



# 计算机科学

COMPUTER SCIENCE

## CA-SFTNet:基于空间特征变换和浓缩注意力机制的皮肤病灶分割模型

张伟, 梁敦英, 周婉婷, 程祥

引用本文

张伟, 梁敦英, 周婉婷, 程祥. [CA-SFTNet:基于空间特征变换和浓缩注意力机制的皮肤病灶分割模型](#)[J]. 计算机科学, 2026, 53(3): 277-286.

ZHANG Wei, LIANG Dunying, ZHOU Wanting, CHENG Xiang. [CA-SFTNet:Skin Lesion Segmentation Model Based on Spatial Feature Transformation and Concentrated Attention Mechanism](#) [J]. Computer Science, 2026, 53(3): 277-286.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [基于Transformer架构的RNA二级结构预测方法](#)

Prediction Method of RNA Secondary Structure Based on Transformer Architecture

计算机科学, 2026, 53(3): 375-382. <https://doi.org/10.11896/jsjcx.250100005>

### [关联策略多特征增强的多目标跟踪](#)

Multi-feature Enhanced Association Strategy for Multi-object Tracking

计算机科学, 2026, 53(3): 231-239. <https://doi.org/10.11896/jsjcx.241100094>

### [语义引导的红外与可见光图像混合交叉特征融合方法](#)

Semantic-guided Hybrid Cross-feature Fusion Method for Infrared and Visible Light Images

计算机科学, 2026, 53(2): 253-263. <https://doi.org/10.11896/jsjcx.250100123>

### [基于Sc-DeepLabV3+模型的铁轨扣件分割方法](#)

Railway Fastener Segmentation Method Based on Sc-DeepLabV3+ Model

计算机科学, 2025, 52(12): 166-174. <https://doi.org/10.11896/jsjcx.241000130>

### [网络协议模糊测试技术研究进展](#)

Survey on Fuzz Testing Techniques for Network Protocols

计算机科学, 2025, 52(11A): 241100173-9. <https://doi.org/10.11896/jsjcx.241100173>

# CA-SFTNet: 基于空间特征变换和浓缩注意力机制的皮肤病灶分割模型

张 伟<sup>1,2,3</sup> 梁敦英<sup>1</sup> 周婉婷<sup>1</sup> 程 祥<sup>1</sup>

1 湖北大学人工智能学院 武汉 430062

2 智能感知系统与安全教育部重点实验室 武汉 430062

3 智慧政务与人工智能应用湖北省工程研究中心 武汉 430062

(zhang\_wei@mail.hubu.edu)

**摘 要** 针对皮肤病灶边缘模糊、毛发等噪声导致的分割病灶区域不完整、病灶特征分布差异较大等问题,基于 U-Net 提出一种结合浓缩注意力机制和残差空间特征变换的皮肤病灶分割算法 CA-SFTNet。首先,在模型下采样过程中进行特征切分,保留皮肤病灶浅层语义信息。其次,在跳跃连接处引入浓缩注意力机制(Condensed Attention Neural Block),使得模型能够聚焦于病灶区域,提高分割精度。最后,在模型尾部加入残差空间特征变换层(Residual Spatial Feature Transformation Layer),增强对皮肤病变图像不同区域的自适应调整能力,提高模型对特征分布差异较大病灶的识别能力。实验在 ISIC2017 和 ISIC2018 数据集上进行,结果表明,CA-SFTNet 在分割性能上优于传统 U-Net,Dice 系数分别达到 93.12% 和 92.36%,比 U-Net 提升 7.15 个百分点和 4.81 个百分点;IoU 值分别为 82.59% 和 82.31%,比 U-Net 提升 6.23 个百分点和 4.45 个百分点。相比 TransUNet 和 Swin-UNet 等拓展算法,Dice 系数提升 2~6 个百分点,IoU 值提升 1.8~4 个百分点。这些结果证明了改进算法在皮肤病变区域分割上的优越性,其能够有效提高分割精度。

**关键词:** 皮肤病变;U-Net;浓缩注意力机制;残差空间特征变换;语义分割

中图分类号 TP391

## CA-SFTNet: Skin Lesion Segmentation Model Based on Spatial Feature Transformation and Concentrated Attention Mechanism

ZHANG Wei<sup>1,2,3</sup>, LIANG Dunying<sup>1</sup>, ZHOU Wanting<sup>1</sup> and CHENG Xiang<sup>1</sup>

1 College of Artificial Intelligence, Hubei University, Wuhan 430062, China

2 Key Laboratory of Intelligent Perception Systems and Security of Ministry of Education, Wuhan 430062, China

3 Hubei Provincial Engineering Research Center for Smart Government Affairs and Artificial Intelligence Application, Wuhan 430062, China

**Abstract** To address issues such as blurry skin lesion boundaries, noise caused by hair, incomplete segmentation of lesion regions, and significant differences in lesion feature distribution, this paper proposes CA-SFTNet, a U-Net-based algorithm integrating a condensed attention neural block and residual spatial feature transformation. Firstly, feature segmentation during downsampling preserves shallow semantic lesion information. Secondly, condensed attention neural block in skip connections enhances focus on lesion regions by adaptively weighting critical features. Finally, a residual spatial feature transformation module is integrated at the network's tail, enabling adaptive adjustment for spatially heterogeneous regions and enhancing recognition of lesions with heterogeneous feature distributions. Experiments conducted on the ISIC2017 and ISIC2018 datasets demonstrate that CA-SFTNet outperforms the conventional U-Net in skin lesion segmentation. Specifically, it achieves Dice coefficients of 93.12% and 92.36%, representing improvements of 7.15 and 4.81 percentage points over U-Net, respectively. The corresponding IoU values are 82.59% and 82.31%, which constitute gains of 6.23 and 4.45 percentage points. Moreover, when compared with state-of-the-art Transformer-based architectures such as TransUNet and Swin-UNet, CA-SFTNet consistently improves the Dice coefficient by 2~6 percentage points and the IoU by 1.8~4.0 percentage points. These results collectively demonstrate the superiority of the proposed method in skin lesion segmentation and its effectiveness in enhancing segmentation accuracy.

**Keywords** Skin lesion, U-Net, Condensed attention neural block, Residual spatial feature transformation, Semantic segmentation

到稿日期:2025-02-12 返修日期:2025-05-14

基金项目:国家自然科学基金(62273135)

This work was supported by the National Natural Science Foundation of China(62273135).

通信作者:梁敦英(986507961@qq.com)

## 1 引言

恶性黑色素瘤是世界上增长最快的癌症之一<sup>[1]</sup>。据美国癌症协会统计,2020年约有100350例新发病例,超过6500例死亡。早期黑色素瘤外观上与痣相似,症状不明显,若病发早期及时干预治疗,五年存活率高达98%;晚期黑色素瘤扩散后,五年存活率低于15%<sup>[2]</sup>。基于深度学习的皮肤病变分割技术通过定量分析病变形态和结构,为医学研究提供数据支持,提高了诊断效率和精准度,降低了误诊和漏诊风险,在辅助医生诊断方面具有重要的临床应用价值。然而,医学图像的特殊性给分割任务带来了三大固有问题:1)病灶-背景边界过渡模糊;2)复杂噪声干扰(毛发、反光、胶体残留等伪影);3)小尺度病灶特征易被深层网络湮灭。这些问题制约了分割结果的准确性,是该技术走向实际应用的重要挑战。

早期基于阈值<sup>[3]</sup>、基于边缘<sup>[4]</sup>、基于区域<sup>[5]</sup>的传统分割算法等,仅能提取表层信息统计特征,这些依赖手工设计特征的方法难以全面刻画图像丰富的内在信息。随着机器学习技术的发展,研究者引入支持向量机(SVM)和随机森林等算法,通过提取纹理和颜色形状等手工设计特征进行分割<sup>[6]</sup>。这一阶段的工作显著提高了分割精度,但手工特征难以充分表达图像的复杂信息,且需要较多的专业领域知识。计算机视觉的飞速发展,使皮肤癌自动分割方法成为研究热点。卷积神经网络(Convolutional Neural Network, CNN)<sup>[7]</sup>在图像特征提取方面取得突出效果,成功取代了传统手工特征标注方法。然而,传统CNN网络的主要任务是将整个图像归类为单一标签,这种方法无法提供医学图像分析中重要的像素级上下文信息<sup>[8]</sup>。

