

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-17

论文引用格式: Liu Yang, Qiu Yunfei, Han Xuefeng. A skin lesion segmentation network with lightweight attention-based feature fusion and boundary-aware constraints[J/OL]. Journal of Image and Graphics, XXXX: 1-17. DOI: 10.11834/jig.260295. (刘扬, 邱云飞, 韩雪峰. 轻量注意力特征融合与边界感知约束的皮肤病灶分割网络[J/OL]. 中国图象图形学报, XXXX: 1-17. DOI: 10.11834/jig.260295.) [DOI: 10.11834/jig.260295]

轻量注意力特征融合与边界感知约束的皮肤病灶分割网络

刘扬¹, 邱云飞¹, 韩雪峰^{2*}

1. 辽宁工程技术大学 软件学院, 辽宁 葫芦岛 123105; 2. 辽宁工程技术大学 鄂尔多斯研究院, 内蒙古自治区 鄂尔多斯 017004

摘要: 目的 面向移动医疗与边缘部署场景下皮肤病灶分割对低运算量与高可靠性的双重需求, 本文聚焦轻量网络在复杂皮肤镜图像中因浅层噪声传递、跨层融合失衡和边界约束不足而导致的结构感知与边界定位退化问题, 提出一种选择性传递、自适应融合和聚焦边界监督的高效分割 (selective transmission, adaptive fusion, and focused boundary supervision for efficient segmentation, SAFE) 范式, 并在轻量主干网络上实例化为轻量注意力特征融合与边界感知约束网络 (lightweight attention-based feature fusion and boundary-aware constraints network, LEAF-UNet)。方法 首先, 在编码器与解码器间的跳跃连接路径入口引入轻量通道重标定机制 (lightweight attention skip, LAS), 利用通道统计信息对进入解码端前的浅层特征进行选择筛选, 从源头抑制毛发、标尺等背景噪声的传播; 其次, 在跨层特征聚合节点设计门控加权融合机制 (gated weighted fusion, GWF), 通过自适应权重分配机制动态协调低层纹理与高层语义的贡献度, 以缓解固定融合策略下的语义细节失衡问题; 最后, 在训练过程中引入边界一致性监督 (boundary consistency loss, BCL), 其利用拉普拉斯算子提取预测掩膜与真实标注的轮廓响应, 并通过显式几何一致性约束强化边界区域学习, 改善低对比度和模糊边界条件下的病灶轮廓刻画能力。结果 在国际皮肤影像协作组织 (International Skin Imaging Collaboration, ISIC) 2017 和 2018 挑战赛数据集 (ISIC2017, ISIC2018) 以及 PH² 数据集上, LEAF-UNet 以 0.053 M 参数量和 0.075 十亿浮点运算 (giga floating-point operations, GFLOPs) 分别取得 89.50%、90.32% 和 94.83% 的 Dice 系数, 并较高效组增强 U 形网络 (efficient group enhanced U-Net, EGE-UNet) 显著降低第 95 百分位豪斯多夫距离, 最大相对降幅约为 30.28%。结论 LEAF-UNet 在保持极低模型复杂度的前提下, 有效缓解了轻量模型在复杂场景下的结构感知退化问题, 为资源受限环境下的皮肤病灶分割提供了具有参考价值的轻量化方案。本研究的数据集与源代码已发布于 Science Data Bank, 访问链接为: <https://doi.org/10.57760/sciencedb.j00240.00278>。

关键词: 医学图像分割; 皮肤病灶分割; 轻量化网络; 注意力特征融合; 边界感知损失

A skin lesion segmentation network with lightweight attention-based feature fusion and boundary-aware constraints

收稿日期: 2026-05-24; 修回日期: 2026-08-12

* 通信作者: 韩雪峰, Email: zyd020311@outlook.com

基金项目: 2024 年度辽宁省教育厅基本科研项目 (自主选题重点项目) 典型液压泵系统维护建模与能耗优化研究 (项目编号: LJ212410147079); 鄂尔多斯市科技“突围”工程“揭榜挂帅”重大项目 (项目编号: JB20251447); 鄂尔多斯市标志性创新团队项目 (项目编号: TD20240005); 鄂尔多斯市重点研发计划项目 (项目编号: YF20250274)。

Supported by: Basic Scientific Research Project of Liaoning Provincial Department of Education (Key Independent Research Project), 2024: Research on Maintenance Modeling and Energy Consumption Optimization of Typical Hydraulic Pump Systems (No. LJ212410147079); Ordos Science & Technology "Breakthrough" Program - Open Competition and Top-Task Major Project (Grant No. JB20251447); Ordos Flagship Innovation Team Project (Grant No. TD20240005); Ordos Key Research and Development Program (Grant No. YF20250274).

Liu Yang¹, Qiu Yunfei¹, Han Xuefeng^{2*}*1. School of Software, Liaoning Technical University, Huludao, 123105, China; 2. Erdos Research Institute, Liaoning Technical University, Erdos 017004, China*

Abstract: Objective Skin lesion segmentation is a fundamental task in computer-aided dermatological analysis, supporting lesion measurement, morphological assessment, risk evaluation, and subsequent clinical decision-making. With the growing demand for mobile healthcare, point-of-care screening, and edge intelligence, segmentation models are expected to achieve high accuracy and reliable boundary delineation under strict constraints on parameters, computation, and memory consumption. However, dermoscopic images often contain hair occlusion, ruler marks, illumination variation, color inhomogeneity, low contrast, and blurred lesion boundaries. These factors are particularly challenging for lightweight networks, whose limited representational capacity makes them vulnerable to shallow-feature noise propagation, imbalanced cross-layer feature fusion, and insufficient boundary supervision. Existing lightweight segmentation methods usually emphasize backbone compression or a single enhancement module. They rarely consider how noisy shallow features, mismatched semantic-detail fusion, and weak contour constraints may form a continuous degradation process. To address this issue, this study proposes a lightweight skin lesion segmentation framework based on selective transmission, adaptive fusion, and focused boundary supervision for efficient segmentation (SAFE), which is instantiated as LEAF-UNet on a lightweight encoder-decoder backbone. **Methods** The proposed method follows a coordinated “transmission-fusion-supervision” design. Instead of increasing network depth or introducing computationally expensive global modeling modules, it performs targeted refinement at three critical information-flow stages in a lightweight U-shaped segmentation architecture. First, a lightweight attention skip (LAS) mechanism is inserted at the entrance of each skip connection between the encoder and decoder. In conventional encoder-decoder networks, shallow features preserve fine spatial details but may also transmit background interference directly to the decoder. This issue is prominent in dermoscopic images, where hair, ruler marks, specular reflections, and texture variations may be captured as strong shallow responses. LAS employs channel-wise global statistical information to estimate the relevance of each channel before shallow features are delivered to the decoder. Through lightweight channel recalibration, informative channels associated with lesion structures are retained, while redundant or noise-dominated channels are suppressed. Therefore, the decoder receives more discriminative shallow representations without changing the spatial resolution of the feature maps or introducing substantial additional computation. Second, a gated weighted fusion (GWF) mechanism is designed for cross-layer feature aggregation. Traditional feature fusion strategies, such as direct concatenation or element-wise addition, assume that shallow texture features and deep semantic features make fixed contributions across different images and spatial locations. This assumption is unsuitable for skin lesions with diverse scales, irregular shapes, low contrast, and ambiguous boundaries. GWF first aligns the shallow and deep features in spatial size and channel dimension. It then generates complementary, normalized gating weights to adaptively determine the contributions of low-level details and high-level semantic information at each spatial location. In regions with strong background texture or irrelevant local structures, the model can increase the contribution of deep semantic features to suppress false responses. In boundary regions or areas with subtle lesion details, the model can preserve more shallow information to avoid detail loss. This adaptive fusion strategy improves the balance between semantic consistency and boundary detail recovery. Third, a boundary consistency loss (BCL) is introduced during training to supplement region-based supervision. Binary cross-entropy loss and Dice loss mainly optimize pixel-wise classification and region overlap, but they may not directly constrain geometric consistency between predicted boundaries and ground-truth contours. BCL uses a fixed Laplacian operator to extract contour responses from the predicted probability map and the ground-truth mask. The difference between these contour responses is incorporated into the training objective as an explicit boundary constraint. This mechanism encourages the network to focus on boundary transition regions and improves contour delineation under low-contrast and blurred-boundary conditions. Importantly, BCL is used only during training. It introduces no additional inference branch, learnable parameters, post-processing operation, or deployment-time computational burden. The three components are designed to be complementary. LAS reduces the propagation of irrelevant shallow noise before feature fusion; GWF adaptively integrates filtered shallow details with deep semantic features; and BCL constrains

the final prediction from the perspective of boundary geometry. Consequently, LEAF-UNet improves segmentation reliability through a lightweight closed-loop design rather than through a substantial increase in model capacity. **Results** Experiments were conducted on the International Skin Imaging Collaboration 2017 and 2018 challenge datasets (ISIC2017 and ISIC2018) and the PH² dermoscopic image dataset. The performance of LEAF-UNet was evaluated using the mean intersection over union, Dice similarity coefficient, and 95th-percentile Hausdorff distance (HD95). Model efficiency was assessed using parameter count and giga floating-point operations (GFLOPs). LEAF-UNet contains only 0.053×10^6 parameters and requires 0.075 GFLOPs, demonstrating its suitability for computation-constrained deployment scenarios. Despite its extremely small model size, LEAF-UNet achieved Dice coefficients of 89.50%, 90.32%, and 94.83% on ISIC2017, ISIC2018, and PH², respectively. Its corresponding mean intersection over union values were 81.26%, 81.76%, and 90.45%. These results demonstrate that the proposed network maintains strong region-overlap performance across datasets with different sample scales, lesion characteristics, and imaging conditions. In comparison with the efficient group-enhanced U-Net (EGE-UNet), which is used as the lightweight backbone baseline, LEAF-UNet improves both overlap-based and boundary-sensitive performance while retaining a comparable parameter count. In particular, the proposed method substantially reduces HD95, with a maximum relative reduction of approximately 30.28%. The improvement is especially evident on the PH² dataset, where LEAF-UNet obtains an HD95 of 9.51 pixels, indicating more accurate and stable lesion contour localization. Ablation experiments further verify the contribution of each component. Introducing LAS alone improves robustness against shallow background interference by filtering skip-connection features before decoding. GWF alone improves the coordination between detailed texture information and high-level semantic representations. BCL alone enhances contour learning by imposing a direct geometric consistency constraint. When the three components are combined, the complete LEAF-UNet obtains the best overall performance, confirming that selective feature transmission, adaptive cross-layer-fusion, and training-time boundary supervision are mutually beneficial. Additional analyses of hyperparameter sensitivity, neighboring-module replacement, and per-case HD95 distributions show that the proposed design provides stable gains without relying on heavy attention modules, complex post-processing, or additional inference-stage branches. **Conclusion** LEAF-UNet is a lightweight skin lesion segmentation network that addresses structural perception degradation in resource-constrained settings through coordinated feature transmission, feature fusion, and boundary supervision. The LAS mechanism suppresses shallow noise propagation, GWF adaptively balances low-level details and high-level semantics, and BCL improves boundary consistency without increasing inference complexity. Experimental results on ISIC2017, ISIC2018, and PH² demonstrate that LEAF-UNet achieves competitive segmentation accuracy and substantially improved boundary localization with only 0.053×10^6 parameters and 0.075 GFLOPs. The proposed framework provides an effective and deployment-friendly solution for skin lesion segmentation in mobile healthcare and edge computing environments. The dataset and source code developed in this study have been released through Science Data Bank and are available at: <https://doi.org/10.57760/sciencedb.j00240.00278>.

