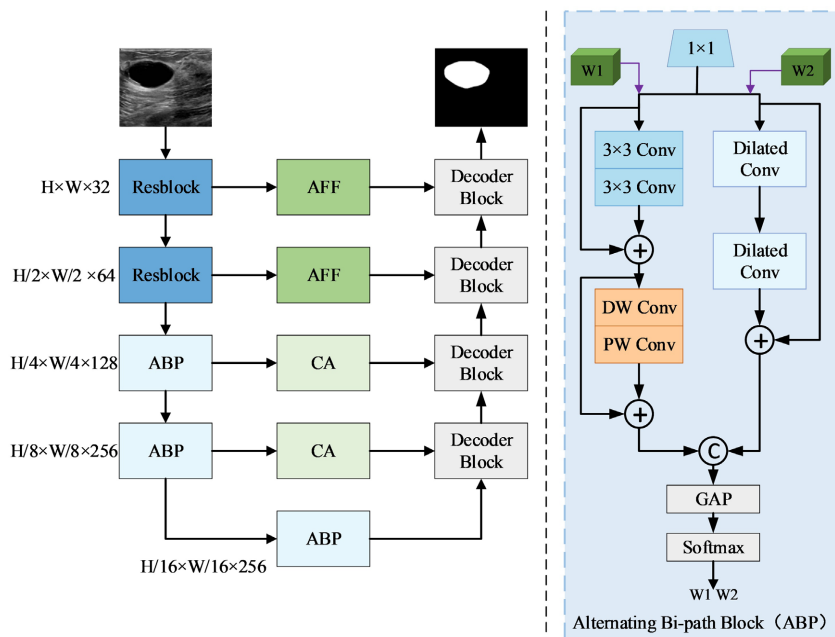


Figure 1. Network structure of SFH-Net

图 1. SFH-Net 网络结构图

SFH-Net 采用经典的编码器-解码器架构,整体结构如图 1 所示。编码器基于交替双路径 ABP 设计,通过标准残差块与轻量化深度可分离卷积块的交替组合实现高效特征提取,同时引入空洞卷积分支以扩展感受野、增强多尺度表征能力。解码器采用双线性插值上采样与 3×3 卷积的级联结构[15],逐级恢复空间分辨率,并通过跳跃连接与编码器特征融合。

针对传统方法对跳跃连接的统一处理忽视了层级差异的问题,本文中提出分层跳跃连接策略 HSC:对噪声敏感的浅层特征,引入自适应频域滤波模块(Adaptive Frequency-domain Filtering, AFF)在频率域选择性抑制斑点噪声;对语义丰富但定位模糊的深层特征,采用坐标注意力模块(Coordinate Attention, CA)通过方向感知的位置编码强化病灶边界定位。最终,分割头输出与输入同分辨率的像素级预测图。

2.2. 交替双路径 ABP 模块(Alternating Bi-Path Block)

编码器作为语义分割特征提取的核心组件[16],其设计需在表达能力与计算效率间寻求平衡。标准卷积建模能力强[17],但深层计算开销大;深度可分离卷积虽能大幅降参[18],但独立表征受限。为兼顾两者优势,本文设计了交替双路径编码器 ABP (如图 1),包含五个经最大池化下采样的特征提取阶段。

针对输入分辨率较高的浅层阶段(Stage 1~2),采用标准残差块[19]提取基础特征,利用恒等映射缓解深层网络的梯度消失。面对语义复杂度剧增的深层阶段(Stage 3~5),本文引入双路径结构以捕获全局上下文信息:主路径将标准卷积与深度可分离卷积交替串联,在控制参数数量的同时维持特征多样性;辅路径则利用空洞卷积,在无额外参数负担下扩展感受野,增强网络对多尺度病灶的捕获能力。

此外,为避免连续同类卷积导致特征同质化,深层阶段引入了阶段间交替策略,即相邻阶段的起始卷积块类型严格相反(如 Stage 3 与 Stage 4 起始类型相异)。该机制促使网络在不同阶段动态切换特征提取视角,有效提升了高层语义特征的互补性与多样性。