Ronneberger等<sup>[9]</sup>提出的基于编解码结构的医学图像分割网络U-Net,由编码器、解码器与跳跃连接3部分组成,通过跳跃连接将浅层特征信息和深层特征信息进行融合,实现了强大的语义表示学习能力,提升了分割准确性。但其深层网络存在两大缺陷:首先,连续下采样导致小病灶和浅层语义特征信息丢失;其次,固定权重的特征融合难以动态抑制噪声干扰。后继研究中,各种U-Net变体结构相继被提出。但这些拓展模型依旧沿用U-Net深层结构,导致深层下采样路径持续丢失浅层细节特征。虽然Attention U-Net在U-Net的跳跃连接处引入了注意力机制,但其注意力门的权重矩阵对所有病灶类型固定,无法针对不同特征分布进行动态调整。

针对这些问题,本文提出一种新的解决方案:引入半实例归一化模块(Half Instance Normalization Block, HIN Block)保留浅层语义特征,经过5层下采样后,通过浓缩注意力机制(Condensed Attention Neural Block, CANB)<sup>[10]</sup>进一步提取浅层语义特征,将每层下采样的浅层语义特征与上采样的深层语义特征进行融合,最后通过残差空间特征变换(Residual Spatial Feature Transform Layer, RSFTLayer),使网络自适应调整和强化病灶区域特征表达,在边界模糊和病灶区域特征不明显的情况下,有效增强网络对细节和边界的关注,提升分割精度。

本文主要工作如下:1)在U-Net下采样阶段引入HIN Block,有效保留梯度信息和局部特征,解决病灶区域与背景

在边界过渡处模糊的问题;2)在跳跃连接处引入CANB注意力机制,增强上下文感知,突出病灶特征,过滤毛发和纹理等干扰信息;3)设计RSFTLayer层,通过空间感知调整特征,更好地适应病灶区域,提升了分割准确性与鲁棒性;4)结合HIN Block、CANB注意力机制以及RSFTLayer,提出CA-SFTNet模型,该模型在ISIC2017和ISIC2018以及武汉市某三甲医院提供的临床数据集上取得了较好的分割效果。

## 2 相关工作

目前,深度学习已成为医学图像处理领域的主流方法。CNN对医学图像噪声具有较强的鲁棒性,能够有效识别目标,提取特征,并实现准确的自动分割<sup>[11]</sup>。Long等<sup>[12]</sup>提出全卷积神经网络(Fully Convolutional Network, FCN),在传统CNN架构基础上进行扩展改进,以适应像素级别的语义分割任务。FCN通过将全连接层替换为全卷积层,并引入反卷积层实现上采样,从而生成与输入图像大小一致的分割结果。Jafari等<sup>[13]</sup>利用卷积神经网络实现皮肤黑色素瘤的分割,通过整体与局部窗口判断像素病变,并采用自适应学习率调整法,可以同时捕捉皮肤黑色素瘤的全局特征(如整体形状和轮廓)和局部细节(如纹理和局部不规则性)。这种综合判断有助于提高黑色素瘤检测和分割的精度。Yuan等<sup>[14]</sup>引入FCN和Jaccard距离损失函数,FCN可以直接输出分割结果的像素级别预测,结合Jaccard损失函数优化策略,模型能够有效改善分割边缘的准确性,尤其在处理皮肤病变的边界时表现更为突出。

CNN和其衍生网络尽管展现出了强大的特征提取能力,但还存在空间信息丢失及像素相关性不足等问题。例如,FCN由于未能充分考虑像素间的相关性,在一定程度上丢失了空间信息,对分割精度产生影响;FCN通过连续卷积和池化操作,在特征图下采样过程中可能会丢失一些细节信息,特别是在处理细小病变时,模型对细节的捕捉能力有所下降。

Zhou等<sup>[15]</sup>在U-Net的基础上提出UNet++,设计带有深度监督的嵌套密集跳跃连接结构来建立图像多尺度特征,更细致地连接浅层网络与深层网络,不再局限于融合相同尺度的特征,而是聚合不同语义规模的信息,在面对复杂背景时能够捕捉更丰富的特征信息。Oktay等<sup>[16]</sup>提出的Attention U-Net,首次实现了注意力机制与U-Net的结合,在U-Net的跳跃连接处加入注意力门(Attention Gate, AG),通过隐式学习抑制皮肤病灶图像内的无关区域,增强病灶区域的显著特征,抑制无关背景。Jha等<sup>[17]</sup>提出的DoubleU-Net,将空洞卷积与迁移学习结合,利用空洞卷积在卷积核之间引入空洞,有效扩大了卷积操作的感受野,使网络能够捕捉更多的上下文信息,尤其在处理大范围图像时能够识别更广泛的特征,提升分割的全局上下文感知能力。Yu等<sup>[18]</sup>提出的BiseNetV2,采用双向边缘增强模块以及上下文信息和细节信息两条分支的设计,分别提取不同层次的特征,能够在捕捉局部细节信息和全局上下文信息之间找到平衡。鉴于Transformer在视觉任务中能够突破CNN的局限性,Chen等<sup>[19]</sup>提出TransUNet,在编码器中并入Transformer通路,提升了模型对长距离依赖关系的建模能力;在解码器中,将自注意力特征与高分辨率

的卷积特征图结合,实现了更精准的定位。Liang 等<sup>[20]</sup>提出融合多尺度 Transformer 的编解码网络 MsF-SegFormer,通过分层编码器提取多分辨率特征,结合多尺度融合模块与通道注意力模块强化空间语义信息依赖,缓解皮肤病变分割中的细节缺失与误分割问题。Cao 等<sup>[21]</sup>将位移窗口思想代入 U-Net,提出 Swin-UNet,位移窗口机制通过滑动窗口提取上下文特征,能够同时学习局部细节特征和全局上下文信息,使得网络能够在处理复杂结构和大范围图像时,捕捉到更丰富的特征。Xu 等<sup>[22]</sup>提出的 DCSAU-Net,结合了主特征守恒策略和紧凑分裂注意力模块,能有效利用低层和高层语义信息;同时,通过通道分割和软注意力机制增强特征表示,有效提升了分割精度。

上述 U-Net 拓展模型虽然在分割精度和效率等方面不断提升,但在细节信息、边界精确度及高低级特征融合等方面的表现仍显不足。例如,TransUNet 和 Swin-UNet 引入全局信息建模机制,使得模型能够关注到图像的广泛区域,但局部细节信息可能被忽略,导致细节信息丢失。DoubleU-Net 由于空洞卷积的卷积核之间有较大间距,可能会丢失一些较小尺度的局部特征,特别是在需要精细空间信息的边界区域,空洞卷积可能导致细节丢失,影响边界精确度。DCSAU-Net 的通道分割将不同的语义信息分配到不同的通道进行特征处理,但是会限制不同层次特征之间的直接交互,特别是在低层和高层特征融合时,如果特征分割过于严格,可能会削弱低级细节信息与高级语义信息之间的联系,影响特征融合的效果。

本文 CA-SFTNet 在增强特征提取的同时保留浅层语义特征信息,提升不同层次特征的融合效率,改善模型对细节的捕捉能力,提高边界精度;通过注意力机制增强模型对病灶区域的识别能力,通过空间特征变换自适应调整强化病灶区域特征的表达,整体提升模型的分割精度、边界细节和鲁棒性。

