

# 基于双路特征的宫颈细胞核分割

崔文成 杨丹 邵虹

(沈阳工业大学信息科学与工程学院 沈阳 110870)

**摘要:** 宫颈细胞核的分割问题研究对宫颈癌筛查诊断具有重要意义,但受边缘模糊及存在干扰物影响给分割任务带来了巨大挑战。针对此问题提出一种基于 DeepLabV3+网络的细胞核分割方法,首先充分利用主干网络的输出进行多尺度特征融合,并引入注意力机制,构建了细胞团分割模型,以减少背景中干扰物对细胞核分割的影响;基于此设计了融合 Transformer 与 ResNet50 的双路特征提取模块,兼顾模型对全局信息的获取及低层上下文特征的敏感度,提高了模型对细胞核与干扰信息的辨别能力。实验结果表明,算法在宫颈细胞核的分割任务中取得了良好的分割效果,均交并比为 0.832 9,较 DeepLabV3+模型提高了 2.33%,且与其他方法相比获得了更优的性能指标。

**关键词:** 宫颈细胞核;多尺度特征融合;Transformer;语义分割;DeepLabV3+

**中图分类号:** TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Cervical nuclear segmentation based on two-path features

Cui Wencheng Yang Dan Shao Hong

(School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

**Abstract:** The study of cervical nucleus segmentation is of great significance for cervical cancer screening and diagnosis, but it brings great challenges to the segmentation task due to the influence of blurring edge and interference. In response to this problem, a nuclear segmentation method based on the DeepLabv3+ network was proposed. Firstly, utilizing the output of the backbone network for multi-scale feature fusion and introducing an attention mechanism, a cell mass segmentation model was built to reduce the effects of interferences in the background on nuclear segmentation. Based on this, a two-path feature extraction module combining Transformer and ResNet50 was designed, which takes into account the sensitivity of the model to global information acquisition and low-level context features, and improves the discrimination ability of the model to nuclei and interference information. The experimental results show that the algorithm has achieved good segmentation results in the task of cervical cell nuclei, and MIoU is 0.832 9, which has increased by 2.33% and obtained better performance indicators compared to other methods.

**Keywords:** cervical cell nuclei; multi-scale features fusion; transformer; semantic segmentation; DeepLabV3+

## 0 引言

国际癌症研究机构数据显示宫颈癌是世界上影响女性健康的第四大常见癌症,我国的发病人数更是占全球的 18.6%,可见宫颈癌是全球健康问题之一<sup>[1]</sup>。宫颈癌潜伏期可长达 8~10 年,症状也不明显,通过预防性接种 HPV 疫苗可对其进行有效一级预防,依托筛查和早诊早治可对其进行二级预防<sup>[2]</sup>。因此,癌变筛查是宫颈癌治疗的关键之一。由于我国人口基数大,病理医生人数匮乏,宫颈癌筛查率远低于全球战略水平<sup>[3]</sup>。液基薄层细胞学检测作为宫颈癌筛查的有效手段,意味着专家阅片准确率与效率的提

升将是满足大量人口周期性筛查的关键。随着计算机视觉的发展,实现机器阅片将有效地减少时间和人力上的消耗,开发高效的自动化细胞分析系统也将加速医学诊断。细胞学检查中,携带遗传信息的细胞核的形状、大小、染色等变化是肿瘤诊断的主要依据,此外核质比也是确定非典型增生程度的标准之一,由此可知细胞核的精准分割是其自动化的关键部分。

由于宫颈细胞图像中存在中性粒细胞、细胞内容物及杂质等干扰细胞核的分割,使用传统分割方法不能充分利用图像深层特征,而使得分割效果有限,难以实现自动化。目前研究中得以广泛应用的细胞核分割方法是全卷积神经

网络方法,比如 FCN<sup>[4]</sup>、SegNet<sup>[5]</sup>、UNet<sup>[6]</sup>、UNet++<sup>[7]</sup>等,可以自动提取图像特征、精度高、误差小。Qing 等<sup>[8]</sup>提出一种基于 Mask R-CNN 改进的核图像分割方法,结合低级语义特征进行模型训练,使用注意力机制为每个低层特征分配权重,使模型学习更具目的性。Shi 等<sup>[9]</sup>提出一种细粒度的多器官细胞核分割模型,除了多尺度卷积、空洞卷积及残差注意模块外,为了进一步提高解码器中特征质量引入 U-Net 策略,将更多的语义信息集成到细节特征中,Dice 系数为 0.949。Tomas 等<sup>[10]</sup>提出一种具有度量空间嵌入的深度学习模型,在不同的多组织显微镜图像中准确地分割了细胞核,深度度量学习对整体学习过程有额外的促进作用,从而提高了分割性能。但基于多尺度特征与注意力机制的引入在提高模型性能的同时增加了模型复杂度。不同与此,宋杰等<sup>[11]</sup>提出的细胞核分割模型,采用级联稀疏卷积与决策树集成学习的方式,能够很好地减少模型的参数量,实现端到端的语义分割。Xiao 等<sup>[12]</sup>提出一种基于极坐标表示的核分割模型,在效率和有效性上都优于目前的技术水平。宫颈细胞核分割任务中,Zhang 等<sup>[13]</sup>提出一种具有双路融合注意力特征的二叉树状模型分割宫颈细胞核,充分利用低级细节信息和高级语义信息,引入双路融合注意力机制选择性地强调重要的特征而抑制次要特征。针对异常宫颈细胞分割精度低的问题,Zhao 等<sup>[14]</sup>提出了一种轻量级特征注意力网络(LFANet),其编码器部分增加了特征层注意模块来精确恢复像素位置,利用全局高层信息作为底层特征的引导,捕捉通道特征的依赖关系,该模型能准确分割宫颈细胞核。但分割宫颈细胞图像在考虑目标边界的分割精度外,更应关注宫颈细胞图像与细胞核特征相似的干扰物,如中性粒细胞、细胞内容物及杂质等,尤其是中性粒细胞极易影响细胞核分割。以 Chen 等<sup>[15]</sup>提出的 DeepLabV3+ 模型为基础分割宫颈细胞图像有以下 3 点优势,1)加入空洞卷积能够扩大适当的感受野捕获多尺度的上下文信息,有利于分割难样本,且不必引入额外的参数而增加计算量。2)采用空间金字塔池化(atrous spatial pyramid pooling, ASPP)模块将不同倍率的空间卷积层串行连接,可以使得模型在多尺度物体上的表现更好<sup>[16]</sup>,利于形态多样的宫颈细胞团分割。3)在解码器部分,融合高层与低层特征有助于提高分割精度,可以比较好地恢复细胞核边界信息。

