计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
[doi: 10.15888/j.cnki.csa.010130] [CSTR: 32024.14.csa.010130]
©中国科学院软件研究所版权所有.
E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

FECNet: 融合特征增强与对比语义引导的皮肤病灶图像分割模型①

张继进, 陈平华, 徐健皓

(广东工业大学 计算机学院, 广州 510006)
通信作者: 陈平华, E-mail: phchen@gdut.edu.cn

摘　要: 皮肤病灶图像分割在智能诊断与疗效评估中具有重要意义, 然而常面临病灶形态复杂、边界模糊及对比度低等挑战. 为此, 提出一种基于混合架构的高效分割模型 FECNet (feature enhancement and contrastive semantic network), 设计多个关键模块提升模型在复杂病灶场景下的结构表达、边界建模和语义判别能力. 特征增强模块强化多层级局部纹理与全局结构的建模能力, 多尺度特征融合模块将增强后的特征进行跨层级整合, 形成更加全面且具有结构一致性的表征, 从而有效应对病灶形态多变的问题; 特征解耦模块显式区分前景、背景与不确定区域, 增强边界表达一致性; 进一步引入对比语义上下文调制模块, 动态建模语义差异, 有效提升前景响应并抑制背景干扰, 从而增强模型在低对比复杂场景下的判别能力. 实验结果表明, FECNet 在多个公开数据集上均取得领先性能, 尤其在结构模糊和对比度低的图像中表现出更高的分割精度与鲁棒性.

关键词: 皮肤病灶分割; 特征增强; 特征解耦; 对比语义上下文调制

引用格式: 张继进,陈平华,徐健皓.FECNet: 融合特征增强与对比语义引导的皮肤病灶图像分割模型.计算机系统应用. http://www.c-s-a.org.cn/1003-3254/10130.html

FECNet: Skin Lesion Image Segmentation Model with Feature Enhancement and Contrastive Semantic Guidance

ZHANG Ji-Jin, CHEN Ping-Hua, XU Jian-Hao

(School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Skin lesion segmentation plays a crucial role in intelligent diagnosis and therapeutic assessment. However, it remains challenging due to the complex lesion morphology, blurry boundaries, and low contrast between lesions and surrounding skin. To address these challenges, we propose an efficient hybrid-architecture segmentation model, termed FECNet (feature enhancement and contrastive semantic network). FECNet integrates several key modules to enhance structural representation, boundary modeling, and semantic discrimination in complex lesion scenarios. A feature enhancement module strengthens the modeling of multi-level local textures and global structures. The multi-scale feature fusion module aggregates enhanced features across different semantic levels, forming a more comprehensive and structurally consistent representation, effectively handling large variations in lesion appearance. The feature decoupling module explicitly separates foreground, background, and uncertainty regions to improve boundary coherence. Furthermore, a contrastive semantic context modulation module dynamically captures semantic discrepancies, enhancing foreground activation while suppressing background interference, thereby improving discriminative ability in low-contrast or visually confusing cases. Experimental results on multiple public datasets demonstrate that FECNet achieves state-of-the-art performance, showing superior segmentation accuracy and robustness, especially on images with fuzzy structures

① 基金项目: 广东省重点领域研发计划 (2023B1111050010)
收稿时间: 2025-10-16; 修改时间: 2025-11-03; 采用时间: 2025-11-17; csa 在线出版时间: 2026-03-13

and low contrast.

Key words: skin lesion segmentation; feature enhancement; feature decoupling; contrastive semantic context modulation

皮肤疾病在全球范围内具有较高的发病率. 据世界卫生组织 (world health organization, WHO) 报告, 超过 1/3 的人群在其一生中将经历不同类型的皮肤病困扰, 既包括致死率较高的恶性黑色素瘤, 也涵盖严重影响生活质量的慢性疾病, 如银屑病与湿疹[1]. 皮肤病种类繁多, 临床表现复杂多样, 其识别与筛查过程高度依赖医生经验, 存在主观性强、效率低等问题, 尤其在基层医疗机构中表现更为突出. 因此, 开展基于影像的皮肤病智能诊断技术研究, 对提升诊断效率和早期干预能力具有重要意义[2].

早期皮肤病诊断主要依赖显微镜等传统工具, 自 20 世纪 80 年代以来, 随着计算机技术的发展, 计算机辅助诊断 (computer-assisted diagnosis, CAD) 逐渐应用于皮肤病学, 有效提升了病灶识别的准确性[3]. 其中, 病灶图像分割作为 CAD 的关键前置步骤, 承担着从原始图像中提取结构化信息的任务, 是实现病变区域精准定位、边界量化及后续诊疗决策的基础[4].

随着深度学习的发展, 基于卷积神经网络 (convolutional neural network, CNN) 的皮肤病灶分割方法取得了显著进展[5]. 然而, 皮肤病灶图像往往同时具有形态尺度多样、边界模糊及低对比度等特征, 因此模型需要在多尺度结构建模、边界敏感性保持以及语义判别鲁棒性这 3 方面具备良好的综合能力. 针对这些挑战, 现有研究主要沿多尺度特征建模、边界增强以及注意力或对比学习调制机制等方向展开改进.

针对形态复杂的病灶分割任务, 典型方法如 U-Net 系列及其后续改进网络 (如 CPFNet[6]、MSRF-Net[7] 等) 通过特征金字塔、空洞卷积或多分支结构扩大感受野, 以实现全局与局部信息的联合建模. 尽管在全局-局部特征互补方面表现突出, 但此类方法的融合策略通常依赖预设尺度层级与固定融合算子, 缺乏对输入内容与语义复杂度的自适应能力. 当病灶形态复杂、细粒度纹理与大尺度结构共存时, 层级特征融合易导致信息冗余与语义失衡, 限制模型对复杂结构的表征精度. 因此, 在增强阶段实现细节与语义的协同建模, 在融合阶段实现内容自适应的多尺度对齐, 是进一步提升复杂病灶分割性能的关键.

针对边界模糊与组织过渡带的问题, 已有研究 (如 Inf-Net[8]、CFA-Net[9]等) 多在网络末端引入边缘监督或反向注意力机制, 通过边缘分支增强模型对边界的响应能力. 虽然这类方法在预测结果层面能有效提升轮廓清晰度, 但由于边界约束主要作用于输出空间, 中间特征层对模糊区域的表征依旧隐含, 模型在灰度渐变区对边界的结构理解不足, 易出现边界断裂或过度平滑的现象. 因此, 在特征生成阶段显式表达边界不确定性, 让模型在学习过程中形成“软边界”几何表示, 成为进一步提高边界感知能力的重要方向.

在低对比度或前后景灰度接近的场景下, 常见方法往往借助注意力机制或对比学习框架来增强前景显著性. 注意力机制通过特征加权突出高响应区域, 对比学习则通过拉近同类特征、拉远异类特征以提高区分度. 这些方法在多数情况下能够提升分割性能, 但其权重或对比目标多由特征自身统计信息生成, 对前景与背景的语义差异利用较为间接. 当病灶与背景灰度相近或纹理相似时, 模型可能受到伪结构影响, 注意力调制不稳定, 导致前景响应减弱. 若能基于语义解耦结果实现条件化调制, 在滤波过程中显式引入“前景增强→背景抑制”的方向性, 将有助于提升模型在低对比场景下的判别鲁棒性.

基于上述分析, 本文提出了一种高效的皮肤病灶分割模型 FECNet (feature enhancement and contrastive semantic network). 模型在保持全局依赖建模优势的同时, 更注重边界敏感性与前后景语义区分能力, 以提升模型在低对比度与模糊区域中的判别鲁棒性. FECNet 的主要工作为以下 3 个方面.

(1) 特征增强模块 (feature enhancement module, FEM) 与多尺度特征融合 (multi-scale feature fusion, MFF) 模块: 针对复杂形态病灶, FEM 通过浅层局部增强路径提升纹理细节感知, 并在深层引入全局语义路径以增强对大尺度与不规则结构的建模能力; MFF 位于增强阶段之后, 将来自不同层级的特征进行空间对齐后进行拼接融合, 在保留各层特征差异性的同时形成更全面的多尺度特征表征, 从而提升模型对复杂病灶形态和结构多样性的适应能力.

(2) 特征解耦模块 (feature decoupling module, FDM): 针对边界模糊问题, FDM 显式划分前景、背景与不确定区域, 从而实现边界区域的显式建模, 提升边界连续性与结构分离能力.