Key words: Medical image segmentation; skin lesion segmentation; lightweight network; attention-based feature fusion; boundary-aware loss

0 引言

医学图像分割是智能辅助诊疗中的基础任务,其目标是从复杂医学影像中准确定位目标结构,为病灶测量、形态分析、风险评估及临床决策提供支撑(Asgari Taghanaki等,2021)。近年来,基于深度学习的方法凭借优越的精度与鲁棒性,已逐步取代依赖手工特征的传统方案,并沿卷积神经网络(convolutional neural networks, CNN)、视觉 Transformer 与视

觉 Mamba 等结构主线持续演进,成为该领域的主流技术(石军等,2025)。其中,皮肤病灶分割直接关系到病灶面积估计、形态对称性判断和边界不规则性评估,其轮廓提取精度会影响后续诊断分析的可靠性(Mirikharaji等,2023)。随着移动医疗、床旁筛查和边缘智能设备的发展,分割模型逐渐从高性能服务器转向资源受限终端,因而需要在低参数量、低计算开销和高鲁棒性之间取得平衡。然而,皮肤镜图像常受毛发遮挡、反光伪影、色素不均、尺度变化和边界模糊等因素干扰,使极轻量模型更易丧失对病

灶结构与轮廓的刻画能力。因此,如何在极低部署成本下保持稳定的结构感知与边界定位能力,成为资源受限皮肤病灶分割的关键问题。

围绕这一问题,现有研究主要从轻量主干设计、跨层特征交互以及边界辅助建模等方向展开探索。Ruan 等人(2023)针对资源受限皮肤病灶分割提出 EGE-UNet,通过高效组增强机制在极低参数规模下取得了较强性能,验证了超轻量分割的可行性。然而,该类方法的设计重点主要在于轻量主干压缩,对浅层噪声传递、跨层融合失衡和边界约束不足之间的连续影响尚未进行显式建模。Munia 等人(2025)提出 Attention-guided Hierarchical Fusion U-Net,通过分层注意力融合增强跨层信息交互;Zhang 等人(2024)在 VM-UNetV2 中通过语义—细节注入和更强的上下文建模改善多尺度特征协同,这类方法能够提升跨层表达能力,但通常更关注融合阶段的特征增强,对浅层特征进入解码端之前的噪声筛选,以及融合阶段低层细节与高层语义的责任分配,仍缺少进一步区分。Xu 等人(2024)进一步提出 LB-UNet,将边界辅助机制引入轻量 U 形网络,以增强模糊轮廓恢复能力,但边界建模更多作为局部分支或辅助约束使用,其与前端特征筛选、跨层融合权重分配之间的协同关系仍有进一步探索空间。总体来看,现有方法虽然分别推动了轻量分割、跨层建模和边界感知的发展,但多数工作仍主要面向单一环节进行增强,或仅在局部耦合两个环节。对于极轻量皮肤病灶分割而言,浅层噪声传递、跨层融合失衡和边界约束不足往往连续发生并相互放大,因此仍缺少一种从资源受限失效机理出发,对特征传递、跨层融合与边界监督进行统一建模的轻量闭环设计。

本文认为,资源受限轻量网络的性能下降并非单纯由容量不足造成,而是源于一条连续失效链:浅层特征在跳跃连接传递中携带背景噪声,受污染的细节特征随后进入跨层融合并干扰语义—纹理协同,而常规区域型损失又难以在训练阶段对模糊边界施加直接约束。基于此,本文提出“选择性传递、自适应融合、聚焦边界监督的高效分割”(selective transmission, adaptive fusion, and focused boundary supervision for efficient segmentation, SAFE)统一范式。SAFE 围绕“传递—融合—监督”形成闭环:首先在跳跃连接入口筛选浅层特征,减少噪声进入解码端;随后依据输入内容动态分配低层细节与高层语

义的融合权重;最后通过训练期边界一致性监督反向约束前两阶段学习,使模型在不增加推理分支的条件下获得更稳定的结构感知与边界定位能力。综上,本文的主要贡献可概括如下。

1) 提出 SAFE 统一范式,将资源受限皮肤病灶分割中的结构感知退化建模为“特征传递污染—跨层融合失衡—边界约束不足”的连续失效链,并构建面向该失效链的轻量闭环修复框架。

2) 在极轻量 U 形主干上构建 LEAF-UNet,通过轻量通道重标定(lightweight attention skip, LAS)与门控加权融合机制(gated weighted fusion, GWF),分别实现跳跃连接处的选择性特征传递和跨层特征的自适应加权融合,以提高复杂背景下的有效特征表达能力。

3) 引入无推理开销的边界一致性监督(boundary consistency loss, BCL),通过固定拉普拉斯算子约束预测掩膜与真实标注之间的轮廓响应一致性,在不增加推理阶段参数量和计算量的前提下增强模型对模糊边界的刻画能力;并在国际皮肤影像协作组织(International Skin Imaging Collaboration, ISIC) 2017 和 2018 挑战赛数据集(ISIC2017、ISIC2018)和 PH²数据集上验证了所提方法的有效性。

1 相关工作

1.1 轻量化与高效医学图像分割网络

随着医学图像分割模型不断追求更高精度,参数量、计算量和推理开销持续增加,限制了其在移动医疗、床旁筛查和边缘设备中的应用(Wang 等, 2022; Qureshi 等, 2023)。针对这一问题, MALUNet(Ruan 等, 2022)、UNeXt(Valanarasu 等, 2022)和 EGE-UNet(Ruan 等, 2023)等轻量方法通过主干压缩、多注意力结构或高效组增强机制降低模型规模,验证了轻量医学分割的可行性;UNETR++(Shaker 等, 2024)、CSWin-UNet(Liu 等, 2025)和 VM-UNetV2(Zhang 等, 2024)等方法则通过 Transformer、状态空间模型或混合建模增强上下文表达。然而,现有研究多集中于主干设计或全局建模。对于极轻量皮肤病灶分割而言,模型退化不仅源于容量受限,还涉及浅层噪声传递、跨层融合失衡和边界约束不足等信息流问题。

1.2 跨层信息传递与多尺度特征融合

跳跃连接是U形网络恢复空间细节的关键机制,但其直接传递的浅层特征也可能携带毛发、反光和色素不均等背景干扰。围绕跨层交互, Attention U-Net (Oktay 等, 2018)、AHF-U-Net (Munia 等, 2025)、VM-UNetV2 和 LB-UNet (Xu 等, 2024) 等方法分别通过注意力门控、分层融合、语义—细节注入或边界辅助信息增强多尺度特征协同。这些方法证明了跨层交互的重要性;例如,李赫等(2024)通过双编码器全局—局部交叉注意力关联局部纹理与全局上下文,并在跳跃连接中引入特征自适应模块以缓解编解码特征差距。但多数方法仍将“传递”和“融合”作为整体处理,较少区分浅层特征进入解码端前的噪声筛选,以及融合阶段低层细节与高层语义的贡献分配。现有边界方法主要包括结构型边界增强和损失型边界约束:前者通过边界分支、细化模块、注意力机制或频域建模显式强化轮廓表达,如 LB-UNet、WA-NET (Hu 等, 2025)、MPBA-Net (Huang 等, 2025) 以及基于边缘信息增强的 U-Net 改进模型 (张蝶等, 2024), 但可能增加推理结构;对于容量有限的轻量模型,受污染的浅层响应若直接进入解码端,会增加后续融合与优化负担。因此,本文将跨层信息流拆分为选择性传递和适应性融合两个环节:前者抑制噪声传播,后者动态协调细节与语义贡献。已有 ECA-Net (Wang 等, 2020) 等通道注意力和门控融合方法为低成本特征选择提供了基础,但如何将其部署到轻量皮肤病灶分割的关键信息流节点,仍有进一步研究空间。

1.3 边界感知分割与轮廓监督

除区域重叠精度外,边界定位质量也是皮肤病灶分割的重要指标。皮肤镜图像中的低对比度、边缘渐变和轮廓不规则等问题,容易使模型在边界附近产生外扩、毛刺或断裂 (Mirikharaji 等, 2023)。现有边界方法主要包括结构型边界增强和损失型边界约束:前者通过边界分支、细化模块或频域建模显式强化轮廓表达,如 LB-UNet、WA-NET (Hu 等, 2025)、MPBA-Net (Huang 等, 2025), 以及在 Swin Transformer 解码端引入多尺度边缘密集模块与基于 Sobel 算子的级联边缘注意力的 Swin-EdgeNet (刘善洁和童孟军, 2026), 但这类设计通常会增加推理阶段的结构与计算开销;后者通过边界损失或轮廓一致性施加几何约束,如 Boundary Loss (Kervadec 等,

2019), 通常不改变推理阶段网络。已有边界损失多作为独立优化项使用,较少与特征筛选和跨层融合机制协同建模,因此对极轻量模型中由浅层噪声传播和融合失衡共同引起的边界偏差,其针对性仍有进一步提升空间。

