

# 基于残差差分卷积的医学图像模糊边缘增强分割<sup>①</sup>



孙继杰<sup>1</sup>, 强彦<sup>1</sup>, 陈琪<sup>2</sup>, 李青<sup>1</sup>, 郭瑛璇<sup>1</sup>, 王彦瑾<sup>1</sup>, 安洋<sup>1</sup>

<sup>1</sup>(中北大学 软件学院, 太原 030051)

<sup>2</sup>(澳门科技大学 创新工程学院, 澳门 999078)

通信作者: 强彦, E-mail: [qiangyan@nuc.edu.cn](mailto:qiangyan@nuc.edu.cn)

**摘要:** 医学影像精准分割常因病灶与邻近组织形态与色彩高度相似而边界模糊, 降低了分割精度与诊断可靠性。本研究提出基于残差差分卷积的模糊边缘增强网络, 设计边缘感知细节提取模块, 通过梯度特征编码强化边缘信息。在瓶颈层引入全局边缘细节双层级路由注意力, 融合上下文实现精准建模; 构建残差边缘感知定位模块, 实现模糊边缘精细定位。实验结果表明, 在结直肠癌、皮肤病变和乳腺超声图像公开数据集上, 该方法显著减少参数量与计算复杂度的同时, 分割性能均稳超现有先进方法。系统消融研究进一步验证各组件在提升边缘检测效能方面的协同作用, 充分展示了该方法提升医学图像分析可靠性的潜在应用价值。

**关键词:** 深度学习; 医学图像分割; 边缘增强; 注意力机制

引用格式: 孙继杰, 强彦, 陈琪, 李青, 郭瑛璇, 王彦瑾, 安洋. 基于残差差分卷积的医学图像模糊边缘增强分割. 计算机系统应用, 2025, 34(12): 111-118. <http://www.c-s-a.org.cn/1003-3254/10034.html>

## Segmentation of Medical Images with Fuzzy Edge Enhancement Based on Residual Difference Convolution

SUN Ji-Jie<sup>1</sup>, QIANG Yan<sup>1</sup>, CHEN Qi<sup>2</sup>, LI Qing<sup>1</sup>, GUO Ying-Xuan<sup>1</sup>, WANG Yan-Jin<sup>1</sup>, AN Yang<sup>1</sup>

<sup>1</sup>(School of Software, North University of China, Taiyuan 030051, China)

<sup>2</sup>(Faculty of Innovation Engineering, Macau University of Science and Technology, Macao 999078, China)

**Abstract:** Accurate medical image segmentation is often challenged by blurred boundaries caused by the high morphological and color similarity between lesions and surrounding tissue, thus compromising both segmentation accuracy and diagnostic reliability. To this end, this study proposes a fuzzy edge enhancement network based on residual differential convolution and designs a feature extraction module for edge awareness detail to enhance edge information via gradient feature coding. Meanwhile, the global edge detail bi-level routing attention is introduced in the bottleneck to integrate contextual information for precise modeling. Furthermore, a residual edge-aware localization module is constructed to enable refined localization of fuzzy edges. The experimental results show that on public colorectal cancer, skin lesion, and breast ultrasound images datasets, the proposed method significantly reduces the parameter count and computational complexity, with even stronger segmentation performance than existing state-of-the-art methods. Systematic ablation studies further validate the synergistic contribution of each component to edge detection, fully showcasing the potential application significance of the proposed method in improving the reliability of medical image analysis.

**Key words:** deep learning; medical image segmentation; edge enhancement; attention mechanism

① 基金项目: 国家自然科学基金面上项目 (62376183); 中国博士后科学基金 (GZC20241586); 山西省高等学校科技创新计划 (2024L181); 山西省基础研究计划 (202403021212184, 202203021212114); 中北大学第二十届研究生科研科技计划 (20242055)

收稿时间: 2025-04-21; 修改时间: 2025-05-28; 采用时间: 2025-07-29; csa 在线出版时间: 2025-10-21

CNKI 网络首发时间: 2025-10-22

## 1 引言

基于深度学习的医学图像分割在计算机辅助诊断 (computer-aided diagnosis, CAD) 系统中至关重要。准确高效的图像分割不仅能准确识别病变区域的特征, 还能为临床诊断和疾病研究提供可靠的定量依据<sup>[1]</sup>。然而目前医学图像分割任务仍然面临边界模糊的严峻挑战。例如, 由于前景和背景之间的像素差异极小, 皮肤病变图像常表现出边界模糊, 从而使准确的边界区分变得复杂。其次, 在结肠息肉图像中, 息肉区域的大小和数量变化很大, 再加上其像素特征与背景区域非常相似, 使得分割任务更加复杂<sup>[2]</sup>。此外超声图像具有灰度对比度低、边界模糊、噪声干扰强的特点, 对分割模型的细节捕捉能力要求较高。

随着医学图像分割技术发展, 针对组织边界模糊导致的分割精度问题, 研究者基于卷积神经网络 (convolutional neural network, CNN) 提出系列创新算法。2015年 U-Net<sup>[3]</sup>通过对称编解码结构提升细节提取能力, 但跳跃连接引入噪声。2018年 UNet++<sup>[4]</sup>采用密集连接实现多层特征融合, 却因缺乏全局建模, 对多变病变分割效果有限。UNeXt<sup>[5]</sup>提出了一种两阶段特征提取策略, 在 MLP 阶段则通过引入 Tok-MLP 实现全局特征建模, 增强了网络对潜在空间表示的捕捉能力, 但缺少细节捕获能力。DCSAU-Net<sup>[6]</sup>整合紧凑型分裂注意力机制融合高低层语义, 但其卷积操作的局部感受野特性仍限制了全局特征建模能力, 影响分割精度与边界细节检测。