(3) 对比语义上下文调制 (contrastive semantic context modulation, CCM) 模块: 针对低对比度场景, 利用解耦后的前/背景语义特征动态生成调制卷积核, 突出目标区域响应、抑制背景干扰, 从而增强模型在灰度接近区域中的判别鲁棒性.

1 相关工作

医学图像分割作为辅助诊断、疗效评估与术前规划等任务的关键基础, 其核心目标在于从复杂图像中精准提取病灶、器官等感兴趣区域. 深度学习技术的兴起使得 CNN 成为主流建模工具. 其中, U-Net[10]通过对称式编码-解码结构与跳跃连接, 有效融合多层级语义与空间信息, 奠定了现代分割网络的基础架构. 此后, UNet++[11]、ResUNet-a[12]等方法在特征路径设计与深层信息建模方面持续优化, 提升了模型对复杂结构的适应能力. 然而, CNN 本质上依赖局部感受野, 难以捕捉远程依赖关系, 尤其在应对结构模糊或尺度变化显著的目标时, 表现出局部建模能力不足的问题. 为缓解上述限制, 研究者开始引入 Transformer 架构, 如 Trans-UNet[13]、Swin-Unet[14]等, 借助自注意力机制提升全局建模能力. 这些方法在语义整合方面取得一定进展, 但由于缺乏空间归纳偏置, 往往难以保持边界的稳定响应. 近期, 状态空间模型 (如 Mamba) 凭借其在长序列建模中的高效性逐渐引起关注, MambaVision[15]更是融合 CNN、Transformer 与 Mamba 构建混合编码器, 为医学图像分割任务提供了一种全新的全局-局部协同建模范式.

皮肤病灶图像分割任务面临更为复杂的结构与纹理特征挑战. 病变区域在图像中可能呈现出细小斑点、曲折条带或大范围扩散等多样形态, 表现出尺度跨度大、边界不规则、纹理剧烈变化等特性. 这对模型的尺度感知与结构表达能力提出更高要求. 为此, 多尺度建模策略被广泛引入. 例如, CPFNet[6]通过上下文金字塔实现多尺度语义融合, MSRF-Net[7]采用主辅双分支扩展感受野结构, MSF-Net[16]利用空洞卷积捕获长程上下文信息. 尽管上述方法在提升模型多尺度建模能力方面有所贡献, 但大多依赖静态融合策略, 缺乏对局部复杂结构的动态适配能力, 难以在小尺度边缘纹理与大尺度语义轮廓之间实现良好平衡, 容易出现目标区域响应不足或边界不连续等问题. 此外, 特征冗余叠加也会加重建模负担, 影响整体分割效果.

除尺度变化外, 皮肤图像中的“软边界”现象亦对模型提出了更高要求. 由于皮肤组织结构常存在自然过渡, 加之成像噪声与灰度模糊等因素干扰, 病灶边界通常呈现出非明确的渐变特征, 导致模型难以准确判定真实轮廓. 为了增强边界感知能力, 部分研究通过显式建模边缘或引入辅助监督机制来提升响应质量. 如 Inf-Net[8]结合边缘与反向注意力机制以提升结构清晰度, BiANet[17]采用双边提取模块强化边界信息一致性, CFA-Net[9]通过独立预测路径生成边界感知特征. 尽管这些策略通过显式引入边界监督来增强模型对边界的敏感性, 但这类方法更多依赖外部约束, 并未从机制层面提升模型自身在模糊区域中抑制不确定性的能力.

进一步地, 皮肤图像还普遍存在前景病灶与背景组织灰度与纹理相似度高的问题. 在光照不均、色素沉积或反射干扰等因素作用下, 前后景对比度显著降低, 致使模型难以区分病变区域, 易出现漏检、误判或前景退化等现象. 为此, 一些研究尝试从注意力机制或对比策略角度增强前景表征. 如 Inf-Net[8]通过反向注意力抑制背景干扰, BASC-Net[18]引入边界感知对比损失强化语义区分性, BDG-Net[19]利用边界分布引导策略改善对弱显著区域的感知能力. 但这类方法多依赖显式边界或局部差异建模, 缺乏对前后景语义对立关系的整体刻画, 尤其在纹理连续或灰度趋同区域易被背景伪结构误导, 削弱了模型的分割准确性与鲁棒性.

2 FECNet 模型结构

如图 1 所示, FECNet 采用四阶段分层架构, 旨在应对皮肤病灶分割中形态复杂、边界模糊和低对比度等挑战. 编码器部分基于 MambaVision 设计, 前两阶段利用卷积块提取高分辨率下的边缘与纹理特征, 后两阶段引入 MambaVision Mixer 与多头自注意力模块, 实现局部感知与全局依赖的协同建模. 为提升模型对复杂病灶的适应性, 本文设计了特征增强模块 (FEM) 与多尺度特征融合 (MFF) 模块. 其中, FEM 采用局部-全局双路径结构, 在浅层突出边缘与纹理, 在深层强化大尺度与模糊结构的表征; MFF 对各阶段增强特征进

行空间与语义对齐, 实现兼具局部细节与全局上下文的统一表达. 针对边界模糊与区域混淆, 模型进一步引入特征解耦模块 (FDM), 将融合特征划分为前景、背景与不确定区域, 并结合互补约束与联合损失, 增强边界连续性与结构判别力. 此外, 对比语义上下文调制模块 (CCM) 基于前后景差异动态生成语义调制卷积核, 突出目标响应、抑制背景干扰, 在灰度接近与弱边界区域显著提升判别能力. 最终, FECNet 采用双解码器结构融合不同语义层次的 CCM 输出, 生成兼具全局一致性与局部精细度的分割结果.

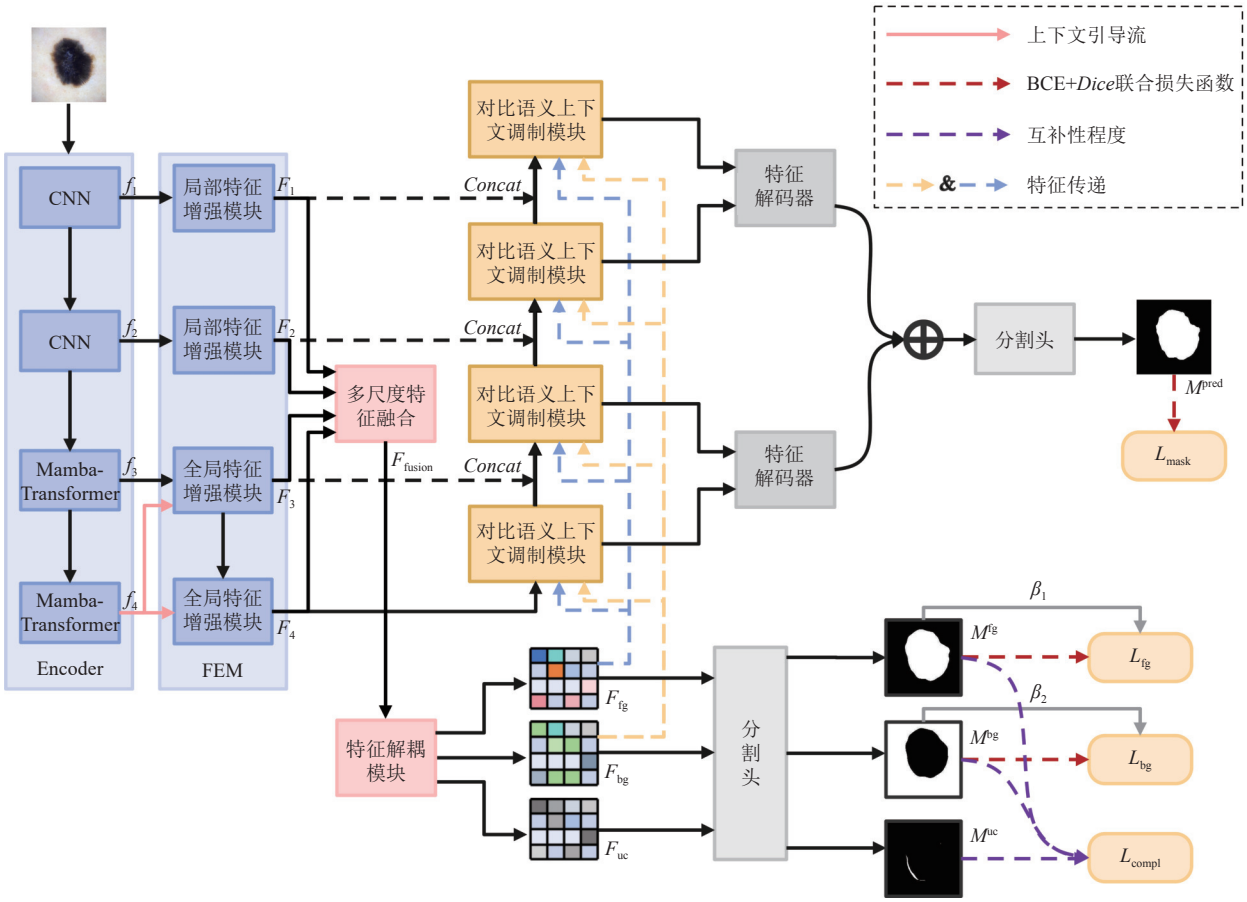


图 1 FECNET 网络总体结构

2.1 特征增强模块

为了应对皮肤病灶在形态复杂与尺度多变上的显著异质性问题, 本文提出的 FECNet 构建于具备分层建模能力的 MambaVision 编码器之上. 皮肤病灶通常表现为尺寸差异大、结构分布不规则、跨尺度形态复杂等特征, 传统卷积网络虽然在局部纹理建模方面具有优势, 但受限于感受野, 难以捕捉大尺度病灶的全局依赖. 而纯粹的注意力机制虽能建模长程关系, 却往往忽视细粒度结构与空间归纳偏置.

