

基于全局特征引导和注意力的宫颈细胞核分割^{*}

董展豪 陈 燕 侯宏花 张鹏程 桂志国

(中北大学信息与通信工程学院 太原 030051)

摘要: 针对宫颈细胞中正常和异常细胞核大小形态各异、细胞图像中存在干扰物影响分割等问题,本文提出一种基于全局特征引导和注意力的分割网络来提高细胞核分割精度。首先以U型网络结构作为主体,引入全局特征引导模块,充分提取各个阶段的特征获得不同层次的全局上下文信息,克服了U型网络单阶段上下文信息提取能力不足的问题以更好地处理不同形状的细胞核,提高边缘分割精度;其次,引入改进的注意门结构抑制图像中干扰物的信息,突出细胞核的信息,提高模型对干扰信息的辨别能力。在Herlev数据集上的实验结果表明,本文的方法能够有效地提高细胞核分割精度,在定量分析中Dice系数为0.9413,相较于其他方法具有一定的优势。

关键词: 宫颈细胞核分割;全局特征引导;注意门;U型网络;巴氏涂片

中图分类号: TN91 文献标识码: A 国家标准学科分类代码: 510.4050

Cervical nuclear segmentation based on global feature guidance and attention

Dong Zhanhao Chen Yan Hou Honghua Zhang Pengcheng Gui Zhiguo

(School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

Abstract: This paper presents a segmentation network that utilizes global feature guidance and attention to enhance the precision of cell segmentation, addressing the variability in size and shape of normal and abnormal nuclei in cervical cells as well as interference in cell images. Firstly, the U-shaped network structure is utilized as the main framework, with the introduction of a global feature guide module to comprehensively extract features at each stage for obtaining global context information at different levels. This overcomes the limitation of insufficient extraction ability of single-stage context information in U-shaped networks, enabling better handling of nuclei with different shapes and improving edge segmentation accuracy. Secondly, an improved attention-gate structure is incorporated to suppress interfering information in images, emphasize nucleus information, and enhance model discrimination against interfering data. Experimental results on the Herlev dataset demonstrate that our proposed method effectively enhances nuclear segmentation precision, achieving a Dice coefficient of 0.9413 in quantitative analysis which presents certain advantages compared to other methods.

Keywords: cervical nuclear segmentation; global feature guidance; attention gate; U-shaped network; Pap smear

0 引言

宫颈癌是全球妇女第四大常见癌症,据世界卫生组织统计,2020年估计有60.4万例新发病例和34.2万例死亡。如果早期诊断并及时治疗,宫颈癌可以治愈。所以对女性来说及时预防并尽早发现宫颈细胞异常尤为重要。目前最常用的早期宫颈癌筛查方法是宫颈细胞涂片筛查,也称巴氏涂片筛查^[1]。此方法需要大量专业的病理医生手动筛查,其过程繁琐,工作量大,长时间工作下容易出错,因此计

算机辅助筛查可以大大提高诊断效率,减轻医生的负担,在该领域具有重大意义。

宫颈细胞学筛查中,细胞核的形状、大小、染色等的变化是诊断细胞是否异常的主要依据,并且Phoulad等^[2]的研究已经证实,细胞核特征在宫颈癌的诊断过程中具有显著价值。宫颈细胞的任何异常变化,都会伴随着特定的细胞核形态或结构的改变,因此本文主要研究宫颈细胞巴氏涂片中细胞核的分割。

传统的图像分割方法根据图像的灰阶分布、色彩以及

收稿日期:2024-04-26

* 基金项目:山西省自然科学基金面上项目(202103021224204)资助

纹理等信息将图像划分为不重叠的区域,使得同一区域呈现相似的特征。Chalfoun 等^[3]通过分析局部像素强度的变化,识别出显著差异区域来标定细胞主体位置,随后应用对比度阈值来界定细胞边缘,实现细胞的分割。Phoulady 等^[4]提出了迭代阈值的方法分割细胞核,先通过形态学方法粗分割出细胞核,然后不断改变阈值直到获得精确的分割结果。Lee 等^[5]利用超像素分割和细胞轮廓细化的方法来分割重叠宫颈细胞。但这些算法大多数只提取了底层的特征信息,无法提取高级的语义信息。在面对复杂的细胞核形状时,往往会由于获得特征信息较少出现误分割或分割不准确等问题。

近几年来,随着深度学习的不断发展,目前广泛应用的细胞核分割网络为全卷积神经网络,如 FCN、U-Net、U-Net++ 等。全卷积网络(fully convolutional network,FCN)是一种用于图像语义分割的深度学习模型^[6],通过移除全连接层并将卷积运算应用于整个网络,能够处理任意尺寸输入图像,实现像素级别的预测输出。刘一鸣等^[7]针对细胞核形态复杂多样的特性,提出了一种整合全卷积神经网络与全连接条件随机场的细胞核分割技术方案。崔文成等^[8]提出利用 DeepLabv3+ 和 Transformer 模型获取多尺度信息来分割宫颈细胞核。Ronneberger 等^[9]提出 U-Net 模型,U-Net 通过创新的跳跃连接机制,将编码器深层的丰富高维特征有效地传递至解码器层,从而在一定程度上有效防止了重要特征信息在处理过程中的损失。正因为这一独特的编解码架构设计,U-Net 模型取得了显著的性能并在医学图像分割领域被广泛应用。Zhou 等^[10]提出的 UNet++ 模型在 U-Net 的基础上通过重构跳跃连接增强多级特征融合,旨在更有效地结合编码器和解码器不同层次的特征表示,从而提高图像分割任务的精度和性能。Gu 等^[11]提出了密集空洞卷积模块和多核池化模块来获得更多的抽象特征和保留更多的空间信息,以提高医学图像分割的性能。Chen 等^[12]将 U-Net 和 Transformer 相结合,提出 TransUNet 增强细节来分割医学图像。Li 等^[13]提出了基于局部注意和全局依赖的宫颈细胞核分割方法,利用通道和空间注意力机制优化了细胞核分割的细节。Zhao 等^[14]提出的 PSPNet 引入金字塔池化模块,通过在模型尾部叠加不同尺度的池化操作,PSPNet 能高效融合多尺度上下文信息,提升了分割的精度和泛化能力。Fu 等^[15]提出的 DANet 融合了空间注意力和通道注意力机制。前者聚焦像素级局部上下文增强边界感知,后者调节通道重要性提炼特征,二者协同作用使模型既能全局理解图像,又能精细处理局部,极大提升分割精度和鲁棒性。Jha 等^[16]将两个 U-Net 堆叠,依次使用两个编码器和两个解码器来捕获更多的语义特征,并利用空间金字塔池化模块提取上下文信息来提高分割精度。Qin 等^[17]提出 U2Net 架构,U2Net 采用两层嵌套的 U 形架构,增强了特征的表达能力,还允许网络在保持高分辨率的同时深入学习,残差 U 型块的设计