1.4 小结与本文定位

综上,现有研究已从轻量主干、跨层交互和边界建模等方面推动医学图像分割发展。轻量模型证明了低参数分割的可行性,跨层融合方法增强了语义与细节协同,边界感知方法改善了复杂轮廓恢复。然而,在资源受限皮肤病灶分割中,模型退化往往不是单一环节问题,而是表现为浅层噪声传递、跨层融合失衡和边界约束不足之间的连续影响。现有方法多针对其中一个环节进行增强,或局部耦合两个环节,仍缺少一种围绕极轻量网络失效链条的统一轻量设计。基于此,本文提出 SAFE 范式,并在极轻量 U 形主干上实例化为 LEAF-UNet,为直观展示 LEAF-UNet 在低参数量、低计算量与低边界定位误差之间的综合权衡优势,图 1 给出了不同模型在三个皮肤病灶数据集上的参数量—十亿次浮点运算 (giga floating-point operations, GFLOPs)—第 95 百分位豪斯多夫距离 (95th-percentile Hausdorff distance, HD95) 对比结果。LEAF-UNet 从特征传递、跨层融合和训练监督三个阶段对资源受限皮肤病灶分割中的结构感知退化进行协同修复。

2 方 法

本文提出 SAFE 统一范式。SAFE 遵循最小必要改造原则,在不重构主干或堆叠复杂模块的前提下,对轻量分割网络中的特征传递、跨层融合和训练监督三个关键节点进行针对性增强。各代表性方法在 SAFE 三个关键环节上的对比如表 1 所示。

2.1 SAFE 总体框架

SAFE 的核心思想是将极轻量分割网络中的结构感知退化归纳为三个关联的信息流节点:特征传递、跨层融合和训练监督。首先,在特征传递阶段,编码器浅层特征具有较高空间分辨率,但也可能包含背景纹理和局部伪响应,因此需要在进入解码端前进行筛选。其次,在跨层融合阶段,低层细节与高层语义对不同尺度和不同边界清晰度的病灶具有不同贡献,固定拼接或相加难以适应样本差异,因此需

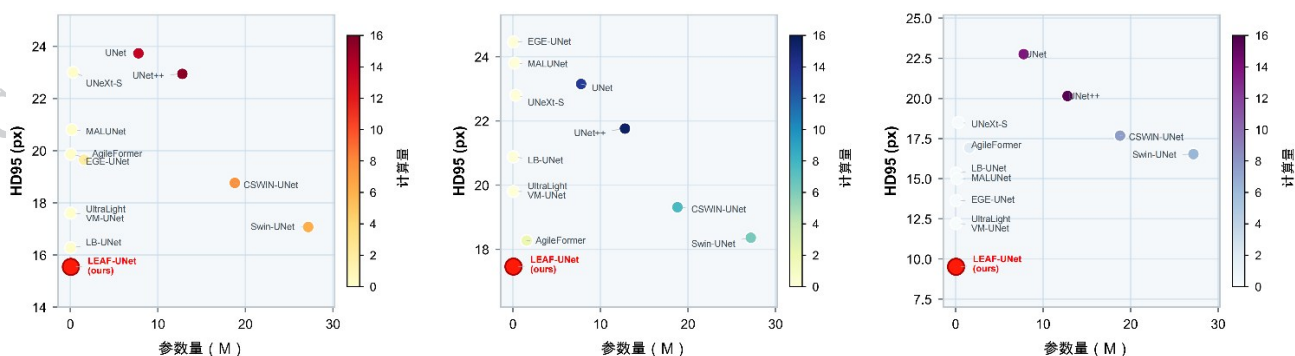


图 1 不同模型在不同数据集上的边界定位误差(HD95)对比结果可视化。图中横坐标表示模型参数量(parameters, Params, 单位: M, 其中 M 表示 10^6), 纵坐标表示第 95 百分位豪斯多夫距离(HD95, 单位: 像素(pixel, px)), 气泡颜色深浅表示计算复杂度(GFLOPs)。图中气泡越靠近左下角且颜色越浅, 表明模型在保持极低部署开销的同时, 取得了更优的边界定位与刻画质量。

((a) ISIC2017 dataset; (b) ISIC2018 dataset; (c) PH² dataset)

Figure. 1 Visualization of comparison results on boundary localization error (HD95) for different models.

要内容自适应的融合权重分配。最后, 在训练监督阶段, 区域型损失主要约束前景区域重叠, 对模糊轮廓的几何一致性约束不足, 因此需要训练期边界约束进行补充。

基于上述分解, SAFE 由选择性特征传递、适应性跨层融合和边界一致性监督三个部分组成。三者并非独立模块的简单叠加, 而是在同一优化目标下形成前向特征筛选、跨层权重分配和训练期边界约束之间的协同关系。

2.2 轻量 U 形主干上的实例化

LEAF-UNet 以 EGE-UNet 作为基础骨架, 保留其编码—解码结构和主要轻量化设计, 以维持原有的低参数量和低计算开销, 具体分割框架如图 2 所示。本文不重构主干网络, 而是在 SAFE 定义的三个信息流节点插入最小改造。具体而言, 轻量通道重标定机制(LAS)被置于编码器输出与对应解码器输入之间的跳跃连接路径, 用于在跨层传递前对浅层特征进行通道级选择性重标定; 门控加权融合机制(GWF)用于融合经 LAS 筛选后的浅层特征与解码端上采样后的深层语义特征, 并替代固定拼接或逐元素相加; 边界一致性监督仅在训练阶段作用于预测输出, 用于补充区域型损失对轮廓几何一致性的约束。

在实现过程中, 当浅层特征与深层特征的空间尺寸或通道数不一致时, 本文先通过上采样和 1×1

卷积进行对齐, 再输入 GWF 计算融合权重。推理阶段仅保留 LAS 和 GWF 的前向路径, 边界一致性

监督不引入额外边界分支、后处理模块或推理计算图。因此, LEAF-UNet 的设计重点不是通过增加模型容量换取性能, 而是在保持极轻量主干优势的同时, 对最容易引起结构误差的传递、融合和监督节点进行针对性修复。

2.3 选择性特征传递: Lightweight attention skip (LAS)

在 U 形分割网络中, 跳跃连接将编码器的高分辨率浅层特征直接输送到解码器, 以弥补细节损失 (Ronneberger 等, 2015)。然而, 在皮肤镜图像中, 浅层特征同时包含大量与病灶无关的干扰模式 (如毛发遮挡、标尺刻度)。对于容量受限的轻量模型而言, 这些干扰一旦未经筛选地进入解码器, 后续融合模块需要同时承担细节恢复和噪声抑制, 容易增加解码阶段的优化负担。为此, 本文在跳跃连接路径上引入轻量通道重标定机制(LAS), 用于在跨层传递前对浅层特征进行轻量级通道筛选。

LAS 用于在跳跃连接特征进入解码端之前进行通道级选择性重标定。设第 l 个编码阶段输出的浅层特征为 $\mathbf{X}_l \in \mathbf{R}^{C_l \times H_l \times W_l}$ 。首先, 对 $\mathbf{X}_l$ 进行全局平均池化, 得到每个通道的全局统计描述:

$$z_{lc} = \frac{1}{H_l W_l} \sum_{i=1}^{H_l} \sum_{j=1}^{W_l} X_{l,c,i,j}, c = 1, \dots, C_l \quad (1)$$

式中, $c \in \{1, 2, \dots, C\}$, i 和 j 分别表示空间维度上的行索引和列索引。其中, $X_{l,c,i,j}$ 表示第 l 层特征图在第 c 个通道、第 (ij) 个空间位置上的响应值, z_{lc} 表示该通道的全局统计描述。将所有通道描述拼接后得

表1 代表性方法在SAFE三个关键环节上的对比

Table 1 Comparison of representative methods across the three key stages of SAFE

方法	来源	主干范式	选择性传递	适应性融合	边界监督	边界开销
UNet	MICCAI 2015	CNN 编解码				—
UNet++	arXiv 2018	CNN/嵌套跳连		部分		—
Attention U-Net	MIDL 2018	CNN+注意力门控	部分*			—
UNeXt	MICCAI 2022	MLP				—
MALUNet	BIBM 2022	轻量 CNN	部分	部分		—
EGE-UNet	MICCAI 2023	组增强轻量 CNN				—
LB-UNet	MICCAI 2024	轻量 CNN		部分	✓	边界/预测分支
UL-VM-UNet	Patterns 2025	轻量 Mamba/SSM				—
VM-UNetV2/AHF	2024 / 2025	Mamba/注意力		✓		—
WA-NET	Sci. Rep. 2025	频域+边界细化		部分	✓	频域/边界分支
LEAF-UNet		EGE轻量 CNN	✓	✓	✓	无额外推理开销

注:UL-VM-UNet表示UltraLightVM-UNet,AHF表示AHF-U-Net,CNN表示卷积神经网络(convolutional neural networks,CNN),MLP表示多层感知机(multilayer perceptron,MLP),SSM表示状态空间模型(state space model,SSM)。MICCAI为医学图像计算与计算机辅助干预国际会议(Medical Image Computing and Computer Assisted Intervention,MICCAI),MIDL为医学影像深度学习会议(Medical Imaging with Deep Learning,MIDL),BIBM为IEEE生物信息学与生物医学国际会议(IEEE International Conference on Bioinformatics and Biomedicine,BIBM)。✓表示此方法具备对应的机制,表示此方法不具备对应的机制,“部分”表示仅覆盖该环节的部分功能,—表示此方法不适用对应的机制。*表示 Attention U-Net 仅对跳跃连接特征进行门控,不单独执行传递前筛选。