MambaVision 针对上述问题提出了一种 CNN+Transformer+Mamba 的混合架构, 在保持高效性的同时兼具局部感知与远程依赖建模能力. 其整体设计为四阶段分层结构 (如图 2): 前两阶段采用基于残差的 CNN 模块, 快速提取高分辨率图像中的纹理、边缘等细粒度特征, 适合小尺度、不规则病灶的建模; 后两阶段引入核心的 MambaVision Mixer 与多头自注意力模块. 其中, MambaVision 对原始 Mamba 进行关键性重构, 用对称一维卷积替代因果卷积以支持并行双向建模, 并增加并行非状态空间分支以保留局部空间归纳偏置. 两个分支分别获取序列依赖与空间特征, 随后在通道维融合, 再通过 Transformer 补充长程语义一致性, 从而实现局部–全局的协同建模.

得益于上述设计, MambaVision 在多个视觉任务中均表现优异, 但在皮肤病灶分割中仍存在局限. 面对

弥漫扩散、纹理断裂、跨尺度结构等高度异质的复杂形态时, 深层特征易出现响应弱化、语义不稳定与区域一致性不足. 为此, 本文设计了特征增强模块 (FEM), 来强化对复杂病灶形态的建模能力. FEM 包含局部增强路径与全局增强路径: 前者嵌入编码器前两阶段, 提升对小尺度、多方向细节结构的感知; 后者部署于后两阶段, 结合动态建模与全局依赖机制, 增强对大尺度扩散、纹理断裂及不规则分布的适应性. 两条路径协同作用, 使 FEM 在多尺度、多层次上显著增强了特征表征能力, 从而提升模型对复杂病灶的整体表达力与鲁棒性.

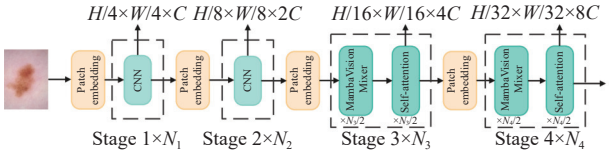


图 2 MambaVision 编码器

2.1.1 局部特征增强模块

病灶形态往往呈多方向延展, 单一路径卷积难以同时捕获横纵向结构特征; 同时, 不同病灶在尺度上存在显著差异, 细粒度小结构与大范围区域并存, 对模型感受野的适应性提出了更高要求. 为此, FECNet 在编码器前两阶段嵌入如图 3 所示的局部特征增强模块 (local feature enhancement module, LFEM), 以提升对空间细节的建模能力. 该模块采用多分支卷积结构结合空洞卷积设计, 增强对多方向、多尺度纹理及边缘信息的感知, 从而提高模型对复杂局部结构的表征精度.

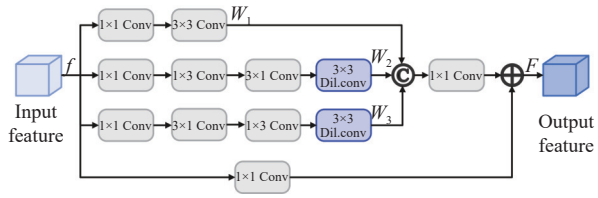


图 3 局部特征增强模块 (LFEM)

LFEM 采用多分支并行卷积结构, 在保持计算效率的同时, 从不同方向和尺度上增强局部特征 $f \in \mathrm{R}^{C\times H\times W}$ (即 f_i, i=1,2) 的表示能力. 其中, 主分支首先通过 1×1 卷积对 f 进行通道压缩, 并接入标准 3×3 卷积以提取基础的局部上下文特征; 另一组方向敏感分支则在对 f 进行通道压缩后分别采用不同顺序的可分离卷积, 其中一条分支先执行 1×3 卷积再执行 3×1 卷积, 另一条分支则先执行 3×1 卷积再执行 1×3 卷积. 虽然两者的理论感受野均等价于 3×3 卷积, 但根据空间可分离卷积理论, 不同的卷积顺序会在中间激活特征上引入方向敏感差异 (可参考 Sobel 算子的可分离实现), 并在 BN 与 *ReLU* 等非线性操作的作用下得到进一步放大, 从而使前一分支更倾向于对垂直边缘与横向纹理响应, 而后一分支则更适合捕获水平边缘与纵向纹理. 此外, 在两条方向敏感分支的末端均叠加了 3×3 空洞卷积 (图 3 中的 Dil.conv), 以在保持参数量可控的同时有效扩大感受野, 从而提升跨尺度结构的建模能力. 具体如式 (1)–式 (3):

$$W_1 = f_{3\times3}^{\mathrm{conv}}\left(f_{1\times1}^{\mathrm{conv}}(f)\right) \tag{1}$$

$$W_2 = f_{3\times3}^{\mathrm{diconv}}\left(f_{3\times1}^{\mathrm{conv}}\left(f_{1\times3}^{\mathrm{conv}}\left(f_{1\times1}^{\mathrm{conv}}(f)\right)\right)\right) \tag{2}$$

$$W_3 = f_{3\times3}^{\mathrm{diconv}}\left(f_{1\times3}^{\mathrm{conv}}\left(f_{3\times1}^{\mathrm{conv}}\left(f_{1\times1}^{\mathrm{conv}}(f)\right)\right)\right) \tag{3}$$

将 3 路特征在通道维度拼接, 以充分保留各分支在方向性与尺度上的互补特征, 使网络能够同时感知水平、垂直及多尺度纹理信息. 然后统一通过 1×1 卷积进行整合, 并与输入特征 f 进行加权残差连接, 最终通过 *ReLU* 激活函数得到模块输出:

$$F = ReLU\left(Conv_{1\times1}\left(Concat(W_1, W_2, W_3)\right)\cdot\alpha + Shortcut(f)\right) \tag{4}$$

其中, α 为可学习缩放系数, $Shortcut(f)$ 为输入特征 f 的 1×1 映射, 用于维度对齐.

该结构在保留边界信息的同时显著增强了模型对细节纹理的响应能力, 有效提升了在复杂小尺度结构下的判别性能.

2.1.2 全局特征增强模块

皮肤病灶形态复杂, 既包含大范围弥漫区域, 也伴随细小结构和不规则边界, 这对模型的语义连续性与结构敏感性提出双重挑战. 现有深层编码器 (如 MambaVision) 虽通过状态空间与注意力机制强化了全局依赖建模, 但缺乏空间归纳偏置, 导致在复杂病灶下难以保持稳定且一致的特征表征.

基于此, 本文在编码器后两阶段设计如图 4 所示的全局特征增强模块 (global feature enhancement module, GFEM), 以弥补深层特征在结构表达与语义统一建模方面的不足. GFEM 包含 3 类互补单元: 深度可分离卷积 (depthwise convolution, DWConv) 引入轻量化的空间归纳偏置, 增强对细粒度纹理与局部结构的敏感性; 核心的 ContMix 动态卷积机制[20]则通过上下文语义引导生成位置自适应卷积核, 在保持局部归纳

偏置的同时增强远程依赖建模, 尤其适合处理大尺度扩散与不规则结构; 卷积前馈网络 (convolutional feed forward network, ConvFFN) 以空间卷积替代 MLP, 实现更稳定的多尺度语义整合. 通过三者的协同作用, GFEM 有效提升了深层特征的语义一致性与结构完整性, 为后续的边界感知与区域判别提供了更加稳健的特征支撑.