促进梯度流动,有助于训练更深的网络,同时避免过高的计算负载,具有广泛的适用性。基于 U-Net 模型的网络结构分割图像都取得了令人满意的分割结果,然而,在其网络的各个阶段中大多只关注了局部特征,对全局上下文信息的有效捕获和利用仍存在不足之处。

对于 U-Net 编码器解码器结构,在编码器的深度学习进程中,所获取的全局上下文信息逐渐向下传输至深层网络,但由于单个阶段的特征提取能力相对有限,这些关键信息可能在传递过程中被逐渐削弱甚至淡化。另一方面,当前阶段采用的简单跳跃连接方式,虽然能够将不同层次的信息进行合并,但它并未对全局信息给予充分重视,仅仅是将局部信息无差别地堆叠合并,这样容易引入低级特征中的无关噪声,进而增加像素误分类的风险^[18]。这对于宫颈细胞的分割结果有很大影响,由于宫颈细胞图像的分割会受到图像中与细胞核特征相似物的干扰,如细胞内容物及杂质等,还会因为细胞病变导致细胞核呈现大小不同、形态各异的形状。这意味着,在分割过程中不仅要求精细化定位细胞边界,还需鉴别并排除那些易混淆、类似细胞核特征的非目标区域,以提升整体分割效果的准确性和可靠性。

为了提高细胞核的分割精度,本文提出了一种基于全局特征引导和注意力的分割网络(global feature guidance and attention network,GFGA-Net)。该网络以残差 U 型网络 Resnet34^[19]作为主干网络,移除了末尾的平均池化层和全连接层,加速网络收敛,避免梯度消失。为了充分识别形态各异的细胞核,使用全局上下文特征引导(global feature guidance,GFG)模块,充分提取并利用各个阶段的特征,保留更多的细节信息。为了有效避免细胞图像中杂质特征对细胞核分割的干扰,采用注意门^[20](attention gate,AG)自动学习关注不同形状和大小的细胞核,抑制背景中的干扰信息,突出细胞核的特征。本研究提出的网络模型具备端到端训练的能力,无需进行额外后处理步骤,在 Herlev 数据集的实验中展现出了出色的性能和竞争力。

1 算法设计

1.1 模型的整体结构

本文所提出方法模型的整体结构如图 1 所示。它是一种典型的编码器-解码器结构,由编码模块、全局上下文特征引导模块(GFG)、注意门模块(AG)和解码模块组成。

模型的左边是编码模块,采用预训练的 Resnet34 残差网络作为主干网络,Resnet34 网络结构如图 2 所示,可以更好地提取图像特征并避免训练时梯度消失和梯度爆炸问题,编码模块由五个阶段组成,每个阶段中最大池化层(max pooling)对特征图进行下采样,每次图像都会减小为原来的一半。解码器是由两个 3×3 的卷积和一个双线性插值上采样组成,解码器在对特征进行恢复的同时在跳跃连接部分利用全局上下文引导模块融合全局特征,并在注意门模块整合前一阶段不同尺度的特征信息,着重关注宫

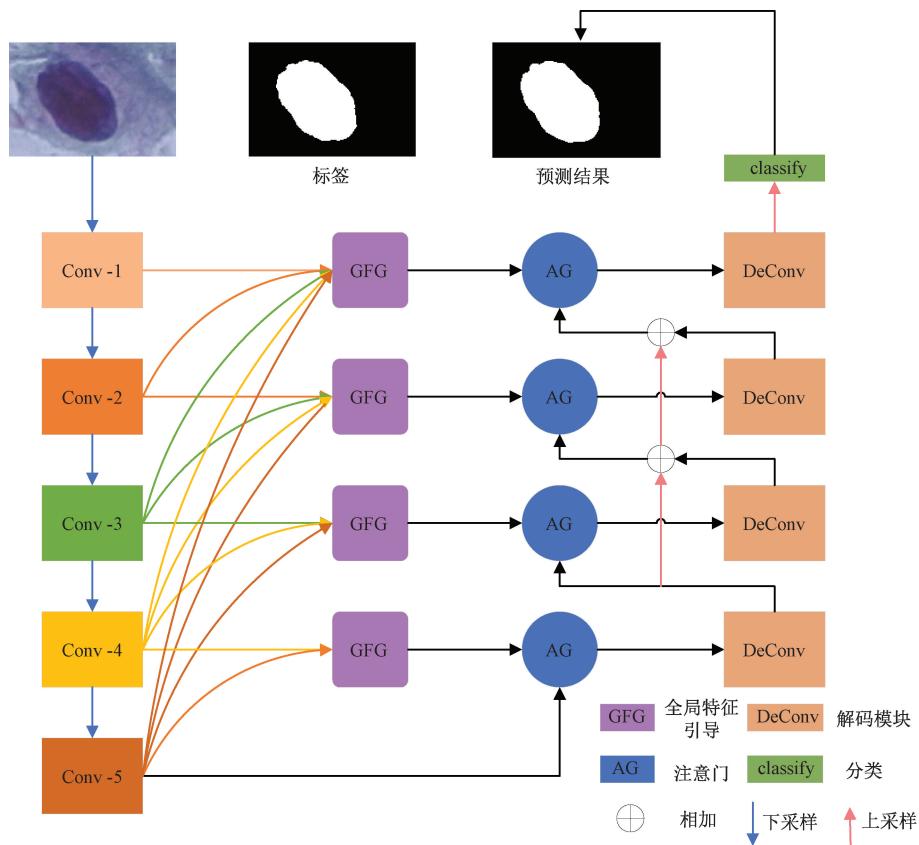


图 1 GFGA-Net 结构图

Fig. 1 GFGA-Net structure

层的名称	输出的大小	Resnet34
Conv-1	128×128	$7 \times 7, 64, \text{stride}2$
Conv-2	64×64	$3 \times 3 \text{maxpool}, \text{stride}2$
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$
Conv-3	32×32	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$
Conv-4	16×16	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$
Conv-5	8×8	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$