Transformer<sup>[7]</sup>在自然语言处理中的全局建模能力推动着其在视觉领域的应用。其强大的全局建模能力可实现精确全局信息建模, 催生了诸多混合 CNN-Transformer 方法。改进 TranUnet<sup>[8]</sup>在跳跃连接中设计跨尺度局部增强模块, 借助半密集融合策略将编码器多层高分辨率细节进行融合, 却可能加剧浅层特征冗余。在 LeViT-UNet<sup>[9]</sup>中, 全局特征捕获模块被引入网络架构, 增强了对图像全局信息的捕捉能力, 但在应对边界模糊问题时, 局部细节描述仍不足。陶惜婷等人<sup>[10]</sup>设计 CNN 与 Transformer 并行, 同时捕获局部细节与全局上下文, 有效平衡皮肤病变的纹理特征与长程依赖关系, 但对高度不规则的病变边界分割效果较差。

XBound-Former<sup>[11]</sup>引入了一种跨尺度边界感知 Transformer, 将边界先验知识与自我注意机制相结合, 以解决模糊边界分割问题, 但在检测关键边界点方面仍

不明确。MGCBFormer<sup>[12]</sup>提出了一种多尺度网格先验和类间边界感知变换器, 在多个公共息肉数据集上实现了较好的结果。但由于使用的卷积核较大, 导致参数推断速度较慢。LM-Net<sup>[13]</sup>通过轻量化多分支模块与局部和全局特征 Transformer 设计, 有效增强了多尺度特征融合与全局和局部语义互补, 但在面对极低对比度模态 (如超声图像) 的边界模糊区域时, 其多尺度特征的深度语义关联能力仍存在提升空间。王朝欣等人<sup>[14]</sup>提出 MSC-LSAM 算法, 采用双编码器结合多尺度交叉注意力和高效通道注意力, 在实验结果上优于主流算法, 但其处理皮肤病等模糊边界时, 存在颜色特征利用不足、纹理细节捕捉有限等问题。U-Net v2<sup>[15]</sup>提出鲁棒高效的 U 型变体, 基于 Encoder-Decoder 架构, 通过深度编码器提取多级特征, 利用 Hadamard 乘积融合高层语义与低层细节, 并通过新颖跳跃连接增强特征表示。但其在处理低对比度模态或复杂解剖结构时, 跨层级特征交互与长距离依赖建模效率有待提升。王骁崑等人<sup>[16]</sup>设计了 MoFormer 网络, 通过编码器中融合 Transformer 学习机制, 扩大上下文感知视野, 提升局部与全局信息的多尺度特征提取能力。该模型在多种医学图像数据集上表现出良好分割精度, 不过模型编码器输出特征图通过跳跃连接送入解码器时存在忽略内在细节的风险。

针对医学图像分割中的边缘模糊与全局依赖问题, 本文提出融合差分卷积的方法, 核心工作包括: (1) 设计边缘感知细节提取模块, 通过梯度特征编码强化边缘信息捕获; (2) 构建残差边缘感知定位模块, 强化模糊边缘的细节定位; (3) 在瓶颈层引入全局细节双层级路由注意力, 融合上下文与边缘细节实现病灶精准建模。通过实验证明该方法在分割精度、边界定位与参数效率上均显著提升。

## 2 方法介绍

图 1 为本文方法示意图, 原始图像首先输入细节特征保留模块 (detailed feature retention, DFR), 通过扩大感受野提取广泛特征并保留细节; 随后特征图经编码器中的边缘感知细节特征提取模块 (edge awareness detail feature extraction block, EADBlock) 处理高级语义信息, 传递至瓶颈层的全局边缘细节双层级路由注意力机制 (global edge detail bi-level routing attention, GEDBAttention) 生成含全局边缘细节的特征图。解码器采用与编码器对应的结构, 同时编码器特征图通过

跳跃连接输入残差边缘感知定位模块 (residual edge-aware localisation block, REALBlock) 增强边缘细节定

位. 解码时融合跳跃连接特征与上一层特征以恢复空间分辨率, 最终由分割头生成预测结果.