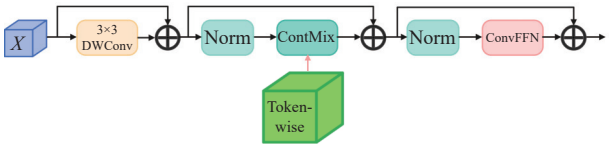


图 4 全局特征增强模块 (GFEM)

该特征增强模块通过局部-全局双路径设计, 从纹理到语义、从细节到上下文, 形成全尺度、全维度的结构建模能力, 为后续融合机制与精细分割提供坚实特征支撑.

2.2 多尺度特征融合模块

不同层级特征在空间分辨率与语义抽象度上存在差异, 单层信息难以全面刻画病灶结构. 为此, 本文提出多尺度特征融合 (MFF) 模块, 在空间与语义两方面对编码器各阶段经 FEM 增强后的特征进行对齐与融合, 从而形成兼具局部细节与全局上下文的统一表征. 该模块不仅有效缓解了病灶形态复杂带来的跨尺度建模难题, 也为后续的特征解耦模块提供了更全面且鲁棒的特征基础.

如图 1 所示, 该模块融合来自前两阶段局部增强路径的细粒度纹理与边界信息, 以及后两阶段全局增强路径的长程语义与上下文信息. 通过整合 4 类特征, 模型能够同时感知小尺度不规则病灶的边缘细节与大尺度模糊结构的上下文表达, 显著提升融合特征的语义完整性与结构敏感性.

在实现过程中, 模块首先对经过特征增强模块输出的增强特征 F_1、F_2、F_3、F_4 进行空间对齐与通道压缩. 考虑到 F_2、F_3、F_4 的空间分辨率逐级降低, 我们采用双线性插值将其上采样至与 F_1 相同的尺度, 得到 $F_2^{\uparrow}$、$F_3^{\uparrow}$、$F_4^{\uparrow}$. 随后, 4 组尺度对齐的特征通过 1×1 卷积压缩通道至统一维度, 分别记为 F'_1、F'_2、F'_3、F'_4, 以降低计算开销并强化关键结构表达.

最终, 4 组压缩后的特征在通道维度拼接后, 再经一层卷积进行融合, 输出统一语义尺度、结构丰富的融合特征张量 F_{fusion}, 表示为:

$$F_{\text{fusion}} = Conv_{1\times1}\left(Concat\left(F'_1, F'_2, F'_3, F'_4\right)\right) \tag{5}$$

该融合特征作为特征解耦模块的唯一输入, 承载了从底层纹理到高层语义的完整结构表达, 有效提升了模型对边缘、纹理及上下文结构的综合感知能力. 该模块的关键价值在于: 弥合深层语义抽象与浅层结构细节之间的表达断层, 显著缓解因单一高层特征输入导致的结构脱敏与边界误判问题, 进一步提升特征解耦阶段的区域建模精度与鲁棒性.

2.3 特征解耦模块

针对皮肤病灶图像中边界模糊的问题, FECNet 设计了特征解耦模块 (FDM), 旨在强化模型对前景、背景与边界过渡区域的区分能力, 从而提升边界结构的判别精度. 如图 1 所示, 该模块接收来自多尺度特征融合模块的输入特征 F_{fusion}, 该特征融合了编码器各阶段提取的局部纹理、长程依赖与上下文语义信息, 为区域划分提供了更为丰富且语义一致的表达. FDM 不依赖显式阈值, 而是通过损失函数驱动网络学习出每个像素在前景、背景、不确定区域间的最优分配, 从而自然形成“前景-过渡-背景”三域划分, 实现边界模糊区域的显式建模. 有效缓解了软边界条件下的结构冲突, 并增强了边界的连续性与定位精度.

对输入特征 F_{fusion}, FDM 通过 3 个卷积分支分别建模前景、背景与不确定区域特征:

$$\begin{cases} F_{\text{fg}} = Conv_{\text{fg}}(F_{\text{fusion}}) \\ F_{\text{bg}} = Conv_{\text{bg}}(F_{\text{fusion}}) \\ F_{\text{uc}} = Conv_{\text{uc}}(F_{\text{fusion}}) \end{cases} \tag{6}$$

并通过 3 个辅助预测头输出对应的概率掩码:

$$\begin{cases} M^{\text{fg}} = \sigma\left(W_{\text{fg}} * F_{\text{fg}}\right) \\ M^{\text{bg}} = \sigma\left(W_{\text{bg}} * F_{\text{bg}}\right) \\ M^{\text{uc}} = \sigma\left(W_{\text{uc}} * F_{\text{uc}}\right) \end{cases} \tag{7}$$

其中, σ 表示 Sigmoid 激活函数, * 表示卷积操作. 理想情况下, 这 3 个掩模的每个像素处的值应该是明确的 0 或 1, 表明像素所属的类别是确定的. 而且, 这 3 个特征图应该表现出互补关系, 即对于任意像素指数 i, 这 3 个掩码 M^{fg}、M^{bg} 和 M^{uc} 应该满足:

$$\begin{cases} M_i^{\text{fg}}, M_i^{\text{bg}}, M_i^{\text{uc}} \in [0,1] \\ M_i^{\text{fg}} + M_i^{\text{bg}} + M_i^{\text{uc}} = 1 \end{cases} \tag{8}$$

因此, 我们的优化目标是让 M^{fg} 和 M^{bg} 分别逼近真实标签 Y 及其取反值 $(1-Y)$. 同时, M^{fg} 和 M^{bg} 应该近似

地满足前述的互补性约束. 考虑到辅助头部的输出 M^{fg}、M^{bg} 和 M^{uc} 实际上是概率分布, 我们设计了一个简单但有效的损失函数来量化它们的互补程度:

$$L_{compl}=\frac{1}{N}\sum_{i=1}^{N}\left(M_i^{fg}\cdot M_i^{bg}+M_i^{fg}\cdot M_i^{uc}+M_i^{bg}\cdot M_i^{uc}\right) \quad (9)$$

其中, M_i^{fg}、M_i^{bg}、M_i^{uc} 表示第 i 个像素分别被预测为前景、背景和不确定区域的概率. 总体而言, L_{compl} 用于强化来自 3 个分支预测结果之间的互补性, 确保每个像素都被唯一地归类为前景、背景或不确定区域之一.

同时, 为了约束前景和背景掩码与真实标签之间的一致性, 采用 BCE+*Dice* 联合损失函数 L_{mask}. 其中, 二值交叉熵 (binary cross entropy, BCE) 用于像素级概率误差度量, *Dice* 损失则提升了对小目标区域的判别能力, 二者定义如下:

$$L_{BCE}(Y,M)=-\frac{1}{N}\sum_{i=1}^{N}(Y_i\log M_i+(1-Y_i)\log(1-M_i)) \quad (10)$$

$$L_{Dice}(Y,M)=1-2\cdot\frac{\sum_{i=1}^{N}M_iY_i}{\sum_{i=1}^{N}M_i+\sum_{i=1}^{N}Y_i} \quad (11)$$

$$L_{mask}(Y,M)=L_{BCE}(Y,M)+L_{Dice}(Y,M) \quad (12)$$

其中, i 表示所有像素的索引, M_i 为预测值, Y_i 为真实标签, N 表示总像素数. 分别对前景和背景区域应用该损失:

$$L_{fg}=L_{mask}\left(Y,M^{fg}\right),\ L_{bg}=L_{mask}\left(1-Y,M^{bg}\right) \quad (13)$$

考虑到医学图像中小病灶区域占比低、识别难度大, 引入预测面积自适应的惩罚因子 β_1 和 β_2, 分别增强前景与背景区域损失权重, 最终的特征解耦部分的损失函数为:

$$\beta_1=\frac{1}{\tanh\left(\frac{\sum_{i=1}^{N}M_i^{fg}}{N}\right)},\beta_2=\frac{1}{\tanh\left(\frac{\sum_{i=1}^{N}M_i^{bg}}{N}\right)} \quad (14)$$

$$L_{Decoupling}=\beta_1\cdot L_{fg}+\beta_2\cdot L_{bg}+L_{compl} \quad (15)$$

通过上述设计, 特征解耦模块有效分离边界模糊区域中的前背景过渡信息, 提升了模型在结构复杂、弱边界条件下的鲁棒性和判别能力.