图 2 Resnet34 结构图

Fig. 2 Resnet34 structure

颈细胞核的特征,抑制无关的噪声信息,最后得到预测的分割结果。

1.2 全局上下文特征引导模块

由于宫颈细胞核常常表现出形态各异、大小各不相同

的特点,需要从不同尺度获取全局上下文信息才能得到比较准确的分割。在 U-Net 网络结构中,设计跳跃连接的目的是将编码阶段得到的高分辨率特征映射与解码阶段经过上采样的特征映射进行融合拼接,然而单纯依赖这种直接

拼接的方式并不能完全滤除低层特征中存在的无关噪声及冗余信息干扰,为了更好地筛选低级特征,需要借助更具指向性的高级语义信息。为此利用如图 3 所示的全局特征引导模块,逐步将全局上下文信息融入到各个阶段。

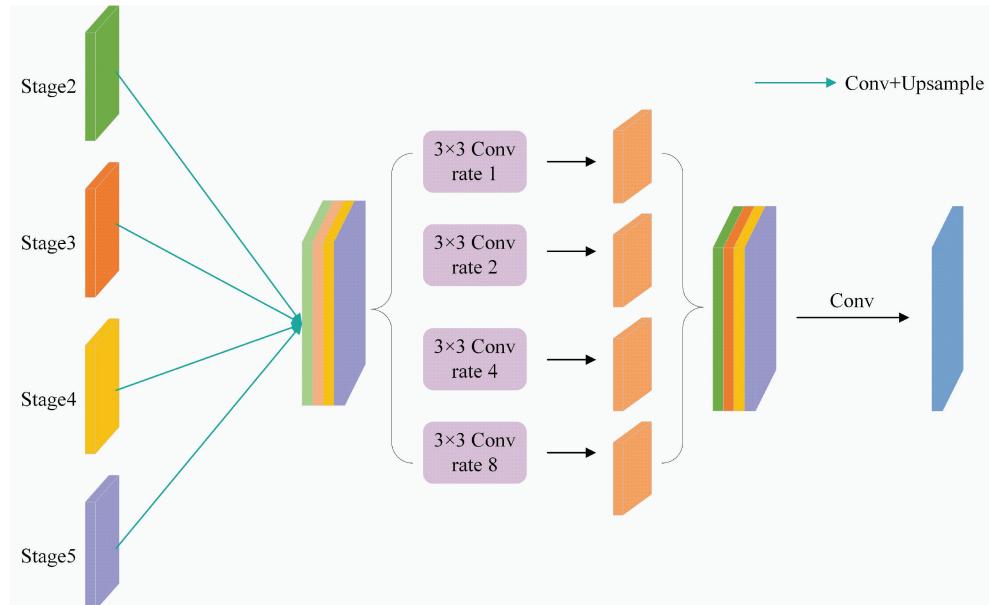


图 3 全局特征引导模块
Fig. 3 Global feature guidance module

在本模块的特征提取步骤中,通过将当前阶段生成的特征图与来自各个更高阶段的特征相结合,以此构建并优化跳跃连接过程。以第二阶段为例,首先,来自第三、四、五阶段的特征图通过一个 3×3 的卷积映射到与第二阶段相同的通道空间,之后对三、四、五阶段映射后的特征上采样到与第二阶段的特征图相同大小并将它们拼接。为了从不同水平特征图中提取多尺度上下文信息,并行使用 4 个不同扩张率的可分离膨胀卷积(dilated convolution)^[21-22],扩张率分别为 1,2,4,8,可分离卷积用于简化模型参数。最后,将得到的 4 个不同尺度的特征图整合并利用常规 1×1 卷积降维得到最终的特征映射。其过程可以表示为:

$$\mathbf{g} = \text{cat}[\mathbf{F}_2, \text{up}(\mathbf{F}_3), \text{up}(\mathbf{F}_4), \text{up}(\mathbf{F}_5)] \quad (1)$$

$$\mathbf{G} = \text{cat}[\mathbf{g} \otimes \text{dilate}(1), \mathbf{g} \otimes \text{dilate}(2),$$

$$\mathbf{g} \otimes \text{dilate}(4), \mathbf{g} \otimes \text{dilate}(8)] \quad (2)$$

$$\mathbf{Y} = \mathbf{G} \otimes \text{Conv}(1,1) \quad (3)$$

其中, cat 为拼接操作, $\mathbf{F}_2, \mathbf{F}_3, \mathbf{F}_4, \mathbf{F}_5$ 表示二、三、四、五阶段的输出, up 为上采样, \otimes 为卷积操作, $\text{dilate}(k)$ 表示膨胀率为 k , $\text{Conv}(1,1)$ 表示卷积核为 1。在之后的三、四阶段,利用同样的方法将深层的全局语义信息逐步引导到不同的阶段。本文网络中使用了四个全局上下文特征引导模块。

原始卷积如图 4 (a)所示,膨胀卷积如图 4 (b)所示:扩张率为 2 的膨胀卷积会在每个元素间仅跳过一个像素,这使得它能够相对精细地感知局部空间变化,适用于捕捉较

