

特征学习和边界引导的皮肤病变图像分割网络

刘梦寒¹, 刘骊^{1,2}, 付晓东^{1,2}, 刘利军^{1,2}, 彭玮^{1,2}

(1. 昆明理工大学信息工程与自动化学院, 云南 昆明 650500; 2. 云南省计算机技术应用重点实验室, 云南 昆明 650500)

摘要: 皮肤病变图像因病变区域边界模糊、尺寸变化大、形状不规则等特性, 显著影响分割算法的准确性。针对该问题, 提出一种结合特征学习和边界引导的皮肤病变图像分割网络。首先, 构建基于卷积神经网络(CNN)和Transformer 的双重特征编码模块, 学习皮肤病变图像的全局-局部特征; 然后, 采用专家标注的真值监督学习指导生成皮肤病变边界窗口, 对病变区域进行边界特征增强; 最后, 通过引入边界引导解码机制, 融合学习到的全局-局部特征和边界增强特征, 进行皮肤病变图像分割, 优化分割结果的边界清晰度和整体精度。在 ISIC 2018 和 ISIC 2016 数据集上的实验结果表明, 该方法在 ISIC 2018 数据集上的交并比(IoU)、特异性和准确率相比 CTO 分别提高 2.88、1.25 和 0.27 个百分点, 在 ISIC 2016 数据集上, Dice 系数和 IoU 分别提高 1.66 和 6.02 百分点, 并在所有评估指标上均取得最优结果。所提分割网络在复杂背景及模糊边界病变的分割任务中表现出卓越性能, 不仅能够精准分割病变区域, 还能有效处理复杂背景下的边界模糊问题, 为皮肤病变图像分割提供了一种高精度的解决方案。

关键词: 皮肤病变图像分割; 特征编码; 特征学习; 边界引导; 边界增强

中图分类号: TP391.41

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0070434

Feature Learning and Boundary Guidance for Skin Lesion Image Segmentation Network

LIU Menghan¹, LIU Li^{1,2}, FU Xiaodong^{1,2}, LIU Lijun^{1,2}, PENG Wei^{1,2}

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, Yunnan, China;

2. Computer Technology Application Key Laboratory of Yunnan Province, Kunming 650500, Yunnan, China)

【Abstract】 Skin lesion images significantly affect the accuracy of segmentation algorithms because of characteristics such as blurred lesion boundaries, large variations in size, and irregular shapes. To address this issue, a skin lesion image segmentation network that combines feature learning and boundary guidance is proposed. First, a dual feature encoding module based on Convolutional Neural Networks (CNN) and a Transformer is constructed to learn the global-local features of skin lesion images. Subsequently, expert-annotated ground truth supervised learning is adopted to guide the generation of the skin lesion boundary windows and enhance the boundary features of the lesion. Finally, by introducing a boundary-guided decoding mechanism, the learned global-local features and boundary-enhanced features are fused to perform skin lesion image segmentation, optimizing the boundary clarity and overall accuracy of the segmentation results. Experimental results on the ISIC 2018 and ISIC 2016 datasets show that the proposed method achieves improvements of 2.88, 1.25, and 0.27 percentage points in Intersection over Union (IoU), specificity, and accuracy, respectively, compared with CTO on the ISIC 2018 dataset. For the ISIC 2016 dataset, the Dice coefficient and IoU are improved by 1.66 and 6.02 percentage points, respectively, achieving optimal results across all evaluation metrics. The proposed segmentation network exhibits excellent performance in segmentation tasks with complex backgrounds and blurred boundaries, accurately segmenting lesion regions and effectively handling boundary blurring issues in complex backgrounds, thereby providing a high-precision solution for skin lesion image segmentation.

【Key words】 skin lesion image segmentation; feature encoding; feature learning; boundary guiding; boundary enhancement

0 引言

皮肤癌^[1]作为最常见的癌症类型之一, 其发病

率呈现稳步上升的趋势, 然而, 若能在早期阶段发现并进行干预, 其 5 年生存率可提高至 90% 以上^[2]。皮肤镜检查作为一种非侵入性的成像技术, 在黑色

基金项目: 国家自然科学基金(62262036, 62362043); 兴滇英才支持计划项目(KKXY202203008)。

作者简介: 刘梦寒(CCF 学生会员), 女, 硕士研究生, 主研方向为计算机视觉、医学图像分割; 刘骊(通信作者), 付晓东, 教授、博士; 刘利军, 副教授、博士; 彭玮, 教授、博士。

收稿日期: 2024-10-08

修回日期: 2025-01-10

E-mail: icall@kust.edu.cn

素瘤的诊断过程中起着关键作用。及时的诊断与有效治疗,特别是针对病变区域的早期切除,能有效地降低皮肤癌的死亡率。皮肤科医生通常依靠视觉检查并识别皮肤病变及潜在的癌症迹象^[3]。然而,正确识别皮肤镜图像既费时又繁琐,并且诊断的准确性易受临床医生主观判断的影响^[4]。计算机辅助诊断技术能有效解决这一问题,而皮肤病变图像的精准确分割成为计算机辅助诊断的关键步骤。

皮肤病变图像分割^[5]旨在将每个像素精确地分类为病变或背景部分,这有助于定位皮肤病变区域。目前,皮肤病变图像分割方法可以分为两大类,即传统分割方法和基于深度学习的方法。传统分割方法主要基于经典图像处理和机器学习算法。文献[6]使用阈值算法确定病变的大致位置,实现病变的快速和精确定位。文献[7]提出了独立直方图追踪算法,对皮肤病变图像进行无监督病变分割。传统方法计算过程相对简单,计算效率较高,具有较好的实时性。然而,这些方法高度依赖手工提取的特征,且由于病变在形态和颜色上的显著差异,导致上述方法在更复杂的数据集上往往表现不佳。

近年来,随着深度学习在医学图像中的快速发展,卷积神经网络(CNN)在皮肤病变图像分割任务中展现出了卓越的性能。U-Net 架构^[8]由于其显著的性能优势,被广泛应用于医学图像分割任务中。为了充分利用多尺度特征,文献[9]提出了一种新的 UNet 3+ 架构,其通过全面的跳跃连接和深度监督机制,实现了精确的医学图像分割。但是,随着网络深度的增加,在降低图像分辨率时会导致空间和上下文信息丢失。为解决上述问题,文献[5]提出了多尺度情境注意网络(MSCA-Net),该网络通过捕获多尺度上下文信息,同时结合 4 种注意力机制生成多尺度输出。文献[10]通过引入多头门控位置交叉自注意力编码器和位置通道注意力机制,结合正则化 Dice 损失函数,提高了皮肤病变图像分割的精度。然而,由于感受野大小的限制及卷积层的特有性质^[11],会导致忽略掉全局背景信息,而在医学图像分割中中程依赖关系尤为重要,上述基于 CNN 的方法在建模远程语义依赖方面仍存在较大局限性,尤其是在需要全局背景信息和细粒度特征的任务中效果不佳。

为了突破 CNN 的局限性,文献[12]引入了视觉 Transformer(ViT),将 Transformer 架构从自然语言处理领域成功迁移到计算机视觉领域,用于处理图像数据。ViT 能够利用多个并行的自注意力头来捕获远程依赖关系。此外,基于 Transformer 的