2.4 对比语义上下文调制模块

由于皮肤病灶图像常受光照不均、色素沉积及组织反射等因素影响, 前景与背景灰度分布相近, 导致整体对比度较低. 针对这一问题, 本文设计了对比语义上下文调制 (CCM) 模块, 其结构如图 5 所示. 该模块的设计灵感主要来源于 OverLoCK 的动态卷积机制[20]与 ConDSeg 中的对比度驱动特征聚合机制[21], 本文在此基础上融合二者优点, 构建出一种兼具结构感知与语义对比能力的上下文增强方式. 具体来说, 将在特征增强模块输出的 4 个特征 F_1、F_2、F_3、F_4 在对比特征 F_{fg} 与 F_{bg} 的引导下实现多层次特征融合.

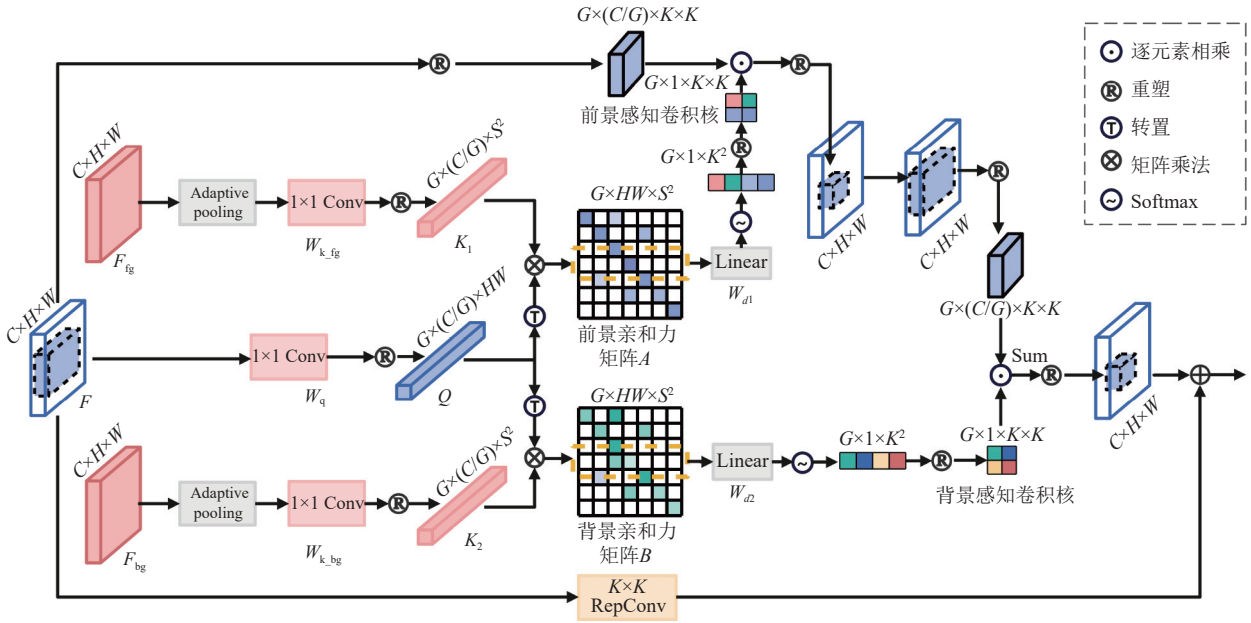


图 5 CCM 结构

如图 5 所示, CCM 接收特征图 $F \in \mathrm{R}^{C\times H\times W}$ (即 F_i, i=1, 2, 3, 4), 并利用解耦模块提供的前景特征 F_{fg} 与背景特征 F_{bg} 作为语义引导信号. 首先, 通过自适应平均池化与 1×1 卷积对 F_{fg}、F_{bg} 这两类特征进行压缩, 提取其全局语义表示:

$$K_1 = Re\left(W_{\mathrm{k_fg}} Pool\left(F_{\mathrm{fg}}\right)\right),\ K_2 = Re\left(W_{\mathrm{k_bg}} Pool\left(F_{\mathrm{bg}}\right)\right) \tag{16}$$

其中, $W_{\mathrm{k_fg}}$ 和 $W_{\mathrm{k_bg}}$ 表示 1×1 卷积层, $Re(\cdot)$ 表示重塑操作, K_1、K_2 表示通过自适应平均池化将 F_{fg}、F_{bg} 聚集到 $S\times S$ 个区域中心. 随后将输入特征图 F 映射为查询特征 $Q = Re\left(W_q F\right)$, 我们将 K_1、K_2 和 Q 的通道均匀划分为 G 组以获得 $\{K_1^g\}_{g=1}^G$、$\{K_2^g\}_{g=1}^G$ 和 $\{Q^g\}_{g=1}^G$, 使得 K_1^g、$K_2^g \in \mathrm{R}^{\frac{C}{G}\times HW}$ 和 $Q^g \in \mathrm{R}^{\frac{C}{G}\times S^2}$, 这些组类似于多头注意力中的“头”. 每个 K_1^g、K_2^g 和 Q^g 被展平为二维矩阵后, 计算其矩阵乘积以获得 G 个前景和背景亲和力矩阵, 分别为:

$$\{A^g\}_{g=1}^G = \left\{(Q^g)^{\mathrm{T}} K_1^g\right\}_{g=1}^G,\ \{B^g\}_{g=1}^G = \left\{(Q^g)^{\mathrm{T}} K_2^g\right\}_{g=1}^G \tag{17}$$

其中, $A^g \in \mathrm{R}^{HW\times S^2}$ 表示第 g 组的前景亲和力矩阵, $B^g \in \mathrm{R}^{HW\times S^2}$ 表示第 g 组的背景亲和力矩阵. 亲和力矩阵 A^g 的第 i 行表示在 Q^g 中的第 i 个特征单元与所有 K_1^g 中特征单元的相似度, 亲和力矩阵 B^g 也同理.

为了生成更强健的特征表示, 我们定义了 G 个空间变化的 $K\times K$ 动态卷积核. 首先使用可学习的线性层 W_{d1}、$W_{d2} \in \mathrm{R}^{S^2\times K^2}$ 来聚合上述按特征单元存储的亲和力矩阵行向量:

$$FK^g = Softmax(A^g W_{d1}),\ BK^g = Softmax(B^g W_{d2}) \tag{18}$$

其中, $Softmax$ 用于对聚合后的亲和度进行归一化. 值得注意的是, 每个卷积位置的动态卷积核是由该位置 (即滑动窗口中心) 对应特征单元的亲和力矩阵行向量生成的. 每一行的 FK^g、BK^g 重新塑形为目标卷积核的形状, 即可在特征图各空间位置生成输入相关的卷积核 (如图 5). 卷积过程中, 特征图 F 的通道也被均匀地划分为 G 个组, 同组通道共享相同的动态卷积核.

为强化语义引导的效果, 首先使用前景感知动态卷积核作用于输入特征图, 突出目标区域响应; 然后再通过背景感知动态卷积核调制上述特征, 进一步抑制冗余干扰并提升结构边界清晰度, 得到“前景增强→背景抑制”的特征图 $F_{\mathrm{fg_bg}}$.

此外, 为了提升模型对模糊区域的感知能力, CCM 进一步引入 *DilatedRepConv* 模块, 作用于原始输入特征 X, 以扩展感受野、丰富通道语义多样性. 最终, CCM 的输出为:

$$F_{\mathrm{out}} = F_{\mathrm{fg_bg}} + DilatedRepConv(X) \tag{19}$$

整体而言, CCM 模块通过语义驱动的前/背景动态调制机制, 实现对模糊边界的重点响应与冗余压制, 显著提升模型在低对比度区域的结构判别能力, 是支撑网络细粒度分割性能的关键组件.

在特征解码阶段, 模型采用并行双解码器结构 (如图 6 所示, SRA、CAM 和 SAM 分别代表空间反向注意力、通道注意力和空间注意力模块), 分别接收来自 CCM 模块的不同语义层级特征. 该设计实现了多层次特征的分层建模与逐级还原, 兼顾浅层细节与深层语义的互补表达. 其中, 低层特征具备高空间分辨率与丰富边缘信息, 有助于局部结构重建; 高层特征具有更强的全局感知能力, 支持区域级语义判别. 每个解码器接收两路特征映射, 分别聚焦于细粒度结构还原与语义完整理解, 生成具有不同语义倾向的中间表征. 两路输出在保持语义独立的基础上经逐元素加法融合, 强化跨层级协同与结构一致性. 融合结果经 Sigmoid 激活输出分割掩码, 实现目标区域的精确预测.