小目标的细节特征。相比常规卷积,它的感受野也有所扩大,但并不像扩张率较大的膨胀卷积那样急剧增大,因此对于小目标的特征提取更加敏感。当扩张率设置为 4 时,卷积核会在每个元素间跳过 3 个像素,这样能够增大每个输出特征点所对应的输入区域,从而更有效地捕获更大范围的空间上下文信息,对于检测较大尺寸的目标效果更佳,因为它可以一次性覆盖更大的区域,增强了模型对较大目标的整体结构特征提取能力。结合不同扩张率的膨胀卷积,网络能够在这一阶段中兼顾不同尺度目标的特征提取。

1.3 注意力模块

注意力机制(attention mechanism)是一种模仿人脑注意力特性的人工智能技术,通常应用于机器学习和深度学习模型中,尤其是自然语言处理、计算机视觉等领域^[23]。其核心思想是让模型能够根据不同的情况动态地分配和调整资源,重点关注输入数据中最相关、最重要的部分,而忽略不太重要的部分。

本文利用基于网格的门控注意力,对局部区域精确关注来提高分割性能。注意门控的示意图如图 5 所示。 x 为编码器传向跳跃连接的特征, g 表示为该层编码器对应的解码器的上一层的输入特征, g 特征图的分辨率是 x 的一半。为了便于操作,对 x 进行下采样处理,使其与 g 拥有相同的分辨率,此时可将 g 视为一种门控信号,滤除 x 中不相关的冗余。接着,分别对 g 和 x 使用 1×1 卷积运算

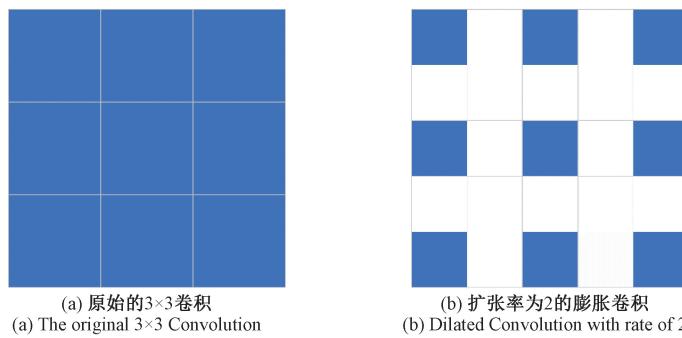


图 4 膨胀卷积

Fig. 4 Dilated Convolution

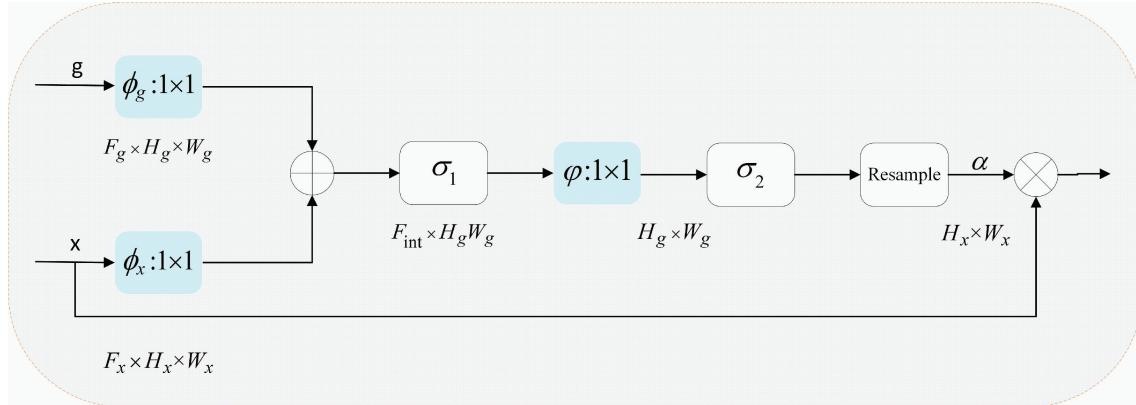


图 5 注意门结构

Fig. 5 Attention gate structure

以调整它们至相同的维度 $\mathbb{R}^{F_{\text{int}}}$, 随后进行逐像素相加的操作。紧随其后, 将叠加后的特征图通过一个 $ReLU$ 激活函数进行非线性变换, 之后再次应用 1×1 卷积进行特征整合, 变换为一张特征图。为了得到归一化的注意力系数, 将整合后的特征图送入 $Sigmoid$ 激活函数进行处理, 并确保其结果值域范围为 $[0, 1]$ 。最后, 对已得到的注意力系数进行双线性插值上采样, 使其分辨率与原始输入特征图 x 匹配, 由此便生成了最终的注意力系数图 $\alpha \in \mathbb{R}^{1 \times H_x \times W_x}$ 。将得到的注意力系数图 α 与输入的 x 相乘来指导特征选择和融合, 强调重要区域, 抑制不相关信息。其过程可以描述为:

$$\alpha = \sigma_2(\varphi(\sigma_1(\phi_x(x)) + \phi_g(g))) \quad (4)$$

$$y = \alpha x \quad (5)$$

其中, ϕ_x 、 ϕ_g 和 φ 为 1×1 卷积, σ_1 和 σ_2 分别为 $ReLU$ 和 $Sigmoid$ 激活函数。

为了充分提升注意门机制对上下文信息的捕获能力, 以及有效地应对在图像上采样处理中可能遇到的细节信息丢失问题, 本文对注意门的输入做了改进: 即将前一阶段 g 网络的输入数据与当前解码器层的输入信息进行深度融合, 如图 6 所示。这样一来, 经过改造后的注意门模块不仅能够充分利用当前层的特征信息, 还能全面获取不

同层次、不同尺度的上下文信息, 从而增强了其对有干扰信息图像的解析能力和区分度。这一方法的引入使得模型在处理图像分割任务时, 能够更好地保持对目标区域边界的刻画及有效辨别出细胞图像中的细胞核相似物和杂质, 提升了模型分割结果的稳定性和分割精度。