双路径输出的有效整合对于发挥不同类型卷积的协同作用至关重要。传统的逐元素相加运算通常假设两条路径具有同等重要性,而通道拼接则又导致参数量翻倍,两者均难以根据输入内容自适应调整各路径的贡献。为此,本文设计了动态融合模块实现自适应加权整合。设主路径输出 F_m ,辅路径输出为 F_a ,融合过程可形式化为:

$$F_{cat} = [F_m; F_a] \in \mathbb{R}^{2C \times H \times W} \quad (1)$$

$$s = \text{GAP}(F_{cat}) \in \mathbb{R}^{2C} \quad (2)$$

$$w = \text{Softmax}(W_2 \cdot \delta(W_1 \cdot s)) \in \mathbb{R}^2 \quad (3)$$

$$F_{out} = w_1 \cdot F_m + w_2 \cdot F_a \quad (4)$$

其中, GAP 为全局平均池化, W_1 、 W_2 为全连接层权重, δ 为 ReLU 激活函数。该融合机制使网络能够根据输入图像的内容特点灵活调整主辅路径的贡献比例,增强了编码器对多样化病灶形态的适应能力。

2.3. 分层跳跃连接策略 HSC (Hierarchical Skip Connection Strategy)

跳跃连接的核心作用是将编码阶段的空间细节信息传递至解码阶段[20],弥补连续下采样过程中不可避免的信息损失。传统 U-Net 及其变体对所有层级的跳跃连接采用统一的直接传递或简单卷积处理策略,这种“一视同仁”的设计忽视了不同层级特征的本质差异。

从特征属性角度分析,浅层特征保留了丰富的纹理细节与边缘信息,空间分辨率为输入的 1/2 至 1/4,对于精确描绘病灶边界具有重要价值。然而,超声图像固有的斑点噪声在浅层特征中最为显著,直接传递可能导致分割结果出现边界伪影和假阳性响应。深层特征经过多次池化和非线性变换后具有较强的语义抽象能力,能够有效区分病灶与正常组织,但空间分辨率已降至输入的 1/16,精确的空间定位能力显

著下降。

针对上述层级差异, 本文提出分层跳跃连接策略 HSC, 根据各层级特征的语义属性施加差异化的增强处理。对于噪声敏感的浅层特征, 设计自适应频域滤波模块在频率域选择性抑制斑点噪声; 对于定位模糊的深层特征, 引入坐标注意力模块通过方向感知的位置编码强化空间定位能力。

2.3.1. 浅层特征: 自适应频域滤波(AFF)

超声图像的斑点噪声源于组织微观结构对超声波的相干散射, 其在频域中主要表现为高频分量的随机扰动[21]。传统的空间域滤波方法(如高斯平滑、中值滤波)在抑制噪声的同时往往模糊了病灶边缘, 难以在去噪与保边界之间取得良好平衡[22]。考虑到噪声与边缘信息在频域中的分布特性存在差异, 本文设计了自适应频域滤波模块 AFF, 通过可学习的频域滤波器选择性抑制噪声频段而保留边缘信息。

首先, 利用二维快速傅里叶变换(FFT)将空间特征映射至频域(如图 2), 得到频谱表征:

$$X = F(F_{in}) \in \mathbb{C}^{C \times H \times W} \quad (5)$$

其中, C 、 H 和 W 分别表示特征图的通道数、高度和宽度。

直接学习全分辨率的频域滤波器需要 $C \times H \times W$ 个复数参数, 不仅计算开销大, 而且容易过拟合。因此, 采用低分辨率滤波器插值策略以实现参数高效的频域滤波。具体而言, 初始化紧凑的可学习滤波器 $W_{lr} \in \mathbb{C}^{C \times h' \times w'}$, 其中 $h' = H/4$, $w' = W/4$ 。在前向传播过程中, 通过双线性插值算子将低分辨率滤波器重构为全分辨率, 随后对频谱进行逐元素调制:

$$Y = X \odot U_{up}(W_{lr}) \quad (6)$$

其中, $\odot$ 表示哈达玛积(Hadamard Product)。

鉴于超声图像的病灶结构与斑点噪声分处低、高频段, AFF 采用“低频调制、高频保留”的差异化策略。利用互补掩膜分离频谱后, 对低频分量通过式(6)的可学习滤波器进行逐元素调制, 以自适应增强关键结构表征; 对于同时包含噪声与微弱边缘的高频分量, 为避免滤波造成不可逆的边界损失, 模块对其不做调制而直接保留。最终将两部分在频域重组并逆变换至空间域, 从而在强化主体结构的同时最大程度保护边缘信息。

2.3.2. 深层特征: 坐标注意力增强(CA)

深层特征经多次下采样后空间分辨率降低、语义判别能力增强, 但空间定位精度不足。通道注意力建模通道依赖时丢失位置信息, 不利于边界定位; 空间注意力虽能生成像素级权重, 但计算复杂且难以捕捉长距离依赖。因此, 本文采用坐标注意力 CA 增强深层特征, 其通过方向感知位置编码, 在保证计算效率的同时捕获长距离空间依赖, 强化网络对病灶边界的定位能力[23]。