模型在皮肤病变图像分割任务中正受到越来越多的关注,如文献[13]提出了一种基于 Transformer 的皮肤病变分割模型,该模型通过全局和局部上下文的建模机制,提高了皮肤镜图像中皮肤病变分割的性能。文献[14]设计了一种自适应位置感知结构,利用多尺度 Transformer 提取皮肤病变图像的特征,优化了不同区域间的细粒度特征分布。Swin-Unet^[15]是一种纯 Transformer 的 U 形架构,通过分层移动窗口提取上下文特征,在多器官分割中优于传统方法。虽然上述方法可以模拟全局上下文表示,优化了细粒度特征分布,但缺乏对局部细节的关注。另外,Transformer 的固定分区策略容易破坏分区线附近的细节信息,特别是在复杂病变边界区域。

由此可见,单一网络架构难以兼顾全局信息和局部细节,在皮肤病变图像分割任务中存在局限性。综上所述,皮肤病变图像分割仍然存在以下难点:

- 1) 皮肤病变尺寸变化大、形状不规则,导致较难泛化到各种形状和大小的病变区域。
- 2) 鉴于皮肤病变图像中病变区域边界所具有的模糊性与复杂性,使得较难精确地定位病变区域的边缘。
- 3) 由于现有方法在解码阶段未引入边界信息作为引导,较难确保皮肤病变图像分割结果中边界的清晰度和准确性。

针对上述难点,本文以皮肤镜图像作为输入,提出特征学习和边界引导的皮肤病变图像分割网络,本文主要贡献如下:

- 1) 构建双重特征编码模块,采用 CNN 和 Transformer 双编码网络结构,有效地捕获皮肤病变图像的全局和局部信息,改善皮肤病变尺寸变化复杂的问题。
- 2) 引入专家标注的地面真值图,采用监督学习进行边界特征增强,生成更准确的病变边界窗口和轮廓,提升形状复杂病变区域的分割精度。
- 3) 提出一种边界引导解码策略,将边界信息融合到 CNN 与 Transformer 中,引导解码过程更加精确地定位和描述病变区域的轮廓,改善传统解码方法在边界保持方面的缺陷,准确识别和区分病变区域与背景,有效避免边界模糊和误判问题。

1 相关工作

CNN 在医学图像分割领域得到了广泛应用,CNN 通过卷积层和池化层来有效捕捉图像中的纹理、形状等关键特征。文献[16]提出的深度残差神经网络 ResNet 进一步简化了深度神经网络的训练

过程,为皮肤病变分割任务提供了新的思路。受此启发,文献[17]开发了一种更深的网络架构和更小的内核,以增强皮肤病变分割的判别能力,其结合来自多个颜色空间的信息,进一步提升了分割性能,并在 ISIC 2017 挑战赛中脱颖而出。文献[18]提出了一种用于皮肤病变分割的轻量级 U-Net 网络,显著降低了计算负载和参数量。

深度学习模型在皮肤病变图像分割中取得了显著成果,但对于全局上下文信息的提取仍是挑战。Transformer 模型^[19]采用并行自注意力机制来捕获长程依赖关系,在自然语言任务中取得空前成功之后,也逐渐应用于多个计算机视觉问题。文献[20]提出了分割转换器(Segmentation Transformer),通过在转换器的每一层中建模全局上下文,获得了更好的图像语义分割效果。近年来,在皮肤病变图像分割中,基于 Transformer 的模型有助于更好地理解图像上下文信息,但由于忽略了边界知识和全局上下文的互补使用,导致未能彻底解决皮肤病变模糊边界的问题。鉴于此,文献[21]通过 3 个专门设计的学习器来捕捉边界知识,解决了皮肤病变分割的边界问题。随着 Transformer 模型在医学领域应用的增多,目前也出现了 CNN 与 Transformer 的混合模型。但是,针对皮肤病变分割任务,边界的定位将直接影响最后分割的准确性,而现有的混合模型未能很好地解决皮肤病变边界模糊情况下分割精度低的问题。

尽管 Transformer 在建模长程依赖关系方面表现出优越性,但直接应用 Transformer,会导致补丁分区策略在分区线穿过边界区域时破坏边界周围的详细信息,特别是对于皮肤病变背景和病变之间对比度非常低的情况,上述方法未将边界信息融合到 CNN 与 Transformer 中,导致在病变区域的分割性能不够理想。因此,本文受文献[22-23]的启发,但区别于现有方法,在混合模型的基础上,为了更好地学习皮肤病变的局部边界信息,进一步引入边界特征增强模块,增强病变区域的边界特征表示。为了更好地分割皮肤病变,该网络结合了 CNN、ViT 与边界引导策略。通过构建双重特征编码模块、边界特征增强模块和边界引导解码模块,实现了对皮肤病变图像分割任务的进一步优化。具体地,首先通过结合 CNN 和 ViT 的双重特征编码模块,提取更全面的皮肤病变的全局-局部特征;其次,通过边界特征增强模块,学习边界周围的局部像素间相关性,在边界周围的窗口内进行自关注来增强局部详细特征表示,使得病变边界在分割过程中更加清晰和易于识别;最终,通过得到的边界增强特征引导解码过程,实现更精准的皮肤病变分割。

2 皮肤病变图像分割网络

本文提出了一种特征学习和边界引导的皮肤病变图像分割网络,该网络结构如图 1 所示。

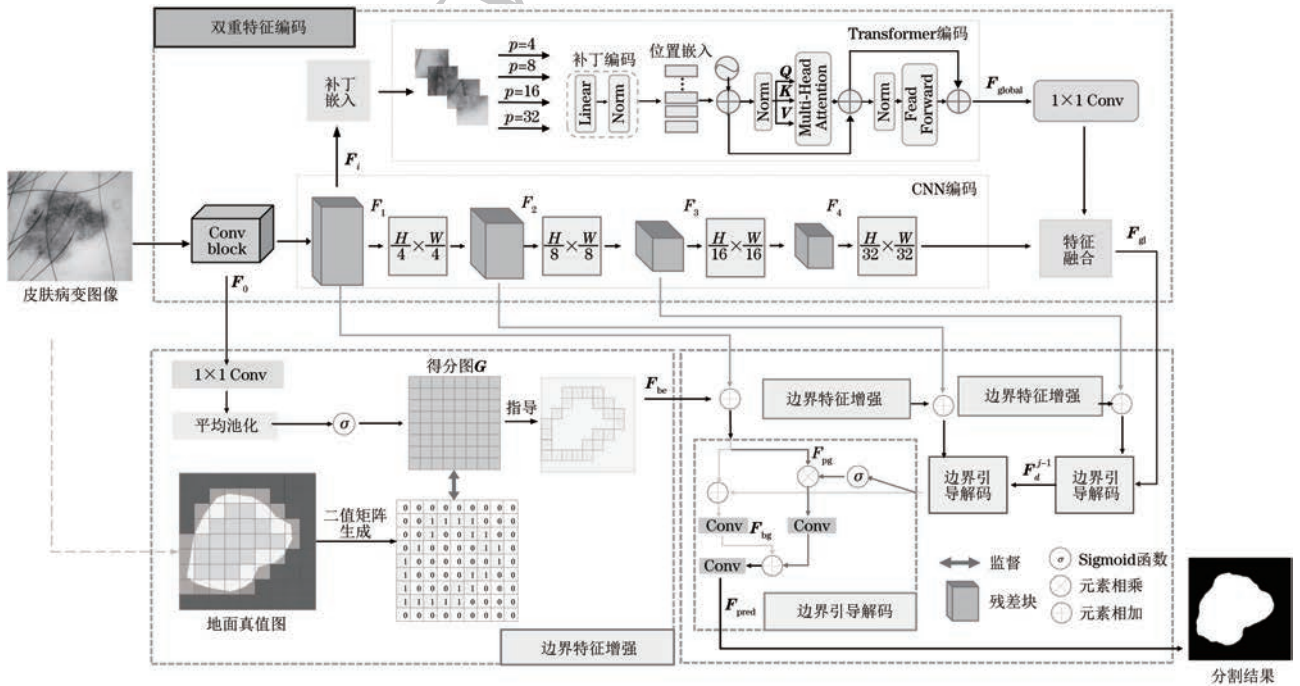


图 1 结合特征学习和边界引导的皮肤病变图像分割网络

Fig. 1 Skin lesion image segmentation network combining feature learning and boundary guidance