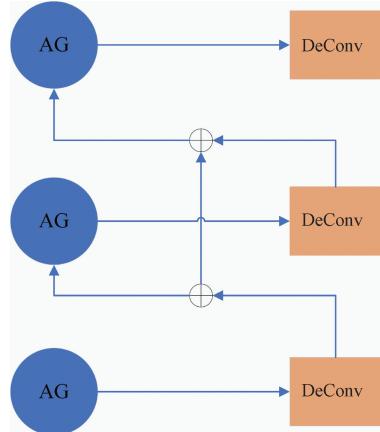


图 6 注意门的输入

Fig. 6 Inputs of AG

1.4 损失函数

图像分割是对图像中所有像素的分类。本文采用图

像分割常用的交叉熵(cross entropy, CE)损失来评估预测值与真实值之间的误差^[24]。交叉熵损失可以表示为:

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N y_i \log p_i + (1-y_i) \log(1-p_i) \quad (6)$$

其中, N 表示像素的数量, p_i 表示预测的结果, y_i 表示标签结果。

2 实验结果与分析

2.1 数据集及预处理

本文使用一个公开的数据集 Herlev^[25] 进行训练验证。数据集由丹麦科技大学(Technical University of Denmark) 和海莱乌大学医院(Herlev University Hospital)联合创建,包含了 917 张高质量的显微镜下拍摄的宫颈细胞涂片图像,这些图像经过病理学家的标注,按照细胞的正常或异常特征进行分类。数据集包含 7 种类别的巴氏涂片,具体分布如表 1 所示。图像的大小各不相同,尺寸从 45×43 到 768×284 ,对于正常的细胞,细胞核往往比较小,细胞质占据细胞图像的绝大部分,而异常细胞细胞核会变大,图像中几乎只剩细胞核。标注的图像中,浅蓝色区域为细胞核,深蓝色区域为细胞质,灰色区域是细胞的兴趣区域(region of interest, ROI),红色区域是整幅图像的背景,如图 7 所示。本文的目的是对细胞核进行分割,因此细胞核为前景,其余部分均设置为背景。

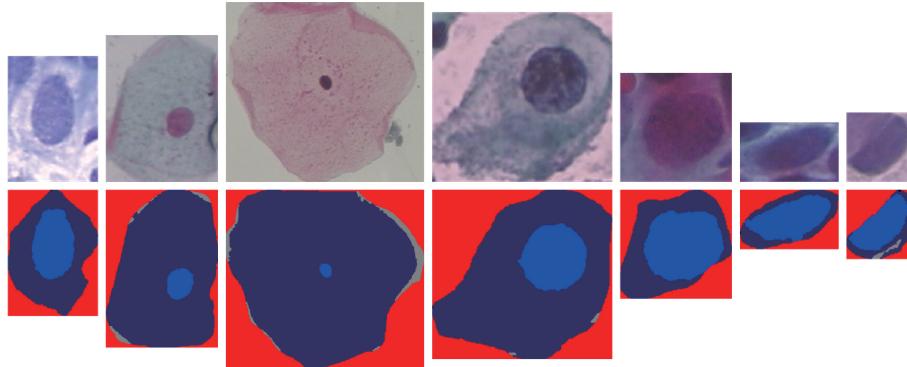


图 7 Herlev7 类示例图像

Fig. 7 Seven classes of Herlev images

表 1 Herlev 数据集

Table 1 Herlev dataset

类别	细胞名称	数量	合计
正常	柱状细胞	98	242
	上皮鳞状细胞	74	
	中层鳞状细胞	70	
异常	轻度非典型增生	182	675
	中度非典型增生	146	
	重度非典型增生	197	
原位癌		150	

由于数据集中的图像大小各异,在输入网络前将图像分辨率统一调整为 256×256 ,大小不足 256×256 的用零填充。此外,Herlev 数据集只有 917 张图像,按照 8 : 1 : 1 的比例分为训练集,验证集和测试集,为了避免网络模型训练出现过拟合,对数据随机水平、垂直翻转扩大量。

2.2 实验细节

本文实验平台为 Windows10 系统,处理器为 Intel i7-10700 CPU,利用 Pytorch 框架和 Python 构建网络结构,并在 NVIDIA GeForce RTX 2070 显卡上进行训练和测试。本文网络采用 Adam 优化器更新网络参数,批处理大

Hospital)联合创建,包含了 917 张高质量的显微镜下拍摄的宫颈细胞涂片图像,这些图像经过病理学家的标注,按照细胞的正常或异常特征进行分类。数据集包含 7 种类别的巴氏涂片,具体分布如表 1 所示。图像的大小各不相同,尺寸从 45×43 到 768×284 ,对于正常的细胞,细胞核往往比较小,细胞质占据细胞图像的绝大部分,而异常细胞细胞核会变大,图像中几乎只剩细胞核。标注的图像中,浅蓝色区域为细胞核,深蓝色区域为细胞质,灰色区域是细胞的兴趣区域(region of interest, ROI),红色区域是整幅图像的背景,如图 7 所示。本文的目的是对细胞核进行分割,因此细胞核为前景,其余部分均设置为背景。

小为 8,训练 epoch 设置为 50,初始学习率设置为 0.001,权重衰减为 0.000 1。

2.3 评价指标

为了更好地评价分割后的结果,本文采用查准率(Precision)、查全率(Recall)和 Dice 系数(Dice coefficient)作为评价指标。其中 Precision 指的是模型识别为细胞核的像素中,真正属于细胞核的实际像素所占的比例,如式(7)所示。查准率越高,意味着模型在标注细胞核区域时出错的概率越低。

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

Recall 指的是模型成功识别出来的细胞核像素数占所有实际存在的细胞核像素数,如式(8)所示。查全率反映了模型识别细胞核的全面性,查全率越高意味着模型的漏检可能性越小。

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

Dice 系数综合考虑了 TP、FP 和 FN,定义式为(9):

$$Dice = \frac{2 | A \cap B |}{| A | + | B |} = \frac{2TP}{2TP + FP + FN} \quad (9)$$