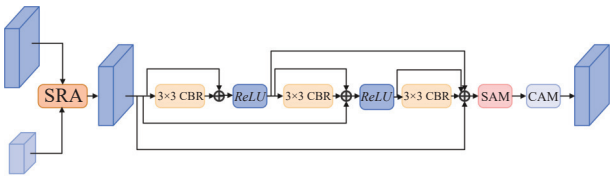


图 6 解码器结构

2.5 损失函数

为全面提升模型在皮肤病灶图像分割中的预测准确性与边界判别能力, FECNet 的损失函数由两部分组成: 特征解耦损失与最终分割掩码的联合损失. 其中, 特征解耦模块用于强化前景、背景与不确定区域之间的判别性, 其损失包括 3 类区域的互补性约束损失及 BCE+*Dice* 联合损失, 第 2.3 节已详细介绍.

在此基础上, 为进一步优化最终分割掩码的整体预测质量, 模型对双解码器融合输出的预测结果也引入了联合监督机制. 具体地, 采用交叉熵损失函数 (cross entropy loss, CE) 与 *Dice* 损失函数 (*Dice* loss) 共同构成最终输出的监督信号. 其中, 交叉熵损失关注像素级

的类别预测误差, 有效优化模型的整体分类能力; 而 *Dice* 损失则对小病灶目标具有更强的敏感性, 提升模型在前景稀疏情况下的分割精度.

最终, 总体损失函数定义如式 (20) 所示:

$$L_{\text{Total}} = L_{\text{Decoupling}} + L_{\text{mask}} \tag{20}$$

该设计通过联合优化区域划分与最终分割输出, 不仅提升了模型对边界模糊区域的鲁棒性, 也增强了整体结构感知与语义一致性, 有效支撑了模型在复杂皮肤病灶图像中的精细化分割性能.

3 实验结果与分析

3.1 数据集

为验证本文所提皮肤病灶分割方法的有效性, 本研究选用国际皮肤图像协作组织 (International Skin Imaging Collaboration, ISIC) 发布的 3 个标准公开数据集: ISIC-2016、ISIC-2017 和 ISIC-2018. 这 3 个数据集均采集自真实临床环境, 涵盖多种类型的皮肤病变图像, 广泛存在边界模糊、结构复杂、前背景对比度低等典型挑战因素, 为皮肤病灶分割模型的性能评估提供了坚实基准.

ISIC-2016 数据集共包含 1 279 张皮肤镜图像, 其中, 900 张用于训练, 379 张用于测试. 该数据集图像质量较高, 但病灶边界模糊严重, 且部分图像前后景对比度低, 对分割模型的边缘定位能力提出了较高要求.

ISIC-2017 数据集包含 2 750 张皮肤镜图像, 原始任务涵盖分类、检测和分割这 3 项任务, 本文仅使用其中的分割部分, 包括 2 000 张训练图像、150 张验证图像和 600 张测试图像. 相比前一版本, 该数据集在病灶类型和成像条件上更为丰富, 部分图像存在毛发遮挡、光照不均和颜色杂乱等干扰因素, 对分割模型的鲁棒性提出了更高要求.

ISIC-2018 数据集采用官方提供的划分方案, 包含 2 594 张训练图像、100 张验证图像和 1 000 张测试图像, 涵盖多种良性与恶性皮肤病灶, 具有明显的形态变异性和低对比度区域, 是评估分割模型精细感知与语义表达能力的重要基准.

选择 ISIC-2016、ISIC-2017 和 ISIC-2018 数据集作为评估平台, 能够在不同层次上检验分割模型在图像边界感知、结构理解和复杂背景适应等方面的综合能力.

3.2 评估指标

为全面评估模型在皮肤病灶分割任务中的性能, 本文采用平均 *Dice* 系数 (mean *Dice* score, mDSC) 与平均交并比 (mean intersection over union, mIoU) 这两种常用的图像分割指标, 分别从区域重叠与边界匹配角度衡量模型的分割准确性.

Dice 系数是一种集合相似度度量函数, 通常用于计算两个样本的相似度, 定义如式 (21):

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \tag{21}$$

其中, A 和 B 分别表示预测区域与真实区域的像素集合. 平均 *Dice* 系数 (mDSC) 指在所有测试样本中对单个样本的 *Dice* 得分取平均值, 反映整体分割性能.

交并比 (intersection over union, *IoU*) 用于衡量预测区域与真实区域的重合程度, 越接近 1 表示分割越准确, 其定义如式 (22):

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{22}$$

与 *Dice* 相比, *IoU* 对预测边界误差更敏感. 平均交并比 (mIoU) 为所有测试样本的 *IoU* 分数平均值, 用于评估模型在整体验证集上的区域覆盖能力.

在医学图像分割任务中, mDSC 更关注结构内部的一致性, 而 mIoU 则在边界定位方面具有更高评价敏感性. 二者联合使用有助于从不同层面全面衡量模型的分割性能.

3.3 实验设置

本实验基于 PyTorch 2.6 深度学习框架, 使用 Python 3.10 进行开发与训练. 实验平台配置为 Intel i9-11900K/F 处理器、64 GB 内存以及 24 GB 显存的 NVIDIA GeForce RTX 3090 GPU. 在数据预处理阶段, 为增强模型的泛化能力与鲁棒性, 对训练图像进行随机旋转与翻转等数据增强操作, 所有输入图像统一调整为 256×256 分辨率. 模型训练过程中, 采用 Adam 优化器[22], 初始学习率设为 1E–4, 批大小设置为 4. 此外, 为提升特征提取能力, 编码器在 ImageNet 数据集上进行了预训练, 以便更有效地迁移通用视觉知识至皮肤病灶分割任务.

3.4 对比实验

为全面评估所提出 FECNet 在皮肤病灶分割任务中的综合性能, 我们在 ISIC-2016、ISIC-2017 和 ISIC-2018 这 3 个公开数据集上, 与 11 种具有代表性的经典

与现代分割模型进行了横向性能对比. 这些方法包括传统 U-Net 系列 (U-Net[10]、UNet++[11]、Att-UNet[23])、基于注意力机制 (FAT-Net[24])、Transformer 模型 (TransUNet[13]、UNeXt[25])、结构优化型轻量模型 (EGE-UNet[26])、扩散建模方法 (BGDiffSeg[27])、状态空间架构 (VM-UNet[28])、边界感知模型 (BiFBA-Net[29]) 以及近期提出的 CCViM[30], 它通过跨层上下文交互机制显式建模多尺度依赖, 进一步优化了边界区域的细节刻画. 上述方法分别代表了近 10 年皮肤图像分割领域的关键方法路径.

3.4.1 ISIC-2016 数据集上的对比实验及分析

该数据集中图像质量较高, 但普遍存在病灶边界模糊、灰度过渡自然等问题, 考验模型对边缘结构的感知能力. 由表 1 可见, FECNet 在 mIoU 和 mDSC 两个指标上均取得了最优成绩. 与 BiFBA-Net[29]相比, FECNet 在 mDSC 上提升了 0.25%, 表明其在边界保持与结构还原方面具有更优能力. 本文表格中加粗字体表示最优结果, ↑表示数值越大性能越优.

表 1 ISIC-2016 数据集上与其他模型性能对比 (%)

模型	年份	指标	
		mIoU↑	mDSC↑
U-Net	2015	80.25	87.81
UNet++	2018	81.84	88.93
Att-UNet	2018	79.70	87.43
TransUNet	2021	84.89	91.26
FAT-Net	2022	85.30	91.60
UNeXt	2022	84.32	91.49
BGDiffSeg	2024	85.52	92.19
BiFBA-Net	2024	87.57	92.86
FECNet (ours)	2025	**87.78**	**93.11**

传统架构如 U-Net 和 UNet++在结构刻画上存在明显瓶颈, TransUNet 等 Transformer 类模型虽然引入全局建模能力, 但对模糊边界的细节处理仍不足. 相比之下, FECNet 借助 FDM 模块对模糊边界区域进行显式解耦并结合 CCM 模块的语义调制, 有效缓解了边界不明确带来的预测偏移问题.

3.4.2 ISIC-2017 数据集上的对比实验及分析

ISIC-2017 数据集包含毛发遮挡、反光干扰与颜色不均等复杂情况, 对模型的鲁棒性与细粒度建模能力提出更高要求. 由表 2 可见, FECNet 在 mDSC 上达到 89.82%、在 mIoU 上达到 80.79%, 整体性能优于大多数主流方法. 与 CCViM 相比, FECNet 在 mDSC 略高 (+0.08%), 而 mIoU 略低 (–0.61%). 这种差异表明模型在预测时更注重边界贴合与区域纯净性, 倾向于保守划分, 从而略微降低区域覆盖. 该特性源于显式特征解耦机制与边界互补约束的共同作用.