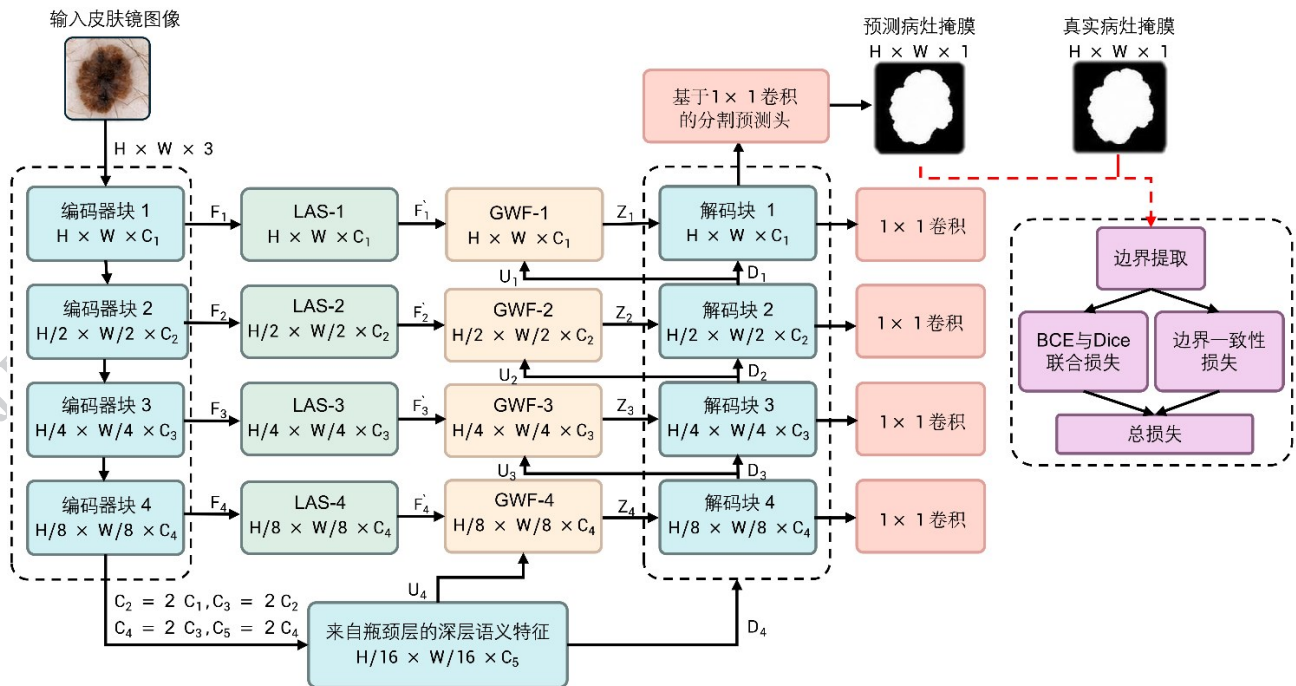


图2 轻量级高效分割LEAF-UNet框架

Figure. 2 Overall framework of the lightweight LEAF-UNet for skin lesion segmentation

到: $z_l = [z_{l,1}, z_{l,2}, \dots, z_{l,C_l}]^T \in \mathbf{R}^{C_l}$ 。随后,采用一维卷积在通道维度上建模局部通道依赖,并通过Sigmoid函数生成通道权重:

$$w_l = \sigma(\text{Conv1D}_{k_l}(z_l)), w_l \in (0,1)^{C_l}. \quad (2)$$

式中, $\sigma(\cdot)$ 表示Sigmoid激活函数, $\text{Conv1D}_{k_l}(\cdot)$ 表示卷积核大小为 k_l 的一维卷积操作。参考高效通道注意力网络(efficient channel attention network, ECA-Net)(Wang等,2020)中的自适应核大小设定, k_l 定义

为: $k_l = \left\lfloor \frac{\log_2(C_l)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}}$ 。式中, $\lfloor \cdot \rfloor_{\text{odd}}$ 表示距离输入值最近的奇数, γ 和 b 为映射超参数, 本文采用 $\gamma = 2, b = 1$ 。最后,将通道权重 w_l 沿空间维度广播,并与原始跳跃连接特征逐通道相乘,得到筛选后的跨层传递特征:

$$\tilde{X}_l = w_l \odot X_l \quad (3)$$

即: $\tilde{X}_{l,c,i,j} = w_{l,c} \cdot X_{l,c,i,j}$ 。式中, $\odot$ 表示逐元素乘法, $w_{l,c}$ 在空间维度上共享。通过上述过程, LAS能够在不改变特征空间分辨率的前提下抑制冗余或噪声通道响应,使进入解码端的浅层特征更加聚焦于与病灶结构相关的有效信息。

LAS仅增加用于建模局部通道依赖的一维卷积,不引入全连接层,也不改变特征图的空间分辨率,因此其额外参数数量和计算开销较低。具体模型复杂度见表5。

LAS与通用通道注意力的设计差异: LAS的目标不是提出一种全新的通道注意力算子,而是将低开销通道重标定约束性地部署在跳跃连接的传递路径上。与将ECA类模块插入编码器或解码器主干以增强整体表征不同, LAS作用于浅层高分辨率特征进入解码端之前,重点在于降低背景纹理和局部伪响应随跳跃连接传播的风险。

需要进一步说明LAS与直接在跳跃连接上部署通用通道注意力(如ECA-Skip)之间的差异。在U形网络中,跳跃连接传递的浅层特征具有两个显著统计特性:其一,浅层特征图的通道间响应方差远大于编码器深层或解码器中间层,表明不同通道中前景有效响应与背景干扰响应的区分度更高,适合进行通道级筛选;其二,浅层特征保留了较高的空间分辨率,使得通道维度上的全局统计量能够更完整地反映图像级前景—背景分布。基于此, LAS选择在跳跃连接入口执行通道重标定,而非在编码器或解码

器主干中使用,其目的是在噪声尚未与深层语义混合之前实施源头拦截。

2.4 适应性跨层融合: Gated weighted fusion (GWF)

现有轻量U形网络通常采用通道拼接或逐元素相加的方式融合浅层细节特征与深层语义特征。这类固定融合策略默认不同层次特征在所有样本和所有空间位置上具有相近贡献,难以适应皮肤病灶在尺度、对比度和边界清晰度上的差异。为此,本文在跨层聚合位置引入门控加权融合机制(GWF),如图3所示,以内容自适应方式分配浅层细节与深层语义的融合权重。

GWF用于在跨层聚合阶段自适应分配浅层细节特征与深层语义特征的贡献。设经LAS筛选后的浅层特征为 $\tilde{X}_l \in \mathbf{R}^{C_l \times H_l \times W_l}$, 解码端上采样后并完成通道对齐的深层语义特征为 $H_l \in \mathbf{R}^{C_l \times H_l \times W_l}$ 。若二者通道数或空间尺寸不一致,则先通过上采样与 1×1 卷积进行对齐:

$$\bar{H}_l = \phi_l(\text{Up}(H_{l+1})) \quad (5)$$

式中, $\text{Up}(\cdot)$ 表示上采样操作, $\phi_l(\cdot)$ 表示 1×1 卷积映射。为简化符号,下文仍用 H_l 表示对齐后的深层语义特征。

首先,将浅层特征和深层特征沿通道维度拼接: $U_l = \text{Concat}(\tilde{X}_l, H_l), U_l \in \mathbf{R}^{2C_l \times H_l \times W_l}$ 。随后,通过轻量 1×1 卷积生成两通道门控响应:

$$G_l = \text{Conv}_{1 \times 1}^{2C_l \rightarrow 2}(U_l), G_l \in \mathbf{R}^{2 \times H_l \times W_l}. \quad (6)$$

式中, $G_{l,1}$ 和 $G_{l,2}$ 分别表示浅层细节分支和深层语义分支的未归一化门控响应。为确保两个分支在每个空间位置的贡献之和为1,对 G_l 在分支维度上执行Softmax归一化:

$$\alpha_l(i,j) = \frac{\exp(G_{l,1,i,j})}{\exp(G_{l,1,i,j}) + \exp(G_{l,2,i,j})} \quad (7)$$

$$\beta_l(i,j) = \frac{\exp(G_{l,2,i,j})}{\exp(G_{l,1,i,j}) + \exp(G_{l,2,i,j})} \quad (8)$$

因此,在任意空间位置 (i,j) 上均满足: $\alpha_l(i,j) + \beta_l(i,j) = 1$ 。最终, GWF根据归一化门控权重进行动态融合: $F_l = \alpha_l \odot \tilde{X}_l + \beta_l \odot H_l$ 。即: $F_{l,c,i,j} = \alpha_l(i,j) \tilde{X}_{l,c,i,j} + \beta_l(i,j) H_{l,c,i,j}$ 。其中, α_l 和 β_l 沿通道维度广播后参与逐元素乘法。与固定拼接或逐元素相加不同, GWF使网络能够根据局部图像内容动态调整低层纹理与高层语义的融合比例,从而缓解轻量

网络在复杂背景、小目标和模糊边界场景中的跨层融合失衡问题。

GWF 仅使用一个轻量 1×1 卷积生成双分支空间权重, 并通过逐元素运算完成特征融合, 因此不会显著增加主干网络的计算负担。实际参数数量和 GFLOPs 见表 5。

GWF 将“如何融合”从固定规则转化为数据驱动的动态决策: 从机制上看, GWF 允许网络在不同空间位置自适应调整浅层细节与深层语义的相对贡献。对于背景纹理较强或边界较清晰的区域, 模型可倾向于提高语义分支权重以抑制局部伪纹理; 对于边界模糊或细节易丢失的区域, 模型可倾向于保留更多浅层细节响应。与单路注意力门控仅对某一路特征进行抑制或增强不同, GWF 同时估计浅层细节和深层语义的贡献, 使二者在每个空间位置上形成归一化融合关系。该设计允许模型根据局部图像内容调整低层纹理与高层语义的融合比例, 从而在较小参数增量下提升跨层融合的适应性。

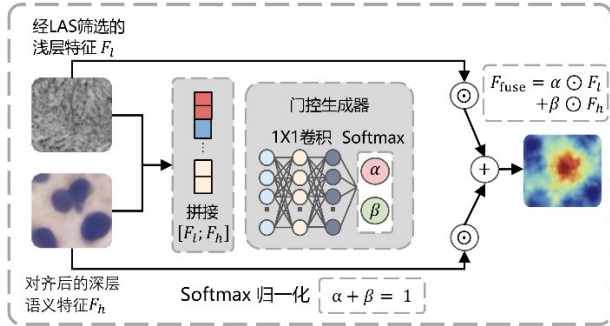


图3 门控加权融合机制(GWF)结构示意图

Figure 3 Schematic Diagram of Gated Weighted Fusion (GWF) Mechanism

2.5 训练期边界一致性监督 (Boundary consistency loss)

二元交叉熵损失 (binary cross-entropy loss, BCE) 和 Dice 损失 (Dice loss) 等区域型损失主要约束前景区域的像素分类与整体重叠程度, 对局部过渡带和模糊轮廓的几何一致性约束相对有限, 这常导致轻量模型的预测结果在边界附近出现局部外扩、毛刺或断裂。为此, 本文在训练阶段引入如图 4 所示的无参数边界一致性监督, 通过固定拉普拉斯算子提取预测掩膜与真实标注的高频轮廓响应并显式约束二者一致, 从而在不增加推理分支的前提下提供额外轮廓约束。对于第 s 个监督尺度, 设预测