给定深层特征 $X \in \mathbb{R}^{C \times H \times W}$, 坐标注意力首先沿水平和垂直两个方向分别进行全局池化, 生成方向感知的特征编码。沿水平方向的池化操作将宽度维度压缩, 保留高度维度的位置信息; 沿垂直方向的池化操作将高度维度压缩, 保留宽度维度的位置信息。具体形式化为:

$$z^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x(h, i) \quad (7)$$

$$z^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x(j, w) \quad (8)$$

其中, h 、 w 分别为水平、垂直方向坐标; H 、 W 为特征图高度、宽度。

上述操作生成了一对方向感知的特征向量 z^h , z^w 。随后, 通过卷积变换与非线性激活生成注意力权

重 g^h 和 g^w ，并对原始特征进行加权激励：

$$Y = X \odot g^h \odot g^w \quad (9)$$

该公式利用广播机制(Broadcasting)将注意力权重映射回原始特征空间，从而使网络能够精准定位病灶边界[24]。

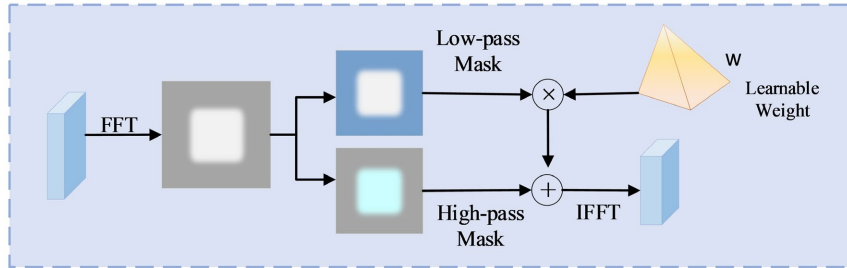


Figure 2. Network architecture of AFF
图 2. AFF 网络结构图

3. 实验结果与分析

3.1. 数据集

为评估所提方法在不同乳腺超声数据分布下的分割性能，本文选取 BUSI 和 UDIAT 两个公开乳腺超声图像数据集进行实验，并在各数据集内部采用 5 折交叉验证以降低小样本条件下单次数据划分所带来的偶然性。

BUSI 数据集由埃及 Baheya Hospital 采集并公开，共包含 780 张乳腺超声图像，平均图像尺寸约为 500×500 像素，图像以 PNG 格式存储，包括正常、良性和恶性三类样本。本文针对病灶分割任务，仅选取其中具有病灶的良性和恶性图像共 647 张。该数据集采用 LOGIQ E9 和 LOGIQ E9 Agile 超声系统采集，并使用 1~5 MHz 线阵探头。UDIAT 数据集采集自西班牙 Sabadell 的 UDIAT Diagnostic Centre，采用 Siemens ACUSON Sequoia C512 超声系统及 17L5 HD 线阵探头，中心频率为 8.5 MHz，共包含 163 张乳腺病灶超声图像，平均图像尺寸约为 760×570 像素，并提供由专业人员标注的病灶分割掩膜。本文实验所采用的 UDIAT 数据版本中包含良性 109 张、恶性 54 张。

从图像特征来看，两个数据集均属于二维 B 模式乳腺超声图像，均包含明显的超声散斑纹理以及不同程度的低对比度和边界模糊问题，因此能够共同用于评价模型对乳腺病灶形态及边界特征的建模能力。但二者在采集中心、超声设备、探头工作频率、图像尺寸以及样本规模等方面存在明显差异。BUSI 样本数量较多而平均图像尺寸相对较小，UDIAT 样本量较少但平均像素尺寸较大；同时，不同设备与采集参数可能导致散斑纹理、组织对比度及病灶边缘表现存在一定的数据分布差异。因此，在两个数据集上进行评价能够考察模型面对不同成像条件时的性能稳定性，但由于实验仍局限于有限的公开数据集，尚不能据此等同于充分验证跨中心临床泛化能力。

3.2. 评价指标

本文采用三项指标进行定量评估，它们分别是 Dice 相似系数、交并比(IOU)和召回率。Dice 相似系数衡量了预测与真实区域的重叠程度：

$$\text{Dice} = \frac{2|P \cap G|}{|P| + |G|} \times 100\% \quad (10)$$

交并比(IoU)定义为预测与真实区域的交并比:

$$\text{IoU} = \frac{|P \cap G|}{|P \cup G|} \times 100\% \quad (11)$$

而召回率则衡量了病灶区域的检出能力:

$$\text{Recall} = \frac{|P \cap G|}{|G|} \times 100\% \quad (12)$$

3.3. 实现细节

本文实验环境配置为 Intel Core Ultra 7 265KF CPU、32 GB 内存及 NVIDIA RTX 5070 Ti GPU。训练数据经随机水平与垂直翻转、旋转及颜色抖动增强。网络训练共计 150 个 epoch, 批大小为 8; 采用 Adam 优化器, 权重衰减为 1×10^{-4} , 且学习率按多项式策略进行衰减。