其中,A 为真实标签,B 为预测分割结果,TP 表示真阳性的样本数量,FN 表示假阴性的样本数量,FP 表示假阳性的样本数量。

2.4 实验结果分析

本文选择 U-Net^[9], Attention U-Net^[20], Double U-Net^[16] 和 MCI-Net^[26] 这几个分割网络和本文改进的网络模型分割结果进行比较。

Attention U-Net 作为经典的医学图像分割模型, 通过在 U-Net 跳跃连接部分添加注意门结构, 旨在突出感兴趣区域, 增强对前景像素的敏感性, 同时有效抑制不相关的背景区域。然而该模型在全局上下文信息的融合方面仍有不足。Double U-Net 是两个 U-Net 结构的堆叠, 通过两个编码器和解码器并采用 ASPP(atrous spatial pyramid pooling)模块来增强复杂场景下语义信息的捕获能力, 由于其出色的性能可以作为图像分割的基线, 但面对有干扰信息的细胞图像时, 其对细胞核分割的精度仍需提高。MCI-Net 为图像分割领域较新的改进模型, 针对分割目标大小不一致的问题, 提出了多尺度上下文提取模块提升模型对不同大小目标的适应能力, 并利用边界矫正块提高边界信息矫正能力, 但其对图像中干扰信息的辨别能力存在不足。

本文提出的模型兼顾了 Attention U-Net 中的注意门机制, 着重突出图像中的细胞核部分, 抑制不相关的部分, 提高了模型的抗干扰能力。同时也具备 Double U-Net 和 MCI-Net 中的多尺度上下文信息提取过程, 本文上下文信息的提取更具有全局性, 利用了每个阶段的输出, 获得的细胞核特征更全面。

表 2 展示了不同分割模型的分割结果。可以看出, 本

文的方法在 3 个指标中均有一定的优势,Dice 指数用于评估预测结果和真实结果的重合度,本文方法在 Dice 指数上明显优于其他方法。相较于 U-Net, Dice 系数提高了 4.06%, Recall 提高了 5.36%, 这主要由于本文在跳跃连接阶段不是将特征无差别的堆叠, 而是充分利用各个阶段的特征, 获取全局上下文的特征引导。此外, 注意门模块的使用让网络更加关注细胞核区域, 抑制图像中的无关信息, 有效提升细胞核边界的分割精度。

表 2 各个模型的评价结果

Table 2 Evaluation results of each model

模型	Dice	Recall	Precision
U-Net	0.904 6	0.909 2	0.924 0
Attention U-Net	0.922 5	0.942 3	0.923 9
Double U-Net	0.940 9	0.958 0	0.935 3
MCI-Net	0.916 6	0.957 9	0.895 5
Ours	0.941 3	0.959 1	0.932 0

图 8 展示了不同模型的可视化分割结果。红线区域是标签分割结果, 蓝线区域是模型预测分割结果, 蓝线和红线的距离越近, 分割效果越好。本文的分割方法蓝线和红线几乎重合, 分割效果最好。U-Net 的分割结果对细胞核整体的分割和边界的分割均有所欠缺, 在第 3 行的细胞图像中没有将完整的细胞核分割出来, Attention U-Net 也出现了对细胞核整体分割不全面的问题, 主要是由于缺乏

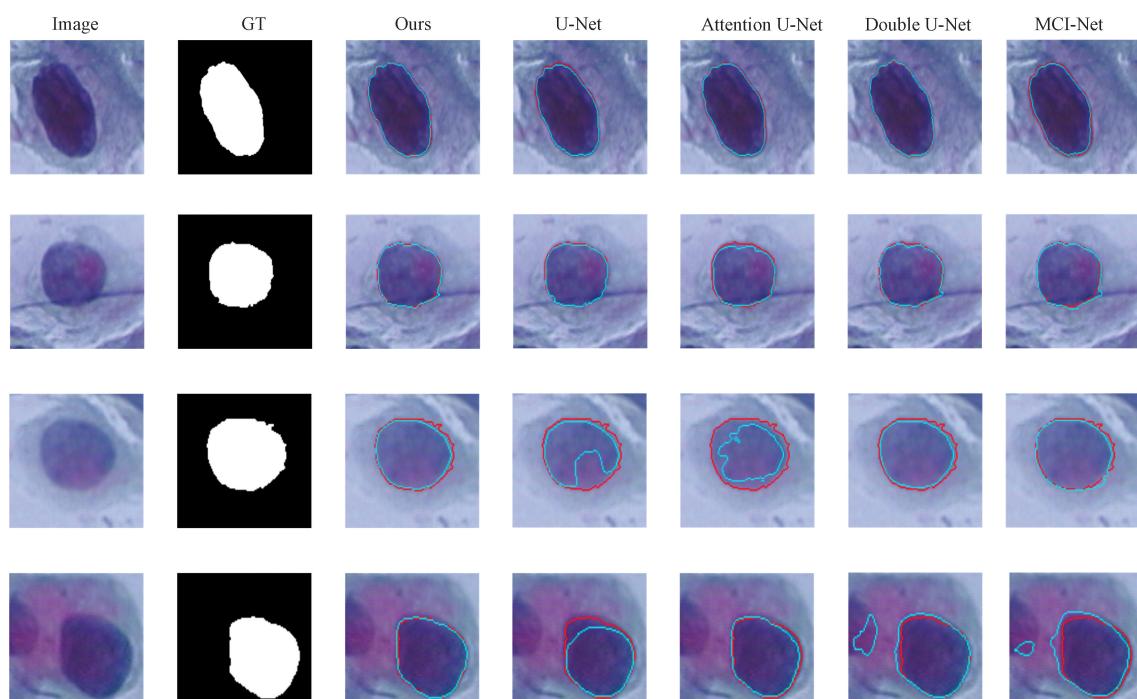


图 8 不同模型分割结果

Fig. 8 Different model segmentation results

全局上下文信息的指导, 特征提取能力较弱, 在第 4 行的细胞图像中存在细胞核相似物的干扰, Double U-Net 和 MCI-Net 这两种方法将图像中的干扰信息也误识别成细胞核, 对细胞核边界分割也存在问题, 而本文的注意门模块着重突出细胞核的信息及边界信息, 对干扰物的影响起到了积极作用, 提高了分割的鲁棒性。