整个模型采用编码器-解码器架构,并采用跳跃连接将编码器的底层特征聚合到解码器中。

1)首先,对通过初始卷积层获取的第一层特征进行双重特征编码,设计 2 条路径:CNN 编码流采用 4 个不同的残差块,生成更高层次的局部特征;以残差块的输出进行补丁嵌入作为 Transformer 编码流的输入,通过 4 个并行的 Transformer 块处理不同尺度的特征,捕获图像补丁之间的局部特征依赖关系和远程特征依赖关系。

2)然后,输入地面真值图,将其划分为多个不重叠的窗口,以区分背景和病变区域。构建二值矩阵作为监督,指导网络选择特征映射中的边界区域,通过初步特征图生成一个窗口得分图,得到边界增强特征。

3)最后,边界引导解码结合增强后的边界特征和多尺度编码特征,分别对前景和背景的特征表示进行增强,完成皮肤病变图像分割。

2.1 双重特征编码

现有的皮肤病变图像分割方法大多依赖于单一的网络架构,而皮肤病变图像存在多尺度变化等特点,单独使用 CNN 编码器无法很好地捕获全局特征。Transformer 能够从全局视角捕捉皮肤病变的全局特征,处理长距离依赖关系,但在处理图像的局部细节时不如 CNN 精细。在皮肤病变分割中,传统方法受限于单一网络架构。为此,本文构建了 CNN 与 Transformer 的双分支编码模块,结合 CNN 对局部细节的精确建模和 Transformer 对全局特征的捕获能力,解决上述问题。由于皮肤病变分割需要高精度的局部细节,同时又不能忽视全局上下文信息,为了更好地泛化到不同形状和大小的病变区域,本文通过并行双分支结构,对局部特征与全局特征进行提取与融合,捕获皮肤病变图像的全局和局部信息。

首先,输入皮肤病变图像 $X \in \mathbb{R}^{H \times W \times C}$,其中,空间分辨率为 $H \times W$,通道数为 C 。第一层 CNN 编码流由 1 个基于 Res2Net^[24]的卷积骨干网络和 4 个残差块组成,每个残差块逐步缩小特征图的空间分辨率,同时提升通道维度,生成多层次特征映射。同时,利用残差连接来缓解深层网络训练中的梯度消失与梯度爆炸问题。通过 4 个残差块的逐层卷积操作处理,生成空间分辨率分别为 $\frac{H}{4} \times \frac{W}{4}$ 、 $\frac{H}{8} \times \frac{W}{8}$ 、 $\frac{H}{16} \times \frac{W}{16}$ 和 $\frac{H}{32} \times \frac{W}{32}$ 的局部特征 $F_i \in \mathbb{R} (i=1,2,3,4)$,以捕获皮肤病变图像中的局部特征依赖关系。

CNN 编码流输出四级特征图,分别为浅层特征 F_1 和深层特征 F_2 、 F_3 、 F_4 ,这些层次化的特征图有助于捕捉从低级细节到高级语义的多层次特征。

为了更好地捕捉皮肤病变图像的整体结构,本文在第二层编码流中引入一个轻量级 ViT 模型作为分支^[10]以提取全局特征。第二层编码流构建了 4 个并行的轻量级 Transformer 模块,并且共享相同的结构框架,将 4 个残差块提取的特征图分别输入到对应大小的 Transformer 模块中。具体地,为了匹配 Transformer 层的输入,首先通过补丁嵌入,将 CNN 输出的浅层特征 F_1 分割成一系列大小均匀的补丁,将其重塑为一个二维(2D)补丁序列 $x_e \in \mathbb{R}^{P^2 \times C}$,其中,图像补丁的个数为 $HW/16P^2$,即输入序列长度,每个补丁的大小为 $P \times P$ 。本文使用 4 个并联变压器块,以不同的比例提供功能补丁,得到大小 P 分别为 4、8、16、32 的特征块;然后,将线性投影应用于每个补丁向量,以获得一维(1D)嵌入图像序列 x_p ,加入具有相同维数可学习参数的位置嵌入,以融合空间先验信息,最后一起反馈至 Transformer 编码层,从而获得输出。

通常情况下,一个典型的 Transformer 层主要由多头自注意力(MHSA)层和前馈网络(FFN)组成。自关注机制作为变压器层的核心组件,是实现各嵌入图像块在输入图像序列上并行更新的关键技术。Transformer 使用 MHSA 来建模不同图像块之间的相互依赖。每个图像块在不同的注意力头上进行计算,这使得模型能够捕捉到不同类型的全局信息,从而增强特征表达能力。MHSA 接收截断的 3 个序列向量(即查询向量 Q 、键向量 K 和值向量 V)作为输入,再计算如下注意力得分 $Y \in \mathbb{R}^{N \times N}$:

$$Y = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

式中: d_k 是键的维度。

经过 MHSA 层处理后,得到一个加权后的值序列,然后 MHSA 层的输出被馈送到 FFN 进行进一步处理,以获得输出的全局特征 F_{global} :

$$F_{\text{global}} = \text{FFN}(Y) \quad (2)$$

式中:FFN 是包含 2 个线性层和 1 个 ReLU 激活函数的前馈网络,可为模型提供非线性变换能力。

由于 Transformer 编码流采用了 4 个并行的 Transformer 块来处理不同大小的特征块,每个块经过编码生成一个全局特征图,分别对应不同尺度的块。将这 4 个不同尺度的全局特征图沿通道维度进行拼接,得到具有多层次语义和全局依赖的全局特征 F_{global} 。为了进一步实现全局特征与局部特征

的有效融合,必须对两者的空间分辨率进行对齐。首先通过双线性插值,将 Transformer 流的全局特征 F_{global} 调整为与 CNN 流深层特征 F_4 相同的空间分辨率;然后,将调整后的全局特征和 CNN 的深层特征在通道维度进行拼接,生成融合特征图,融合后的特征图通过卷积操作进行进一步处理,得到双重特征编码模块的最终输出 F_{gl} 。该融合特征同时包含局部细节和全局上下文信息,为后续解码和分割任务提供了高质量的输入特征。

2.2 边界特征增强

皮肤病变图像中病变区域边界具有模糊性和复杂性,传统分割方法主要关注全局特征的提取和分类,忽略了边界区域特有的局部特征,而这些局部特征对于准确界定病变边界至关重要。因此,本文设计了边界特征增强模块,旨在精细化地捕捉皮肤病变图像中边界区域的特征,增强边界区域周围的局部信息。本文采用一种监督学习方法来增强边界特征,该方法引入了专家标注的地面真值图作为监督信号,用于生成病变边界窗口。边界特征增强的核心是在边界窗口内进行自注意力(SA)^[25],通过强化边界特征,使得模型更精确地定位病变区域的边缘。