### 3 本文方法

CA-SFTNet 模型是基于 U-Net 进行优化而来,其网络结构如图 1 所示。

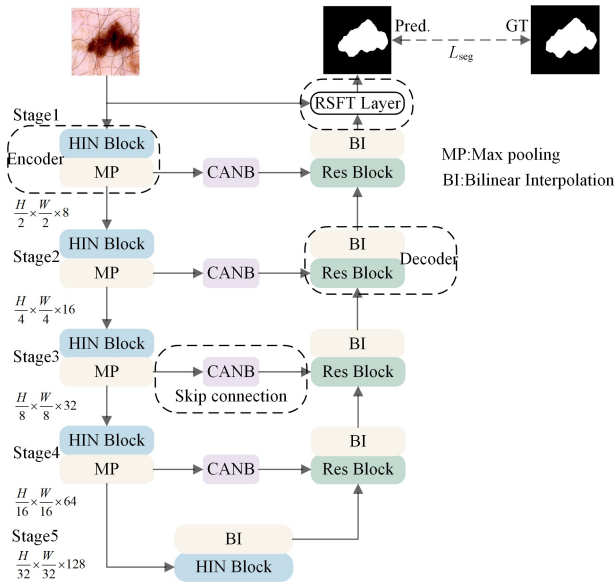


图 1 CA-SFTNet 结构

Fig. 1 Structure of CA-SFTNet

可以看出,CA-SFTNet 由编码器(Encoder)、跳跃连接(Skip Connection)、解码器(Decoder)和 RSFTLayer 层 4 部分组成,逐步提取深层语义特征并恢复高分辨率分割图像。整个模型结构基于 U-Net,下采样包含 5 个编码器,每一个编码器里添加一个 HIN Block 模块和规格为  $2 \times 2$  的最大池化层;上采样同样包含 5 个解码器,每一个解码器里包含一个残差模块(Res Block)和双线性插值(Bilinear Interpolation, BI)模块对图像进行恢复;编码器与解码器通过 CANB 模块进行特征映射连接;最后通过引入的 RSFTLayer 层输出皮肤病变分割图。

皮肤病变区域通常具有不同大小、形状和纹理,在编码器下采样过程中,浅层特征语义信息逐渐丢失,比如病灶边缘和纹理等,这些正是皮肤病变分割易被忽视的特征。为了保留浅层特征语义信息,提高皮肤病灶边界分割的效果,CA-SFTNet 加入 HIN Block 模块,通过对部分通道进行实例归一化(Instance Normalization, IN)<sup>[23]</sup>增强模型对细粒度特征的表达能力;对于另一部分通道,则保留整体上下文信息。在 CA-SFTNet 跳跃连接部分,引入 CANB 浓缩注意力机制,动态调整特征权重,聚焦病灶区域,抑制无关噪声特征。针对解码器上采样恢复图像分辨率导致特征信息丢失的问题,CA-SFTNet 添加残差模块(Res Block),通过显式学习残差信息补偿丢失细节。不同患者的皮肤图像的病灶特征分布和强度可能存在显著差异,为了适应差异化特征,CA-SFTNet 引入 RSFTLayer 层,通过条件特征生成一个参数对,动态调整输出特征。RSFTLayer 层结合了条件信息和特征图信息来提升特征的适应性,在面对特征差异较大的皮肤病灶图像时能够更好地适应病灶区域,提升模型分割的准确性和鲁棒性,增强模型的泛化能力。

#### 3.1 编码器与 HIN Block 模块

编码器的主要功能是提取输入图像的特征信息。U-Net 的编码器通过简单的卷积层对图像进行特征提取,利用最大池化操作降低特征图的分辨率。在此过程中,多次池化和降维会丢失重要的浅层语义特征信息。为了解决这些问题,CA-SFTNet 在编码器中引入 HIN Block 模块。皮肤病灶图像分割属于低级视觉任务,通常涉及高分辨率、小区域目标的处理。这些任务中,图像块批量之间差异较大,传统的批量归一化(Batch Normalization, BN)<sup>[24]</sup>会因批量间特征分布存在差异而降低性能。与 BN 不同,HIN Block 模块引入 IN,使归一化在训练和推理阶段具有一致性,减少了批量差异带来的不稳定性,提升了模型的泛化能力;通过平衡全局和局部特征归一化,能适应病灶图像的细粒度特征。这种设计不仅提升了模型对皮肤病灶区域分割的性能,还保留了浅层语义信息,增强了模型提取皮肤病灶边缘浅层信息的能力。

从图 2 可以看出,HIN Block 模块首先将输入特征  $F_{in} \in \mathbb{R}^{C_{in} \times H \times W}$  通过  $3 \times 3$  卷积生成中间特征  $F_{mid} \in \mathbb{R}^{C_{out} \times H \times W}$ ,其中  $C_{in}$  和  $C_{out}$  是 HIN Block 模块的输入和输出通道数。然后,将  $F_{mid}$  特征分割为两部分 ( $F_{mid_1} / F_{mid_2} \in \mathbb{R}^{C_{out}/2 \times H \times W}$ ),第一部分  $F_{mid_1}$  是将 IN 进行归一化处理,并使用可学习的线性参数,再在通道维度上将其与  $F_{mid_2}$  进行拼接。HIN Block 模块在一半通道上应用 IN,另一半通道则保持上下文信息,这种设计更

有利于保留浅层语义特征。完成拼接操作后,将特征传递给 $3 \times 3$ 的卷积层和两个 Leaky ReLU 激活层,输出带有残差特征的 $R_{out} \in \mathbb{R}^{C_{out} \times H \times W}$ 。最后, HIN Block 模块将 $R_{out}$ 与 $F_{in}$ 相加,得到 $F_{out}$ 。

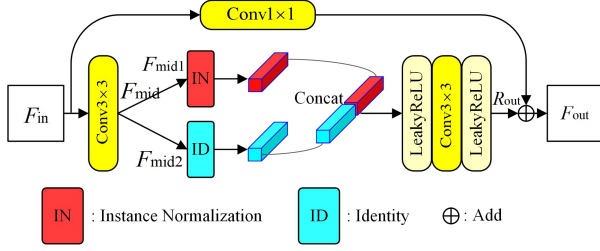


图 2 HIN Block 结构图

Fig. 2 Structure diagram of HIN Block

### 3.2 解码器与 Res Block 模块

解码器的主要功能是恢复图像分辨率。传统解码器依赖上采样和跳跃连接恢复细节信息,但仅靠简单的卷积层可能无法充分利用编码器传递的高分辨率特征。解码器结构较深时,深层梯度传递路径较长,部分特征信息在传递过程中容易被丢失。为了解决此问题,CA-SFTNet 在上采样模块中加入残差网络(Residual Network)<sup>[25]</sup>,通过引入跨层连接残差块,将前一层特征直接与输出进行加和处理,以有效减轻梯度消失问题。残差网络通过叠加操作保留多层特征信息,有效实现全局语义信息与局部细节的融合,提升多尺度目标的分割能力。Res Block 模块提供直接梯度传递路径,使深层网络训练更加稳定,加快模型收敛速度。从图 3 可以看到,Res Block 模块首先将输入特征 $F_{in} \in \mathbb{R}^{C_{in} \times H \times W}$ 通过两次 $3 \times 3$ 卷积

和两个 Leaky ReLU 激活层,得到 $R_{out}$ ,又将输入特征 $F_{in} \in \mathbb{R}^{C_{in} \times H \times W}$ 经过一个 $1 \times 1$ 卷积生成的结果与 $R_{out}$ 相加,得到 $F_{out}$ 。

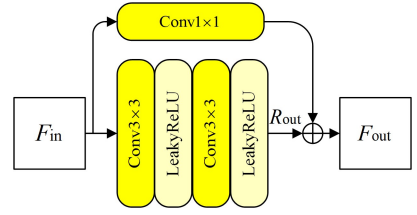


图 3 Res Block 结构图

Fig. 3 Structure diagram of Res Block

### 3.3 CANB 注意力机制

U-Net 的跳跃连接有效地将编码器和解码器特征信息结合到一起,有助于恢复高分辨率细节。然而,传统 U-Net 的跳跃连接直接进行特征拼接,没有考虑特征的重要性,特征融合单一;并且缺乏精细的特征增强,对于小目标与毛发干扰的病灶,特征信息难以充分捕获,影响了分割精度。为了解决这些问题,CA-SFTNet 对跳跃连接进行结构优化,通过改进跳跃连接模块,引入一种包含特征聚合、注意力计算和特征恢复的三范式 CANB 注意力机制,实现更高效的特征融合和特征病灶区域增强。如图 4 所示,CANB 首先将通道和空间特征聚合为浓缩特征,即超像素特征。由于特征涉及通道和空间两个维度,CANB 依次对它们进行通道和空间注意力计算,以便在两个维度上完全捕捉全局依赖关系。最后,CANB 恢复空间和通道维度,以使输出的超像素特征分辨率和通道数与输入的像素特征一致。