2.5 消融实验

为了进一步验证本文全局特征引导模块(G)和注意门模块(A)的有效性, 设计了表 3 的消融实验, 以 U-Net 作为基本网络(Base)。将全局引导模块添加到网络, 3 个评价指标都有了实质性改善, 其中召回率提高了 1.34%, 由此可得采用不同尺寸感受野的并行结构更有利于全面捕捉和整合全局信息, 提高细胞核的分割精度; 将注意门模块添加到网络, Dice 系数从 0.904 6 提高到 0.922 5, 提高了 1.98%, 精确率提高了 3.64%, 证明了这两个模块对模型的性能提升的有效性。

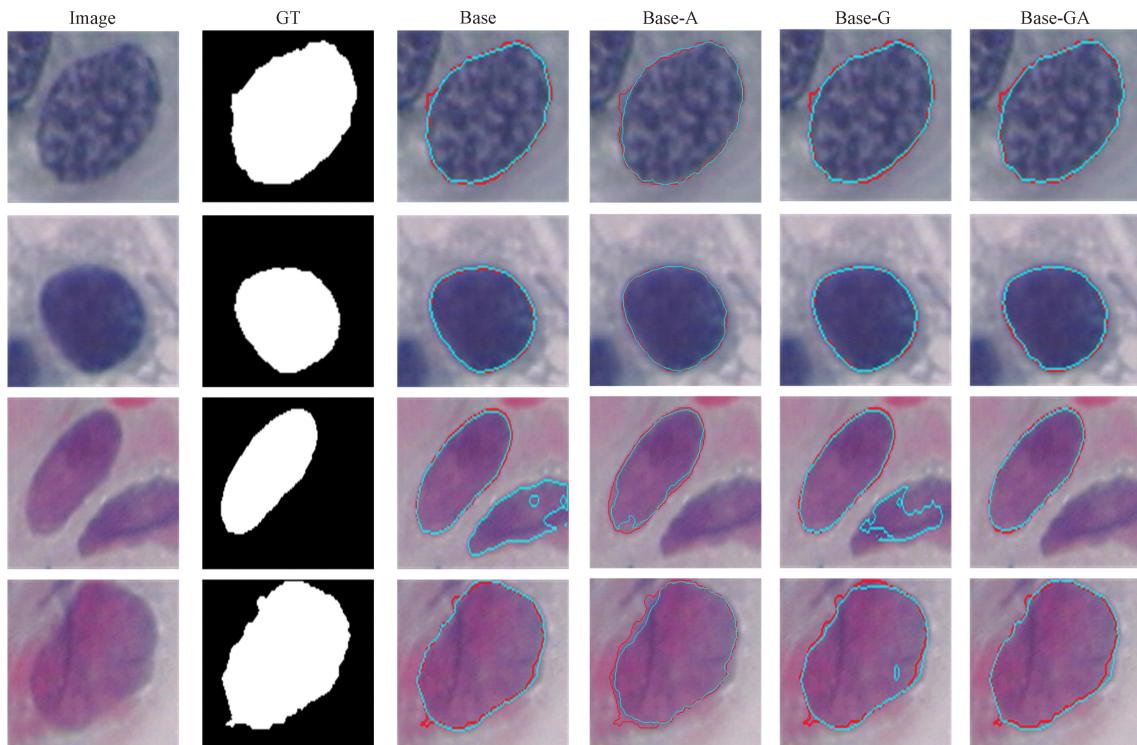


图 9 消融实验可视化结果

Fig. 9 Visualization results of ablation experiment

网络, Base-GA 得到了最佳的分割结果, 从定性角度证明了两个模块的有效性。

3 结 论

本文搭建的宫颈细胞核分割模型以 U 型网络为基础, 针对正常和异常宫颈细胞核表现为大小形态各异影响分割的问题, 引入全局特征引导模块, 在跳跃连接阶段整合各个阶段的特征, 并利用膨胀卷积从不同感受野提取信

表 3 消融实验评价结果

Table 3 Results of ablation experiment evaluation

模型	G	A	Dice	Precision	Recall
Base			0.904 6	0.909 2	0.924 0
Base-G	√		0.911 6	0.915 5	0.936 4
Base-A		√	0.922 5	0.942 3	0.923 9
Base-GA	√	√	0.941 3	0.959 1	0.932 0

图 9 展示了消融实验的可视化结果, Base-G 表示只引入全局特征引导模块的结果, Base-A 表示只引入注意门模块的结果, Base-GA 表示本文方法的结果。观察分割结果可得 Base-A 相比于 Base 可以有效辨别出细胞核和其干扰物, 表明注意门模块对干扰物的影响起到了积极作用, 但对细胞核边缘的分割并不是十分精确。Base-G 相比于 Base 提高了细胞核边缘分割的精度, 但其仍然存在误分割的情况, 抗干扰能力需进一步提高, 将两个模块同时引入

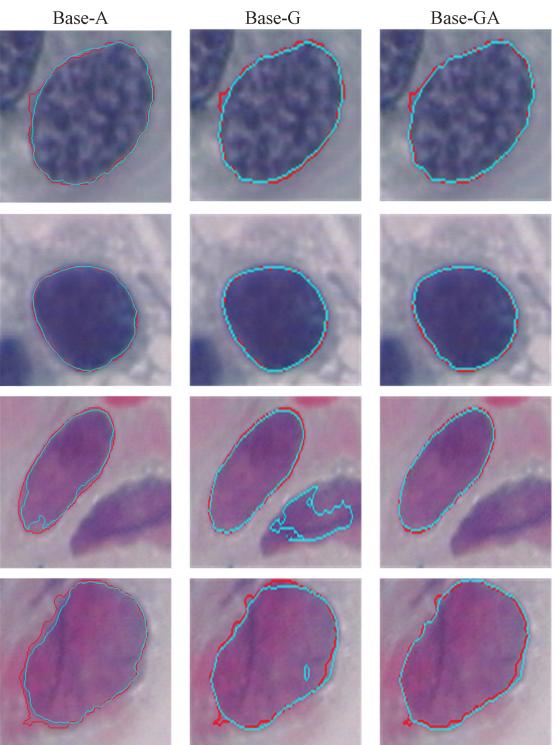


图 9 消融实验可视化结果

Fig. 9 Visualization results of ablation experiment

息, 提高模型对细胞核边界的分割性能; 同时引入注意门模块, 使模型更加注重细胞核区域, 抑制不相关区域, 减少细胞图像中杂质对分割的影响。本文模型在 Herlev 数据集上进行实验, 实验结果表明本文方法相较于其他方法具有一定的优势。