首先,进行地面真值图的划分与二值矩阵的构建。基于输入的地面真值图,将其细分为多个不重叠的窗口。为了区分哪些窗口内同时存在背景和病变区域,构建一个二值矩阵。具体地,如果一个窗口中同时含有背景和病变区域,该窗口在二值矩阵中对应的值被设置为 1;反之,如果窗口仅包含单一类型的区域(全为背景或全为病变),则值为 0。为了有效地监督并生成边界窗口,将生成的二值矩阵用作监督信号,根据二值矩阵中不同窗口的得分生成一个分数图,该分数图指导网络选择边界窗口。然后,进行分数图的生成与边界窗口的选择。在监督过程中,输入第一层卷积得到的特征图 F_0 ,通过 1×1 的卷积层处理,将特征图还原为一个通道 $C^{1 \times 1}$,以简化后续操作,再使用自适应平均池化操作,将处理后的特征图划分为 $\frac{h}{4} \times \frac{w}{4}$ 个非重叠窗口,在不同的空间位置上独立地分析图像特征。最后,对 Sigmoid 激活后的窗口得分图 $G \in \mathbb{R}$ 进行预测,得分图 G 计算为:

$$G = O(P^{\frac{h}{4} \times \frac{w}{4}}(C^{1 \times 1}(F_0))) \quad (3)$$

式中: O 和 P 分别表示 Sigmoid 函数和自适应平均合并操作。

分数图中每个窗口的得分代表该窗口包含病变

边界的可能性,分数更高的窗口更有可能包含边界。为了区分边界窗口和非边界窗口,本文中的阈值被设置为 0.5,所有得分大于 0.5 的窗口都被认为是潜在的边界窗口。

最后,引入自注意力机制并进行边界特征增强。在确定了得分大于阈值的窗口后,进一步对这些窗口内的像素进行自注意力操作,使模型能够更加聚焦于边界特征,并在这些关键区域内实现更为精细的特征提取和表示。通过分数图指导网络并定位病变边界窗口,最后生成边界增强特征 F_{be} 。

2.3 边界引导解码

在皮肤病变图像分割中,边界的准确性是评估分割性能的关键。然而,传统的解码方法未特别关注边界的保持,导致分割结果在边界区域可能出现模糊或错位。为解决这一问题,本文提出了一种边界引导解码方法,该方法在解码过程中以边界增强特征作为引导,确保在最终皮肤病变图像分割结果中能保持边界的清晰度和准确性。将第 2.2 节边界特征增强模块中得到的边界特征作为先验,以增强编码器生成的图像特征的代表能力。边界引导解码器需要 2 个输入:一是边界增强特征 F_{be} 与来自双重特征编码中的相应局部特征 F_i ,两者进行通道级联;二是前一解码器层 F_d^{j-1} 的特征。边界引导解码采用 2 条处理路径,分别针对前景和背景的特征表示进行增强,从而为后续的像素分类提供丰富且全面的特征表示,有效区分前景和背景的特征。

在前景路径中,为了直接融合 2 个输入中与前景相关的特征,模型采取简单而直接的方法:将这 2 个输入沿其通道维度进行连接,再送入一系列连续的卷积-批处理归一化-ReLU 激活函数层的组合(Conv-BN-ReLU block)进行处理,最终输出优化后的前景特征图 F_{pg} 。对于背景路径,为了选择性地强调与背景相关的信息,本文设计了一个背景注意力组件,具体如下:

$$F_{\text{bg}} = \text{Convs}((1 - \sigma(F_d^{j-1})) \odot F_i) \quad (4)$$

式中:Convs 为 3 层 Conv-BN-ReLU; σ 为 Sigmoid 函数; $\odot$ 表示元素的乘法; $(1 - \sigma(F_d^{j-1}))$ 为背景注意力图,它是通过 Sigmoid 函数应用于前一解码器层的特征图计算而得出的前景注意力图。

然后,从 1 中减去前景注意力图得到背景注意力图。通过这种设计,抑制特征图中的前景区域,从而凸显背景区域。随后,沿着通道维度将前景特征 F_{pg} 、背景特征 F_{bg} 和之前的解码器特征 F_d^{j-1} 进行串联,保留前景和背景各自的独特信息,增强整体图像表征的丰富性和准确性。最后,使用 Softmax 激活函数

和 1×1 卷积层来完成像素级预测,获得最终的皮肤病变分割结果 F_{pred} 。

2.4 损失函数

由于皮肤病变图像分割是一个像素级的二元分类任务,为了优化分割性能,本文定义交叉熵损失 (L_{CE})和平均交集并集损失 (L_{mIoU})的加权组合作为主要的损失函数,这 2 个损失函数共同作用于模型训练,旨在提高分割的准确性和边界的清晰度,具体定义如下:

$$L_{\text{CE}} = -\frac{1}{N} \sum_{i=1}^N (z_i \log_a(\hat{z}_i) + (1-z_i) \log_a(1-\hat{z}_i)) \quad (5)$$

$$L_{\text{mIoU}} = 1 - \frac{\sum_{i=1}^N (z_i * \hat{z}_i)}{\sum_{i=1}^N (z_i + \hat{z}_i - z_i * \hat{z}_i)} \quad (6)$$

式中: z_i 和 $\hat{z}_i$ 分别是第 i 个像素的地面真值和预测标签; N 是图像的总像素数。为了解决前景和背景像素类别不平衡的问题,本文还引入了基于 Dice 系数的 Dice 损失 (L_{Dice}),定义如下:

$$L_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N (z_i * \hat{z}_i)}{\sum_{i=1}^N (z_i + \hat{z}_i)} \quad (7)$$

最终,总损失由交叉熵损失 L_{CE} 、平均交集并集损失 L_{mIoU} 的加权和与 Dice 损失组成,通过调整权重系数,平衡不同损失函数对模型训练的影响。总损失计算如下:

$$L_{\text{all}} = \sum_i^L (L_{\text{CE}} + L_{\text{mIoU}}) + \alpha L_{\text{Dice}} \quad (8)$$

式中: L 是 BOM 表的数量,在本文中设置为 3; α 是权重系数,在本文中设置为 3,以平衡损失。

3 实验结果与分析

3.1 实验设置与评估标准

为了验证本文方法的有效性,在 2 个公共皮肤病变分割数据集 ISIC 2016^[26] 和 ISIC 2018^[27] 上进行实验,这 2 个数据集由国际皮肤成像合作组织 (ISIC) 提供,包含了大量的皮肤病变图像和详细的标注信息,并且覆盖了各种皮肤疾病和病变图像。其中,ISIC 2016 数据集包含 1 279 幅 RGB 皮肤病变图像,训练数据包括 900 幅图像,测试数据包含 379 幅图像,用于衡量使用训练数据开发的系统的性能,训练集和测试集均由专业皮肤科医生生成,确保了数据的准确性和可靠性。ISIC 2018 数据集包

含 2 594 幅 RGB 皮肤病变图像,与文献[20]的设置一样,本文将这些数据随机划分为 1 815 幅作为训练集、259 幅作为验证集、其余 520 幅作为测试集,并在 ISIC 2018 数据集上执行 5 次交叉验证。

实验基于 PyTorch 1. 8. 1 框架实现,使用 Adam 优化器优化模型,在 100 个 Epoch 中设置初始学习率为 0. 000 3。默认批处理大小设置为 32,图像大小为 256×256 像素,以适应模型的输入要求。所有实验均在 NVIDIA RTX 3090 的 GPU 上进行,以确保实验硬件环境的一致性和高效性。