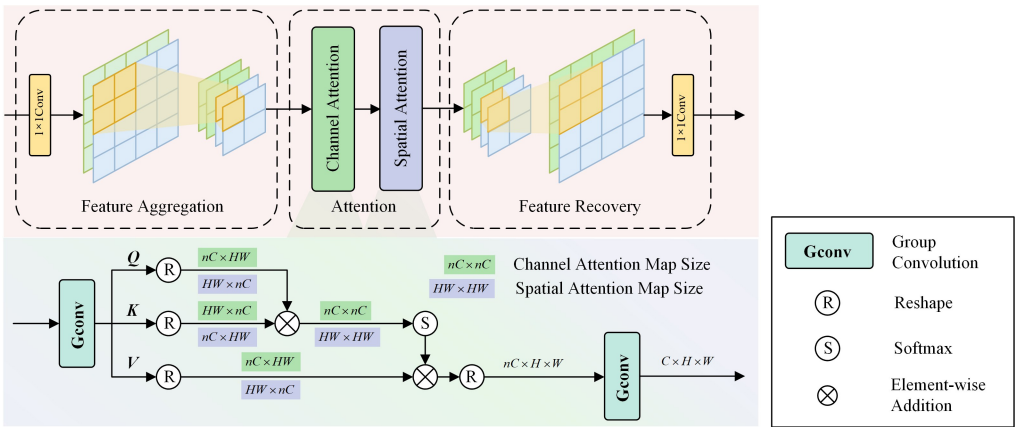


图 4 Condensed Attention Neural Block 注意力机制网络结构图

Fig. 4 Structure diagram of Condensed Attention Neural Block

#### 3.3.1 特征聚合与恢复

皮肤病灶图像中通常含有大量的冗余信息,包括空间背景、通道相关性、任务无关噪声及跨维度耦合冗余。例如,病灶仅占有小部分区域,而背景区域的像素(健康区域)对分割没有贡献,但占据大量计算资源;图像采集时残留的毛发、反光等噪声,部分良性病灶(如对称性色素痣)在空间上呈现镜像对称性,对称区域的信息重复度高。针对这些冗余信息,CANB 通过特征聚合,在适当减少冗余的同时尽可能保留重要的特征信息,并在注意力计算之后,在特征恢复阶段恢复通道和空间的特征分布。

CANB 采用自适应方式进行特征聚合与恢复,即自适应地学习聚合有用信息特征,并通过网络恢复特征分布。

在多通道特征图中,不同的通道可能提取到相似的特征,通道聚合可减少冗余的通道特征。给定输入特征 $F \in \mathbb{R}^{H \times W \times C}$ ,其中 $H \times W$ 表示空间分辨率, $C$ 为通道数,CANB 沿着通道维度聚合特征以获得通道压缩特征 $\tilde{F} \in \mathbb{R}^{H \times W \times C'}$ 。

$$\tilde{F} = \Phi^{CA}(F) \quad (1)$$

其中, $C' = C/r_c$ , $r_c$ 为通道聚合因子。为了不改变空间分辨率,CA 引入一个逐点卷积作为 $\Phi^{CA}(\cdot)$ ,以自适应地在通道

域聚合信息特征。

图像中的相邻像素通常会包含相似的信息,尤其在图像中的某些区域,像素的局部信息重复较多。通过聚合操作对空间信息中相邻像素的相似特征进行压缩,避免重复计算这些冗余信息。CA 沿着空间维度对特征进行聚合,得到空间压缩特征  $\hat{F} \in \mathbb{R}^{H' \times W' \times C'}$  :

$$\hat{F} = \Psi^{SA}(\tilde{F}) \quad (2)$$

其中,  $H' = H/S$ ,  $W' = W/S$  和  $C' = C'/r_s$ ,  $r_s$  为空间聚合因子。简言之,  $\Psi^{SA}(\cdot)$  将空间特征从  $S \times S \times 1$  的大小聚合到  $1 \times 1 \times r_s$  的压缩特征中,从而获得大小为  $H/S \times W/S \times C'/r_s$  的  $\hat{F}$ 。同时,CANB 引入逐通道分组卷积,其输入和输出通道数量分别为  $C'$  和  $C''$ ,核大小为  $S \times S$ ,步长为  $S$ ,以实现空间特征聚合。通过这种方式,可在扩展通道充分保留聚合空间特征,抑制空间噪声。因此,可以获得通道和空间域的超像素特征,用于在低维空间中捕获超像素级依赖性。

为了在通道域和空间域中恢复注意力计算后的特征分布,CANB 采用聚合过程的逆过程实现这一目标。具体来说,它首先恢复空间特征,然后恢复通道特征。

$$\bar{F} = \Phi^{CR}(\Psi^{SR}(\Theta(\hat{F}))) \quad (3)$$

其中,  $\Theta(\hat{F})$  是经过注意力计算后的特征,  $\Psi^{SR}(\cdot)$  和  $\Phi^{CR}(\cdot)$  分别是空间和通道特征的恢复函数。为了将空间特征从  $H' \times W' \times C'$  大小恢复到  $H \times W \times C'$ ,  $\Psi^{SR}(\cdot)$  使用了一个  $1 \times 1$  的逐通道分组卷积,输入和输出通道分别为  $C'$  和  $C'/S^2$ ,然后通过像素重排(Pixel-shuffle)<sup>[26]</sup>来恢复  $H \times W$  的空间分辨率。为了恢复特征通道,  $\Phi^{CR}(\cdot)$  使用一个点卷积,输入和输出通道分别为  $C'$  和  $C$ 。因此,可在保持空间分辨率和通道数与输入特征  $F$  相同的情况下,获得最终特征  $\bar{F}$ 。

### 3.3.2 通道和空间注意力

皮肤癌分割任务通常涉及复杂纹理、模糊边界和不均匀的颜色分布以及毛发等噪声干扰,尤其在高分辨率医学图像中,局部细节对准确分割至关重要。CANB 的设计通过通道注意力(Channel Attention)和空间注意力(Spatial Attention)联合作用,实现关键区域特征的高效提取。通道注意力识别最有区分性的特征通道,关注哪些特征对皮肤病变区域分割最为重要(如边缘、颜色特征)。空间注意力关注空间上的显著区域,尤其是皮肤病变区域所在的具体位置,忽略无关背景(如健康皮肤区域)。这种两级注意力机制可以更准确地捕捉肿瘤的形态学特征,尤其在病灶区域边界模糊的情况下,能提升分割的准确度。

为了全面捕捉两个维度上超像素级别的全局依赖关系,按顺序在超像素特征  $\hat{F} \in \mathbb{R}^{H' \times W' \times C'}$  上执行通道注意力和空间注意力,即先利用通道注意力捕捉通道维度的依赖关系,之后通过空间注意力捕捉空间维度的依赖关系。由此,可在通道和空间领域全面捕捉全局依赖关系。

如图 5 所示,通道注意力模块首先接收输入特征图,然后通过深度可分离卷积将通道压缩后的特征在空间维度进行降维,输出  $n$  个特征通道,得到压缩后的特征图。通道注意力矩阵通过分组卷积生成  $Q$ (查询)、 $K$ (键)、 $V$ (值)特征矩阵。

最后,压缩的特征图通过通道注意力矩阵更新特征,生成通道增强后的特征图。整个过程可以表示为:

$$X_{CA} = \text{Conv} \left( \text{Reshape}^{-1} \left( \text{Softmax} \left( \frac{Q \cdot K^T}{|Q| |K|} \right) V \right) \right) \quad (4)$$

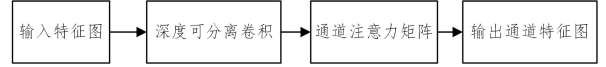


图 5 通道注意力模块

Fig. 5 Channel attention module

如图 6 所示,空间注意力模块通过输入增强的通道特征图与分组卷积生成空间注意力矩阵更新特征,得到空间增强的特征图;再对特征图像素进行重排(Pixel-shuffle)来恢复空间分辨率,获得通道和空间均增强的特征图。这一过程可表示为:

$$X_{SA} = \text{Conv} \left( \text{Reshape}^{-1} \left( \text{Softmax} \left( \frac{Q^T \cdot K}{|Q| |K|} \right) V^T \right) \right) \quad (5)$$