本文采用二元交叉熵损失、Dice 损失与 IoU 损失的加权组合:

$$\mathcal{L} = 0.4\mathcal{L}_{\text{BCE}} + 0.3\mathcal{L}_{\text{Dice}} + 0.3\mathcal{L}_{\text{IoU}} \quad (13)$$

其中, $\mathcal{L}_{\text{BCE}}$ 为像素级二元交叉熵损失, $\mathcal{L}_{\text{Dice}} = 1 - \text{Dice}$, $\mathcal{L}_{\text{IoU}} = 1 - \text{IoU}$ 。三种损失的组合使模型在像素级分类、区域重叠度和边界精度上实现均衡优化。

3.4. 对比实验结果

本文选取 8 种代表性方法进行对比, 包括经典的 U-Net 及其变体 Attention U-Net、U-Net++, 基于 Transformer 的 TransUNet 和 Swin-UNet, 以及近年的 ACC-UNet [25]、SF-UNet [26] 和 MA-UNet [27]。实验结果见表 1 和表 2。

实验结果表明, 相较于基线 U-Net, SFH-Net 在 BUSI 数据集上的 Dice 和 IoU 分别提升 5.13% 和 5.01%, 参数量降低 46.3%。与次优模型 MA-UNet 相比, SFH-Net 以仅其 22.6% 的参数量, 在 BUSI 和 UDIAT 上使 Dice 分别进一步提升 1.69% 和 1.33%。此外, 受限于 BUSI 的小样本特性, 纯 Transformer 架构(如 Swin-UNet)表现欠佳, 而 SFH-Net 凭借 CNN 与空频域协同优势展现出更佳的小样本学习能力。

Table 1. Performance comparison of different methods on the BUSI dataset

表 1. BUSI 数据集上各方法的性能对比

Method	Params (M)	Dice (%)	IoU (%)	Recall (%)
U-Net	24.44	73.60	65.01	76.09
Attention U-Net	25.80	74.15	65.82	76.85
U-Net++	26.90	74.92	66.25	77.40
TransUNet	93.16	75.20	66.80	78.50
Swin-UNet	27.17	67.10	56.80	65.40
ACC-UNet	16.79	75.45	66.80	79.60
SF-UNet	28.91	76.50	67.95	79.25
MA-UNet	58.04	77.04	68.11	79.20
SFH-Net (Ours)	13.12	78.73	70.02	80.54

Table 2. Performance comparison of different methods on UDIAT
表 2. UDIAT 数据集上各方法的性能对比

Method	Params (M)	Dice (%)	IoU (%)	Recall (%)
U-Net	24.44	80.52	68.15	81.34
Attention U-Net	25.80	81.45	69.23	82.10
U-Net++	26.90	82.68	70.85	83.45
TransUNet	93.16	80.15	67.54	80.92
Swin-UNet	27.17	72.86	58.42	74.15
ACC-UNet	16.79	83.12	71.45	83.89
SF-UNet	28.91	83.56	72.18	84.05
MA-UNet	58.04	83.82	72.65	84.62
SFH-Net (Ours)	13.12	85.15	74.64	83.12

针对在 UDIAT 上 Recall 略低于 MA-Unet，但在 BUSI 上取得最高 Recall 的差异，主要归因于成像特性：BUSI 斑点噪声严重，SFH-Net 通过 AFF 与 CA 模块协同有效抑制干扰并强化定位，大幅降低了漏检率；而 UDIAT 低对比度与边界模糊的特性，使得模型在决策时更倾向于平衡 Precision 与 Recall。总体分割可视化(图 3)证实，SFH-Net 对各类不规则及边界模糊病灶均具备优越的刻画能力。

为探究 SFH-Net 高精度的内在机理，特征可视化(图 4、图 5)证实了核心模块的有效性：AFF 显著抑制了浅层特征中的背景异常高频响应，在滤除散射伪影的同时保留了高信噪比的轮廓细节；CA 模块则通过方向感知机制，在深层生成贯穿病灶的主体带状注意力权重，有效弥补了连续下采样导致的空间坐标丢失。两者的协同强化了抑噪与定位优势，保障了复杂背景下的精准分割。