基于以上研究,本文提出一种宫颈细胞核分割方法,主要工作包括两部分:1)为了降低背景中性粒细胞等干扰物对细胞核分割的影响,在 DeepLabV3+ 网络的基础上,构建了一种基于多尺度特征和注意力机制的细胞团分割模型,使用特征金字塔模块融合多尺度特征有利于保存更多细节信息,采用卷积块注意力模块抑制干扰信息的影响。2)通过此细胞团分割模型得到细胞区域后,有效去除背景中的干扰信息,以提取感兴趣区域后的图像为基础,融合 Transformer 与 ResNet50 分支的特征构建了细胞核分割

模型,根据 Transformer 分支对全局信息及低层上下文特征的敏感性,模型能够有效辨别细胞核与干扰特征。

本文的主要贡献包括:1)采用特征金字塔模块融合了主干网络的中层特征,增强了模型对宫颈细胞团边界的分割精度。2)使用了两次混合注意力机制,将卷积块注意力模块应用于编码器的输出特征,使模型关注于细胞核细节特征而非干扰信息。3)在主干网络中融合基于 Transformer 的特征提取,并通过一种双路特征融合模块突出了模型的上下文建模能力,使融合特征表示更紧凑。

## 1 本文算法

针对宫颈细胞图像中存在干扰物影响细胞核分割的问题,在 DeepLabV3+ 模型结构的基础上,构建了一种细胞团分割模型以去除背景区域的中性粒细胞,并利用细胞团分割结果进行感兴趣区域提取作为下一步细胞核分割的测试集。另外,构建具有辨别细胞核与干扰物能力的核分割模型,完成对宫颈细胞核的分割。

### 1.1 基于多尺度特征的细胞团分割

针对宫颈细胞图像中细胞团形状多样且边界不规则,以及存在形态特征与细胞核相似的干扰物问题,在分析了以 ResNet50 为主干网络的 DeepLabV3+ 模型结构的基础上,充分利用主干网络的不同尺度特征优化解码器部分所应用的高层特征,并增加注意力机制,它缓解了干扰物对细胞团分割精度的影响,同时提高了模型对细胞团边界的分割性能,尤其在细胞重叠的凹点处更为明显。结合多尺度特征融合和注意力机制的宫颈细胞团分割模型结构如图 1 所示,包括如下两方面优化:1)在主干网络 ResNet50 中的 4 层 block 输出的多尺度特征 block1~block4 中,将 block2 和 block3 与 ASPP 模块在编码器部分的输出利用特征金字塔结构(feature pyramid network, FPN)融合在一起,替代原来网络模型中的高层特征;2)在解码器部分将 block1 经过  $1 \times 1$  卷积块降维后的特征与 FPN 结构输出的特征均经过卷积块注意力模块处理。

DeepLabV3+ 中,ResNet50 的 4 层 block 输出了多尺度特征 block1~block4,其中 block1 作为编码器输出的低层特征输入到解码器部分;block4 是 ASPP 模块的输入,其经过 ASPP 模块后输出的特征图(ASPP\_out)作为编码器输出的高层特征也输入到解码器部分。在此基础上,利用 FPN 结构将 block2、block3 和 ASPP\_out 进行特征融合,其结构如图 2 所示。FPN 结构将 block2、block3 和 ASPP\_out 依次经过  $1 \times 1$  卷积块统一降维为 256 层,方便上采样后与大尺度特征依次融合,最后再次利用  $1 \times 1$  卷积块降维为原 ASPP\_out 通道数输出。在细胞团分割中,采用 ResNet-50 作为主干网络训练的 DeepLabV3+ 模型得到的分割结果存在细胞团边界分割不准确的问题,尤其是在重叠细胞的凹点处较明显。通过多尺度特征的融合,浅层的大尺度特征和深层的小尺度特征可以在识别出分割对象的