概率图 $\hat{Y}^{(s)} \in [0, 1]^{H_s \times W_s}$, 真实标签为 $Y^{(s)} \in \{0, 1\}^{H_s \times W_s}$ 。第 s 个监督尺度上的基础区域损失由二元交叉熵损失与 Dice 损失构成, 记为 $\mathcal{L}_{r,s}^{(s)}$ 两类损失均采用标准定义。

为提取轮廓高频响应, 本文采用固定四邻域拉普拉斯卷积核:

$$K_{\Delta} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (9)$$

预测掩膜与真实标签的边界响应分别表示为:

$$B_p^{(s)} = |K_{\Delta} * \hat{Y}^{(s)}|, B_g^{(s)} = |K_{\Delta} * Y^{(s)}|, \quad (10)$$

式中, $*$ 表示二维卷积操作, $|\cdot|$ 表示逐元素绝对值。基于二者的差异, 边界一致性损失定义为:

$$\mathcal{L}_{b,s}^{(s)} = \frac{1}{N_s} \|B_p^{(s)} - B_g^{(s)}\|_1. \quad (11)$$

第 s 个尺度上的联合损失为:

$$\mathcal{L}^{(s)} = \mathcal{L}_{r,s}^{(s)} + \gamma \mathcal{L}_{b,s}^{(s)}, \quad (12)$$

式中, γ 为边界一致性损失权重。最终总损失为各尺度损失的加权和:

$$\mathcal{L}_{total} = \sum_{s=1}^S \omega_s \mathcal{L}^{(s)}, \quad (13)$$

式中, S 表示深度监督尺度数, ω_s 表示第 s 个尺度损失的权重系数。推理阶段仅保留主干网络、LAS 和 GWF 的前向路径, BCL 不会引入额外边界分支、后处理模块或推理计算图。

需要说明的是, BCL 仅在训练阶段参与损失计算与反向传播, 不引入额外可学习参数, 也不改变推理阶段的网络结构。边界权重 γ 是训练超参数, 需通过验证集选择, 其敏感性将在实验部分进一步分析。该监督项的作用是为 BCE 和 Dice 等区域型损失补充轮廓几何约束, 引导模型在学习区域覆盖的同时提高预测边界与真实边界之间的一致性。

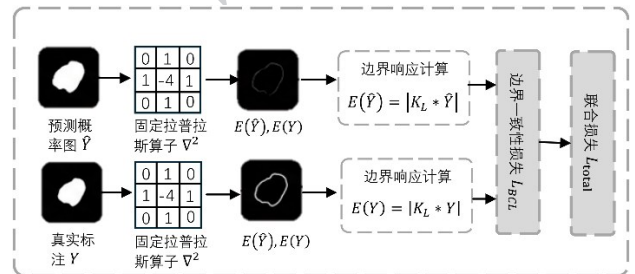


图4 无参数边界一致性监督模块结构示意图

Figure 4 Schematic Diagram of Parameter-Free Boundary Consistency Supervision Module

2.6 总体优化目标与复杂度讨论

综上所述,LEAF-UNet通过LAS、GWF与边界一致性监督,将SAFE范式的三个原则无缝嵌入特征传递、跨层融合和优化阶段。其核心设计摒弃了重型模块的堆叠,转而在信息最易出错的节点进行最小必要改造。

3 实验

3.1 数据集与评价指标

本文在ISIC2017(Codella等,2017)、ISIC2018(Codella等,2019)和PH²(Mendonça等,2013)三个公开皮肤病灶分割数据集上评估LEAF-UNet。对于ISIC2017和ISIC2018,本文遵循EGE-UNet及相关工作的常用划分方式,以保证与轻量基线和现有方法的可比性。PH²数据集仅包含200张高分辨率皮肤镜图像,样本规模较小且更易受划分波动影响,因此本文在固定随机种子下划分训练集、验证集和测试

集,并在实验设置中保持所有模型一致。

本文采用平均交并比(mean intersection over union, mIoU)和Dice相似系数(Dice similarity coefficient, DSC)评价区域重叠性能,并采用第95百分位豪斯多夫距离(95th-percentile Hausdorff distance, HD95)评价边界定位误差。HD95取预测边界与真实边界双向距离的95百分位数,能够降低少量离群边界点的影响。由于各数据集未提供统一的物理像素间距,本文统一以像素(pixel, px)为单位计算HD95。mIoU和DSC越高表示区域分割性能越好,HD95越低表示边界定位越准确。

3.2 实验协议与实施细节

为保证实验结果的公平性与可复现性,本文对EGE-UNet、各消融变体和LEAF-UNet在统一环境下重新训练与测试。均在单张图形处理器(graphics processing unit, GPU)上完成。模型基于PyTorch实现,主要实验参数设置如表2所示。

表2 实验参数设置
Table 2 Experimental parameter settings

类别	参数项	设置
实验平台	计算平台	AutoDL
实验硬件	GPU	单张 NVIDIA RTX 4090
输入设置	输入分辨率	256 × 256
数据增强	增强方式	随机水平翻转、随机垂直翻转、随机旋转
优化设置	优化器	AdamW
学习率	初始学习率	1e-3
学习率调度	最小学习率	1e-5
训练设置	训练轮数	300 轮
训练设置	批大小	8
基础损失	损失函数	BCE 损失 + Dice 损失

3.3 对比方法与选取依据

具体而言,本文的对比方法并非简单罗列,而是围绕“经典分割基线—高容量全局建模方法—资源受限轻量化方法”三个层级进行选择,以全面评估LEAF-UNet在分割精度、边界定位能力和模型复杂度之间的综合表现。

第一类为经典CNN分割基线,包括U-Net(Ronneberger等,2015)和UNet++(Zhou等,2018)。U-Net

是医学图像分割中最具代表性的编码—解码式基线模型,UNet++则通过嵌套跳跃连接增强多尺度特征交互,二者可用于衡量本文方法相对于经典U形结构的性能改进。

第二类为基于Transformer或混合全局建模的高容量方法,包括Swin-UNet(Cao等,2022)、AgileFormer(Qiu等,2026)和CSWin-UNet(Liu等,2025)等。这类方法具有较强的全局上下文建模能力,代

表了医学图像分割中高表达能力模型的发展方向。将其纳入对比,有助于评估 LEAF-UNet 在极低参数数量和计算量条件下,能否通过信息流优化获得接近甚至优于复杂模型的分割效果。

第三类为面向资源受限场景的轻量化先进方法,包括 UNeXt-S (Valanarasu 和 Patel, 2022)、MALU-Net (Ruan 等, 2022)、EGE-UNet (Ruan 等, 2023)、LB-UNet (Xu 等, 2024) 和 UltraLight VM-UNet (Wu 等, 2025) 等。这类方法与本文研究目标最为接近,均关注模型轻量化、边缘部署或高效医学图像分割。其中,EGE-UNet 是本文 LEAF-UNet 的基础骨架,因此与其对比能够直接反映 SAFE 范式中 LAS、GWF 和 BCL 三个改造环节带来的增益;LB-UNet 引入边界相关设计,与本文关注的边界一致性监督具有较强关联;UltraLight VM-UNet 则代表了近期视觉状态空间模型在轻量医学分割中的应用趋势。通过与这些方法进行比较,可以更直接地验证 LEAF-UNet 在同类轻量化方法中的有效性和竞争力。

此外,为保证实验公平性,本文对 EGE-UNet、各消融变体和 LEAF-UNet 在统一实验环境下重新训练和测试;对于其他对比方法,则尽量采用公开论文或公开实现中的标准设置,并在相同数据集和评价指标下进行比较。评价指标同时包含 mIoU、DSC 等区域重叠指标,以及 HD95 边界定位误差指标,从而避免仅依据单一精度指标评价模型性能。通过上述补充,修订稿能够更清晰地说明本文对比实验的基线选择依据、技术覆盖范围和公平性设置。

3.4 对比实验

表 3 给出了 LEAF-UNet 与代表性分割模型在 ISIC2017、ISIC2018 和 PH² 数据集上的定量比较。总体来看,LEAF-UNet 在保持低模型复杂度的同时,在三个数据集上均取得了较稳定的区域重叠和边界定位表现。具体而言,LEAF-UNet 的参数量约为 0.053M,计算量约为 0.075GFLOPs,在 ISIC2017、ISIC2018 和 PH² 上的 HD95 分别为 15.54、17.46 和 9.51,说明该方法在低复杂度条件下仍能保持较好的轮廓定位能力。

与轻量基线 EGE-UNet 相比,LEAF-UNet 不仅提升了 mIoU 和 DSC,也在三个数据集上降低了 HD95。这表明,所提方法在增强区域覆盖能力的同时,对边界稳定性也具有一定改善。尤其在 PH² 数据集上,LEAF-UNet 取得 90.45/94.83 的 mIoU/DSC 和 9.51

的 HD95,说明其在小规模外部数据集上仍保持了较好的区域一致性与边界保真度。

图 5 和图 6 以青绿色、洋红色和金黄色分别表示预测结果与真实标注的重叠区域、过分割区域和欠分割区域。其中,青绿色区域越完整,且洋红色、金黄色区域越少,表明分割结果与真实标注的一致性越高。在毛发遮挡、边界模糊及病灶形态不规则等复杂样本中,部分对比方法存在较明显的局部外扩或病灶主体缺失,表现为边界附近洋红色或金黄色误差区域增多。相比之下,LEAF-UNet 的青绿色重叠区域相对更完整,边界附近的过分割与欠分割区域较少,预测轮廓与真实标注更为接近。该结果与表 3 中 LEAF-UNet 较低的 HD95 结果相一致,说明所提方法的性能提升不仅体现在区域重叠指标上,也体现在病灶边界定位与轮廓恢复质量的改善上。

综合表 3 和定性结果可见,LEAF-UNet 在极低参数量与计算量下取得了较均衡的精度—效率表现。不同于依赖更强主干或更高计算开销的方法,LEAF-UNet 通过协同优化特征传递、跨层融合与边界监督,在轻量约束下改善了区域分割与边界定位的平衡。

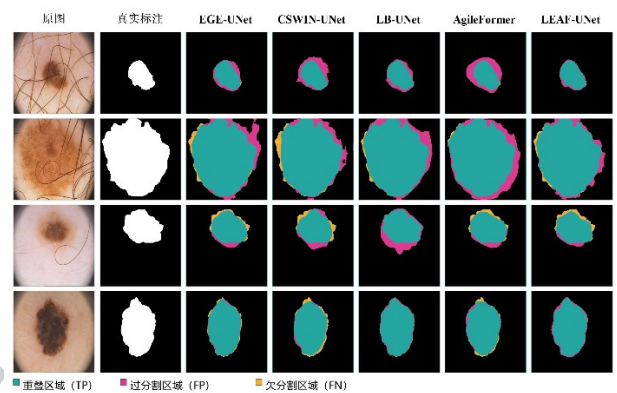


图 5 不同模型在 PH² 数据集上的分割结果对比图

Figure 5 Qualitative comparison of segmentation results produced by different models on the PH² dataset.