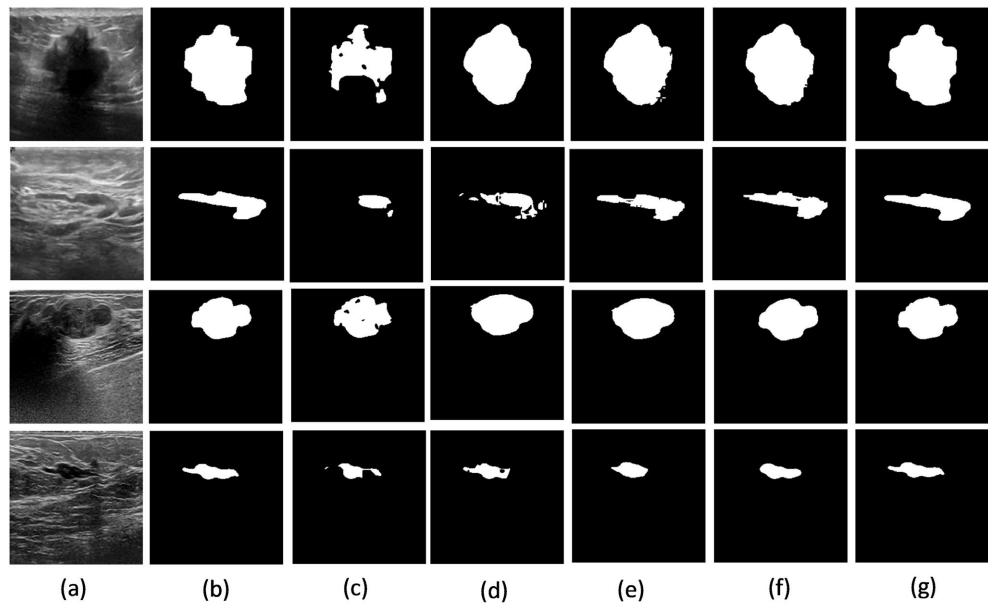


Figure 3. Segmentation results of different algorithms on the BUSI and UDIAT Datasets. (a) Original images; (b) Ground truth (GT); (c) Swin-UNet; (d) U-Net; (e) ACC-UNet; (f) MA-UNet; (g) SFH-Net

图 3. 不同算法在 BUSI 和 UDIAT 上的分割结果 (a) 原图; (b) GT; (c) Swin-UNet; (d) U-Net; (e) ACC-UNet; (f) MA-UNet; (g) SFH-Net

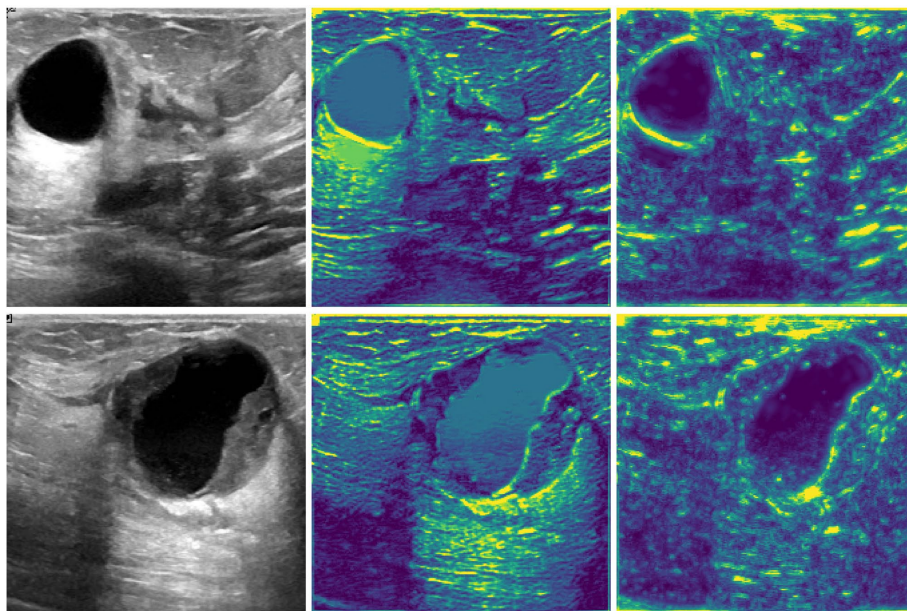


Figure 4. Visual comparison of shallow features before and after processing by the AFF model
图 4. AFF 模块处理前后的浅层特征可视化对比