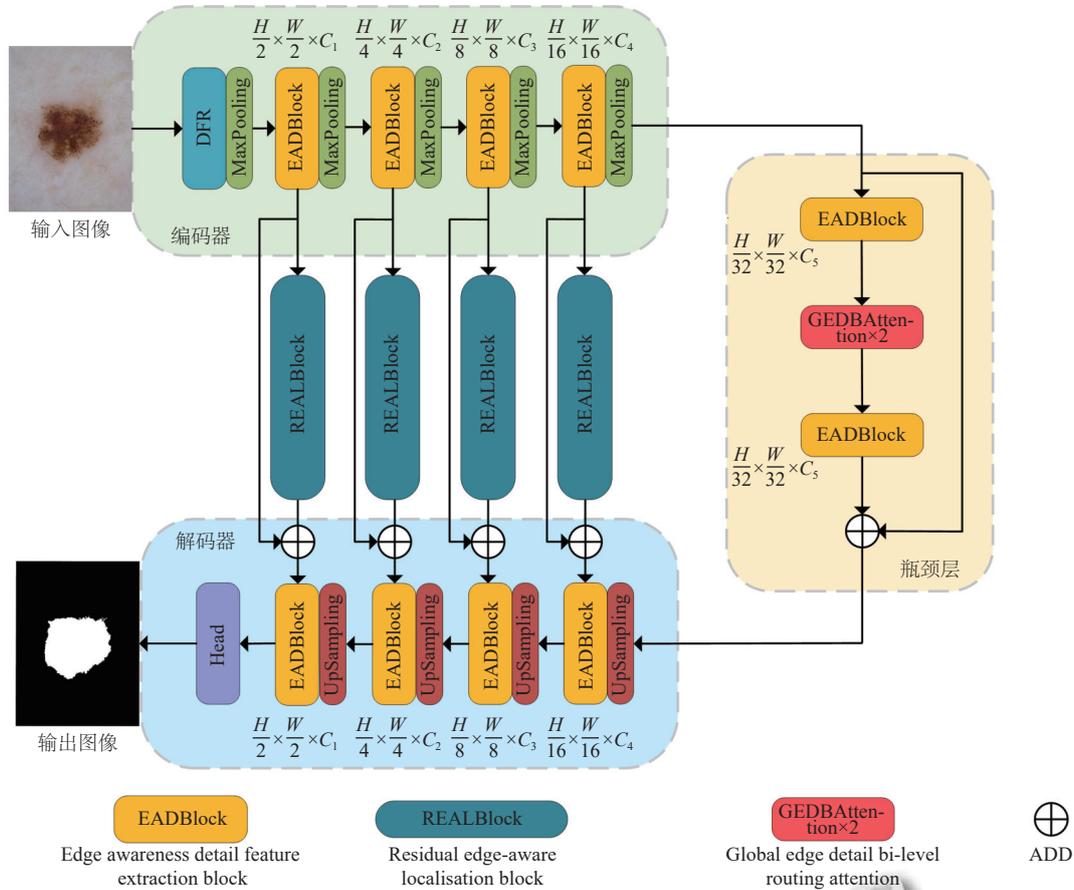


图1 本文分割算法

### 2.1 边缘感知细节特征提取模块

在医学图像分割中, 捕捉边界和梯度信息至关重要. 为了增强该能力, 本文引入了混合差分卷积, 并设计边缘感知细节特征提取模块. 如图2所示, 边缘感知细节特征提取模块核心是混合差分卷积 (mixed differential convolution, MDC), 它集成了标准卷积和4种像素差分卷积: 中心差分卷积 (central difference convolution, CDC)、自适应差分卷积 (adaptive difference convolution, ADC)、水平差分卷积 (horizontal difference convolution, HDC) 和垂直差分卷积 (vertical difference convolution, VDC). 这种设计改进了边缘和细节提取. 此外为减少并行运行4个卷积所增加的参数和推理时间, 本文采用权重融合策略. 这种方法整合了多个差分卷积的优势, 增强了模型捕捉各种边缘和梯度信息的能力. 具体来说, 4个差分卷积核相加形成一个等效卷积核. 给定输入特征图  $F_{IN}$ ,  $F_{IN}$  首先应用普通卷积, 然

后合并4个差分卷积核的输出, 构建最终核  $K_{SUM}$ . 然后使用混合差分卷积的权重和偏置计算输出  $F_{OUT}$ , 如式(1)和式(2)所示:

$$F_{OUT} = Conv(F_{IN}) \times K_{SUM} \quad (1)$$

$$K_{SUM} = \sum_{i=1}^4 K_i \quad (2)$$

### 2.2 细节特征保留

在深度特征图中, 小病变和模糊边界往往会在反复下采样过程中丢失, 从而降低分割精度. 为此, 本文提出了一种细节特征保留方法, 如图3所示. 原始图像输入首先经过MDC处理, 生成富含边缘细节的浅层特征图. 然后, 这些图进入具有扩展感受野卷积核大小为  $7 \times 7$  的深度卷积, 在此过程中, 使用  $1 \times 1$  的点式卷积整合通道信息, 每次操作后都会应用ReLU和批量归一化. 与传统的卷积堆叠不同, 深度卷积存在梯度消失的

风险并增加了模型的复杂性,而本文在深度卷积之前加入了残差连接模块,以防止梯度损失.通过将深度可分离卷积与残差连接相结合,该方法保留了边缘细节,减少了计算开销和内存需求,提高了模型效率和稳定性.

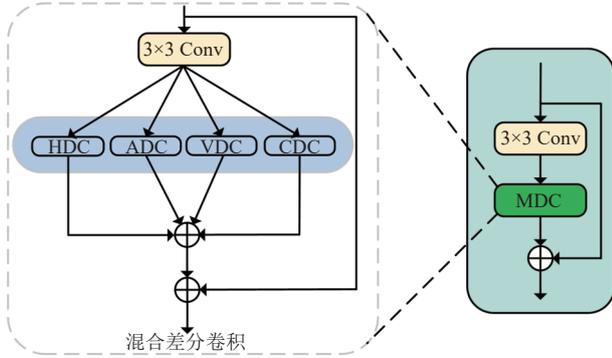


图2 边缘感知细节特征提取模块

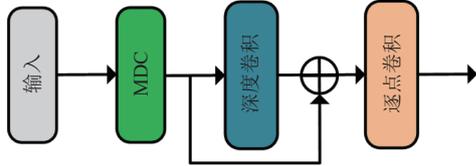


图3 细节特征保留

### 2.3 全局边缘细节双层级路由注意力机制

针对医学影像中病灶形态多样化的挑战,本文提出了一种高效的全局边缘细节双层级路由注意力机制

制,相较于传统视觉Transformer中全局自注意力机制的高计算复杂度问题,本方法通过动态路由策略实现稀疏化全局建模.如图4(a)所示,该模块由MDC与双路由注意力计算<sup>[17]</sup>两部分构成:首先通过具有残差校准的MDC强化深层语义特征的边缘-细节表征能力,随后在瓶颈层注入动态路由注意力进行全局上下文建模.