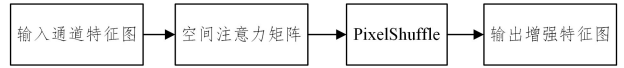


图 6 空间注意力模块

Fig. 6 Spatial attention module

在特征聚合过程中,可能会丢失部分信息,但 CANB 属于聚焦式注意力机制,其通过动态权重分配和隐含在模型内部中的残差连接,能够补偿一定的信息损失,提升关键特征的显著性。其动态权重分配机制对应式(4)和式(5)所描述的多头注意力计算过程。该过程先对查询向量  $Q$  与键向量  $K$  进行点积运算,计算特征间的相似度;随后通过 Softmax 归一化,将相似度转化为归一化注意力权重;最终注意力权重矩阵根据不同输入特征的相似度差异,自适应地动态生成。这些权重矩阵对  $V$  进行加权求和,从而实现特征的动态聚焦和加权,使模型能够自适应地调整特征的重要性,增强网络对关键信息的聚焦能力。

引入 CANB 注意力机制可使模型的编码器和解码器特征在融合时更加高效,特别是在处理有毛发等噪声干扰的病灶或对对比度低的皮肤病灶图像时,CANB 可以显著提升分割的准确性。

### 3.4 RSFTLayer 层

在皮肤病灶分割任务中,病灶区域形态多样,常伴随边界模糊、小病灶嵌套或特征分布不规则等问题。RSFTLayer 层是一种结合先验条件动态调整特征图的创新方法,其核心思想是通过条件调制(Condition Modulation)引入空间特征变换,增强深度网络对复杂任务的适应能力。

如图 7 所示,RSFTLayer 层基于先验条件  $\Psi$  学习一个映射函数  $M$ ,该函数通过神经网络多层卷积和非线性激活操作,生成一对动态调制参数  $(\gamma, \beta)$ 。这对参数分别用于调整特征图的幅度和偏移,从而实现逐像素的仿射变换。映射函数  $M$  可以表示为:

$$M: \Psi \rightarrow (\gamma, \beta) \quad (6)$$

$\gamma$  和  $\beta$  与特征图  $F$  维度相同,分别表示调制幅度和偏移量。通过这对参数,特征图变换过程可以表示为:

$$SFT(F|\gamma, \beta) = \gamma * F + \beta \quad (7)$$

其中,  $F$  为输入特征图;  $\gamma$  用于增强目标区域特征的幅度,抑制

背景噪声,动态调整特征图  $F$  的通道幅度; $\beta$  通过逐像素偏移进一步对调整后的特征图进行偏移补偿,修复压缩或降采样导致的空间细节损失,精确调整病灶区域的特征分布。相比传统静态卷积操作,RSFTLayer 层能够根据输入条件信息动态调整特征图,有效捕捉复杂输入细节(如多形状、大小和纹理特征等),并结合全局语义信息,提升对目标区域的识别能力。

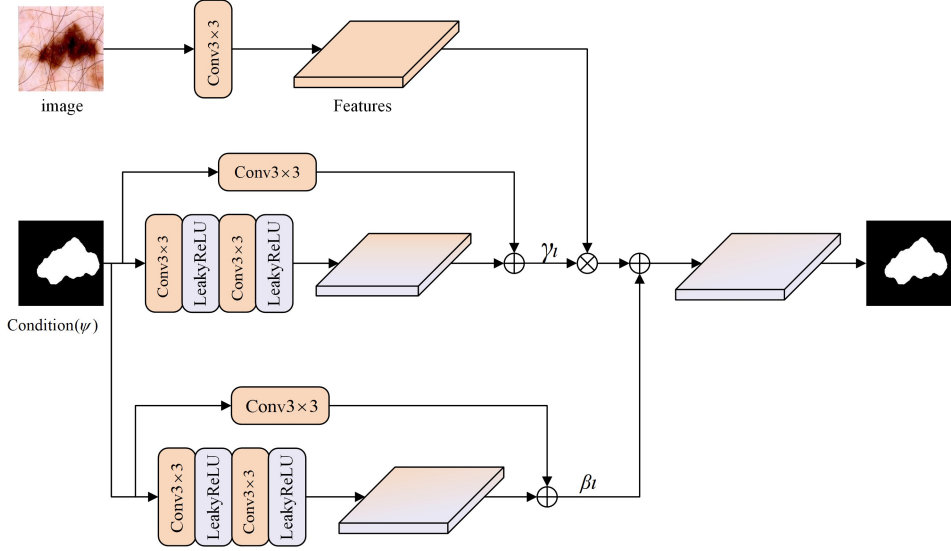


图7 残差空间特征变换网络结构图

Fig. 7 Structure diagram of residual space feature transformation network

从残差角度来看,RSFTLayer 层进一步优化了特征图的动态调整过程。通过残差连接的方式,将基准特征直接引入调制参数生成结果中,避免了过度拟合或信息丢失的问题。同时,这种设计可以提升特征变换的稳定性,增强模型对复杂任务的鲁棒性。

RSFTLayer 层通过动态调制机制,将 U-Net 输出的条件信息(如分割概率图)嵌入特征变换过程,根据先验信息生成适配的参数对  $(\gamma, \beta)$ ,实现特征图的动态调整。此外,RSFT-Layer 层能够结合局部特征和全局语义,在复杂场景下更精准地识别和分割多样化的病灶区域,展现出在医学影像分割中的强大优势。

## 4 实验与结果分析

### 4.1 实验设置

#### 4.1.1 数据集与预处理

为了客观评价 CA-SFTNet 模型的性能,在国际皮肤成像协会(The International Skin Imaging Collaboration, ISIC)发布的 ISIC2017 和 ISIC2018 两个公共数据集以及武汉市某三甲医院提供的临床数据集上进行综合实验。ISIC2017 数据集包含 2000 张训练图片、150 张验证图片和 600 张测试图片。ISIC2018 数据集包含 2594 张训练图片、100 张验证图片和 1000 张测试图片。三甲医院提供的数据集包含日光性角化病、基底细胞癌、脂溢性角化病和鲍温病这 4 种临床病症。3 个数据集的每张图片都有经过专家手动标注的分割标准图像。本实验将所有图像预处理为  $256 \times 256$  的尺寸。

#### 4.1.2 实验环境与参数配置

所有实验均基于 PyTorch 深度学习框架,在 Windows 操

在具体实现上,RSFTLayer 层的映射函数  $M$  通过一系列卷积层生成调制参数  $(\gamma, \beta)$ 。如图 7 所示,生成  $\gamma$  和  $\beta$  的两条路径具有相同的网络结构,但它们的卷积层权重参数和偏置参数均独立,因此两条路径输出的结果也是完全独立的。这种设计可以确保  $\gamma$  和  $\beta$  在调整幅度和偏移特性时互不干扰,各自发挥最优作用。

作系统下完成,实验环境基于 Intel I5-12600KF 处理器,显卡为 NVIDIA GeForce RTX 4070Super,显存为 12 GB,运行内存为 32GB。训练过程中使用 Adam 算法优化模型,初始学习率为  $1 \times 10^{-4}$ ,Batchsize 为 4,总共训练 200 个 epoch。

#### 4.1.3 损失函数

损失函数(Loss Function)是衡量模型预测结果与真实结果之间差异的重要指标,本文使用 Dice Loss 损失函数。其公式如下:

$$Loss_{Dice} = 1 - \frac{2 \times |A \cap B|}{|A| + |B|} \quad (8)$$

其中, $A$  是预测的区域, $B$  是真实的区域, $|A \cap B|$  是预测区域和真实区域的交集部分的面积。

#### 4.1.4 评价指标

本文采用 Dice 相似系数(Dice Similarity Coefficient, DSC)、准确率(Accuracy, ACC)和交并比(Intersection over Union, IoU)来评估模型针对皮肤镜图像的分割性能。使用参数量(Params)和浮点运算量(GFLOPs)衡量模型的计算复杂度。DSC 指标衡量分割标签和预测结果的相似度,取值为  $0 \sim 1$ ,其值越高,表示预测结果与标签的契合度越高。ACC 计算正确分类像素的数量与总像素数量间的比值,准确率越高,模型判断病变和正常区域的能力越强。IoU 表示预测值与真实值之间的交并比。