为了定量评估分割性能,采用 5 种广泛使用的评价指标,即骰子相似系数 (Dice, D_{Dice})、交并比 (IoU, I_{IoU})、灵敏度 (Se, S_{Se})、特异性 (Sp, S_{Sp}) 和准确率 (Accuracy, A)。各指标具体计算如下:

$$S_{\text{Se}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (9)$$

$$S_{\text{Sp}} = \frac{N_{\text{TN}}}{N_{\text{TN}} + N_{\text{FP}}} \quad (10)$$

$$A = \frac{N_{\text{TP}} + N_{\text{FN}}}{N_{\text{TP}} + N_{\text{TN}} + N_{\text{FP}} + N_{\text{FN}}} \quad (11)$$

$$D_{\text{Dice}} = \frac{2 \cdot N_{\text{TP}}}{2 \cdot N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}}} \quad (12)$$

$$I_{\text{IoU}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}}} \quad (13)$$

式中: N_{TP} (真阳性) 和 N_{TN} (真阴性) 分别表示正确分割的皮肤病变像素数和背景像素数; N_{FP} (假阳性) 表示被错误标记为皮肤病变像素的背景像素; N_{FN} (假阴性) 表示被错误预测为背景像素的皮肤病变像素。

3.2 实验结果分析

3.2.1 定量对比分析

本文方法在 ISIC 2018 和 ISIC 2016 数据集上的实验结果如表 1 所示。从中可以看出: 本文方法在 2 个数据集上的交并比均大于 0. 87, 有力证明了模型在皮肤病变图像分割中的高精度优势, 并且模型的误判率较低; 特异性分别达到了 0. 981 8 和 0. 974 1, 表明本文方法在识别非病变区域时具有较高的识别率, 能够准确区分病变区域; 准确率分别达到了 0. 959 6 和 0. 969 2, 进一步验证了本文方法在皮肤病变图像识别中具有较高的精度。

表 1 本文方法在公开数据集上的性能分析

Table 1 Performance analysis of the method in this paper on public datasets

数据集	Dice	IoU	Se	Sp	Accuracy
ISIC 2018	0. 886 7	0. 873 8	0. 904 9	0. 981 8	0. 959 6
ISIC 2016	0. 935 5	0. 912 0	0. 955 4	0. 974 1	0. 969 2

为了验证本文方法的分割性能,与其他 12 种主流的皮肤病变图像分割方法进行定量和定性对比分析,对比方法包括 U-Net^[8]、TransUNet^[28]、FAT-Net^[29]、ADF-Net^[30]、RA-Net^[31]、CTO^[22]、SUTrans-NET^[32]、CCT^[33]、HDS-Net^[34]、文献[35]方法、Swin-Unet^[15]和 RMMLP^[36]。其中,TransUNet、FAT-Net、CTO、ADF-Net、SUTrans-NET、HDS-Net、CCT、RMMLP 都使用了混合编码器。本文方法与其他方法在 2 个数据集上的对比结果如表 2 所示(加粗字体为各列最优结果,下划线字体为次优结果)。可以看出,本文方法在提升交并比和特异性上效果显著,得分均优于其他方法,说明总体分割性能较其他方法具有一定优势。本文方法在交并比和特异性上的提升,得益于全局与局部特征的信息融合、边界区域的细粒度增强以及分割解码过程中对边界信息的动态引导。本文方法的交并比达到了 87.38%,比 U-Net 高 10.05 百分点。由于本文方法通过双重特征编码学习了皮肤病变的局部细节和全局特征,Dice 系数和灵敏度相比 U-Net 分别提高了 3.22 和 2.49 百分点,交并比、特异性、准确率相比 CTO 分别提高了 2.88、1.25 和 0.27 百分

点。实验结果表明,本文方法能够准确提取多种形态规模的皮肤病变特征。Swin-Unet 使用纯 Transformer 架构,在医学图像分割任务中表现优异,在 ISIC 2018 数据集上,本文方法的 Dice 系数、特异性、准确率相比 Swin-Unet 分别提升了 0.40、3.22、1.43 百分点。相比之下,本文方法使用双重特征编码模块来提取不同尺度图像的全局特征,更高效地提取了多尺度图像特征。此外,通过引入双重特征编码和边界特征增强,本文方法在 ISIC 2018 和 ISIC 2016 数据集上的交并比和特异性相比于性能次优的方法均有提升,并且在 ISIC 2016 数据集上,Dice 系数、交并比、灵敏度、特异性、准确率均达到了最优结果。与 U-Net 相比,Dice 系数、交并比、灵敏度、特异性、准确率分别提高了 4.71、9.36、5.38、0.85 和 2.26 百分点。模型在 ISIC 2016 数据集上的性能结果普遍高于 ISIC 2018 数据集,这主要得益于 ISIC 2016 数据集背景噪声较少且病变区域清晰,有助于模型更高效地拟合病变区域特征。实验结果说明,本文网络通过结合皮肤病变边界特征和全局-局部信息,实现了对边界模糊病变区域的精准分割,在皮肤病变图像分割中具有较好的效果。

表 2 本文方法与其他方法在 2 个数据集上的性能对比

Table 2 Performance comparison between the method in this paper and other methods on two datasets										%
方法	ISIC 2018					ISIC 2016				
	Dice	IoU	Sp	Se	Accuracy	Dice	IoU	Sp	Se	Accuracy
U-Net	85.45	77.33	96.97	88.00	94.04	88.84	81.84	96.56	90.16	94.66
TransUNet	88.88	81.13	<u>97.89</u>	90.08	95.94	92.12	85.40	96.19	93.69	95.49
FAT-Net	89.03	82.02	96.99	91.00	95.78	91.59	85.30	96.02	92.59	96.04
ADF-Net	90.82	—	97.41	<u>92.34</u>	96.70	92.89	—	96.41	<u>94.45</u>	96.53
RA-Net	93.25	—	94.16	93.63	95.84	92.94	—	<u>96.79</u>	92.27	<u>96.70</u>
CTO	91.20	<u>84.50</u>	96.93	90.31	95.69	91.89	85.18	—	—	—
SUTrans-NET	88.58	81.63	94.02	91.93	93.33	91.43	85.48	96.19	93.66	95.90
文献[35]方法	—	82.10	—	—	93.50	—	—	—	—	—
CCT	<u>92.72</u>	<u>86.81</u>	—	—	—	<u>93.21</u>	<u>87.70</u>	—	—	—
HDS-Net	89.80	83.00	97.40	89.40	—	91.40	85.10	95.90	93.70	—
Swin-Unet	88.27	—	94.96	—	94.53	91.43	—	94.48	—	95.52
RMMLP	91.98	85.40	—	—	—	92.62	86.32	—	—	—
本文方法	88.67	87.38	98.18	90.49	<u>95.96</u>	93.55	91.20	97.41	95.54	96.92

尽管本文方法在多个指标上表现优异,但在 ISIC 2018 数据集上的部分实验性能指标略低于对比方法,这主要是由于部分图像伪影样本比例较高、病变特征复杂性高以及小病变区域占比高,增加了分割任务的难度。在处理此类复杂病变时,模型可能难以准确捕捉病变区域的细微特征,从而导致方法在 ISIC 2018 数据集上的 Dice 系数和灵敏度表现

不佳。小病变区域的特征表达弱,全局注意力机制可能在小病变捕获上注意力分散,这些复杂样本对模型提出了更高的鲁棒性和特征捕获要求,进一步影响了模型的灵敏度和准确率指标。

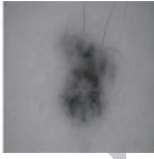
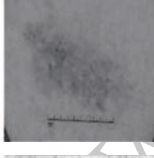
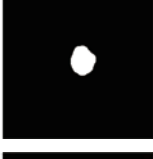
3.2.2 定性对比分析