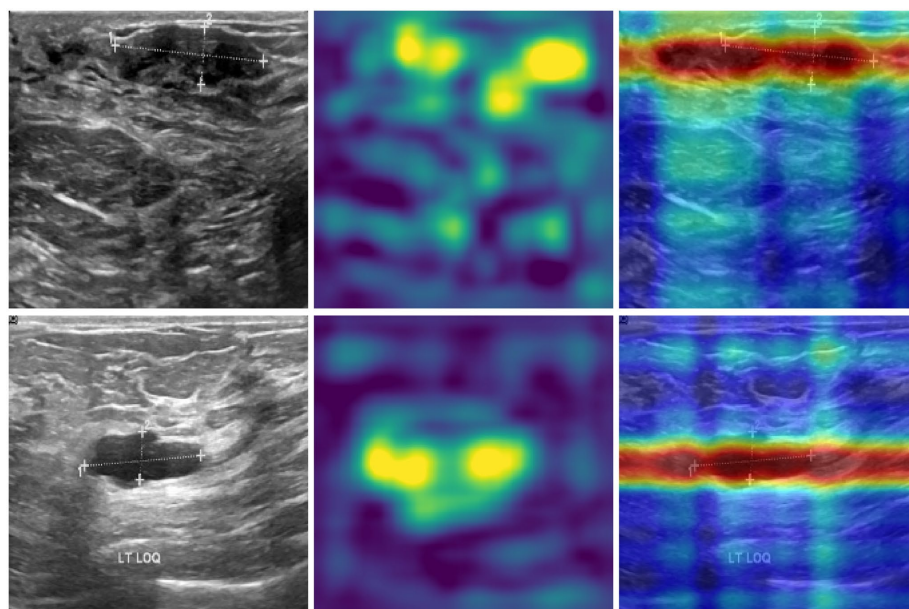


Figure 5. Visualization of deep features and spatial attention weights generated by the CA model
图 5. 原始深层特征与 CA 模块空间注意力权重的可视化

3.5. 消融实验

为验证 SFH-Net 各核心模块及分层设计策略的有效性, 本文在 BUSI 数据集上进行了系统的消融实验, 其结果见表 3~5。

3.5.1. 编码器结构设计分析

首先, 本文探讨了交替双路径编码器(ABP)的引入及其最佳部署策略, 结果如表 3 所示。相较于 U-

Net 基线，ABP 的引入带来了显著的性能提升。实验进一步对比了“全阶段 ABP”与本文的“浅层残差 + 深层 ABP”混合策略。结果显示，全 ABP 配置虽优于基线，但不及混合策略(Dice 为 77.75%，IoU 为 68.93%)。原因在于，浅层特征富含高分辨率纹理，过早引入大感受野的空洞卷积与复杂双路径易破坏局部细节并引入冗余；而标准残差块更适于浅层高效提取。这证实了非对称编码器设计在捕获多尺度上下文与保留局部细节间的平衡优势。

针对深层 ABP 主路径的卷积组合方式(表 4)，对比实验表明：全标准卷积串联(Serial-Std)因模型容量大易拟合斑点噪声而泛化受限；全深度可分离卷积串联(Serial-DW)因跨通道关联不足致使精度下降；同等参数下的并联结构(Parallel)则缺乏深层非线性叠加。相比之下，本文的交替串联结构在不增参的前提下加深了特征传播路径，使标准卷积的通道深度融合与深度可分离卷积的空间高效提取形成逐层协同，在计算效率与特征表征深度之间取得最优平衡。

Table 3. Ablation study results on the effectiveness of encoder layering
表 3. 编码器分层有效性消融实验结果

Model	Encoder1, 2	Encoder3, 4	Dice (%)	IoU (%)	Params (M)
U-Net	—	—	73.60	65.01	24.44
SFH-Net	ABP	ABP	77.75	68.93	13.42
Ours	Resblock	ABP	78.73	70.02	13.12

Table 4. Ablation study results of ABP module main path combination methods
表 4. ABP 模块主路径组合方式消融实验结果

主路径结构	Params (M)	Dice (%)	IoU (%)
Serial-std	15.62	78.25	69.53
Serial-dw	10.62	76.58	67.76
Parallel	13.12	77.89	69.12
Alternating (Ours)	13.12	78.73	70.02

3.5.2. 分层跳跃连接策略分析

Table 5. Ablation study results on the effectiveness of hierarchical skip connections
表 5. 分层跳跃连接有效性消融实验结果

Model	Skip 1, 2	Skip 3, 4	Dice (%)	IoU (%)	Params (M)
U-Net	—	—	73.60	65.01	24.44
SFH-Net (Encoder)	—	—	74.82	64.15	12.30
SFH-Net (Encoder)	AFF	—	76.95	66.82	12.90
SFH-Net (Encoder)	—	CA	77.41	67.53	12.52
SFH-Net (Encoder)	CA	CA	76.69	68.12	12.55
SFH-Net (Encoder)	AFF	AFF	76.23	66.05	13.50
SFH-Net (Reverse)	CA	AFF	76.12	66.25	12.93
SFH-Net (Ours)	AFF	CA	78.73	70.02	13.12