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (9)$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

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

其中,  $TP$  表示真阳性,代表病灶区域被正确分割;  $TN$  表示真阴性,代表正常区域被正确分割;  $FP$  表示假阳性,代表正常区域被误分割为病灶区域;  $FN$  表示假阴性,代表病灶区域被误分割为正常区域。

#### 4.2 对比实验与结果分析

为了验证 CA-SFTNet 模型的有效性,根据上述参数设置和环境配置,选取 U-Net, UNet++, Attention-UNet, DoubleU-Net, BiSeNetV2, TransUNet, Swin-UNet 和 DCASU-Net 语义分割基础网络和医学分割方法在 ISIC2017 和 ISIC2018 两个皮肤病公开数据集上进行对比实验。不同网络的分割性能指标如表 1 和表 2 所列。

表 1 ISIC2017 数据集上各模型的分割效果数据

模型	DSC/%	IoU/%	ACC/%	Params	GFLOPS
U-Net	85.97	76.36	94.07	$7.800 \times 10^6$	$1.378 \times 10^{10}$
UNet++	86.23	77.47	94.22	$9.160 \times 10^6$	$3.490 \times 10^{10}$
Attention-UNet	86.32	77.66	94.37	$8.730 \times 10^6$	$1.674 \times 10^{10}$
DoubleU-Net	84.17	77.04	95.83	$2.929 \times 10^7$	$8.620 \times 10^{10}$
BiSeNetV2	83.66	75.70	95.25	$4.950 \times 10^6$	<b><math>9.460 \times 10^9</math></b>
TransUNet	86.82	80.40	95.87	$1.925 \times 10^7$	$2.057 \times 10^{10}$
Swin-UNet	82.95	76.17	93.41	$2.717 \times 10^7$	$1.595 \times 10^{10}$
DCASU-Net	87.33	80.79	96.22	<b><math>2.590 \times 10^6</math></b>	$1.105 \times 10^{10}$
CA-SFTNet	<b>93.12</b>	<b>82.59</b>	<b>96.33</b>	$6.170 \times 10^6$	$1.127 \times 10^{10}$

表 1 结果表明,在 ISIC2017 数据集上,CA-SFTNet 在 DSC, IoU 和 ACC 指标上性能较好,分别达到 93.12%, 82.59% 和 96.33%。U-Net 和 UNet++ 等模型是基于 U 型结构构建的,可捕获局部上下文信息和定位像素,但缺乏全局信息。CA-SFTNet 在跳跃连接处引入 CANB 注意力机制,建立长距离特征依赖关系,可获取丰富的全局信息;并且 CANB 中的通道注意力和空间注意力使模型能够更加聚焦于重要的区域或通道,提高了细节捕捉能力。CA-SFTNet 与医学分割基础模型 U-Net 相比,各项指标分别提升了 7.15 个百分点、6.23 个百分点和 2.26 个百分点。与次优网络 DCASU-Net 相比,虽然 DCASU-Net 引入了深度空间通道注意力机制,但是 CA-SFTNet 不仅通过空间和通道注意力机制增强了特征提取,还通过 RSFTLayer 层进行条件特征变换,这种自适应的特征调整能力使得 CA-SFTNet 在处理细节和特征分布差异较大的图像时具有更大优势。CA-SFTNet 相比 DCASU-Net,在 DSC 指标上提高了 5.79 个百分点, IoU 指标提高了 1.8 个百分点, ACC 指标提高了 0.11 个百分点。相较于其他对比模型,CA-SFTNet 在 3 个指标中均达到最高,表明该模型能较为准确地分割皮肤病灶区域和正常区域,其预测结果与真实标签高度相似,并且该模型得益于 HIN Block 模块对浅层特征的保留,对病灶区域模糊边界的分割更加准确,可以更好地捕获细节信息和一些关键特征,提升分割精度。从 Params 指标和 GFLOPs 指标来看,相比 U-Net 与 Transformer 结合的 TransUNet 模型以及 Swin Transformer 与 U-Net 结合的 Swin-UNet 模型,CA-SFTNet 模型的参数量要小得多。CA-SFTNet 的计算复杂度比 U-Net 减少了  $2.510 \times 10^9$ , 浮点运算量是 TransUNet 的 54.7%。表 2 列出了各模型在 ISIC2018 数据集上的测试结果。

表 2 ISIC2018 数据集上各模型的分割效果数据

模型	DSC/%	IoU/%	ACC/%	Params	GFLOPS
U-Net	87.55	77.86	94.05	$7.800 \times 10^6$	$1.378 \times 10^{10}$
UNet++	87.83	78.32	94.09	$9.160 \times 10^6$	$3.490 \times 10^{10}$
Attention-UNet	87.32	78.43	94.12	$8.730 \times 10^6$	$1.674 \times 10^{10}$
DoubleU-Net	85.97	78.63	94.63	$2.929 \times 10^7$	$8.620 \times 10^{10}$
BiSeNetV2	85.19	77.67	94.46	$4.950 \times 10^6$	<b><math>9.460 \times 10^9</math></b>
TransUNet	88.37	79.09	94.37	$1.925 \times 10^7$	$2.057 \times 10^{10}$
Swin-UNet	85.02	78.74	92.19	$2.717 \times 10^7$	$1.595 \times 10^{10}$
DCASU-Net	88.80	80.13	95.12	<b><math>2.590 \times 10^6</math></b>	$1.105 \times 10^{10}$
CA-SFTNet	<b>92.36</b>	<b>82.31</b>	<b>95.99</b>	$6.170 \times 10^6$	$1.127 \times 10^{10}$

从表 2 可以看出,CA-SFTNet 在 ISIC2018 数据集上的 DSC 指标达到 92.36%, IoU 指标达到 82.31%, ACC 指标达到 95.99%, 比基础模型 U-Net 分别提升了 4.81 个百分点、4.45 个百分点和 1.94 个百分点; 比次优模型 DCASU-Net 在各项指标上分别提升了 3.56 个百分点、2.18 个百分点和 0.87 个百分点。CA-SFTNet 的分割预测结果更加接近真实掩码图像。

#### 4.3 消融实验

为了探讨本文引入的各个模块的有效性,在 ISIC2018 数据集上进行了消融实验。实验结果如表 3 所列。

表 3 消融实验结果

Table 3 Results of ablation study

						(%)
HIN Block	CANB	RSFTLayer	DSC	IoU	ACC	
			87.55	77.86	94.05	
✓			88.61	79.65	94.61	
	✓		89.42	80.21	94.82	
		✓	90.86	80.86	95.03	
✓	✓		89.86	81.68	95.72	
✓		✓	90.66	81.83	95.22	
	✓	✓	91.86	81.06	95.62	
✓	✓	✓	92.36	82.31	95.99	

从表 3 可以看出,单独加入 HIN Block 相较于基础网络,在 DSC, IoU 和 ACC 指标上分别提升 1.06 个百分点、1.79 个百分点和 0.56 个百分点,说明 HIN Block 的特征切分能够增强多层特征之间的信息传递和融合,改善模型特征的表达能力。单独加入 CANB 注意力机制,3 个指标分别提升 1.87 个百分点、2.35 个百分点和 0.77 个百分点。CANB 的特征聚合和特征恢复除了可以减少冗余特征的计算,还能够抑制不相关特征(如毛发等噪声)的影响,增强关键特征通道,提升分割精度,尤其是 IoU 和 DSC 指标有较大提升,表明 CANB 在提升准确性和细粒度特征提取方面有显著贡献。单独加入 RSFTLayer 层时,3 个指标分别提升 3.31 个百分点、3 个百分点和 0.98 个百分点。RSFTLayer 层通过残差学习与空间特征变换机制强化自适应能力,有效提升恢复细节信息的能力。3 个模块都加入时,3 个指标均得到了最大提升,分别提升 4.81 个百分点、4.45 个百分点和 1.94 个百分点。这表明 HIN Block、CANB 注意力机制和 RSFT Layer 层 3 个模块的协同作用显著提升了网络的整体性能,能够有效分割皮肤病变区域。