如图4(b)所示,双层级路由注意力是一种稀疏注意力机制,它利用自我注意和双路由机制动态选择最相关的相邻特征区域.其核心思想是将输入特征图划分为多个区域,每个区域自适应地选择最相关的相邻区域,同时过滤掉不相关的区域.具体而言,给定输入  $X \in R^{H \times W \times C}$ ,将其划分为  $S \times S$  个非重叠区域,每个区域包含  $\frac{HW}{S^2}$  个特征向量.然后将  $X$  重构为  $X^r \in R^{S^2 \times \frac{HW}{S^2} \times C}$ ,再根据式(3)计算线性映射:

$$Q = X^r \times W^Q, K = X^r \times W^K, V = X^r \times W^V \quad (3)$$

在区域级路由选择阶段,通过对区域平均特征  $Q^r$  和  $K^r \in R^{S^2 \times C}$ ,再计算  $Q$  和  $K^r$  的点积,得到矩阵  $A^r \in R^{S^2 \times S^2}$ ,其中  $A^r$  用来衡量区域的相关性,计算方法如式(4)所示:

$$A^r = Q^r (K^r)^T \quad (4)$$

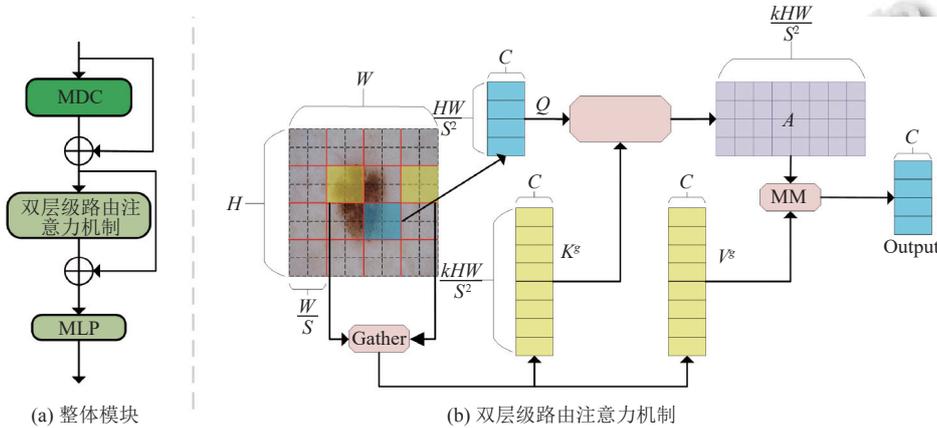


图4 全局边缘细节双层级路由注意力机制

接下来,在粗粒度级别对  $A^r$  进行滤波操作,动态筛选出每个区域前  $k$  个最相关邻域形成索引矩阵  $I^r \in N^{S^2 \times k}$ ,计算过程如式(5)所示:

$$I^r = \text{topkIndex}(A^r) \quad (5)$$

随后,依据路由索引  $I^r$ ,聚集相关键值特征  $K^s$  和  $V^s$ ,如式(6)和式(7):

$$K^s = \text{gather}(K, I^r) \quad (6)$$

$$V^s = \text{gather}(V, I^r) \quad (7)$$

最后,通过注意力权重计算实现特征校准,得到特征图输出,如式(8)所示:

$$\text{Output} = \text{Softmax}\left(\frac{Q \times (K^s)^T}{\sqrt{C}}\right) \times V^s \quad (8)$$

## 2.4 残差边缘感知定位模块

在医学图像分割任务中,模型需精准聚焦于病变区域,但病灶尺度差异显著、对比度低及边界模糊等问题严重阻碍了定位精度.现有方法常采用扩张卷积或 Transformer 架构以扩大感受野,但此类设计在增加参数量与计算复杂度的同时,易在特征图中引入高频噪声干扰.为此,本文提出残差边缘感知定位模块,其核心架构如图 5 所示,通过边缘梯度强化与注意力协同机制来优化模糊区域分割.残差边缘注意力增强模块首先通过 MDC 执行边缘细化,随后,通道注意力(channel attention, CA)模块通过通道维度重标定强化病灶敏感特征,而空间注意力(spatial attention, SA)模块则通过空间权重映射聚焦模糊区域的结构<sup>[18]</sup>.两模块输出的注意力特征图经点乘运算实现跨维度交互,显著提升弱对比区域的响应强度.最终通过 MDC 单元对增强后的特征进行边缘细节整合,并采用残差连接融合深层语义与浅层边缘信息.该设计通过差分卷积的定向边缘检测与注意力机制的特征筛选,在避免复杂计算的同时,显著提升了模型对模糊病变区域的边界定位能力.

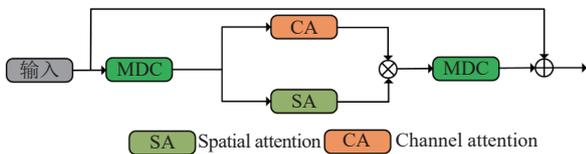


图 5 残差边缘感知定位模块

## 3 实验及结果分析

### 3.1 实验准备

#### 3.1.1 数据集

本文使用公开数据集 ISIC2018<sup>[19]</sup>、CVC-ClinicDB<sup>[20]</sup>和 BUSI<sup>[21]</sup>,并按照训练集、验证集、测试集比例为 7:1:2 进行划分.