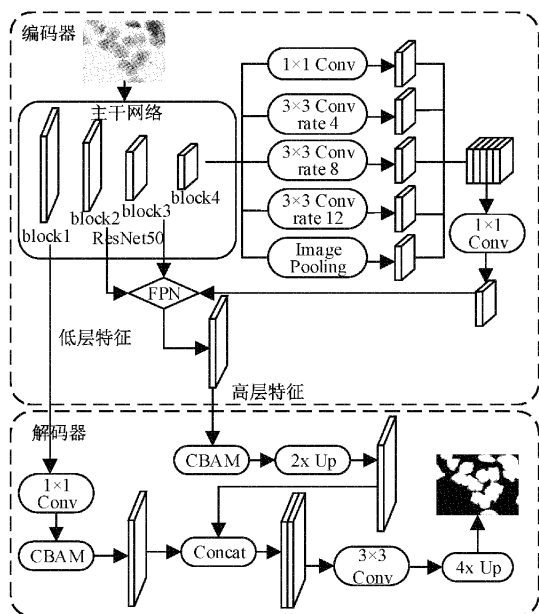


图1 细胞团分割模型结构

同时兼顾模型对细节信息的学习能力<sup>[17]</sup>。由此融合多尺度特征可同时把握更多全局信息和局部信息,增强了模型对浅层特征所提供的轮廓、边缘等局部细节信息的学习,从而提高细胞团分割模型的分割性能。

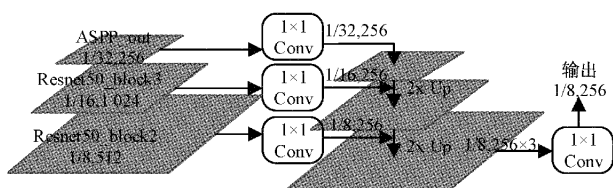


图2 FPN结构

此外,由于在宫颈细胞图像中存在不同类型的干扰物,且其形态特征与细胞核相似,甚至仅凭肉眼辨别时也存在难度,如中性粒细胞和杂质,尤其是部分附着在宫颈细胞边缘的中性粒细胞。这些干扰物应属背景部分,但易被误识别为细胞核而分割到细胞区域,影响细胞团分割效果。通过将卷积块注意力模块(convolutional block attention module, CBAM)分别作用于低层特征 block1 和 FPN 输出的多尺度特征,能够增强模型对目标区域的注意力,并降低其对于干扰物的关注<sup>[18]</sup>,从而通过关注重要信息并抑制干扰信息,提高模型对细胞团的表征能力。CBAM 是一种混合注意力机制,包含通道和空间两个独立的子模块,分别进行通道与空间维度的注意力,且采用先通道再空间的顺序串行组合,其结构如图3所示。以 block1 和 FPN 结构的输出作为输入特征,CBAM 沿着特征映射的通道和空间的两个维度注入注意力映射,并将注意力与输入特征映射相乘,以自适应地细化输入特征映射。

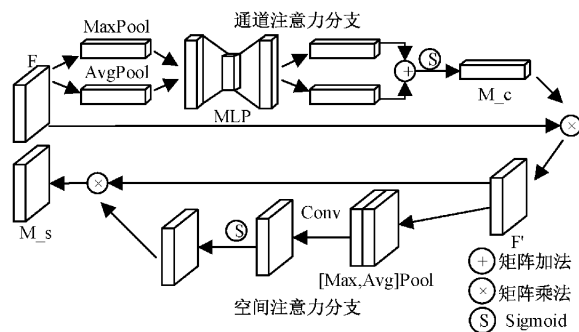


图3 CBAM结构

### 1.2 融合双路特征的细胞核分割

针对图像中宫颈细胞内部存在细胞内容物和中性粒细胞干扰的问题,在细胞团分割模型的基础上,将基于 ResNet50 和 Transformer 的特征提取网络并行运行,构建了细胞核分割模型,其结构如图4所示,主干网络由两个平行的分支组成,它们对信息的处理方式不同,其中 CNN 分支感受野逐渐增大,从局部到全局进行编码;Transformer 分支,从全局自我注意力开始,最后恢复局部细节。两个分支的特征采用双路特征融合(BiFusion)模块融合在一起。提出的分支并行方法有两点优势:首先,它兼顾 Transformer 对全局上下文建模的能力,并利用 CNN 分支弥补模型捕获细节的局限性;其次,利用 BiFusion 模块使得 CNN 和 Transformer 的不同特征的融合表示强大而紧凑。

尽管 CNN 在众多医学图像分割任务中取得了无与伦比的性能,但在获取全局上下文信息方面的能力有限且缺乏效率。由于现有研究中通过产生更大的感受野来获取全局信息,需要连续向下采样并叠加足够深的卷积层,同时带来了3个缺点:1) 深层网络的训练受逐渐减小的特征重用问题影响,其中低级特征被连续的乘法淘汰;2) 当空间分辨率逐渐降低时,对密集预测任务举足轻重的局部信息被丢弃,比如语义分割;3) 训练参数大,深度网络在医学图像数据集小的情况下往往不稳定,容易过拟合。但近两年 Transformer 在计算机视觉领域的应用为此提供了新的机会。