3.5 消融实验

为验证 LEAF-UNet 各组成部分的作用,本文在统一实验协议下进行了单模块和双模块组合消融,结果见表 4。整体来看,LAS、GWF 和 BCL 对应的是轻量网络中的不同退化节点:LAS 主要作用于跳

跃连接传递阶段,GWF 主要作用于跨层融合阶段,BCL 则在训练阶段提供边界几何约束。

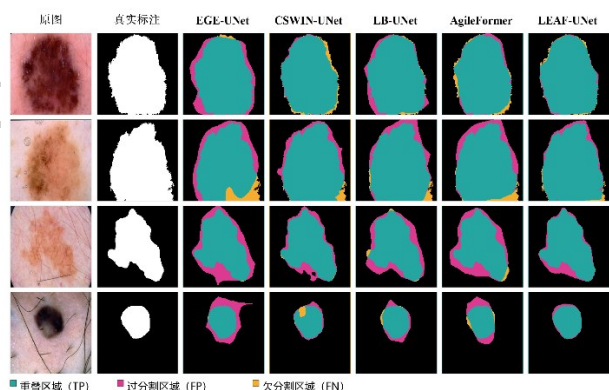


图6 不同模型在ISIC2018数据集上的分割结果对比图
Figure 6 Qualitative comparison of segmentation results produced by different models on the ISIC2018 dataset.

从单模块结果看, LAS在三个数据集上均降低了HD95, 说明跳跃连接前的特征筛选有助于抑制背景纹理和伪响应传播; GWF单独引入时, 在 PH^2 上取得了更优的 mIoU/DSC 与 HD95, 但在 ISIC2018 上 DSC 略低于基线。这一现象与 GWF 的作用机制一致: GWF 通过内容自适应权重协调低层细节与高层语义, 其收益依赖于跳跃连接特征已较为干净; 当浅层特征仍混杂背景噪声(如 ISIC2018 中较多的毛发与反光样本)而未经 LAS 前置筛选时, 动态加权可能将部分噪声细节误判为有效响应, 从而在区域指标上产生轻微波动。这也解释了为何 LAS 与 GWF 必须协同: 只有先在传递阶段抑制噪声, 融合阶段的自适应加权才能稳定发挥作用。BCL 单独引入后同样带来一定边界改善, 但其主要作用是训练期几何约束, 难以完全替代前向特征筛选和跨层融合增强。

双模块组合进一步体现出不同阶段之间的互补性。LAS+GWF 同时覆盖传递和融合两个前向环节, 能够在减少浅层噪声输入的基础上进一步优化语义—细节分配; GWF+BCL 则结合动态融合和边界监督, 在 PH^2 等边界较敏感的数据集上表现出较好的轮廓定位能力; LAS+BCL 则说明前端噪声筛选与训练期边界约束之间也存在一定协同。

完整 LEAF-UNet 同时引入 LAS、GWF 和 BCL, 在三个数据集上取得最稳定的综合表现, 说明其性能提升并非来自单一模块的孤立作用, 而是来自“传递—融合—监督”三个阶段的协同修复。由此, 消融结果支持本文的核心设计: 在极轻量网络中, 结构感知退化需要通过多个关键节点的轻量化联合干预来

缓解。

3.6 超参数敏感性分析

在联合优化目标中, 边界一致性损失权重 γ 控制区域重叠约束与边界几何约束之间的平衡。为分析 γ 对模型性能的影响, 本文在三个数据集上分别测试不同权重设置, 并记录 mIoU 与 HD95 的变化趋势, 如图 7 所示。整体来看, 适当的边界权重有助于改善区域分割和轮廓定位, 但过大的边界权重可能削弱模型对病灶主体区域的稳定学习, 导致性能下降。

不同数据集上的最优 γ 并不完全一致, 说明边界监督强度需要与数据分布和标注边界特性相匹配。对于边界噪声、低对比度或局部遮挡较多的数据集, 过强的边界项可能使模型过度关注高频响应; 而在边界相对清晰的数据集上, 适度增强轮廓约束可能进一步降低 HD95。因此, 本文在各数据集上通过验证集选择 γ , 并将其作为训练期超参数进行调节。该结果也说明, BCL 虽然不增加推理阶段参数量和计算量, 但其训练期权重仍需要进行合理设置。

3.7 近邻模块替换实验分析

为进一步排除性能提升仅来自已有注意力或边界损失算子的可能, 本文在相同 EGE-UNet 骨干上进行了近邻模块替换实验(表 5)。在传递模块方面, 将 ECA-Net 的通道注意力直接插入跳跃连接(ECA-Skip)与 LAS 进行对比; 在融合模块方面, 将注意力门控(attention gate, AG)与 GWF 进行对比; 在边界监督方面, 将经典边界损失(boundary loss, BL)与 BCL 进行对比。结果表明, 各近邻方法均能在一定程度上改善基线, 但其收益主要局限于单一环节: ECA-Skip 能降低 HD95 但区域指标提升有限, 注意力门控能改善融合, 但对边界质量的作用不如 GWF 在 PH^2 上稳定, 边界损失能够直接引入边界相关约束, 但在本文实验设置下, 其区域指标和 HD95 改善不如 BCL 稳定, 说明针对极轻量皮肤病灶分割任务, 简单替换边界损失并不能完全替代“传递—融合—监督”协同设计。相比之下, LEAF-UNet 在三个数据集上取得了最稳定的 mIoU/DSC 与 HD95 综合表现, 说明 SAFE 的收益并非来自某一算子的简单替换, 而是来自传递筛选、融合协调和边界约束之间的跨阶段协同。

表 3 不同模型在三个皮肤病灶分割数据集上的性能比较
Table 3 Performance comparison of different models on three skin lesion segmentation datasets

模型	参数量 (M) ↓	计算量 (GFLOPs) ↓	ISIC2017 mIoU/ DSC ↑	HD95 (px) ↓	ISIC2018 mIoU/ DSC ↑	HD95 (px) ↓	PH ² mIoU/ DSC ↑	HD95 (px) ↓
U-Net(Ronneberger 等, 2015)	7.78	13.64	76.85/ 86.96	23.73	77.86/ 86.99	23.15	78.90/ 87.65	22.76
UNet++ (Zhou 等, 2018)	12.79	15.50	77.67/ 87.43	22.94	78.31/ 87.83	21.76	78.90/ 88.26	20.15
CSWin-UNet (Liu 等, 2025)	18.79	7.66	80.37/ 88.83	18.76	79.90/ 88.76	19.31	86.34/ 92.14	17.68
Swin-UNet (Cao 等, 2022)	27.17	6.20	77.05/ 87.92	17.07	78.43/ 88.25	18.36	86.54/ 90.17	16.53
AgileFormer (Qiu 等, 2026)	1.57	1.95	79.05/ 88.30	19.65	79.00/ 88.59	18.27	87.35/ 91.72	16.92
LB-UNet (Xu 等, 2024)	0.038	0.098	80.13/ 88.86	16.27	81.37/ 89.54	20.88	88.93/ 93.81	15.34
EGE-UNet(Ruan 等, 2023)	0.053	0.072	79.81/ 88.77	19.86	80.94/ 89.46	24.46	88.06/ 93.65	13.64
UNeXt-S (Valanarasu 和 Patel, 2022)	0.32	0.10	78.26/ 87.80	23.0	79.29/ 88.53	22.8	85.80/ 92.30	18.5
MALUNet(Ruan 等, 2022)	0.177	0.085	78.78/ 88.13	20.8	80.30/ 89.05	23.8	87.00/ 93.00	15.0
UltraLight VM-UNet (Wu 等, 2025)	0.049	0.060	80.30/ 89.10	17.6	81.20/ 89.70	19.8	89.10/ 94.10	12.2
LEAF-UNet	0.053	0.075	81.26 0.21 89.50 0.19	15.54 0.14	81.76 0.10 90.32 0.11	17.46 0.05	90.45 0.07 94.83 0.13	9.51 0.08

注: ↑表示数值越大越优, ↓表示数值越小越优; mIoU/DSC 表示平均交并比和 Dice 系数, 单位均为%; HD95 单位为像素; 粗体表示最优结果。±后面的数字表示五次实验后的方差。

3.8 边界误差分布分析

为进一步分析 HD95 改善是否来自整体边界误差分布的下降, 本文统计了不同模型在测试集上的逐样本 HD95 分布, 结果如图 8 所示。LEAF-UNet 在三个数据集上的箱体位置均低于 EGE-UNet, 且四分位范围更为集中, 说明完整模型不仅降低了平均边界误差, 也减少了高误差样本带来的分布长尾。

具体而言, 在 ISIC2017 数据集上, LEAF-UNet 相较 EGE-UNet 的 HD95 中位数降低约 2.4px, 四分位距缩小约 6.2px; 在 ISIC2018 数据集上, 中位数降低约 4.7px, 四分位距缩小约 8.2px; 在 PH² 数据集上, 中位数降低约 4.8px, 四分位距缩小约 4.4px。这表

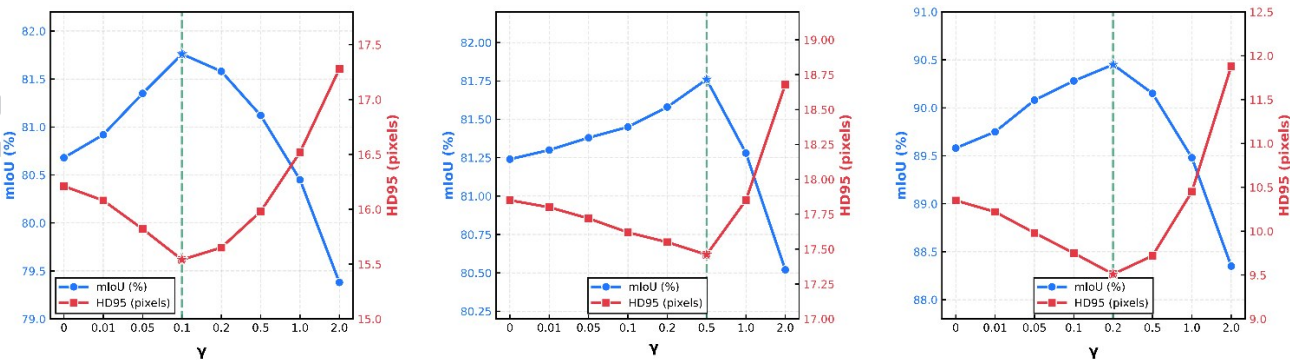
明所提方法在不同数据分布下均能提升边界定位的稳定性, 尤其是在 ISIC2018 和 PH² 上表现出更明显的分布收缩趋势。