#### 3.1.2 实验环境与参数设置

本文所有实验均在配备 Linux 操作系统的工作站上进行,采用 PyTorch 框架,其中 CPU 型号为 Intel Xeon(R) Platinum 8269CY,显卡型号为 NVIDIA GeForce GTX A5000,内存大小为 128 GB.

鉴于上述 3 个数据集在尺寸、分辨率和病变特征方面存在显著差异,为了使实验条件标准化,本文将输入图像的大小调整为 256×256,并使用数据增强技术增

强训练集,以提高分割性能.在训练过程中,采用 Adam 优化器,学习率为 0.001,训练迭代次数为 200 次,批量大小为 8.采用 BCE 和 Dice 损失相结合的方法进行训练,其中具体的损失函数定义如式 (9) 所示.其中,  $y^{\Delta}$  表示预测值,  $y$  表示目标值.

$$L = 0.5BCE(y^{\Delta}, y) + Dice(y^{\Delta}, y) \quad (9)$$

### 3.2 评价指标

本文采用骰子系数 (Dice similarity coefficient, Dice)、平均交并比 (mean intersection over union, mIoU)、准确度 (accuracy, ACC).

- **Dice**: 骰子系数,预测分割与真实分割区域空间重叠程度,对边缘模糊、小尺寸病灶的分割误差敏感,是评估肿瘤、息肉等不规则结构吻合度的核心指标,如式 (10) 所示:

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (10)$$

- **mIoU**: 平均交并比,计算病变组织与正常组织预测区域与真实标注的交并比均值,如式 (11) 所示:

$$mIoU = \frac{TP}{TP + FP + FN} \quad (11)$$

- **ACC**: 准确度,计算整个图像预测准确性,包括病变区域和正常区域,如式 (12) 所示:

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \quad (12)$$

其中,  $TP$  指实际属于病变的像素被正确识别为病变区域;  $TN$  表示实际为正常组织的像素被准确归类为背景;  $FN$  对应实际病变像素被误判为背景,反映漏分割风险;  $FP$  则为正常组织像素被错误标记为病变,提示过分割问题.

### 3.3 对比实验

为验证本文模型的有效性,将其与 9 种代表性方法进行比较,包括 U-Net<sup>[3]</sup>、UNet++<sup>[4]</sup>、DeepLabv3+<sup>[22]</sup>、UNeXt<sup>[5]</sup>、TransAttUnet<sup>[23]</sup>、SwinPA-Net<sup>[24]</sup>、DCSAU-Net、LM-Net<sup>[13]</sup>和 U-Net v2<sup>[15]</sup>.

如表 1 所示,在 ISIC2018 数据集上,本文与 9 种方法进行了比较.由于噪声过滤和模糊边界处理不足,U-Net 和 UNet++ 表现不佳. UNeXt 的参数量和计算复杂度较低,且使用了移位和标记化的 MLP 块,但由于缺乏全局特征建模,因此表现不佳. TransAttUnet 虽然通过变换器进行了全局建模,但在处理模糊的皮损边

界时遇到了困难,从而导致错误.使用 Swin Transformer 模块和 LPA 模块的 SwinPA-Net 实现了 88.94% 的 *mIoU* 和 94.18% 的 *Dice*. DCSAU-Net 的 *mIoU* 和 *Dice* 分别达到了 89.90% 和 94.68%, 这要归功于其更深的结构和紧凑的分割注意力模块. 本文的方法相比其他方法在 3 项指标中取得了最优结果, 同时保持了较低的数量和计算复杂度. 同时本文的 *ACC* 也达到了 98.84%, 优于所有其他方法. LM-Net 凭借轻量化设计实现 90.52% *mIoU* 与 94.31% *Dice*, 但多尺度特征融合灵活性有限, 在复杂皮损边界处理中存在细节丢失风险. U-Net v2 受限于传统跳跃连接的长距离依赖建模不足, *mIoU* 与 *Dice* 分别为 89.85%、94.02%, 在极低对比度区域易出现边界分割偏移, 两项方法指标均低于本文方法. 图 6 前两行显示了 ISIC2018 数据集上的

分割结果, 本文方法在提取边界变化明显和毛发遮挡区域的边缘细节方面更为有效, 即使在毛发遮挡的情况下也能准确识别皮损轮廓.

表 1 ISIC2018 数据集对比实验结果

方法	Params (M)	FLOPs (G)	<i>mIoU</i> (%)	<i>Dice</i> (%)	<i>ACC</i> (%)
U-Net <sup>[3]</sup>	31.04	41.91	87.33	93.13	97.42
UNet++ <sup>[4]</sup>	9.16	27.73	88.27	93.79	97.50
DeepLabv3+ <sup>[22]</sup>	54.70	15.95	88.22	93.79	97.46
UNeXt <sup>[5]</sup>	<b>1.48</b>	<b>0.44</b>	86.21	92.55	97.32
TransAttUnet <sup>[23]</sup>	22.65	67.97	82.55	90.57	96.51
SwinPA-Net <sup>[24]</sup>	102.59	23.06	88.94	94.18	97.56
DCSAU-Net <sup>[6]</sup>	<u>2.60</u>	5.29	89.90	<u>94.68</u>	97.77
LM-Net <sup>[13]</sup>	5.41	4.76	<u>90.52</u>	94.31	<u>98.23</u>
U-Net v2 <sup>[15]</sup>	26.03	5.42	89.85	94.02	98.01
Ours	6.26	2.61	<b>91.39</b>	<b>95.45</b>	<b>98.84</b>