在验证了编码器结构的有效性后, 本文进一步探究了不同跳跃连接增强策略对分割性能的影响, 结果如表 5 所示。实验表明, 单独在浅层引入 AFF 或深层引入 CA 均有性能收益, 但全局单一策略(全 AFF 或全 CA)效果不佳: 深层语义特征不适合频域滤波, 而浅层噪声会严重干扰注意力机制的聚焦。值得注意的是, 反向分层配置(浅层 CA + 深层 AFF)取得了最低的 Dice, 较本文策略低 2.61%, 这深刻揭示了特征属性匹配的关键性: 在噪声主导的浅层强行使用注意力会放大干扰, 而在语义主导的深层使用频域滤波则破坏结构。最终, 本文提出的正向分层策略(浅层 AFF + 深层 CA)在仅增加 0.82 M 参数的情况下实现了 3.91% 的性能提升, 充分证实了依据特征层级属性定制增强策略的科学性与优越性。

4. 结论

本文针对乳腺超声图像中斑点噪声干扰与病灶边界模糊问题, 提出了一种基于空间 - 频率域分层特征增强的轻量化分割网络 SFH-Net。网络采用交替双路径编码器, 通过标准卷积、深度可分离卷积与空洞卷积的协同设计, 在控制参数规模的同时增强多尺度特征表征能力; 针对浅层特征易受噪声干扰的问题, 引入自适应频域滤波模块, 对频率成分进行自适应调制; 针对深层特征空间定位信息不足的问题, 引入坐标注意力机制增强病灶区域的位置感知能力。实验结果表明, SFH-Net 以 13.12 M 的参数量, 在 BUSI 和 UDIAT 数据集上分别取得 78.73% 和 85.15% 的 Dice 系数以及 70.02% 和 74.64% 的 IoU, 在所比较方法中取得了较好的整体分割性能, 表明所提出的分层特征增强策略能够有效兼顾病灶分割精度与模型参数效率。

尽管 SFH-Net 在 BUSI 和 UDIAT 两个公开数据集上取得了较好的实验结果, 当前研究仍存在一定局限性。一方面, 现有实验仅涉及两个公开乳腺超声数据集, 且主要采用数据集内部的交叉验证方式进行评价, 尚不能充分覆盖不同医疗中心、超声设备、成像协议以及患者群体所带来的数据分布差异, 因此模型的真实跨中心泛化能力仍需进一步验证; 另一方面, 对于极端低对比度、边界高度模糊或伴有明显声影干扰的病灶, 分割结果仍存在改进空间。未来将进一步引入来自不同医疗中心和不同设备的多样化外部数据, 开展跨数据集独立测试和多中心外部验证, 以系统评估模型的鲁棒性与泛化能力; 同时探索弱监督或半监督学习策略以降低对像素级标注的依赖, 并进一步研究该分层增强范式在甲状腺结节等其他超声图像分割任务中的适用性。

参考文献

- [1] Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., *et al.* (2024) Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 74, 229-263. <https://doi.org/10.3322/caac.21834>
- [2] Berg, W.A. (2008) Combined Screening with Ultrasound and Mammography vs Mammography Alone in Women at Elevated Risk of Breast Cancer. *JAMA*, 299, 2151-2163. <https://doi.org/10.1001/jama.299.18.2151>
- [3] Loizou, C.P. and Pattichis, C.S. (2015) Despeckle Filtering for Ultrasound Imaging and Video. 2nd Edition, Morgan & Claypool Publishers.
- [4] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Springer, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [5] Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N. and Liang, J. (2018) UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In: Stoyanov, D., *et al.*, Eds., *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 3-11. https://doi.org/10.1007/978-3-030-00889-5_1
- [6] Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., *et al.* (2020) UNet 3+: A Full-Scale Connected Unet for Medical Image Segmentation. *ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4-8 May 2020, 1055-1059. <https://doi.org/10.1109/icassp40776.2020.9053405>
- [7] Oktay, O., Schlemper, J., Le Folgoc, L., *et al.* (2018) Attention U-Net: Learning Where to Look for the Pancreas. arXiv: 1804.03999