表 2 ISIC-2017 数据集上与其他模型性能对比 (%)

模型	年份	指标	
		mIoU↑	mDSC↑
U-Net	2015	76.98	86.99
TransUNet	2021	78.42	87.90
FAT-Net	2022	78.36	87.86
UNeXt	2022	76.57	86.74
EGE-UNet	2023	79.81	88.77
BGDiffSeg	2024	79.73	88.72
VM-UNet	2024	80.23	89.03
CCViM	2025	**81.40**	89.74
FECNet (ours)	2025	80.79	**89.82**

在复杂背景下, 模型仍保持较高的 mDSC, 说明 FEM 模块尤其是 LFEM 路径对毛发和纹理类干扰具有显著抑制作用, 而 CCM 模块则通过前后景差异引导, 进一步增强了对低对比区域的响应能力.

3.4.3 ISIC-2018 数据集上的对比实验及分析

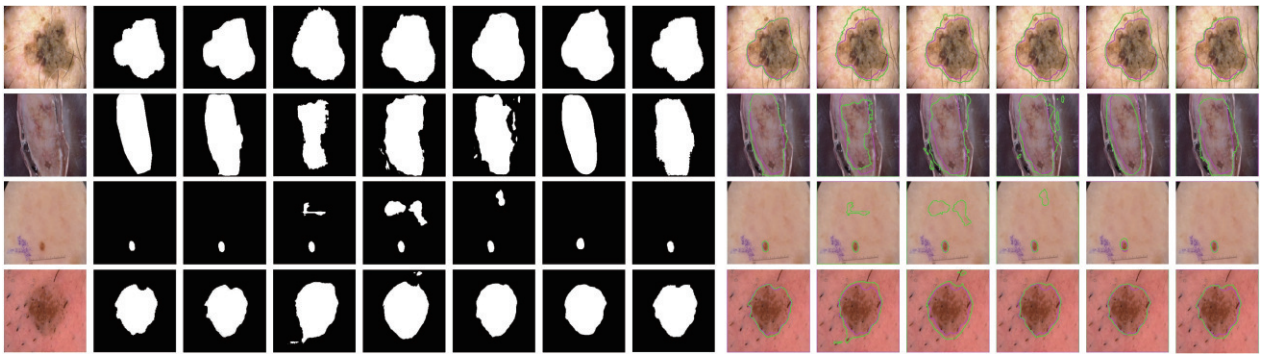
ISIC-2018 是 3 组数据中最具挑战性的集合, 病灶形态复杂、边界不规则且多病种共存. 实验结果如表 3 所示, FECNet 以 83.68% 的 mIoU 和 90.30% 的 mDSC 在所有对比方法中表现最佳, 优于 CCViM 与 VM-UNet 等新一代模型. FECNet 在 mIoU 方面的优势, 说明其在大区域结构建模与前景预测准确率上具有显著提升. 这主要归功于 FEM 模块的双路径结构: GFEM 强化了长程语义建模, 提升了对大范围病灶的统一建模能力; LFEM 则兼顾边缘细节刻画, 使得边界部分不会因结构增强而丢失局部特征. 此外, 多尺度特征融合模块 MFF 有效整合了不同深度的结构信息, 为解耦与最终预测提供了兼顾细节与全局的支持.

表 3 ISIC-2018 数据集上与其他模型性能对比 (%)

模型	年份	指标	
		mIoU↑	mDSC↑
U-Net	2015	77.86	87.55
UNet++	2018	78.31	87.83
Att-UNet	2018	78.43	87.91
TransUNet	2021	79.74	88.73
FAT-Net	2022	82.02	89.03
UNeXt	2022	78.01	87.64
EGE-UNet	2023	80.94	89.46
BGDiffSeg	2024	80.22	89.02
VM-UNet	2024	81.35	89.71
CCViM	2025	81.92	90.06
FECNet (ours)	2025	**83.68**	**90.30**

图 7 展示了各方法在 ISIC-2018 数据集上的可视化分割结果. 图 7(c)–(h) 为分割可视化效果图; 图 7

(i)–(n) 展示了各种模型在原始皮肤图像标注的轮廓, 紫色轮廓表示真实标注,绿色轮廓表示预测的分割结果. 可以看出, U-Net 与 U-Net++对病灶区域的刻画较为粗糙, 常出现边界外溢或凹陷填平, 难以准确反映真实形态; TransUNet 能较好保留整体轮廓, 但在细粒度边界处存在不连续与偏移, 细小结构还原不足; EGE-UNet 能准确捕获主体区域, 但在形态复杂样例中易出现过度平滑, 导致边缘细节模糊; VM-UNet 的整体轮廓较接近真实标注, 但在尖角和细微结构处仍存在轻微缺失与模糊. 相比之下, FECNet 在多种典型场景下均表现出更优的分割质量, 能够在复杂纹理、低对比度及毛发干扰下保持清晰边界与完整形态, 精准描绘不规则及微小病灶, 并有效抑制伪响应. 其预测结果在形态完整性、边界精度与细节还原方面最接近真实标注, 充分体现了模型在复杂皮肤病灶分割任务中的稳定性与泛化能力.



(a) Input (b) Ground truth (c) FECNet (ours) (d) U-Net (e) UNet++ (f) TransUNet (g) EGE-UNet (h) VM-UNet (i) FECNet (ours) (j) U-Net (k) UNet++ (l) TransUNet (m) EGE-UNet (n) VM-UNet

图 7 不同模型在 ISIC-2018 数据集上的分割可视化效果

3.4.4 综合性能分析

从 3 组数据集的对比结果可见, FECNet 在 mDSC 指标上具有更高的稳定性与结构一致性表达能力, 适用于对边界清晰度和病灶保真性要求较高的任务场景; 在 mIoU 指标上, FECNet 在高形变与复杂纹理条件下展现出优越的区域覆盖性能, 但在背景不确定性较强的样例中预测略显保守.

FECNet 能取得优异性能, 主要得益于以下设计.

(1) 边界建模增强: FDM 模块通过显式特征解耦强化了对模糊过渡区域的刻画能力, 并结合互补性损失对边界分布进行有效约束.

(2) 结构增强机制: FEM 模块融合局部与全局双路径, 有效应对病灶在形态与尺度上的高度异质性.

(3) 对比引导调制: CCM 模块在低对比区域动态引导特征响应, 显著缓解前背景模糊导致的预测误差.

(4) 多尺度语义融合: MFF 模块实现多层特征的统一表达, 在结构建模与语义感知之间建立协同作用.

总体而言, FECNet 无需依赖复杂后处理或外部先验, 凭借模块化的内生设计在多类典型与复杂场景中均表现稳健, 展现出良好的泛化性与应用潜力.

3.5 消融实验

为了验证各模块在 FECNet 中对性能提升的具体贡献, 本文在 ISIC-2018 数据集上设计了两组系统性的消融实验, 分别从整体模块组装过程与特征增强路径两方面进行评估.

首先, 通过构建 4 组逐步扩展的模型变体, 评估 FDM、CCM、MFF 和 FEM 模块在最终性能中的独立与协同价值. 我们在 ISIC-2018 数据集上设计了 4 组逐步递进的消融实验. 所有实验均基于统一的 Mamba-Vision 编码器与解码器框架构建, 以确保结果具有可比性. 具体实验设置如下.

(1) Baseline: 仅包含四阶段 MambaVision 编码器与双分支解码器, 用作对比基线.

(2) Baseline+FDM+CCM: 在基线基础上引入 FDM 与 CCM 模块, 其中 FDM 接收编码器第 4 阶段特征.

(3) Baseline+FDM+CCM+MFF: 在上述基础上进一步引入 MFF 模块作为解耦模块的输入, 以提升其多尺度结构感知能力.

(4) FECNet (完整模型): 在第 3 组设置的基础上加

入特征增强模块 (FEM), 构成完整网络架构.

从表 4 的结果可以看出, 随着模块的逐步引入, 模型在 mIoU 和 mDSC 两项指标上持续提升. 其中, FDM 与 CCM 的组合有效强化了边界模糊区域的结构建模能力, CCM 进一步增强了模型在低对比场景中的前背景区分能力; MFF 的引入则提升了融合特征的多尺度感知能力, 使模型对复杂结构具备更强的语义完整性表达; 最终, 加入 FEM 后的完整模型在两个指标上均达到最优, 验证了局部-全局联合建模对提升模型鲁棒性与感知能力的关键价值.