从模块作用看, EGE+LAS+GWF 的箱体整体低于 EGE-UNet, 说明前向阶段的浅层特征筛选与跨层自适应融合本身已能改善边界预测质量; 而完整 LEAF-UNet 进一步取得更低的中位数和更稳定的分布, 表明边界一致性监督与前向表征增强之间具有互补作用。值得注意的是, EGE+BCL 在 PH² 上的分布高于基线, 说明在小样本数据集上单独引入边界约束可能使优化过度关注局部轮廓响应, 从而削弱整体边界稳定性。

表 4 LEAF-UNet各组成模块在三个数据集上的消融实验
Table 4 Ablation study of LEAF-UNet components on three datasets

	LAS	GWF	BCL	ISIC2017 mIoU/DSC ↑	HD95 (px) ↓	ISIC2018 mIoU/DSC ↑	HD95 (px) ↓	PH ² mIoU/DSC ↑	HD95 (px) ↓
(1)	EGE-UNet			79.81/88.77	19.86	80.94/89.46	24.46	88.06/93.65	13.64
(2)	✓			79.90/89.07	18.53	80.90/89.61	22.43	88.11/93.50	12.97
(3)		✓		79.94/89.29	18.08	80.97/89.01	21.13	89.16/94.09	11.21
(4)			✓	79.78/89.14	18.23	81.02/89.76	18.59	89.19/93.47	13.42
(5)	✓	✓		80.68/89.30	16.21	81.24/89.82	18.85	89.58/94.35	10.35
(6)	✓		✓	80.05/89.32	15.86	81.35/89.94	18.37	89.72/94.72	10.28
(7)		✓	✓	80.42/89.25	17.92	81.17/90.09	17.59	89.55/94.51	10.05
(8)	✓	✓	✓	81.26/89.50	15.54	81.76/90.32	17.46	90.45/94.83	9.51

注:✓表示引入对应模块,空白表示未引入;↑表示数值越大越优,↓表示数值越小越优;粗体表示最优结果



(a)ISIC2017数据集 (b)ISIC2018数据集 (c)PH²数据集
((a)ISIC2017dataset; (b)ISIC2018dataset; (c)PH²dataset)

图 7 不同边界损失权重 γ 对模型在三个数据集上分割性能(mIoU与HD95)的影响

Figure. 7 Impact of different boundary loss weights γ on model segmentation performance (mIoU and HD95) across three datasets.

3.9 困难样本分层分析

为进一步验证 LEAF-UNet是否主要改善结构感知更易退化的困难样本,为保证困难样本划分的可复现性,本文给出四类困难属性的定量定义。给定输入图像 I 和真实标注 Y ,病灶面积比例定义为:

$$r_{\text{area}} = \frac{|Y|}{H_0 W_0}, \quad (14)$$

其中, $|Y|$ 表示前景像素数量。区域对比度定义为:

$$r_{\text{con}} = \frac{|\mu_{\text{fg}} - \mu_{\text{bg}}|}{\sigma_{\text{fg}} + \sigma_{\text{bg}} + \epsilon}, \quad (15)$$

其中, μ_{fg} 和 μ_{bg} 分别表示前景与背景区域的平均亮度值, σ_{fg} 和 σ_{bg} 分别表示对应标准差。轮廓复杂度定义为:

$$r_{\text{cmp}} = \frac{|\partial Y|}{2\sqrt{\pi|Y|} + \epsilon}, \quad (16)$$

其中, $|\partial Y|$ 表示病灶边界长度。该指标越大,说明轮廓相对于同面积圆形目标越不规则。边界梯度定义为:

$$r_{\text{grad}} = \frac{1}{|\partial_{\rho} Y|} \sum_{p \in \partial_{\rho} Y} \|\nabla I(p)\|_2, \quad (17)$$

其中, $\partial_{\rho} Y$ 表示真实边界附近半径为 ρ 的邻域, $\nabla I(p)$ 表示图像在位置 p 处的梯度幅值。本文将 r_{area} 位于最低 25% 的样本定义为小病灶子集,将 r_{con} 位于最低 25% 的样本定义为低对比度子集,将 r_{cmp} 位于最高 25% 的样本定义为复杂边界子集,将 r_{grad} 位于最低 25% 的样本定义为模糊边界子集。所有阈值均仅根据测试集真实标注和输入图像统计得到,不依赖模型预测结果。

可以看到,在所有数据集和困难子集上,LEAF-UNet相较于 EGE-UNet 的 HD95降低幅度均为正值,

表 5 近邻模块替换实验

Table 5 Replacement experiment of neighboring modules

模型设置	传递模块	融合模块	边界监督	参数量(M)	计算量 (GFLOPs)	ISIC2017		ISIC2018		PH ²	
						mIoU/ DSC	HD95	mIoU/ DSC	HD95	mIoU/ DSC	HD95
EGE-UNet	原始跳跃连接	原始融合	BCE+Dice	0.0529	0.072	79.81/ 88.77	19.86	80.94/ 89.46	24.46	88.06/ 93.65	13.64
EGE+ECA-Skip	ECA-Skip	原始融合	BCE+Dice	0.0530	0.073	79.28/ 88.44	16.92	79.48/ 88.57	18.65	87.20/ 93.16	15.72
EGE+Attention Gate	原始跳跃连接	AG	BCE+Dice	0.0536	0.077	80.02/ 88.90	19.38	80.78/ 89.37	23.85	88.72/ 94.02	12.35
EGE+Boundary Loss	原始跳跃连接	原始融合	BL	0.0529	0.072	78.42/ 87.90	17.55	79.05/ 88.30	20.52	84.35/ 91.51	16.80
EGE+LAS	LAS	原始融合	BCE+Dice	0.0531	0.073	79.90/ 89.07	18.53	80.90/ 89.61	22.43	88.11/ 93.50	12.97
EGE+GWF	原始跳跃连接	GWF	BCE+Dice	0.0532	0.074	79.94/ 89.29	18.08	80.97/ 89.01	21.13	89.16/ 94.09	11.21
EGE+BCL	原始跳跃连接	原始融合	BCL	0.0529	0.072	79.78/ 89.14	18.23	81.02/ 89.76	18.59	89.19/ 93.47	13.42
LEAF-UNet	LAS	GWF	BCL	0.0533	0.075	81.26/ 89.50	15.54	81.76/ 90.32	17.46	90.45/ 94.83	9.51

注:AG表示注意力门控(attention gate, AG),BL表示边界损失(boundary loss, BL),粗体表示最优结果。

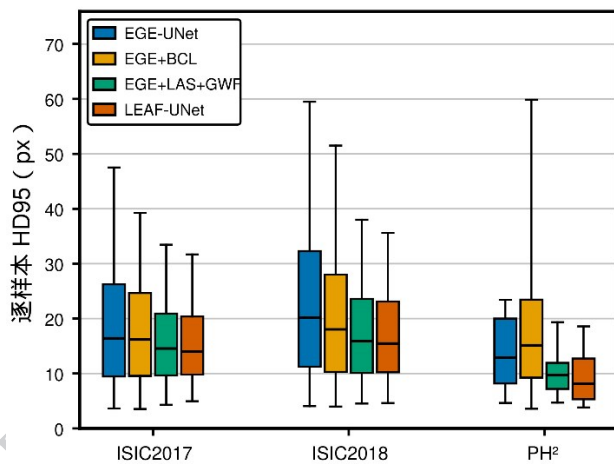


图 8 不同模型在 ISIC2017、ISIC2018 和 PH²数据集上的逐样本 HD95 分布

Figure. 8 Per-case HD95 distributions of different models on ISIC2017, ISIC2018, and PH².

说明其边界定位收益不仅体现在整体测试集上,也能够稳定迁移到结构感知更易退化的困难样本中。四类困难样本中,模糊边界和复杂边界的平均改善最为明显,HD95 分别降低 9.26 px 和 8.14 px,表明所提边界一致性监督对低对比度过渡区域和不规则轮廓具有较强约束作用;小病灶和低对比度样本也

分别获得 7.22 px 和 6.47 px 的改善,说明 LAS 与 GWF 有助于减少背景噪声传递并增强细粒度结构保留能力。从数据集看,ISIC2018 的平均改善最大,达到 10.47 px,其中模糊边界子集降低 12.69 px,进一步表明 LEAF-UNet 在复杂背景和边界不确定场景下具有更好的鲁棒性。上述结果进一步支持了 SAFE 在“传递—融合—监督”三个环节协同修复结构感知退化的有效性。

4 总结

本文面向资源受限皮肤病灶分割任务,提出 SAFE 统一范式,并在极轻量 U 形主干上实例化为 LEAF-UNet。该方法将轻量网络中的结构感知退化归纳为特征传递污染、跨层融合失衡和边界约束不足三个连续环节,并分别通过 LAS、GWF 和训练期边界一致性监督进行轻量化修复。实验结果表明,LEAF-UNet 在 ISIC2017、ISIC2018 和 PH²数据集上能够在保持低参数量和计算量的同时,提升区域重叠精度并降低 HD95,说明所提方法有助于改善轻量模型在复杂背景和模糊边界条件下的结构感知能力。

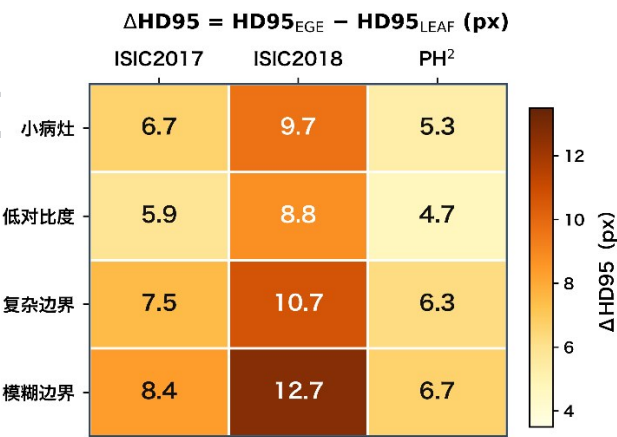


图9 不同数据集与困难样本类型下的边界误差下降幅度热力图

Figure. 9 Heatmap of boundary error reduction across different datasets and hard sample types.