为了进一步证明 3 个模块对 CA-SFTNet 模型的有效性,逐一去除 3 个模块进行消融实验。当去除 RSFTLayer 层时,

3个指标分别提升2.31个百分点、3.82个百分点和1.67个百分点,与去除前进行对比,3个指标分别下降2.5个百分点、0.63个百分点和0.27个百分点,其中DSC指标影响最大,说明RSFTLayer自适应调整能力对提升模型分辨相似病灶图像的能力有较大提升。当去除CANB注意力机制时,3个指标分别提升3.11个百分点、3.97个百分点和1.17个百分点,其中ACC指标受影响最大,说明CANB聚焦病灶的能力能够提升模型识别病灶区域的准确率。当去除HIN Block模块时,3个指标分别提升4.31个百分点、3.2个百分点和1.57个百分点,与去除前进行对比,3个指标分别下降0.5个百分点、1.25个百分点和0.37个百分点,其中IOU指标受影响最大,说明HIN Block保留的浅层特征提高了边界分割的准确性,使预测分割区域面积增加,提升了分割精度。

## 5 可视化案例分析

为了更直观地反映CA-SFTNet相较于其他算法的优越性,图8展示了9种网络在ISIC2017数据集上的主观分割结果。其中,image为皮肤病灶原始图像,mask为专家标注的病灶区域。每一行为9种网络对同一原始图像的分割结果,每一列为同一模型对不同图像的分割效果。病灶图像(a)中,右半部分病灶区域较为明显,而区域下方的边界较为模糊,因为

HIN Block保留了较多的浅层语义信息,CA-SFTNet相比U-Net和UNet++等U-Net变体模型,能够将病灶下方的模糊边界分割得更为准确。可以看出,相比于Swin-UNet和TransUNet等与其他模型结合的拓展模型,CA-SFTNet对左上方模糊边界的分割更精准。病灶图像(b)和(c)中的病灶区域都是左半边较为清晰,右半边病灶特征不明显。得益于RSFTLayer层对病灶分布不均匀的自适应调整机制,CA-SFTNet分割的病灶区域比其他8种模型都更加准确,说明CA-SFTNet对病灶区域不明显图像的分割性能更优。图像(d)中存在毛发遮挡,并且右下角有一部分皮肤颜色和病变皮肤颜色类似,8种对比模型在分割这张图像时右下角区域均存在分割错误。Swin-UNet将病灶左边和右下边的皮肤都识别为病灶区域。CA-SFTNet的分割结果与真实掩码图像非常接近,说明CANB注意力机制的引入使得模型能够更加聚焦于病灶区域,有效过滤干扰噪声信息,能够很好地分割病灶区域。图像(e)和(f)中病灶区域均不明显,U-Net,UNet++,Attention-UNet和Swin-UNet不能完成区域分割。CA-SFTNet对于此图边界的分割也存在一定偏差,但仍能明显看出其分割结果是最接近标注结果的。图像(g)中病灶区域明显,但是边界较模糊。CA-SFTNet能够将病灶区域较为准确地分割出来,虽然边缘细节分割效果略逊于TransUNet,但仍明显优于其他模型。

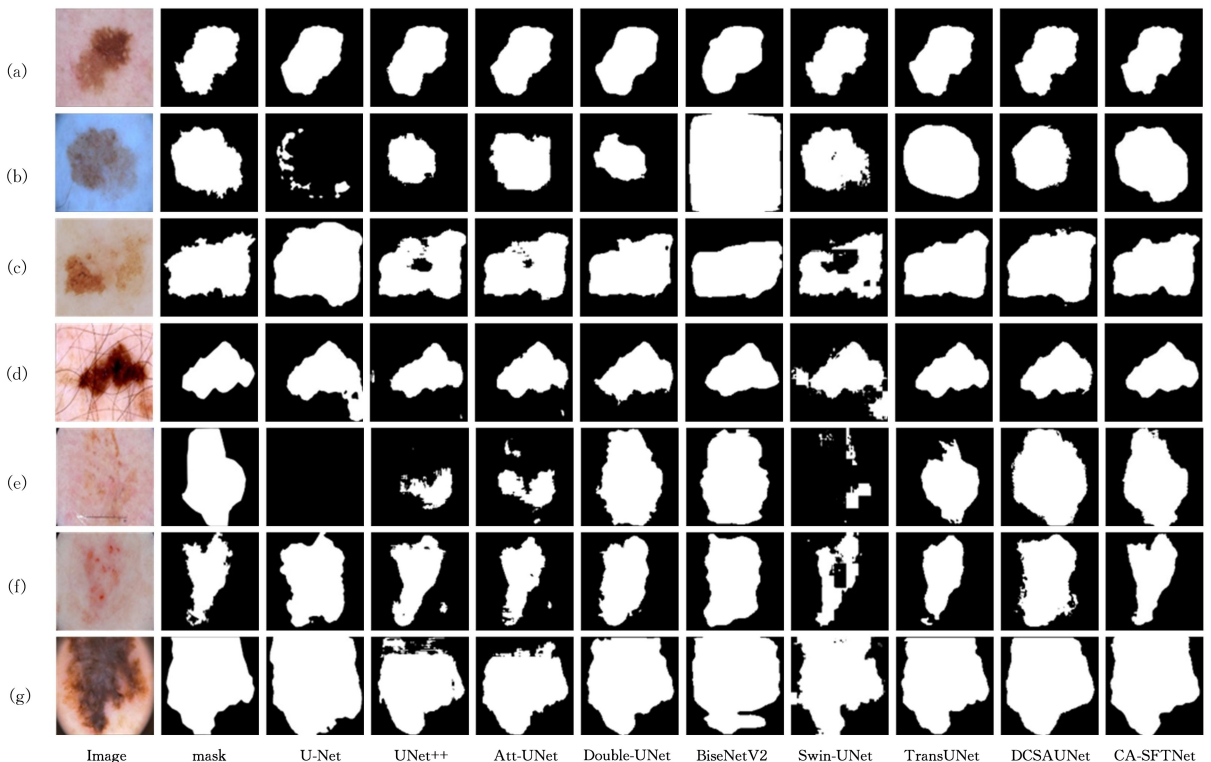


图8 不同拓展模型在ISIC2017数据集上的分割结果图

Fig. 8 Segmentation result plots of different expansion models on the ISIC2017 dataset

为了进一步明确模型在临床应用中的有效性,基于武汉市某三甲医院提供的4类皮肤病灶临床数据集进行了验证实验,图9展示了可视化实验结果。鉴于院方要求及患者隐私保护的需要,所有实验结果均以专家标注的掩码作为参考。其中,图9(a)、图9(b)的图像为日光性角化病,该病灶通常边

界较模糊,且病灶区域特征不够明显,与正常皮肤纹理较为相似。从图中可以看出,CA-SFTNet能够准确地分割出病灶区域;而其余6种模型均未能有效识别该区域,如Double-UNet分割的病灶区域要比实际病灶大得多,其他模型的分割表现也欠佳。这是因为CA-SFTNet设计的RSFTLayer层能够针

对病灶分布不均的特性进行自适应调整,从而提升模型的识别准确率。图 9(c)、图 9(d)的图片为基底细胞癌,其病灶特征区域较为明显,边界也较为清晰。从图中可以看出,所有实验模型均能识别出病灶的大致区域,但在边界细节还原能力上存在差异。其中,CA-SFTNet 的分割结果与真实掩码图像的边界最为吻合,证明模型引入 HIN Block 结构后在保留浅层语义特征信息方面具有显著优势,边界识别能力较强,优于其他拓展模型。图 9(e)、图 9(f)的图像为脂溢性角化病,其病灶特征在不同患者间存在较大的个体差异,其中图 9(e)的