本文网络在 ISIC 2018 和 ISIC 2016 数据集上部分图像的可视化分割结果如表 3 所示。通过对比

分析可以看出,本文方法能够准确分割 ISIC 2018 数据集中多种复杂形态的皮肤病变图像,不仅能够准确分割出病变区域,还能维持边界的基本走向。此外,针对 ISIC 2018 数据集中存在毛发、手术标记等伪影的皮肤病变图像,本文网络也能够准确分割出皮肤病变区域。在不同成像环境下,本文网络可以在各种病灶大小和形状上准确分割出病变的真实区域,特别是对于模糊边界的病变,即使在病变与周围健康皮肤之间对比度特别低的情况下,本文网络仍然可以很好地区分皮肤病变背景,得到清晰的病变边界,获得较好的皮肤病变图像分割结果。

表 3 本文方法分割结果

Table 3 Segmentation results of the method in this paper

数据集	输入图像	分割结果
ISIC 2016		
		
		
ISIC 2018		
		
		

为进一步验证本文方法对皮肤病变图像分割的有效性,将其与 U-Net 及两种混合模型在 ISIC 2018 和 ISIC 2016 数据集上进行可视化对比,结果如图 2 所示。其中,TransUNet 是第一个基于 Transformer 的医学图像分割框架,其将图像特征转为序列来编码全局上下文,利用 U 形混合架构设计获得低级 CNN 特征;CTO 结合 CNN、ViT 和显

式边界检测操作,利用中间特征映射来合成高质量的边界监督掩码。由于本文提出了边界特征增强模块和边界引导解码模块,与 U-Net、TransUNet、CTO 相比,本文网络在分割精度和边界处理上均具有一定的优势。在定性对比实验中,选择了几个病变尺寸小、形状不规则、存在伪影和边界模糊的测试样本,并分析了不同模型的分割结果。图 2 第一列是输入图像,第二列是输入图像对应的真值,从中可以看出,在分割形态规模较大的病变图像时,大多数方法会错误地分割类似于病变区域的正常皮肤区域,从而导致过度分割和分割不足。相较于其他 3 种网络,本文网络能更加准确地分割出各种形状和大小的病变。当面对模糊的病变边界时,如第四行和第五行的病变边界,本文网络由于构建了边界特征增强模块和边界引导解码模块,可以准确地捕获长距离依赖关系和边界细节信息,获得了良好的分割性能。第二行和第五行的图像有背景噪声,例如头发和手术标记,大多数方法都能实现相对准确的分割,但本文网络提出的边界引导解码模块能够实现最精确的分割,准确识别此类病变的区域像素。实验结果表明,本文网络的分割结果更接近地面真值,同时有效改善了模糊病变的皮肤病变分割效果,提高了皮肤病变图像的分割精确性。

为了进一步验证边界增强机制对分割性能的贡献,实验选取了一张边界模糊并且病变区域与正常皮肤区域对比度低的挑战样本进行可视化分析,结果如图 3 所示。U-Net 仅采用局部卷积操作,分割结果存在边界噪声较多、整体形态不完整的问题;TransUNet 尽管引入了 Transformer 架构,在边界连续性上有所改善,但仍难以避免小面积误分割问题,尤其在细节复原上表现不足;CTO 虽然结合了混合编码器的优势,在边界分割精度上较优,但在复杂背景和细粒度病变区域的复原能力上仍有不足。相比之下,本文方法的分割结果边界更为清晰、完整,同时有效减少了背景噪声的干扰,病变区域的整体轮廓得到了更好的复原。本文方法通过引入边界增强机制,在模糊病变边界的分割任务中取得了显著的性能提升。边界特征增强模块通过自注意力机制聚焦病变边界区域的特征,动态建模上下文信息,显著提高了边界精度。同时,利用边界窗口的细粒度特征捕获,有效抑制了伪影和背景噪声的干扰。从图 3 中可以看到,本文方法避免了传统方法中背景误分割与分割区域破碎的问题,充分证明了边界增强机制在复杂病变分割中的有效性。