注: 加粗字体表示最优结果, 下划线表示次优结果

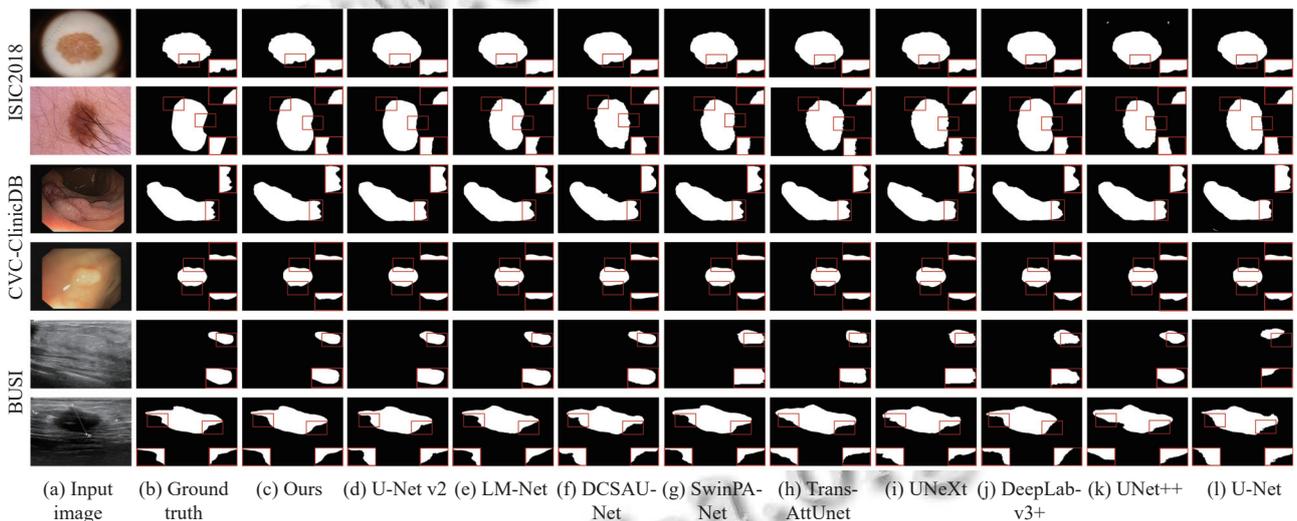


图 6 本文方法与其他 9 种方法在 ISIC2018、CVC-ClinicDB、BUSI 上的可视化结果

在 CVC-ClinicDB 数据集上的实验结果如表 2 所示, 本文方法展现出显著优势. 与次优模型 DCSAU-Net 相比, 本方法在 *mIoU*、*Dice* 和 *ACC* 指标上分别达到 94.43%、97.12% 和 99.38%, 这 3 项指标均居首位. 这表明所提方法对边缘模糊区域具有较强鲁棒性, 且在小型息肉及复杂形态病灶分割任务中表现突出. 本文方法能精准聚焦目标区域并有效抑制背景噪声干扰, 显著降低误分割风险.

图 6 第 3、4 行展示了 CVC-ClinicDB 数据集的可视化结果, 可见本文方法在边缘细节模糊区域仍能保持完整分割效果, 而其他方法在同类场景下则出现分割断裂或区域缺失现象. 特别值得注意的是, 本文方法

对微小病灶及形态复杂的病变区域展现出优异的形态保持能力.

在 BUSI 数据集上的对比实验结果如表 3 所示, 本文方法在乳腺超声图像分割任务中展现出显著优势. 与次优模型 LM-Net 相比, 本文方法的 3 项指标分别提升至 85.94%、91.04%、98.53%, 在灰度对比度低、边界模糊且噪声密集的超声模态下实现了更高的分割精度. 这表明模型通过边缘感知细节提取与全局细节双层级注意力机制, 有效增强了对低信噪比图像中模糊病灶的边界捕获能力, 尤其在微小乳腺病变及腺体结构复杂区域表现出更强的特征区分度. 图 6 第 5、6 行显示了 BUSI 数据集上的可视化结果, 在因噪声干扰出

现分割断裂的场景中, 本文方法仍能保持连续的病变轮廓和清晰的区域划分, 显著抑制了背景噪声对分割结果的干扰. 对于超声图像中常见的低对比度微小钙化灶, 本文方法实现了病灶形态的精准还原, 验证了其在高挑战性医学影像任务中的泛化能力.

表2 CVC-ClinicDB 数据集对比试验结果 (%)

方法	<i>mIoU</i>	<i>Dice</i>	<i>ACC</i>
U-Net <sup>[3]</sup>	92.16	95.79	99.13
UNet++ <sup>[4]</sup>	92.29	96.01	99.21
DeepLabv3+ <sup>[22]</sup>	92.72	96.16	99.24
UNeXt <sup>[5]</sup>	92.22	95.92	99.19
TransAttUnet <sup>[23]</sup>	92.94	96.27	99.24
SwinPA-Net <sup>[24]</sup>	93.63	96.71	99.29
DCSAU-Net <sup>[6]</sup>	93.83	96.79	99.31
LM-Net <sup>[13]</sup>	<u>93.90</u>	<u>96.85</u>	<u>99.32</u>
U-Net v2 <sup>[15]</sup>	93.50	96.60	99.30
Ours	<b>94.43</b>	<b>97.12</b>	<b>99.38</b>

表3 BUSI 数据集对比试验结果 (%)