实际患处在眼角,图 9(f)的实际患处在眼睛下方,且均有睫毛干扰。实验结果表明,CA-SFTNet 所分割出的病灶区域与真实掩码高度一致,证明模型引入的 CANB 注意力机制能够有效聚焦病灶区域,并抑制噪声干扰。图 9(g)的图像为鲍温病,是原位鳞癌的一种。该病灶的特征不够明显,边界模糊,与正常皮肤的视觉差异较小,因此分割任务更具挑战性。实验结果表明,尽管所有模型在该病灶的分割表现均不尽理想,但 CA-SFTNet 的分割结果仍与真实掩码最为接近,进一步证明了其在复杂病灶分割任务中的有效性。

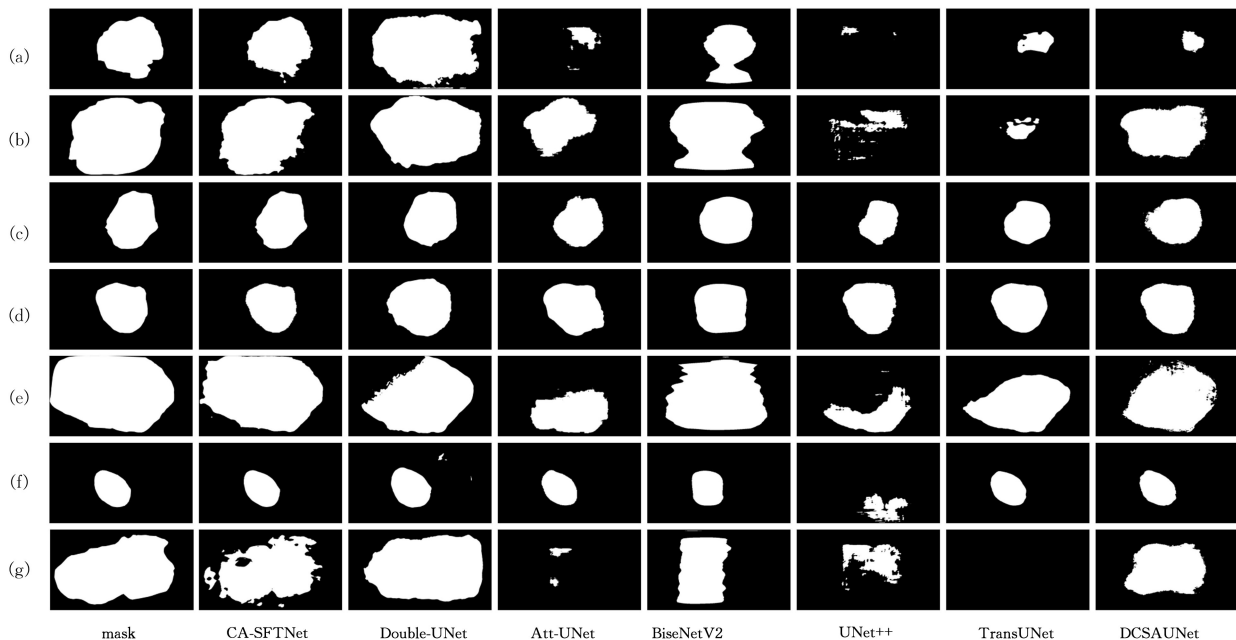


图 9 不同拓展模型在临床数据集上的分割结果图

Fig. 9 Segmentation result plots on clinical datasets of different expansion models

综上所述,CA-SFTNet 在多种皮肤病灶的分割任务中均表现出较强的适应性和准确性,验证了其在临床应用中的潜在价值。

**结束语** 针对传统 U-Net 模型在分割皮肤病灶图像时,因病灶边缘模糊、毛发等噪声干扰、病灶特征分布差异较大导致的分割区域不完整等问题,提出了基于空间特征变换和浓缩注意力机制的 CA-SFTNet 模型。该模型在下采样中使用特征切分保留浅层语义特征信息,有效避免信息丢失;而在上采样中使用残差模块叠加多层次特征信息,通过优化跳跃连接结构使模型特征融合更充分,减少噪声干扰,利用残差空间特征变换层,实现特征分布的动态调整,从而提升模型分割的精度,为皮肤病变临床辅助诊断应用提供必要的性能保障。在未来工作中,针对医疗现场算力资源限制问题,将对 CA-SFTNet 模型进行轻量化研究,减少模型参数量及浮点运算,实现模型边缘化部署应用。

## 参考文献

[1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA: A Cancer Journal for Clinicians*, 2021, 71(3): 209-249.

[2] FONTANILLAS P, ALIPANAHI B, FURLOTTE N A, et al.

Disease risk scores for skin cancers[J]. *Nature Communications*, 2021, 12(1): 160.

[3] YEN J C, CHANG F J, CHANG S. A new criterion for automatic multilevel thresholding[J]. *IEEE Transactions on Image Processing*, 1995, 4(3): 370-378.

[4] TREMEAUA, BOREL N. A region growing and merging algorithm to color segmentation[J]. *Pattern Recognition*, 1997, 30(7): 1191-1203.

[5] ADAMS R, BISCHOF L. Seeded region growing[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(6): 641-647.

[6] HEARST M A, DUMAIS S T, OSUNA E, et al. Support vector machines[J]. *IEEE Intelligent Systems and Their Applications*, 1998, 13(4): 18-28.

[7] FUKUSHIMA K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. *Biological Cybernetics*, 1980, 36(4): 193-202.

[8] SIDDIQUE N, PAHEDING S, ELKIN C P, et al. U-net and its variants for medical image segmentation: A review of theory and applications[J]. *IEEE Access*, 2021, 9: 82031-82057.

[9] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// *Medi-*

- cal Image Computing and Computer-assisted Intervention-MIC-CAI 2015: 18th International Conference. Springer, 2015: 234-241.
- [10] ZHAO H, GOU Y, LI B, et al. Comprehensive and delicate: An efficient transformer for image restoration[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023:14122-14132.
- [11] AZAD R, AGHDAM E K, RAULAND A, et al. Medical image segmentation review: The success of u-net[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12):10076-10095.
- [12] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015:3431-3440.
- [13] JAFARI M H, NASR-ESFAHANI E, KARIMI N, et al. Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma[J]. International Journal of Computer Assisted Radiology and Surgery, 2017, 12:1021-1030.
- [14] YUAN Y, CHAO M, LO Y C. Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance[J]. IEEE Transactions on Medical Imaging, 2017, 36(9): 1876-1886.
- [15] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKSH N, et al. Unet++: A nested u-net architecture for medical image segmentation[C]// Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop. Springer, 2018: 3-11.
- [16] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention u-net: Learning where to look for the pancreas[J]. arXiv:1804.03999, 2018.
- [17] JHA D, RIEGLER M A, JOHANSEN D, et al. Doubleu-net: A deep convolutional neural network for medical image segmentation[C]// 2020 IEEE 33rd International Symposium on Computer-based Medical Systems (CBMS). IEEE, 2020: 558-564.
- [18] YU C, GAO C, WANG J, et al. Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation [J]. International Journal of Computer Vision, 2021, 129: 3051-3068.
- [19] CHEN J, LU Y, YU Q, et al. Transunet: Transformers make strong encoders for medical image segmentation [J]. arXiv: 2102.04306, 2021.
- [20] LIANG L M, ZHOU L S, YIN J, et al. Fusion multi-scale Transformer skin lesion segmentation algorithm[J]. Journal of Jilin University(Engineering and Technology Edition), 2024, 54(4): 1086-1098.
- [21] CAO H, WANG Y, CHEN J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation [C]// European Conference on Computer Vision. Cham: Springer, 2022: 205-218.
- [22] XU Q, MA Z, NA H E, et al. DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation [J]. Computers in Biology and Medicine, 2023, 154: 106626.
- [23] ULYANOV D. Instance normalization: The missing ingredient for fast stylization[J]. arXiv:1607.08022, 2016.
- [24] IOFFE S. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. arXiv:1502.03167, 2015.
- [25] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:770-778.
- [26] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1874-1883.



**ZHANG Wei**, born in 1979, associate professor, master's supervisor, is a member of CCF (No. Y8013M). His main research interests include computer vision, image processing and artificial intelligence.



**LIANG Dunying**, born in 2000, post-graduate, is a member of CCF (No. W0351G). His main research interests include image processing and so on.

(责任编辑:柯颖)