Transformer 分支的设计遵循典型的编码器-解码器结构。具体来说,首先将输入图像  $x \in R^{H \times W \times 3}$  平均分成  $N$  小块(patches),如式(1)所示,其中  $S$  设为 16。

$$N = \frac{H}{S} \times \frac{W}{S} \quad (1)$$

然后将 patches 平铺,传递到输出维数为  $D_0$  的线性嵌入层,得到原始嵌入序列  $e \in R^{N \times D_0}$ 。为了利用空间先验,在  $e$  中加入了一个可学习的相同维度的位置嵌入,得到的嵌入信息  $z_0 \in R^{N \times D_0}$  是 Transformer 编码器的输入,该编码器包含  $L$  层的多头自注意力(multi-headed self-attention, MSA)和多层感知器,其前后各分别应用层归一

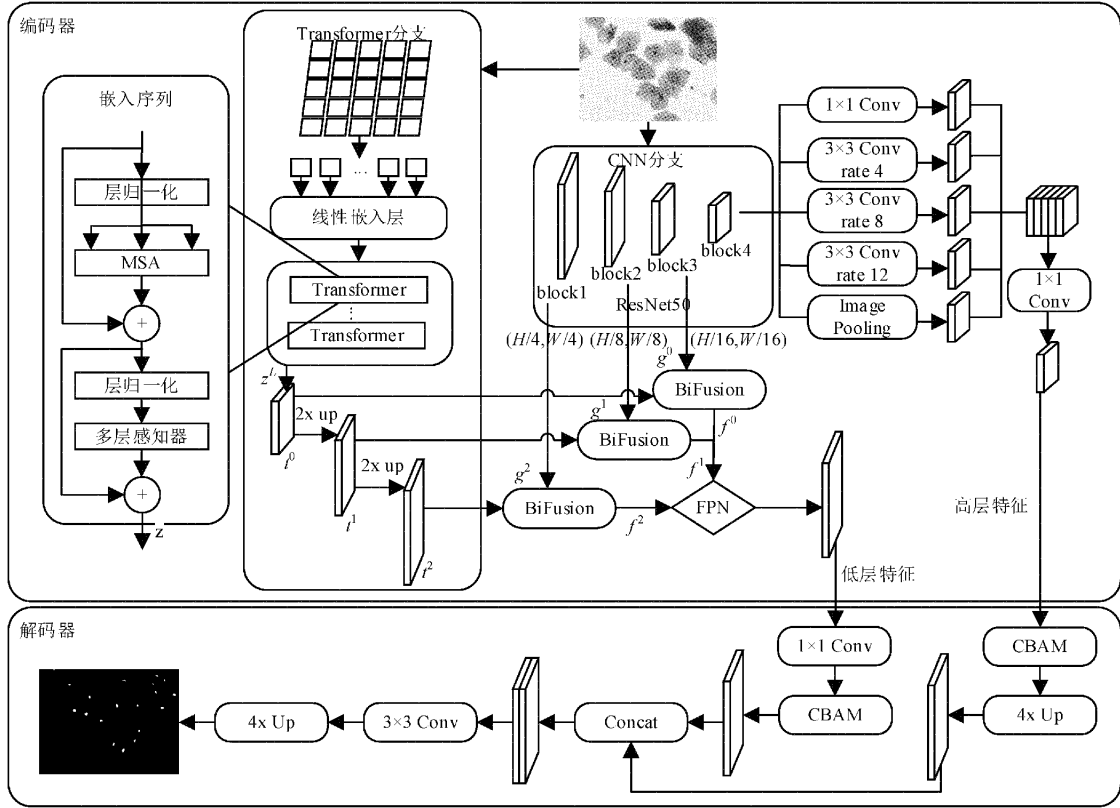


图 4 细胞核分割模型结构

化和残差连接。自注意力机制(SA)是 Transformer 的核心原理,它通过在每一层聚合全局信息来更新每个嵌入 patch 的状态,如式(2)所示。

$$SA(z_i) = \text{softmax}\left(\frac{q_i k^T}{\sqrt{D_h}}\right)v \quad (2)$$

其中,  $[q, k, v] = zW_{qkv}, W_{qkv} \in R^{D_0 \times 3D_h}$  为映射矩阵, 向量  $z_i \in R^{1 \times D_0}$  与  $q_i \in R^{1 \times D_h}$  分别为  $z$  和  $q$  的第  $i$  行。MSA 是 SA 的扩展,它将多个 SA 连接起来,并将维度映射回  $R^{D_0}$ 。对最后一层 Transformer 的输出进行层归一化得到编码序列  $z_L \in R^{N \times D_0}$ 。可学习的位置信息的嵌入更好的利用了图像的局部信息相关性,且利用了注意力机制对全局信息进行建模,保证了 Transformer 对全局上下文信息的提取能力。对于解码器部分,使用渐进式上采样方法,首先将  $z_L$  重构回  $t^0 \in R^{\frac{H}{16} \times \frac{W}{16} \times D_0}$ , 可将其视为具有  $D_0$  通道的二维特征图。然后,使用两个连续的上采样卷积层来恢复空间分辨率,分别得  $t^1 \in R^{\frac{H}{8} \times \frac{W}{8} \times D_1}$  和  $t^2 \in R^{\frac{H}{4} \times \frac{W}{4} \times D_2}$ , 保存  $t^0, t^1, t^2$  不同尺度的特征图,与 CNN 分支对应的特征图进行后期融合。

虽然 Transformer 擅长对全局上下文进行建模,但它在捕获细粒度细节方面存在局限性,特别是对于医学图像,这是因其在建模局部信息时缺乏空间诱导偏差,为此尝试将 ResNet50 与 Transformer 结合。为了有效地结合