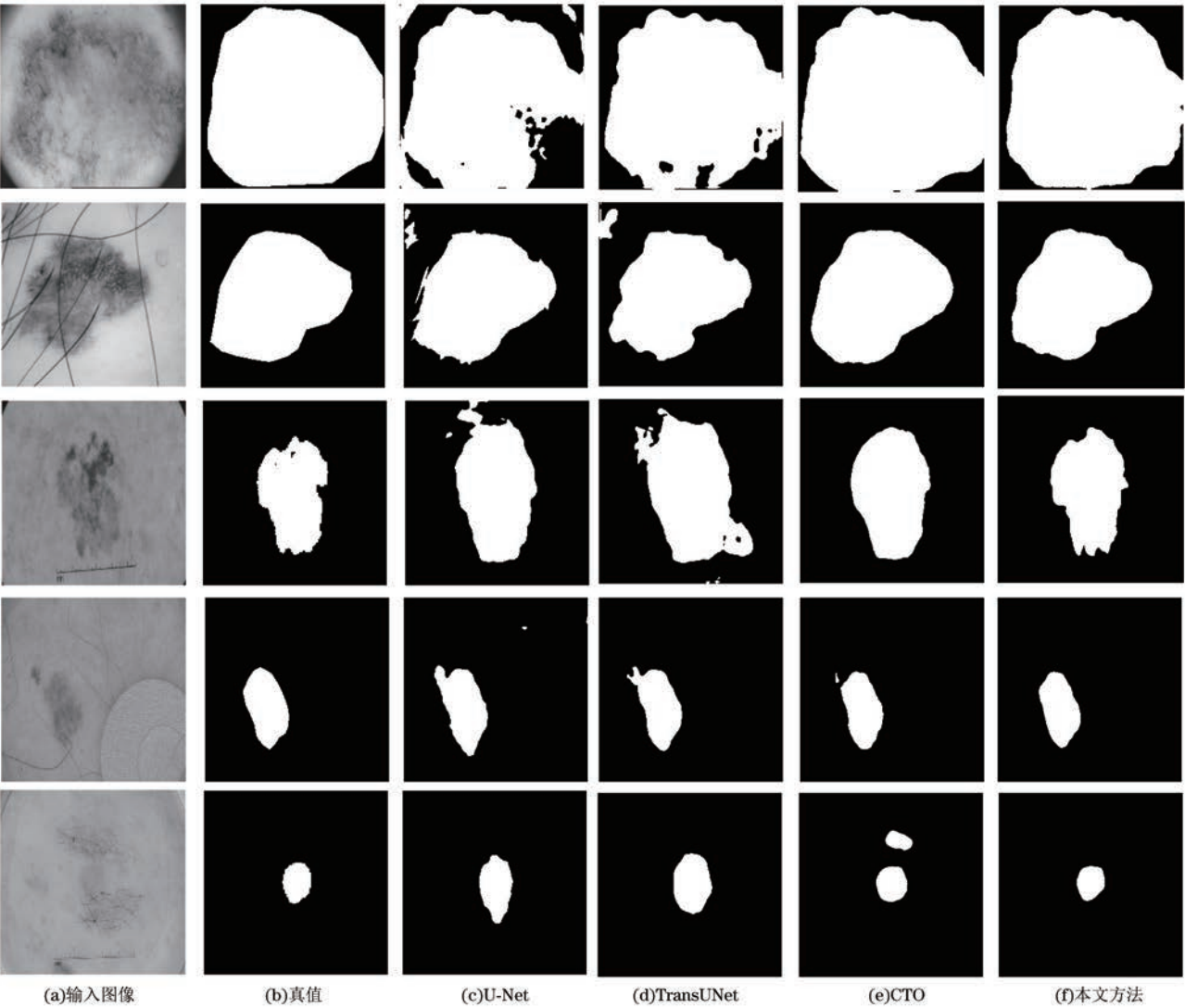


图 2 不同方法的分割结果对比

Fig.2 Comparison of segmentation results of different methods

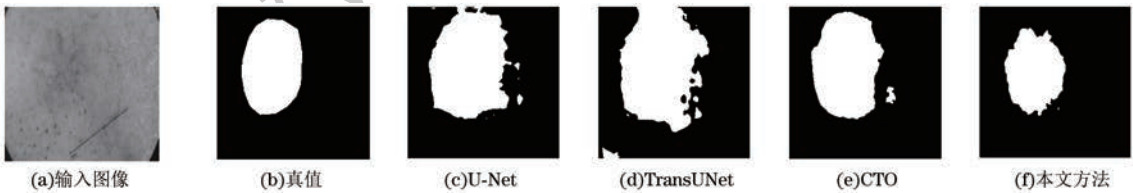


图 3 模糊病变图像分割结果

Fig.3 Segmentation results of blurred lesion images

3.2.3 复杂度分析

为了进一步评估方法的复杂度,将本文方法与 ViT^[12]、TransUNet^[28]、ResUNet++^[37] 进行比较,并且在相同的计算环境和相同的数据扩展上运行,评估 Params、FLOPs 两个指标。表 4 所示为模型在 ISIC 2018 数据集上的复杂度对比结果。本文模型参数量达到 109.10×10^6 ,表明模型具有一定的复杂度,同时具有更强的表示能力和潜在的学习复杂特征的能力,这是因为设计了双重特征编码模块来同时学习全局和局部特征,以及边界特征增强模

块来提高分割的精确度。本文模型计算复杂度达到了次优结果,比混合模型 TransUNet 和基于 CNN

表 4 不同方法在 ISIC 2018 数据集上的复杂度对比

Table 4 Comparison of complexity of different methods on the ISIC 2018 dataset

方法	Params/ 10^6	FLOPs/ 10^9
ViT	110.62	26.91
ResUNet++	87.00	95.00
TransUNet	<u>105.32</u>	38.52
本文方法	109.10	<u>27.76</u>

的 ResUNet++ 表现更好,也保证了模型能够精确地捕捉和分割出皮肤病变区域。虽然本文网络在复杂度上有所牺牲,但其在分割精度和特征表示能力上具有一定优势。

3.3 模型消融实验

本文通过设计 6 组消融实验,采用 Dice 系数、交并比、灵敏度、特异性、准确率 5 个指标作为评价指标,在 ISIC 2018 和 ISIC 2016 数据集上进一步验证双重特征编码模块(DFC)、边界特征增强模块(BFE)、边界引导解码模块(BGD)的有效性。本文模型将 Res2Net-50 作为基础网络(baseline),不包含上述模块,将边界引导解码替换为普通卷积解码器,并依次进行 6 组消融实验,即 B、B+DFC、B+BFE、B+DFC+BFE、B+BFE+BGD 以及本文完整方法(B 代表主干网络 Res2Net-50),这 6 组实验分别代表仅添加主干网络、在主干网络基础上添加一个模块、在主干网络基础上添加两个模块以及本文完整方法。消融实验结果如表 5 所示,在 ISIC 2018 数据集上,使用主干网络分割皮肤病变的 Dice 系数为 0.865 0,而加入 B+DFC 和 B+BFE 的 Dice 系数分别提升至 0.876 3 和 0.874 5,这说明每个模块都

能有效提升分割结果。进一步地,B+DFC+BFE 和 B+BFE+BGD 的 Dice 系数分别为 0.884 0 和 0.885 1,优于只加入一个模块的情况。在使用所有模块的情况下,本文方法的 Dice 系数为 0.886 7,这说明结合两个模块比单独使用一个模块能获得更优的结果。添加 3 个模块后,所有指标均达到最高。DFC 的引入能有效提升各项指标,这得益于 DFC 能够融合两个编码器在各个阶段的特征,更好地捕获皮肤病变图像的全局和局部信息,提高网络的特征学习能力,完成更全面的特征提取。通过引入 BFE 辅助皮肤病变图像分割任务,能够有效提高网络在各个数据集上的性能表现,说明本文提出的边界特征增强模块能强化网络对病变边界区域的感知能力。在解码过程中引入边界信息作为引导,可以准确识别整体的病变区域,关注前景和背景之间的关系,确保分割结果的边界更加准确和清晰,提高分割结果的精度。总之,结合双重特征编码模块、边界特征增强模块、边界引导解码模块,能够有效处理皮肤病变区域尺寸变化大、形状和边界不规则、边界模糊的病变区域分割不准确等问题。

表 5 在 ISIC 2018 和 ISIC 2016 数据集上的消融实验结果

Table 5 Ablation experiment results on the ISIC 2018 and ISIC 2016 datasets

方法	ISIC 2018					ISIC 2016				
	Dice	IoU	Se	Sp	Accuracy	Dice	IoU	Se	Sp	Accuracy
B	0.865 0	0.852 8	0.875 5	0.963 2	0.946 2	0.910 9	0.879 7	0.945 1	0.953 6	0.959 2
B+DFC	0.876 3	0.861 5	0.885 7	0.974 7	0.948 9	0.916 1	0.898 3	0.950 5	0.961 3	0.962 7
B+BFE	0.874 5	0.861 0	0.884 5	0.975 8	0.948 4	0.914 7	0.899 3	0.948 9	0.962 0	0.963 6
B+DFC+BFE	0.884 0	0.870 7	0.896 6	0.980 2	0.955 7	0.929 1	0.908 2	0.954 5	0.965 5	0.964 5
B+BFE+BGD	0.885 1	0.864 4	0.903 6	0.981 3	0.953 3	0.931 8	0.907 8	0.952 9	0.969 3	0.966 0
本文方法	0.886 7	0.873 8	0.904 9	0.981 8	0.959 6	0.935 5	0.912 0	0.955 4	0.974 1	0.969 2

4 结束语

为提高皮肤病变边界模糊图像的分割精度,本文提出一种特征学习和边界引导的皮肤病变图像分割网络。通过双重特征编码,增强各种形态规模的皮肤病变的全局-局部特征提取能力。边界特征增强通过计算皮肤病变的边界窗口,更精确地定位病变区域的边缘。实验结果显示,边界引导解码通过边界增强特征引导网络进行分割,能够有效保持边界的清晰度和准确性,提高分割结果的精度。虽然本文方法可以有效改善模糊边界的皮肤病变图像分割效果,在多个评估指标上具有优势,但是由于双分支编码结合了 CNN 和 Transformer 两种网络结构,增加了模型的复杂性和参数量。后续研究将围

绕上述工作深入展开,通过优化模型架构来降低计算复杂度,实现更好的资源利用效率,进一步提升皮肤病变图像分割模型的性能和实用性。同时,未来将重点优化伪影鲁棒性和小病变分割性能,提升模型在实际场景中的应用价值。