方法	<i>mIoU</i>	<i>Dice</i>	<i>ACC</i>
U-Net <sup>[3]</sup>	79.06	87.82	97.73
UNet++ <sup>[4]</sup>	80.52	88.51	97.85
DeepLabv3+ <sup>[22]</sup>	81.50	88.23	97.96
UNeXt <sup>[5]</sup>	82.71	89.17	98.25
TransAttUnet <sup>[23]</sup>	82.51	90.14	98.06
SwinPA-Net <sup>[24]</sup>	83.02	89.51	98.07
DCSAU-Net <sup>[6]</sup>	84.50	90.30	98.39
LM-Net <sup>[13]</sup>	<u>85.21</u>	<u>90.83</u>	<u>98.45</u>
U-Net v2 <sup>[15]</sup>	84.89	90.62	98.40
Ours	<b>85.94</b>	<b>91.04</b>	<b>98.53</b>

### 3.4 消融实验

为了评估本文方法的特征提取能力及其模块的有效性, 在 CVC-ClinicDB 数据集上进行了消融实验, 使用 *mIoU* 和 *Dice* 作为指标, 并以 U-Net 作为基线. 如表 4 所示. 模型 2 将 U-Net 编码器和解码器的卷积结构替换为 EADBlock, *mIoU* 和 *Dice* 分别提升至 93.29% 和 96.41%. 模型 3 将 U-Net 的第 1 个编码器替换为 DFR 块, 通过 DFR 块的特征保留能力, *mIoU* 和 *Dice* 均有所提升. 为进一步验证关键模块的有效性, 在模型 4 和模型 5 中分别将 GEDBAttention 和 REALBlock 模块加入 U-Net. 实验表明, *Dice* 系数分别提升 0.21% 和 0.45%. 融入 GEDBAttention 后, 模型的全局建模与边缘细节捕捉能力增强, 使模型能充分捕获全息边缘信息. 融入 REALBlock 后, 模型在病灶区域定位方面表现更优, 尤其在形态不规则、边界模糊的病灶区域. 上

述结果全面验证了 GEDBAttention 的全局边缘建模能力与 REALBlock 的边界感知定位增强. 这种增强不仅提升了模型分割精度, 还强化了模型在复杂场景中对外边缘区域的感知能力.

表4 本文所提出方法的消融实验 (%)

模型	消融类型				结果	
	EADBlock	DFR	GEDBAttention	REALBlock	<i>mIoU</i>	<i>Dice</i>
1 (U-Net)	—	—	—	—	92.16	95.79
2	√	—	—	—	93.29	96.41
3	—	√	—	—	92.42	96.12
4	—	—	√	—	92.27	96.00
5	—	—	—	√	92.65	96.24
6	√	√	—	—	93.58	96.68
7	√	√	√	—	93.63	96.71
8	√	√	—	√	94.08	96.94
9 (Ours)	√	√	√	√	<b>94.43</b>	<b>97.12</b>

本文模型通过集成 DFR、EADBlock、GEDBAttention 和 REALBlock (模型 6–8) 验证了多模块的协同效应, 实验显示模型 8 的 *mIoU* 和 *Dice* 较原始 U-Net 分别提升 1.92% 和 1.15%. 最终模型 9 通过全模块集成实现最优性能, 在 *mIoU*、*Dice* 等指标上达到峰值. 得益于模块间的协同机制, 网络同时强化全局特征建模与局部边缘细化能力, 显著提升对模糊边界目标的处理鲁棒性, 验证了本文方法在分割精度上的综合优势.

## 4 结论与展望

本文提出了一种高效的医学图像分割网络, 旨在解决医学图像中存在的模糊边界问题. 该网络在编码器-解码器结构中集成了边缘感知细节提取模块, 将残差边缘感知定位模块与全局边缘细节双层级路由注意力机制相结合. 这种设计能有效聚合上下文语义和局部边界信息, 解决病变类型不同、对比度低和病变形状复杂等难题. 在 ISIC2018、CVC-ClinicDB 和 BUSI 数据集上进行的大量实验证明了该方法在分割性能和特征学习方面的显著优势. 总体而言, 这一方法为医学影像分割问题提供了一种极具前景的解决方案, 对推动医学影像分析技术的发展具有重要意义.

### 参考文献

- Ma JY, Tang LF, Fan F, et al. SwinFusion: Cross-domain long-range learning for general image fusion via Swin Transformer. IEEE/CAA Journal of Automatica Sinica, 2022, 9(7): 1200–1217. [doi: 10.1109/JAS.2022.105686]
- Paluru N, Dayal A, Jenssen HB, et al. Anam-Net: Anamorphic depth embedding-based lightweight CNN for segmentation of anomalies in COVID-19 chest CT images.

- IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(3): 932–946. [doi: [10.1109/TNNLS.2021.3054746](https://doi.org/10.1109/TNNLS.2021.3054746)]
- 3 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention. Munich: Springer, 2015. 234–241. [doi: [10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)]
- 4 Zhou ZW, Rahman Siddiquee MM, Tajbakhsh N, *et al.* UNet++: A nested U-Net architecture for medical image segmentation. Proceedings of the 4th International Workshop on Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada: Springer, 2018. 3–11. [doi: [10.1007/978-3-030-00889-5\\_1](https://doi.org/10.1007/978-3-030-00889-5_1)]
- 5 Valanarasu MJM, Patel VM. UNeXt: MLP-based rapid medical image segmentation network. Proceedings of the 25th International Conference on Medical Image Computing and Computer-assisted Intervention. Singapore: Springer, 2022. 23–33. [doi: [10.1007/978-3-031-16443-9\\_3](https://doi.org/10.1007/978-3-031-16443-9_3)]
- 6 Xu Q, Ma ZC, He N, *et al.* DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation. Computers in Biology and Medicine, 2023, 154: 106626. [doi: [10.1016/j.combiomed.2023.106626](https://doi.org/10.1016/j.combiomed.2023.106626)]
- 7 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010. [doi: [10.5555/3295222.3295349](https://doi.org/10.5555/3295222.3295349)]
- 8 雍俊杰. 基于 TransUnet 中跳跃连接层多角度特征加强的医学图像分割网络 [硕士学位论文]. 成都: 四川师范大学, 2024. [doi: [10.27347/d.cnki.gssdu.2024.001538](https://doi.org/10.27347/d.cnki.gssdu.2024.001538)]
- 9 Xu GP, Zhang X, He XW, *et al.* LeViT-UNet: Make faster encoders with Transformer for medical image segmentation. Proceedings of the 6th Chinese Conference on Pattern Recognition and Computer Vision. Xiamen: Springer, 2023. 42–53. [doi: [10.1007/978-981-99-8543-2\\_4](https://doi.org/10.1007/978-981-99-8543-2_4)]
- 10 陶惜婷, 叶青. 融合 CNN 和 Transformer 的并行双分支皮肤病灶图像分割. 计算机应用研究, 2024, 41(8): 2554–2560. [doi: [10.19734/j.issn.1001-3695.2023.10.0600](https://doi.org/10.19734/j.issn.1001-3695.2023.10.0600)]
- 11 Wang JC, Chen F, Ma YX, *et al.* XBound-Former: Toward cross-scale boundary modeling in Transformers. IEEE Transactions on Medical Imaging, 2023, 42(6): 1735–1745. [doi: [10.1109/TMI.2023.3236037](https://doi.org/10.1109/TMI.2023.3236037)]
- 12 Xia Y, Yun HJ, Liu YJ, *et al.* MGCBFormer: The multiscale grid-prior and class-inter boundary-aware Transformer for polyp segmentation. Computers in Biology and Medicine, 2023, 167: 107600. [doi: [10.1016/j.combiomed.2023.107600](https://doi.org/10.1016/j.combiomed.2023.107600)]
- 13 Lu ZK, She CY, Wang W, *et al.* LM-Net: A light-weight and multi-scale network for medical image segmentation. Computers in Biology and Medicine, 2024, 168: 107717. [doi: [10.1016/j.combiomed.2023.107717](https://doi.org/10.1016/j.combiomed.2023.107717)]
- 14 王朝欣, 杨汶汶, 戎泽, 等. 基于 MSC-LSAM 的多尺度交叉超声医学图像分割方法. 数据采集与处理, 2025, 40(2): 469–484. [doi: [10.16337/j.1004-9037.2025.02.015](https://doi.org/10.16337/j.1004-9037.2025.02.015)]
- 15 Peng YP, Chen DZ, Sonka M. U-Net v2: Rethinking the skip connections of U-Net for medical image segmentation. Proceedings of the 22nd IEEE International Symposium on Biomedical Imaging (ISBI). Houston: IEEE, 2025. 1–5. [doi: [10.1109/ISBI60581.2025.10980742](https://doi.org/10.1109/ISBI60581.2025.10980742)]
- 16 王骁崑, 邢树礼, 毛国君. 多尺度特征融合的轻量化 Transformer 医学图像分割研究. 中国生物医学工程学报, 2025, 44(2): 165–173. [doi: [10.3969/j.issn.0258-8021.2025.02.004](https://doi.org/10.3969/j.issn.0258-8021.2025.02.004)]
- 17 Zhu L, Wang XJ, Ke ZH, *et al.* BiFormer: Vision Transformer with bi-level routing attention. Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023. 10323–10333. [doi: [10.1109/CVPR52729.2023.00995](https://doi.org/10.1109/CVPR52729.2023.00995)]
- 18 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich: Springer, 2018. 3–19. [doi: [10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1)]
- 19 Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data, 2018, 5: 180161. [doi: [10.1038/sdata.2018.161](https://doi.org/10.1038/sdata.2018.161)]
- 20 Bernal J, Sánchez FJ, Fernández-Esparrach G, *et al.* WMDOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. Computerized Medical Imaging and Graphics, 2015, 43: 99–111. [doi: [10.1016/j.compmimag.2015.02.007](https://doi.org/10.1016/j.compmimag.2015.02.007)]
- 21 Al-Dhabyani W, Gomaa M, Khaled H, *et al.* Dataset of breast ultrasound images. Data in Brief, 2020, 28: 104863. [doi: [10.1016/j.dib.2019.104863](https://doi.org/10.1016/j.dib.2019.104863)]
- 22 Chen LC, Zhu YK, Papandreou G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich: Springer, 2018. 833–851. [doi: [10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)]
- 23 Chen BZ, Liu YS, Zhang Z, *et al.* TransAttUnet: Multi-level attention-guided U-Net with Transformer for medical image segmentation. IEEE Transactions on Emerging Topics in Computational Intelligence, 2024, 8(1): 55–68. [doi: [10.1109/TETCI.2023.3309626](https://doi.org/10.1109/TETCI.2023.3309626)]
- 24 Du H, Wang JZ, Liu M, *et al.* SwinPA-Net: Swin Transformer-based multiscale feature pyramid aggregation network for medical image segmentation. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(4): 5355–5366. [doi: [10.1109/TNNLS.2022.3204090](https://doi.org/10.1109/TNNLS.2022.3204090)]

(校对责编: 王欣欣)