来自 ResNet50 和 Transformer 的编码特征构建了 BiFusion 模块,其结构如图 5 所示。将 ResNet50 的 block1~block3 的输出特征  $g^0, g^1, g^2$  与 Transformer 分支解码器下采样的不同尺寸特征图  $t^0, t^1, t^2$  逐一融合。

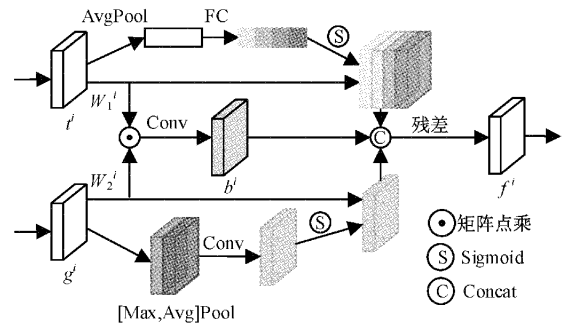


图 5 BiFusion 模块结构

BiFusion 模块中,融合特征  $f^i (i = 0, 1, 2)$  的计算过程如式(3)所示。

$$\begin{aligned} \hat{i} &= \text{ChannelAttn}(t^i) & \hat{g} &= \text{SpatialAttn}(g^i) \\ \hat{b} &= \text{Conv}(t^i W_1^i \odot g^i W_2^i) & f^i &= \text{Residual}([\hat{b}, \hat{i}, \hat{g}]) \end{aligned} \quad (3)$$

其中,  $W_1^i \in R^{D_i \times L_i}, W_2^i \in R^{C_i \times L_i}, \odot$  代表矩阵点乘, Conv 是一个  $3 \times 3$  的卷积层。采用 SE-Block 实现对 Transformer 分支输出特征的通道注意力,促进其全局信

息的提取;采用CBAM的空间注意力部分作为空间滤波器处理CNN分支输出特征,增强局部细节并抑制无关区域。然后,通过矩阵点乘操作对来自两个分支的特征之间的细粒度交互信息进行建模。最后,将交互特征 $\hat{b}^i$ 和分支特征 $\hat{i}^i$ 、 $\hat{g}^i$ 连接,并经过一个残差块,由此生成的新特征 $f^i$ 有效地捕获当前位置的全局和局部上下文信息。由此可知,BiFusion模块可以在捕获全局信息的同时,保存模型对低层上下文特征的敏感度;也在特征提取过程中有效地连接ResNet50和Transformer的编码特征,从而提高模型对细胞核与干扰物的辨别能力。此外,区别于细胞团分割模型,细胞核分割模型将BiFusion模块的输出特征 $f^0$ 、 $f^1$ 及 $f^2$ 利用FPN结构融合在一起,替代原模型block1特征作为低层特征输入到解码器部分,而非高层特征,因为融合后的大分辨率特征更利于局部信息的上下文建模。

## 2 实验与结果分析

### 2.1 实验细节

#### 1) 数据集

本文使用SIPaKMeD数据集<sup>[19]</sup>的宫颈细胞图像作为实验数据,其图像均来自于巴氏染色的正常和病变的宫颈鳞状细胞,但每张图像中仅有部分细胞含有细胞核和细胞质的标注信息,因此在病理医生的帮助下使用Labelme软件手工制作了更全面精准的数据集标签。文中使用了141张分辨率为2048 pixel×1536 pixel的RGB图像,为了增加数据的泛化能力和鲁棒性,对每张图像进行了水平、垂直、中心点翻转、随机修剪、平移、旋转、高斯噪声干扰,对比度变换等8倍的数据扩充。

#### 2) 实验环境

实验使用Window10系统,基于PyTorch框架以Python作为开发语言,使用32核2.8 GHz AMD EPYC 7543处理器进行模型的训练和测试,并利用NVIDIA GeForce RTX 3090显卡加速模型运算。

#### 3) 模型训练

实验数据包括564张带有标注数据的训练、验证、测试集,按照7:2:1的比例划分。输入网络的图像设置为512 pixel×512 pixel的尺寸大小。训练过程分为2个阶段:首先固定主干网络,使用预训练模型的权重训练分割模型,批处理大小设为8,训练50轮,初始学习率为0.0005,采用Adam优化器进行参数更新;然后将主干网络解除冻结,参考前面轮次权重联合训练,批处理大小设为2,同样训练50轮,初始学习率衰减0.1倍,设为0.00005。此外通过实验验证DeepLabV3+模型中的空洞空间金字塔池化模块采用4个卷积块设置膨胀因子为[1,4,8,12]时性能最佳。

为了更好地确定图像中的语义和真实细节,本文将Dice损失与交叉熵损失结合起来共同训练分割模型,如

式(4)~(6)所示。

$$Loss_{Dice} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (4)$$

$$Loss_{CE} = -[y \log \hat{y} + (1 - y) \log(1 - \hat{y})] \quad (5)$$

$$Loss = Loss_{Dice} + Loss_{CE} \quad (6)$$

其中, $|X|$ 和 $|Y|$ 分别表示分割的真实标签和预测结果, $|X \cap Y|$ 表示集合X和Y的交集, $y$ 代表二分类中模型预测正确的概率值。