- [8] Zhang, Z., Liu, Q. and Wang, Y. (2018) Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, **15**, 749-753. <https://doi.org/10.1109/lgrs.2018.2802944>
- [9] Gu, Z., Cheng, J., Fu, H., Zhou, K., Hao, H., Zhao, Y., *et al.* (2019) Ce-Net: Context Encoder Network for 2D Medical Image Segmentation. *IEEE Transactions on Medical Imaging*, **38**, 2281-2292. <https://doi.org/10.1109/tmi.2019.2903562>
- [10] Chen, J., Lu, Y., Yu, Q., *et al.* (2021) TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv: 2102.04306
- [11] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., *et al.* (2023) Swin-Unet: UNET-Like Pure Transformer for Medical Image Segmentation. In: Karlinsky, L., Michaeli, T. and Nishino, K., Eds., *Computer Vision—ECCV 2022 Workshops*, Springer, 205-218. https://doi.org/10.1007/978-3-031-25066-8_9
- [12] Wang, H., Cao, P., Wang, J. and Zaiane, O.R. (2022) Uctransnet: Rethinking the Skip Connections in U-Net from a Channel-Wise Perspective with Transformer. *Proceedings of the AAAI Conference on Artificial Intelligence*, **36**, 2441-2449. <https://doi.org/10.1609/aaai.v36i3.20144>
- [13] Tang, F., Wang, L., Ning, C., Xian, M. and Ding, J. (2023) CMU-Net: A Strong ConvMixer-Based Medical Ultrasound Image Segmentation Network. 2023 *IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, Cartagena, 18-21 April 2023, 1-5. <https://doi.org/10.1109/isbi53787.2023.10230609>
- [14] Xu, K., Qin, M., Sun, F., Wang, Y., Chen, Y. and Ren, F. (2020) Learning in the Frequency Domain. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 14-19 June 2020, 1740-1749. <https://doi.org/10.1109/cvpr42600.2020.00181>
- [15] 陈柯炎, 刘巧红, 韩啸翔, 等. 结合频域先验和特征增强的心脏图像分割方法[J]. 激光与光电子学进展, 2024, 61(10): 367-376.
- [16] 刘峰. 基于信息迁移的半监督心脏磁共振图像分割[D]: [硕士学位论文]. 烟台: 烟台大学, 2025.
- [17] Howard, A.G., Zhu, M., Chen, B., *et al.* (2017) MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv: 1704.04861
- [18] 杨军, 张金影, 康玥. 基于自注意力机制的高分遥感影像语义分割[J]. 哈尔滨工程大学学报, 2025, 46(2): 344-354.
- [19] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [20] 陈誥劫. 基于改进 U-Net 的医学影像分割方法研究[D]: [硕士学位论文]. 广州: 广东财经大学, 2025.
- [21] Joel, T. and Sivakumar, R. (2018) An Extensive Review on Despeckling of Medical Ultrasound Images Using Various Transformation Techniques. *Applied Acoustics*, **138**, 18-27. <https://doi.org/10.1016/j.apacoust.2018.03.023>
- [22] Esakkirajan, S., Vimalraj, C.T., Muhammed, R. and Subramanian, G. (2013) Adaptive Wavelet Packet-Based De-Speckling of Ultrasound Images with Bilateral Filter. *Ultrasound in Medicine & Biology*, **39**, 2463-2476. <https://doi.org/10.1016/j.ultrasmedbio.2013.07.009>
- [23] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 19-25 June 2021, 13713-13722. <https://doi.org/10.1109/cvpr46437.2021.01350>
- [24] Wu, H., Chen, S., Chen, G., Wang, W., Lei, B. and Wen, Z. (2022) Fat-Net: Feature Adaptive Transformers for Automated Skin Lesion Segmentation. *Medical Image Analysis*, **76**, Article ID: 102327. <https://doi.org/10.1016/j.media.2021.102327>
- [25] Ibtehaz, N. and Kihara, D. (2023) ACC-UNet: A Completely Convolutional UNet Model for the 2020s. In: Greenspan, H., *et al.*, *Medical Image Computing and Computer Assisted Intervention (MICCAI 2023)*, Springer, 692-702. https://doi.org/10.1007/978-3-031-43898-1_66
- [26] Zhou, Z.H., He, A., Wu, Y., Yao, R., Xie, X. and Li, T. (2024) Spatial-Frequency Dual Domain Attention Network for Medical Image Segmentation. arXiv: 2406.07952v2
- [27] Li, H., Ren, Z., Zhu, G., Liang, Y., Cui, H., Wang, C., *et al.* (2025) Enhancing Medical Image Segmentation with MA-UNet: A Multi-Scale Attention Framework. *The Visual Computer*, **41**, 6103-6120. <https://doi.org/10.1007/s00371-024-03774-9>