在本研究中, 选用的 Herlev 数据集专注于细胞核分割, 其特性为每幅细胞学图像仅包含单个细胞核, 数据集构造比较简单, 训练出的模型在面对更为复杂的应用

场景时,其适用范围和泛化能力较为有限。因此,为进一步提升模型的通用性和有效性,后续的研究将在包含更多复杂细胞核的数据集上进行。

参考文献

- [1] SACHAN P L, SINGH M, PATEL M L, et al. A study on cervical cancer screening using pap smear test and clinical correlation [J]. Asia-Pacific Journal of Oncology Nursing, 2018, 5(3): 337-341.
- [2] PHOULADY H A, ZHOU M, GOLDGOF D B, et al. Automatic quantification and classification of cervical cancer via adaptive nucleus shape modeling [C]. Proceedings of the 2016 IEEE International Conference on Image Processing. Piscataway: IEEE, 2016: 2658-2662.
- [3] CHALFOUN J, MAJURSKI M, PESKIN A, et al. Empirical gradient threshold technique for automated segmentation across image modalities and cell lines[J]. Journal of Microscopy, 2015, 260(1): 86-99.
- [4] PHOULADY H, GOLDGOF D, HALL L, et al. A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images[J]. Computerized Medical Imaging and Graphics, 2017, 59: 38-49.
- [5] LEE H, KIM J. Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and Cell-Wise contour refinement[J]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2016, 1367-1373.
- [6] 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.
- [7] 刘一鸣, 张鹏程, 刘祎, 等. 基于全卷积网络和条件随机场的宫颈癌细胞学图像的细胞核分割[J]. 计算机应用, 2018, 38(11): 3348-3354.
- LIU Y M, ZHANG P CH, LIU Y, et al. Segmentation of cervical nuclei based on fully convolutional network and conditional random field [J]. Journal of Computer Applications, 2018, 38(11): 3348-3354.
- [8] 崔文成, 杨丹, 邵虹. 基于双路特征的宫颈细胞核分割[J]. 电子测量技术, 2023, 46(6): 129-136.
- CUI W CH, YANG D, SHAO H. Cervical nuclear segmentation based on two-path features [J]. Electronic Measurement Technology, 2023, 46(6): 129-136.
- [9] 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.
- [10] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHSH 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.
- [11] GU Z, CHENG J, FU H, et al. CE-Net: Context encoder network for 2D medical image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 38(10): 2281-2292.
- [12] CHEN J, LU Y, YU Q, et al. TransUNet: Transformers make strong encoders for medical image segmentation [J]. ArXiv preprint arXiv: 2102.04306, 2021.
- [13] LI G, SUN C, XU C, et al. Cervical cell segmentation method based on global dependency and local attention[J]. Applied Sciences, 2022, 12(15): 7742.
- [14] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881-2890.
- [15] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3146-3154.
- [16] 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). Rochester, USA, 2020, 558-564.
- [17] QIN X, ZHANG Z, HUANG C, et al. U2-Net: Going deeper with nested U-structure for salient object detection [J]. Pattern Recognition, 2020, 106: 107404.
- [18] FENG S, ZHAO H, SHI F, et al. CPFNet: Context pyramid fusion network for medical image segmentation [J]. IEEE Transactions on Medical Imaging, 2020, 39(10): 3008-3018.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Washington, DC: IEEE Computer Society, 2016, 770-778.
- [20] OKTAY O, SCHLEMPER J, FOLGOC L, et al. Attention U-Net: Learning where to look for the pancreas [J]. ArXiv preprint arXiv: 1804.03999, 2018.
- [21] YU F, KOLTUN V. Multi-scale context aggregation

- by dilated convolutions [J]. ArXiv preprint arXiv: 1511.07122, 2015.
- [22] CHOLLET F. Xception: Deep learning with depthwise separable convolution [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA, 2017, 1800-1807.
- [23] GUO M, XU T, LIU J, et al. Attention mechanisms in computer vision: A survey [J]. Computational Vision Media, 2022, 8(3): 331-368.
- [24] JADON S. A survey of loss functions for semantic segmentation [C]. Proceedings of the 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). IEEE, 2020: 1-7.
- [25] ZHANG L, LU L, NOGUES I, et al. DeepPap: Deep convolutional networks for cervical cell classification[J]. IEEE Journal of Biomedical and Health Informatics, 2017, 21(6): 1633-1643.
- [26] XIE X, PAN X, SHAO F, et al. MCI-Net: Multi-scale context integrated network for liver CT image segmentation [J]. Computers and Electrical Engineering, 2022, 101: 108085.

作者简介

董展豪,硕士研究生,主要研究方向为深度学习、图像处理。

E-mail:dongangoo@163.com

陈燕,博士,讲师,主要研究方向为医学图像处理。

E-mail:nucchenyan@126.com

侯宏花,博士,副教授,主要研究方向为信号与信息处理、医学图像处理。

E-mail:hhgqb@163.com

张鹏程,博士,副教授,主要研究方向为精准放射治疗剂量计算及方案优化。

E-mail:zhangpc198456@163.com

桂志国(通信作者),博士,教授,主要研究方向为三维CT、医学图像处理与重建、模式识别。

E-mail:gzgtg@163.com