#### 4) 评价指标

为评估本文分割模型的分割准确度,选择类别平均像素准确率(mean pixel accuracy, MPA)和平均交并比(mean intersection over union, MIoU)<sup>[20]</sup>作为评价指标。其中,MPA代表预测结果中正确分类的像素与总预测值的比例,完美分割该值为1,定义如式(7)所示。

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij}} \quad (7)$$

MIoU计算所有类别的IoU均值,IoU计算真实分割类别与预测的分割类别这两个集合的交集与并集的重合比例,如式(8)所示。

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (8)$$

其中,数据有 $k$ 类,这里均取2, $P_{ii}$ 表示第 $i$ 类被正确分类的像素数量, $P_{ij}$ 表示第 $i$ 类被错误分到第 $j$ 类的像素数量。

## 2.2 实验结果

### 1) 细胞团分割消融实验

为了验证多尺度特征融合(F)与注意力机制(A)对于细胞分割起到了积极的作用,依次进行消融实验来验证所选用结构的有效性,结果如表1所示,其中DL模型表示以ResNet50为主干网络的原DeepLabV3+模型,DL\_F表示单独使用多尺度特征融合的模型,DL\_A表示单独使用CBAM的模型,DL\_FA表示共同使用的模型。

表1 细胞团分割消融实验性能评价结果

模型	F	A	MPA	MIoU
DL	×	×	0.921 0	0.896 4
DL_F	√	×	0.944 1	0.903 5
DL_A	×	√	0.958 9	0.914 7
DL_FA	√	√	<b>0.962 2</b>	<b>0.917 8</b>

观察表1,对比DL和DL\_F的结果可发现多尺度特征融合能有效提升细胞团分割模型性能;对比DL和DL\_A的结果可以验证卷积块注意力模块的引入也可以提升模型性能。模型DL\_FA的MIoU为0.9178,相比于原模型

DL 提高了 0.021 4,可知提出的架构对细胞团分割是有效的,且比单独使用两个结构的效果更优。此外,如图 6 所

示,展示了各模型的细胞团分割结果,为了便于观察其细节,选择了图像中的典型区域分析。

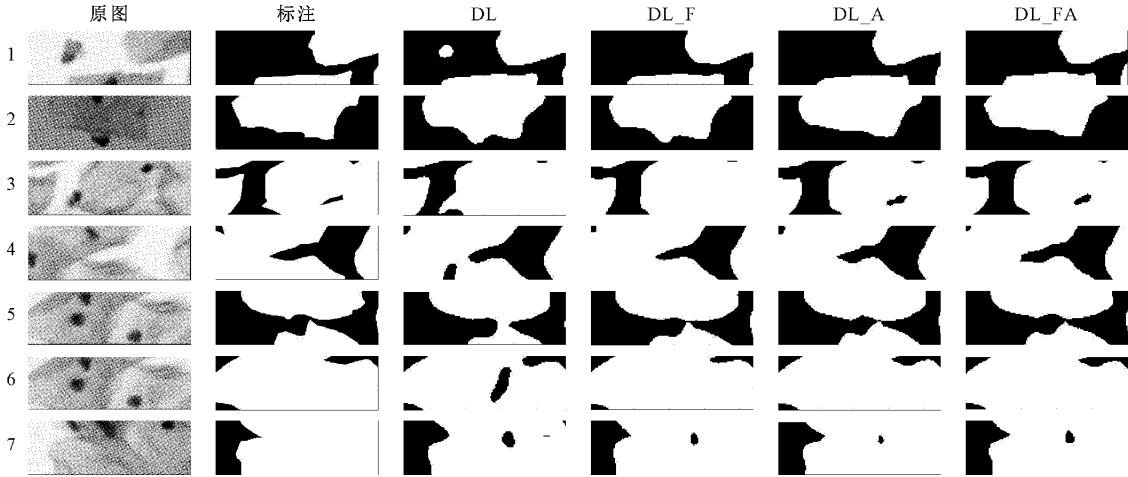


图 6 不同模型的细胞团分割结果

观察图 6 中的分割结果,整体来说,优化后模型分割精度更高,尤其是在最后一列的测试结果中,模型对边缘与细节的分割更准确。第 1 行图像背景存在干扰物,DL 模型中被误分割入细胞团,但在 DL\_FA 模型中得到了准确分割;第 2、3 行区别于第 1 行中干扰物位置,中性粒细胞均附着在细胞边缘,相比前者与细胞核的形态、颜色以及在细胞中位置信息更为相近,使用 DL 及 DL\_F 均被误分割,但对比 DL\_A 模型的结果可发现其被识别为背景区域中,可见注意力机制对干扰物的影响起到了积极作用,能够有效辨别细胞核与具有相似特征的干扰物,尤其针对于附着在细胞边缘的中性粒细胞干扰,且 3 行中 DL 模型并未识别出的 3 个细胞重叠形成的三角区域也被分割准确;第 4、5 行中可观察到 DL 模型在多处细胞重叠处存在欠分割现象,分割精度差,对比 DL 与 DL\_F 的结果可知使用多尺度特征融合能够有效提高细胞边缘的分割精度,其中第 5 行细胞间未重叠,DL 模型的分割结果粘连在了一起,但 DL\_F 及其后续模型均分割较准确;第 6、7 列存在过分割的情况,由于细胞自身或与其他细胞重叠部分颜色更深,细胞未重叠且透明度低部分的特征不明显问题尤为突出,第 6 行中 DL\_FA 模型将 DL 模型误分割部分分割正确,但第 7 行中分割错误,仍有优化空间。