参考文献

- [1] SIEGEL R L, GIAQUINTO A N, JEMAL A. Cancer statistics, 2024 [J]. CA: A Cancer Journal for Clinicians, 2024, 74(1): 12-49.
- [2] GE Z Y, DEMYANOV S, CHAKRAVORTY R, et al. Skin disease recognition using deep saliency features and multimodal learning of dermoscopy and clinical images[EB/OL]. [2024-08-05]. https://link.springer.com/chapter/10.1007/978-3-319-66179-7_29.
- [3] ALGARIN Y A, MCCULLUM C, PATEL V A. 33256 skin

- cancer screening practices among dermatologists: a survey study[J]. *Journal of the American Academy of Dermatology*, 2022, 87(3): AB203.
- [4] 赵宏, 王泉. 基于 Swin-Transformer 的黑色素瘤图像病灶分割研究[J]. *计算机工程*, 2024, 50(8): 249-258.
ZHAO H, WANG X. Study on lesion segmentation of melanoma images based on Swin-Transformer[J]. *Computer Engineering*, 2024, 50(8): 249-258. (in Chinese)
- [5] SUN Y H, DAI D W, ZHANG Q N, et al. MSCA-Net: multi-scale contextual attention network for skin lesion segmentation[J]. *Pattern Recognition*, 2023, 139: 109524.
- [6] CELEBI M E, IYATOMI H, SCHAEFER G, et al. Localization of lesions in dermoscopy images using ensembles of thresholding methods[EB/OL]. [2024-08-05]. https://link.springer.com/chapter/10.1007/978-3-540-92957-4_95.
- [7] GOMEZ D D, BUTAKOFF C, ERSBOLL B K, et al. Independent histogram pursuit for segmentation of skin lesions[J]. *IEEE Transactions on Biomedical Engineering*, 2008, 55(1): 157-161.
- [8] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/1505.04597>.
- [9] HUANG H M, LIN L F, TONG R F, et al. UNet 3+: a full-scale connected UNet for medical image segmentation[C]// *Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Washington D. C., USA: IEEE Press, 2020: 1055-1059.
- [10] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [11] 李大湘, 杨福杰, 刘颖, 等. 融入交叉注意力编码的皮肤病变分割网络[J]. *光学精密工程*, 2024, 32(4): 609-621.
LI D X, YANG F J, LIU Y, et al. Skin lesion segmentation network with cross-attention coding[J]. *Optics and Precision Engineering*, 2024, 32(4): 609-621. (in Chinese)
- [12] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16 × 16 words: transformers for image recognition at scale[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2010.11929>.
- [13] CHEN F, WANG J C, MAGNIER B, et al. G2LL: global-to-local self-supervised learning for label-efficient transformer-based skin lesion segmentation in dermoscopy images[C]// *Proceedings of the 20th IEEE International Symposium on Biomedical Imaging (ISBI)*. Washington D. C., USA: IEEE Press, 2023: 1-5.
- [14] XIN C, LIU Z F, MA Y Z, et al. Transformer guided self-adaptive network for multi-scale skin lesion image segmentation[J]. *Computers in Biology and Medicine*, 2024, 169: 107846.
- [15] CAO H, WANG Y Y, CHEN J, et al. Swin-Unet: Unet-like pure transformer for medical image segmentation[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2105.05537>.
- [16] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Washington D. C., USA: IEEE Press, 2016: 770-778.
- [17] YUAN Y D, LO Y C. Improving dermoscopic image segmentation with enhanced convolutional-deconvolutional networks[J]. *IEEE Journal of Biomedical and Health Informatics*, 2019, 23(2): 519-526.
- [18] LI J, WANG J W, LIN F W, et al. DSEUNet: a lightweight UNet for dynamic space grouping enhancement for skin lesion segmentation[J]. *Expert Systems with Applications*, 2024, 255: 124544.
- [19] VASWANI A, NOAM S, NIKI P, et al. Attention is all you need[EB/OL]. [2024-08-05]. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.
- [20] ZHENG S X, LU J C, ZHAO H S, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Washington D. C., USA: IEEE Press, 2021: 6877-6886.
- [21] WANG J C, CHEN F, MA Y X, et al. XBound-former: toward cross-scale boundary modeling in transformers[J]. *IEEE Transactions on Medical Imaging*, 2023, 42(6): 1735-1745.
- [22] LIN Y, ZHANG D, FANG X, et al. Rethinking boundary detection in deep learning models for medical image segmentation[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2305.00678>.
- [23] GUO Q Q, FANG X Y, WANG L B, et al. LGANet: local-global augmentation network for skin lesion segmentation[C]// *Proceedings of the 20th IEEE International Symposium on Biomedical Imaging (ISBI)*. Washington D. C., USA: IEEE Press, 2023: 1-5.
- [24] GAO S H, CHENG M M, ZHAO K, et al. Res2Net: a new multi-scale backbone architecture[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(2): 652-662.
- [25] FU J, LIU J, TIAN H J, et al. Dual attention network for scene segmentation[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Washington D. C., USA: IEEE Press, 2020: 3141-3149.
- [26] GUTMAN D, CODELLA N C F, CELEBI E, et al. Skin lesion analysis toward melanoma detection: a challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC)[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/1605.01397>.
- [27] CODELLA N, ROTEMBERG V, TSCHANDL P, et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the International Skin Imaging Collaboration (ISIC)[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/1902.03368>.
- [28] CHEN J N, LU Y Y, YU Q H, et al. TransUNet: transformers make strong encoders for medical image segmentation[EB/OL]. [2024-08-05]. <https://arxiv.org/abs/2102.04306>.
- [29] WU H S, CHEN S H, CHEN G L, et al. FAT-Net: feature adaptive transformers for automated skin lesion segmentation[J]. *Medical Image Analysis*, 2022, 76: 102327.
- [30] HUANG Z W, DENG H M, YIN S C, et al. ADF-Net: a novel adaptive dual-stream encoding and focal attention decoding network for skin lesion segmentation[J]. *Biomedical Signal Processing and Control*, 2024, 91: 105895.
- [31] NAVEED A, NAQVI S S, IQBAL S, et al. RA-Net: region-

- aware attention network for skin lesion segmentation [J]. Cognitive Computation, 2024, 16(5): 2279-2296.
- [32] LI Y Q, TIAN T H, HU J, et al. SUTrans-NET: a hybrid transformer approach to skin lesion segmentation [J]. PeerJ Computer Science, 2024, 10: e1935.
- [33] XU Z J, GUO X Y, WANG J. Enhancing skin lesion segmentation with a fusion of convolutional neural networks and transformer models [J]. Heliyon, 2024, 10(10): e31395.
- [34] XUE Y, CHEN X Y, LIU P, et al. HDS-Net: achieving fine-grained skin lesion segmentation using hybrid encoding and dynamic sparse attention [J]. PLoS One, 2024, 19(3): e0299392.
- [35] 贵向泉, 张馨月, 李立. 高分辨率皮肤黑色素瘤图像的两阶段式分割算法 [J]. 计算机工程, 2023, 49(11): 267-274.
- GUI X Q, ZHANG X Y, LI L. Two-stage segmentation algorithm of high resolution skin melanoma image [J]. Computer Engineering, 2023, 49 (11): 267-274. (in Chinese)
- [36] JI C, DENG Z H, DING Y, et al. RMMLP: rolling MLP and matrix decomposition for skin lesion segmentation [J]. Biomedical Signal Processing and Control, 2023, 84: 104825.
- [37] JHA D, SMEDSRUD P H, RIEGLER M A, et al. ResUNet++: an advanced architecture for medical image segmentation [C] // Proceedings of the IEEE International Symposium on Multimedia (ISM). Washington D.C., USA: IEEE Press, 2020: 2250-2255.

文字编辑 吴云芳

栏目编辑 宋 圆