表 4 不同模块的消融实验 (%)

模块				指标	
Baseline	FDM+CCM	MFF	FEM	mIoU↑	mDSC↑
√	×	×	×	81.10	88.24
√	√	×	×	82.64	89.35
√	√	√	×	83.04	89.85
√	√	√	√	83.68	90.30

为进一步验证局部特征增强模块 (LFEM) 与全局特征增强模块 (GFEM) 在应对皮肤病灶形态多变、结构模糊等挑战中的实际作用, 本文在 ISIC 2018 数据集上设计了针对上述两个增强路径的消融实验. 通过逐步移除 GFEM、LFEM 及两者同时移除, 评估其对分割性能的具体贡献, 特别是对小尺度不规则病灶与大范围模糊病灶的适应能力.

从表 5 的结果中可见, 无论移除哪个增强路径, 性能均有明显下降, 表明二者在提升模型感知能力方面均发挥了重要作用. 具体而言, GFEM 对 mDSC 的贡献略高于 LFEM, 说明其在全局建模与语义一致性方面作用更为关键; 而 LFEM 对 mIoU 提升更明显, 展现了其在捕捉边缘细节与小结构病灶方面的重要性. 两者协同时形成互补增强, 显著提升了模型的鲁棒性与边界感知能力.

表 5 不同特征增强模块的消融实验 (%)

模块		指标	
LFEM	GFEM	mIoU↑	mDSC↑
√	√	83.68	90.30
√	×	83.32	90.03
×	√	83.26	90.12
×	×	83.04	89.85

综上, 两组消融实验从整体结构到细粒度路径, 全面验证了 FECNet 各模块设计的必要性与有效性, 为模型在临床复杂病灶图像中的泛化能力提供了坚实支撑.

4 结论与展望

本文针对皮肤病灶图像分割中常见的形态复杂、边界模糊及前背景对比度低等问题, 提出了一种融合特征增强与对比语义引导机制的分割框架 FECNet. 该方法结合 CNN、Transformer 与状态空间模型的优势, 并通过 FEM、MFF、FDM 与 CCM 等模块协同设计, 提升了模型在局部细节刻画与全局语义建模方面的平衡性. 在 ISIC-2016、ISIC-2017 和 ISIC-2018 这 3 个公开数据集上的实验结果表明, FECNet 在分割精度和边界保持等方面较部分现有方法取得了一定提升. 需要说明的是, 本研究仍存在局限, 如在部分场景下区域覆盖能力相对保守, 且模型效率仍有改进空间. 未来工作可在轻量化设计、跨模态信息融合和实际临床适配性方面进一步探索, 以提升模型的泛化性和应用价值.

参考文献

1 Karimkhani C, Dellavalle RP, Coffeng LE, *et al*. Global skin disease morbidity and mortality: An update from the global burden of disease study 2013. JAMA Dermatology, 2017, 153(5): 406–412. [doi: 10.1001/jamadermatol.2016.5538]

2 Tschandl P, Rinner C, Apalla Z, *et al*. Human-computer collaboration for skin cancer recognition. Nature Medicine, 2020, 26(8): 1229–1234. [doi: 10.1038/s41591-020-0942-0]

3 Litjens G, Kooi T, Bejnordi BE, *et al*. A survey on deep learning in medical image analysis. Medical Image Analysis, 2017, 42: 60–88. [doi: 10.1016/j.media.2017.07.005]

4 Codella NCF, Gutman D, Celebi ME, *et al*. 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). Proceedings of the 15th International Symposium on Biomedical Imaging. Washington: IEEE, 2018. 168–172.

5 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 3431–3440.

6 Feng SL, Zhao HM, Shi F, *et al*. CPFNet: Context pyramid fusion network for medical image segmentation. IEEE Transactions on Medical Imaging, 2020, 39(10): 3008–3018. [doi: 10.1109/TMI.2020.2983721]

7 Srivastava A, Jha D, Chanda S, *et al*. MSRF-Net: A multi-scale residual fusion network for biomedical image segmentation. IEEE Journal of Biomedical and Health Informatics, 2022, 26(5): 2252–2263. [doi: 10.1109/JBHI.2021.3138024]

8 Fan DP, Zhou T, Ji GP, *et al*. Inf-Net: Automatic COVID-19

lung infection segmentation from CT images. IEEE Transactions on Medical Imaging, 2020, 39(8): 2626–2637. [doi: 10.1109/TMI.2020.2996645]

9 Zhou T, Zhou Y, He KL, *et al*. Cross-level feature aggregation network for polyp segmentation. Pattern Recognition, 2023, 140: 109555. [doi: 10.1016/j.patcog.2023.109555]

10 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.

11 Zhou ZW, Rahman Siddiquee M, 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. Granada: Springer, 2018. 3–11.

12 Diakogiannis FI, Waldner F, Caccetta P, *et al*. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 162: 94–114. [doi: 10.1016/j.isprsjprs.2020.01.013]

13 Chen JN, Lu YY, Yu QH, *et al*. TransUNet: Transformers make strong encoders for medical image segmentation. arXiv:2102.04306, 2021.

14 Cao H, Wang YY, Chen J, *et al*. Swin-Unet: Unet-like pure Transformer for medical image segmentation. Proceedings of the 2023 European Conference on Computer Vision. Tel Aviv: Springer, 2023. 205–218.

15 Hatamizadeh A, Kautz J. MambaVision: A hybrid mamba-Transformer vision backbone. Proceedings of the 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2025. 25261–25270.

16 Shao DG, Ren LF, Ma L. MSF-Net: A lightweight multi-scale feature fusion network for skin lesion segmentation. Biomedicines, 2023, 11(6): 1733. [doi: 10.3390/biomedicines11061733]

17 Zhang Z, Lin Z, Xu J, *et al*. Bilateral attention network for RGB-D salient object detection. IEEE Transactions on Image Processing, 2021, 30: 1949–1961. [doi: 10.1109/TIP.2021.3049959]

18 Kou WX, Marshall H, Chiu B. Boundary-aware semantic clustering network for segmentation of prostate zones from T2-weighted MRI. Physics in Medicine & Biology, 2024, 69(17): 175009.

19 Qiu ZH, Wang ZC, Zhang MM, *et al*. BDG-Net: Boundary distribution guided network for accurate polyp segmentation. Proceedings of SPIE 12032, Medical Imaging 2022: Image Processing. San Diego: SPIE, 2022. 792–799.

20 Lou M, Yu YZ. OverLoCK: An overview-first-look-closely-next ConvNet with context-mixing dynamic kernels. Proceedings of the 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2025. 128–138.

21 Lei MQ, Wu HC, Lv XH, *et al*. ConDSeg: A general medical image segmentation framework via contrast-driven feature enhancement. Proceedings of the 39th AAAI Conference on Artificial Intelligence. Philadelphia: AAAI Press, 2025. 4571–4579.

22 Kingma DP, Ba J. Adam: A method for stochastic optimization. Proceedings of the 3rd International Conference on Learning Representations. San Diego: ICLR, 2015.

23 Oktay O, Schlemper J, Le Folgoc L, *et al*. Attention U-Net: Learning where to look for the pancreas. arXiv:1804.03999, 2018.

24 Wu HS, Chen SH, Chen GL, *et al*. FAT-Net: Feature adaptive Transformers for automated skin lesion segmentation. Medical Image Analysis, 2022, 76: 102327. [doi: 10.1016/j.media.2021.102327]

25 Valanarasu JMJ, 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.

26 Ruan JC, Xie MY, Gao JS, *et al*. EGE-UNet: An efficient group enhanced UNet for skin lesion segmentation. Proceedings of the 26th International Conference on Medical Image Computing and Computer Assisted Intervention. Vancouver: Springer, 2023. 481–490.

27 Guo YL, Cai QL. BGDiffSeg: A fast diffusion model for skin lesion segmentation via boundary enhancement and global recognition guidance. Proceedings of the 27th International Conference on Medical Image Computing and Computer Assisted Intervention. Marrakesh: Springer, 2024. 150–159.

28 Ruan JC, Li JC, Xiang SC. VM-UNet: Vision Mamba UNet for medical image segmentation. ACM Transactions on Multimedia Computing, Communications and Applications, 2025.

29 Yuan FN, Peng YH, Huang QH, *et al*. A bi-directionally fused boundary aware network for skin lesion segmentation. IEEE Transactions on Image Processing, 2024, 33: 6340–6353. [doi: 10.1109/TIP.2024.3482864]

30 Zhu Y, Zhang D, Lin Y, *et al*. Merging context clustering with visual state space models for medical image segmentation. IEEE Transactions on Medical Imaging, 2025, 44(5): 2131–2142. [doi: 10.1109/TMI.2025.3525673]

(校对责编: 李慧鑫)