2)细胞核分割消融实验

为验证双路特征融合对细胞核分割的作用,基于不同主干网络的模型进行消融实验,性能评价的对比结果如表 2 所示,其中 DL\_R、DL\_T 分别表示使用 ResNet50、Transformer 作为主干网络的模型,DL\_FT 表示主干网络融合两分支特征的模型。

在表 2 中,DL\_R 及 DL\_T 的 MPA 和 MIoU 值均高于 DL,但 MIoU 提升不明显,主要是因为对于细胞核分割任务,分割效果受干扰物影响更大,干扰物的误分割易造成

表 2 细胞核分割消融实验性能评价结果

模型	ResNet50	Transformer	MPA	MIoU
DL	×	×	0.914 5	0.809 6
DL_R	√	×	0.927 3	0.826 6
DL_T	×	√	0.925 6	0.814 6
DL_RT	√	√	<b>0.937 6</b>	<b>0.832 9</b>

假阳性升高影响评分,但也验证了多尺度特征融合与注意力机制能够提升细胞核分割模型性能,不过单独使用 Transformer 进行特征提取时得分更低;对比 DL\_RT 与 DL\_R、DL\_T 的结果,MPA 和 MIoU 均有提升,说明通过融合 Transformer 分支增强对全局上下文建模的能力有利于提升模型对细胞核的识别能力,不过模型 DL\_RT 的 MIoU 为 0.832 9,相比于原模型虽有提高,相比于细胞团的分割精度却稍低,有待进一步提升。具体地各模型上细胞核的进一步分割结果,如图 7 所示。

图 7 中主要包括两种情况,1)是细胞团内核的周围存在中性粒细胞干扰,在第 1、2 行由于干扰信息特征更明显,尽管 DL 模型造成了误分割,但是使用 DL\_R 模型解决其干扰;第 3 行核下方的干扰物 DL、DL\_R 及 DL\_T 模型均被误分割,但使用 DL\_RT 模型进一步优化后分割正确,同时也验证了 Transformer 融合的效果;第 4 行中与第 3 行相似,但最后未分割正确,不过误分割像素占比减少,还有优化空间。2)是细胞内容物干扰,最后一行明显观察到细胞内环境更复杂,核周围被大面积细胞内容物包围,且 DL 模型误分割大部分内容物区域,而 DL\_R 模型降低了误分割面积却未分割出细胞核,DL\_T 模型分割结果与其相似,最后一列同样存在误分割,但同时微小部分的核区域得到正确分割,虽有效果,但相较中性粒细胞,模型对细胞内容物的识别能力较差。

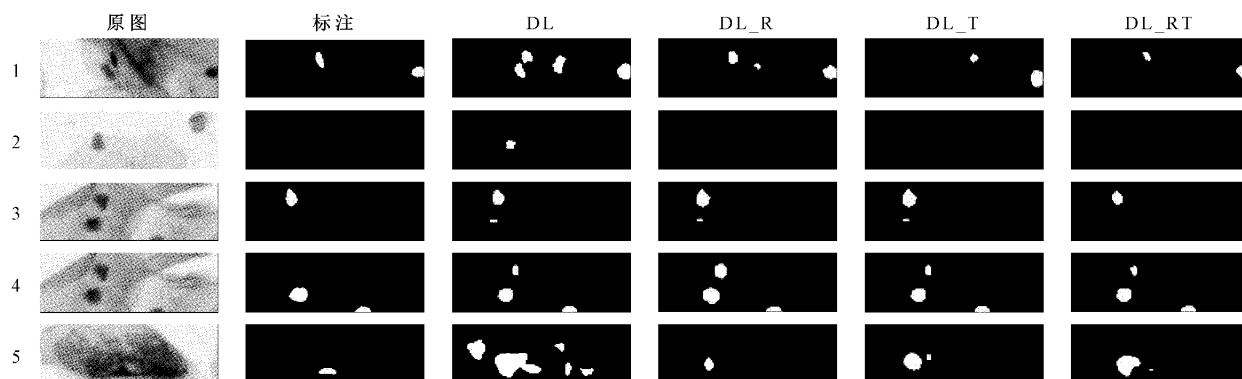


图7 不同模型的细胞核分割结果

### 3) 不同细胞核分割算法的结果分析

为了更好地验证模型的性能,将本文提出的细胞核分割算法与经典的卷积神经网络 FCN、SegNet 及针对医学图像分割的 UNet 和 UNet++ 网络进行对比。使用本文数据集分别训练以上 4 种网络模型, batch-size 设为 4 迭代 100 次,分别计算每个模型的 MPA 与 MIoU 值,结果如表 3 所示。

表3 不同算法细胞核分割结果对比

模型	MPA	MIoU
FCN	0.835 7	0.756 7
SegNet	0.844 2	0.760 9
UNet	0.863 5	0.804 5
UNet++	0.919 2	0.822 1
本文方法	<b>0.937 6</b>	<b>0.832 9</b>