本文仍存在一定局限。首先,当前验证主要集中在皮肤镜分割任务,SAFE在其他医学分割任务和不同轻量骨干上的泛化能力仍需进一步评估。其次,本文主要从参数量和GFLOPs角度讨论部署开销,尚缺少真实边缘硬件上的延迟、吞吐和能耗测试。未来工作将进一步开展跨骨干、跨任务和真实设备部署实验,并结合更细粒度的边界误差分析,验证SAFE在复杂资源受限场景中的稳定性与可迁移性。

参考文献(References)

Asgari Taghanaki S, Abhishek K, Cohen J P, Cohen-Adad J and Hamarneh G. 2021. Deep semantic segmentation of natural and medical images; a review. *Artificial Intelligence Review*, 54(1): 137-178 [DOI: 10.1007/s10462-020-09854-1].

Cao H, Wang Y, Chen J, Jiang D, Zhang X, Tian Q, et al. 2022. Swin-Unet: Unet-like pure transformer for medical image segmentation// *Computer Vision - ECCV 2022 Workshops*. Cham: Springer: 205-218 [DOI: 10.1007/978-3-031-25066-8_9].

Codella N C F, Gutman D, Celebi M E, Helba B, Marchetti M A, Dusza S W, et al. 2017. Skin lesion analysis toward melanoma detection: a challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC) [EB/OL]. [2026-08-11].
<https://arxiv.org/abs/1710.05006> [DOI: 10.48550/arXiv. 1710.05006].

Codella N, Rotemberg V, Tschandl P, Celebi M E, Dusza S, Gutman D, et al. 2019. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the International Skin Imaging Col-

laboration (ISIC) [EB/OL]. [2026-08-11].
<https://arxiv.org/abs/1902.03368> [DOI: 10.48550/arXiv. 1902.03368].

Hu G, Zhu W, Liao X and Li Q. 2025. WA-NET: enhanced boundary-aware segmentation of skin lesions via frequency-spatial feature fusion and attention-guided edge refinement. *Scientific Reports*, 15: 41598 [DOI: 10.1038/s41598-025-25583-3].

Huang X, Ma Y, Mei X, Wu Z, Sun M and She Q. 2025. Lesion boundary detection for skin lesion segmentation based on boundary sensing and CNN-transformer fusion networks. *Artificial Intelligence in Medicine*, 167: 103190 [DOI: 10.1016/j.artmed.2025.103190].

Kervadec H, Bouchtiba J, Desrosiers C, Granger E, Dolz J and Ben Ayed I. 2019. Boundary loss for highly unbalanced segmentation// *Proceedings of the 2nd International Conference on Medical Imaging with Deep Learning*. London: PMLR, 102: 285-296.

Li H, Liu J J and Xiao L. 2024. Dual-encoder global-local cross-attention network for medical image segmentation. *Journal of Image and Graphics*, 29(11): 3462-3475 (李赫, 刘建军, 肖亮. 2024. 融合交叉注意力与双编码器的医学图像分割. *中国图象图形学报*, 29(11): 3462-3475) [DOI: 10.11834/jig.230705].

Liu S J and Tong M J. 2026. Swin-EdgeNet: medical image segmentation with multi-scale edge enhancement based on transformer [J/OL]. *Journal of Image and Graphics*, XXXX: 1-14 (刘善洁, 童孟军. 2026. Swin-EdgeNet:基于Transformer的多尺度边缘增强医学图像分割模型 [J/OL]. *中国图象图形学报*, XXXX: 1-14 [DOI: 10.11834/jig.250592].

Liu X, Gao P, Yu T, Wang F and Yuan R Y. 2025. CSWin-UNet: transformer UNet with cross-shaped windows for medical image segmentation. *Information Fusion*, 113: 102634 [DOI: 10.1016/j.inffus.2024.102634].

Mendonça T, Ferreira P M, Marques J S, Marcal A R S and Rozeira J. 2013. PH²: a dermoscopic image database for research and benchmarking//2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Osaka: IEEE: 5437-5440 [DOI: 10.1109/EMBC.2013.6610779].

Mirikharaji Z, Abhishek K, Bissoto A, Barata C, Avila S, Valle E, et al. 2023. A survey on deep learning for skin lesion segmentation. *Medical Image Analysis*, 88: 102863 [DOI: 10.1016/j.media.2023.102863].

Munia A A, Abdar M, Hasan M, Jalali M S, Banerjee B, Khosravi A, et al. 2025. Attention-guided hierarchical fusion U-Net for uncertainty-driven medical image segmentation. *Information Fusion*, 115: 102719 [DOI: 10.1016/j.inffus.2024.102719].

Oktay O, Schlemper J, Le Folgoc L, Lee M J, Heinrich M P, Misawa K, et al. 2018. Attention U-Net: learning where to look for the pancreas [EB/OL]. [2026-08-11].
<https://arxiv.org/abs/1804.03999> [DOI: 10.48550/arXiv. 1804.03999].

Qiu P, Yang J, Kumar S, Ghosh S S and Sotiras A. 2026. AgileFormer;

- spatially agile and scalable transformer for medical image segmentation. *Biomedical Signal Processing and Control*, 112: 108842 [DOI: 10.1016/j.bspc.2025.108842].
- Qureshi I, Yan J, Abbas Q, Shaheed K, Riaz A B, Wahid A, et al. 2023. Medical image segmentation using deep semantic-based methods: a review of techniques, applications and emerging trends. *Information Fusion*, 90: 316-352 [DOI: 10.1016/j.inffus.2022.09.031].
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//*Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*. Cham: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28].
- Ruan J, Xiang S, Xie M, Liu T and Fu Y. 2022. MALUNet: a multi-attention and light-weight UNet for skin lesion segmentation//2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Las Vegas: IEEE: 1150-1156 [DOI: 10.1109/BIBM55620.2022.9995040].
- Ruan J, Xie M, Gao J, Liu T and Fu Y. 2023. EGE-UNet: an efficient group enhanced UNet for skin lesion segmentation//*Medical Image Computing and Computer Assisted Intervention - MICCAI 2023*. Cham: Springer: 481-490 [DOI: 10.1007/978-3-031-43901-8_46].
- Shaker A, Maaz M, Rasheed H, Khan S, Yang M H and Khan F S. 2024. UNETR++: delving into efficient and accurate 3D medical image segmentation. *IEEE Transactions on Medical Imaging*, 43 (9): 3377-3390 [DOI: 10.1109/TMI.2024.3398728].
- Shi J, Wang T T, Zhu Z Q, Zhao M F, Wang B X and An H. 2025. Deep learning-based medical image segmentation methods. *Journal of Image and Graphics*, 30(6): 2161-2186 (石军, 王天同, 朱子琦, 赵敏帆, 王炳勋, 安虹. 2025. 基于深度学习的医学图像分割方法综述. *中国图象图形学报*, 30(6): 2161-2186) [DOI: 10.11834/jig.240467].
- Valanarasu J M J and Patel V M. 2022. UNeXt: MLP-based rapid medical image segmentation network//*Medical Image Computing and Computer Assisted Intervention - MICCAI 2022*. Cham: Springer: 23-33 [DOI: 10.1007/978-3-031-16443-9_3].
- Wang Q, Wu B, Zhu P, Li P, Zuo W and Hu Q. 2020. ECA-Net: efficient channel attention for deep convolutional neural networks//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE: 11534-11542 [DOI: 10.1109/CVPR42600.2020.01155].
- Wang R, Lei T, Cui R, Zhang B, Meng H and Nandi A K. 2022. Medical image segmentation using deep learning: a survey. *IET Image Processing*, 16(5): 1243-1267 [DOI: 10.1049/ipr2.12419].
- Wu R, Liu Y, Ning G, Liang P and Chang Q. 2025. UltraLight VM-UNet: parallel Vision Mamba significantly reduces parameters for skin lesion segmentation. *Patterns*, 6 (11): 101298 [DOI: 10.1016/j.patter.2025.101298].
- Xu J and Tong L. 2024. LB-UNet: a lightweight boundary-assisted UNet for skin lesion segmentation//*Medical Image Computing and Computer Assisted Intervention - MICCAI 2024*. Cham: Springer: 361-371 [DOI: 10.1007/978-3-031-72114-4_35].
- Zhang D, Huang H, Ma Y, Huang B C and Lu W P. 2024. Prostate MR image segmentation network with edge information enhancement. *Journal of Image and Graphics*, 29(3): 755-767 (张蝶, 黄慧, 马燕, 黄丙仓, 陆炜平. 2024. 基于边缘信息增强的前列腺MR图像分割网络. *中国图象图形学报*, 29(3): 755-767) [DOI: 10.11834/jig.230338].
- Zhang H, Qiu D W, Feng Y B and Liu J. 2022. Improved U-Net models and its applications in medical image segmentation: a review. *Laser & Optoelectronics Progress*, 59(2): 55-71 (张欢, 仇大伟, 冯毅博, 刘静. 2022. U-Net模型改进及其在医学图像分割上的研究综述. *激光与光电子学进展*, 59(2): 55-71) [DOI: 10.3788/LOP202259.0200005].
- Zhang M, Yu Y, Jin S, Gu L, Ling T and Tao X. 2024. VM-UNET-V2: rethinking Vision Mamba UNet for medical image segmentation//*Bioinformatics Research and Applications: 20th International Symposium, ISBRA 2024*. Singapore: Springer: 335-346 [DOI: 10.1007/978-981-97-5128-0_27].
- Zhou Z, Siddiquee M M R, Tajbakhsh N and Liang J. 2018. UNet++: a nested U-Net architecture for medical image segmentation//*Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer: 3-11 [DOI: 10.1007/978-3-030-00889-5_1].

作者简介

刘扬,女,硕士研究生,主要研究方向为医学图

像分割. E-mail:3334917591@qq.com

邱云飞,男,教授,主要研究方向为数据挖掘和智能信息处

理. E-mail:7415575@qq.com

韩雪峰,男,博士,研究员,主要从事智慧矿山建设的研究与

应用. E-mail:zyd020311@outlook.com