从表 3 可以看出, SegNet 在 FCN 基础上加入了编解码结构, 明显提高了准确率; UNet++ 在 UNet 基础上加入了密集的跳跃连接, 准确率也有提高; DL\_RT 相比于其他全卷积网络及编解码结构融合了 Transformer 模型, MIoU 进一步提升至 0.832 9, 比较优的 UNet++ 网络高出 0.108。

## 3 结 论

文中分别搭建了宫颈细胞团及细胞核的分割模型。以 DeepLabV3+ 网络为基础, 通过融合多尺度特征提高模型对细胞团边界的分割性能, 并引入注意力机制降低中性粒细胞等的干扰, 使细胞团分割模型更加专注于待分割的关键区域。此外利用 Transformer 利于全局上下文建模的特点提取到更充分的全局与局部特征, 结合 ResNet50 的特点, 使核分割模型能够有效辨别细胞核与干扰物而得到更好的训练, 从而进一步提升了细胞核分割精度。综上, 细胞团分割的 MIoU 为 0.917 8, 细胞核分割的 MIoU 为 0.832 9, 能够较准确地分割出所有宫颈细胞及其核区域。但是对于细胞重叠程度高的难样本仍需要进一步提高; 针

对宫颈细胞重叠的问题, 如何利用细胞核与细胞间位置关系对当前结果后处理也是接下来研究的重点。

本文设计的宫颈细胞核分割方法可以应用于的自动化细胞分析系统, 辅助宫颈癌诊断, 为接下来的细胞图像内容分析和理解与细胞分类提供了必要的基础准备。但所提出的细胞核分割算法主要基于采用巴氏染色且具有中性粒细胞作为干扰物的宫颈鳞状细胞图像, 当待分割宫颈细胞图像采用 H&E 等其他方式染色时, 需重新训练分割模型且将影响分割结果; 另外, 当出现新的干扰物或因细胞癌变程度严重导致图像环境更复杂时, 模型分割能力也将下降。

## 参考文献

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLO-BOCAN 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] 张韶凯, 赵方辉, 乔友林. 中国宫颈癌防治研究 20 年历程与成就 [J]. 中华流行病学杂志, 2020, 41(6): 809-812.
- [3] 张梅, 包鹤龄, 王丽敏, 等. 2015 年中国宫颈癌筛查现状及相关因素分析 [J]. 中华医学杂志, 2021, 101(24): 1869-1874.
- [4] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [5] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [6] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]. Medical Image Computing and

- Computer-Assisted Intervention, 2015: 234-241.
- [7] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHS N, et al. UNet ++: Redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 39(6): 1856-1867.
- [8] QING Z, ZHU Q, HUANG X, et al. U-shaped feature extractor used on Mask RCNN for cell nuclei image segmentation [J]. Journal of Physics: Conference Series, 2020, 1646(1): 012069.
- [9] SHI T, LI C, XU D, et al. Fine-grained histopathological cell segmentation through residual attention with prior embedding [J]. Multimedia Tools and Applications, 2022, 81(5): 6497-6511.
- [10] TOMAS I, AGNE P, KRISTINA S. Enhancing multi-tissue and multi-scale cell nuclei segmentation with deep metric learning [J]. Applied Sciences, 2020, 10(2): 615-630.
- [11] 宋杰, 肖亮, 练智超. 级联稀疏卷积与决策树集成的病理图像细胞核分割方法[J]. 自动化学报, 2021, 47(2): 378-390.
- [12] XIAO W, JIANG Y, YAO Z, et al. Polar representation based cell nucleus segmentation in non-small cell lung cancer histopathological images [J]. Biomedical Signal Processing and Control, 2021, 70(2): 103028.
- [13] ZHANG J, LIU Z, DU B, et al. Binary tree-like network with two-path fusion attention feature for cervical cell nucleus segmentation [J]. Computers in Biology and Medicine, 2019, 108(1): 223-233.
- [14] ZHAO Y, FU C, XU S, et al. LFANet: Lightweight feature attention network for abnormal cell segmentation in cervical cytology images [J]. Computers in Biology and Medicine, 2022, 145(1): 105500.
- [15] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 801-818.
- [16] 陈彦彤, 李雨阳, 陈伟楠, 等. 基于深度语义分割的遥感图像海面舰船检测研究 [J]. 仪器仪表学报, 2020, 41(1): 233-240.
- [17] 赵文慧, 杨霄, 孟丽洁. 充分利用多尺度特征改进 UNet 实现皮肤病变分割 [J]. 电子测量技术, 2022, 45(2): 110-116.
- [18] 何晓云, 许江淳, 陈文绪. 基于改进 U-Net 网络的眼底血管图像分割研究 [J]. 电子测量与仪器学报, 2021, 35(10): 202-208.
- [19] PLISSITI M E, DIMITRAKOPOULOS P, SFIKAS G, et al. Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images [C]. IEEE International Conference on Image Processing, Greece; IEEE, 2018: 3144-3148.
- [20] 赵俊强, 李骥, 王昌, 等. DA-GAN 肺结节分割网络研究 [J]. 中国测试, 2022, 48(9): 118-124.

### 作者简介

崔文成, 硕士, 副教授, 主要研究方向为智能信息处理等。

E-mail: 624618764@qq.com

杨丹(通信作者), 研究生, 主要研究方向为图像处理与识别等。

E-mail: 2279373382@qq.com

邵虹, 博士, 教授, 主要研究方向为智能信息处理等。

E-mail: 576022085@qq.com