doi:10.3772/j.issn.1002-0470.2024.05.006

基于边界分布注意引导的结直肠腺体分割网络①

凡振邦②* 石淑玲* 马 悦** 李 胜③*

(* 浙江工业大学信息工程学院 杭州 310023)

(** 浙江爱达科技有限公司 杭州 310012)

摘 要 从结直肠组织病理图中自动精确地分割腺体轮廓对结直肠病理诊断有极大帮助,然而,腺体之间间隙狭小且不同等级的腺体具有形态变异性,准确地分割出每个腺体实例具有很大的挑战。为此,本文提出了一种基于注意力的边界引导网络用于腺体分割。具体来说,本方法在边界分支中使用理想的边界图进行监督,引入全局特征整合模块提取腺体边界,并输入后续的解码阶段辅助腺体分割。通过多尺度注意力模块提取多尺度上下文信息,增大模型的感受野。提出边界注意融合模块补充边界细节信息,进一步细化分割结果,得到最终的腺体分割图。所提出模型的有效性在公开的结直肠腺癌数据集 GlaS 上得到了验证,取得了优于其他网络的性能。

关键词 腺体分割; 结直肠癌; 病理图; 深度学习

结直肠癌是最常见的癌症之一[1],对其进行正确分级可以指导医生制定合适的治疗计划。而结直肠癌分级的关键标准就是肠道腺体的形态,包括腺体的形状结构以及排列情况[2]。腺体的分割能为医生提供更明确的感兴趣区域,更好地判断腺体形态[3]。但人工对腺体进行标注费时且易受主观因素影响,利用计算机辅助技术可以大幅提高医生工作的效率,并能得到更加客观真实的结果[4-5]。为了辅助医生诊断,需要有效的图像分割方法从数字化组织切片中准确分割出腺体。

目前,基于深度学习的计算机辅助技术已经可以实现快速的图像分割[6-8],但设计一种高精度的腺体分割方法仍具有挑战性:不同分化程度腺体具有异质性,良性腺体呈较为规则的椭圆形,而恶性腺体形态变异、边界模糊;细胞间相互作用力导致相邻腺体之间具有邻近性,组织边缘间隙狭小。为解决这些难点,Chen 等人[9-10]提出了一种多任务学习的深度轮廓感知网络,采用改进的全卷积网络(fully convolutional network,FCN)[11],在网络后半部分使用一条边界分支用于腺体边界的分割,通过双分支模型结合腺体轮廓和腺体实例的互补信息,集成多层次的上下文特征,取得了较好的分割精度。Xu 等人[12-13]提出了一种深度多通道框架,分别用于像素级的腺体分割、腺体区域建议和边界检测,并在最后进行信息融合得到腺体分割预测图。Graham 等人[14]提出了一种最小信息损失扩张网络,通过一个对象级的不确定性评分舍弃具有高不确定性的预测,能够在特征提取过程中保留最大的信息。Raza 等人[15]提出了多输入多输出网络融合不同通道之间的信息,结合深度监督技术处理大小和形状可变的腺体细胞。王红玉等人[16]设计了一种双路径网络分别用于提取上下文特征和空间特征,获得更大的感受野和空间信息,增强网络对腺体形态学特征的学习能力。Wen 等人[17]提出了一种基于 Gabor 编码器和级联压缩的双注意力机制网络,捕获多尺度通道和空间信息,并设计了一个新的损失函数进

① 浙江省重点研发计划(2020C03074)资助项目。
② 男,1997 年生,硕士生;研究方向:医学图像处理;E-mail: 15839768963@163.com。
③ 通信作者,E-mail: shengli@zjut.edu.cn。
(收稿日期:2022-11-08)

行类平衡。Rastogi 等人[18]针对恶性腺体难以分割的挑战提出了一种编解码结构网络,通过跳跃连接捕获上下文信息,并且利用形态学运算进行后处理,细化原始预测结果。

以往的很多方法通过多分支策略或引入边界损失强调边界的重要性,弥补网络流中丢失的腺体边界细节信息,细化腺体图片中的模糊边界,取得了优异的性能,这证实增强边界特征表达在腺体分割中能有效地提升分割的准确性。现有的边界增强方法通过共享编码器参数隐式地强化模型边界表征能力,或通过在网络浅层引出卷积支路作为边界提取分支显式引入边界信息,类似的方法可以减少参数量,但由于缺乏全局信息在提取边界特征时仍有不足。并且,使用逐元素相加进行两分支融合并不能自适应地调节腺体图片中不同特征的权重,网络很难区分腺体边界的远近像素点,在处理距离很近的腺体边界分割任务时会有较大的局限性。基于以上思想,本文提出了优化的实例-边界双分支网络,改进了网络对边界的提取能力并提出一种新的融合模块自适应地选择边界分支输出的特征。方法在公开数据集上取得了最好的综合性能指标结果,证明了方法的有效性。

1 方 法

1.1 ABG-Net 模型架构

本文提出的双分支网络——基于注意力的边界引导网络(attention and boundary guided network, ABG-Net)结构如图 1 所示,网络主体用于分割腺体实例,采用类似 Unet 的编解码结构。编码器阶段使用预训练的主干网络减少过拟合,为了更好地处理良恶性腺体之间形态和尺度的差异,本文在解码器末端引入多尺度注意力(multi-scale attention, MA)模块整合网络中的深层特征,提升网络的上下文信息提取能力。本文在边界分支中使用理想的边界图通过边界损失进行监督,引入全局特征整合模块(global feature integration, GFI)从网络浅层提取腺体边界,再输入后续的解码器分支中辅助腺体分割,以增强网络对模糊腺体边界的捕获能力。解码阶段本文通过上采样和卷积操作逐步恢复图像的像素尺寸。最后,引入边界注意融合(boundary-attention fusion, BF)模块关注腺体的空间信息并自适应地调整权重,在融合腺体实例特征和边界特征时强调腺体之间狭小的间隙,细化分割结果,经过一系列上采样

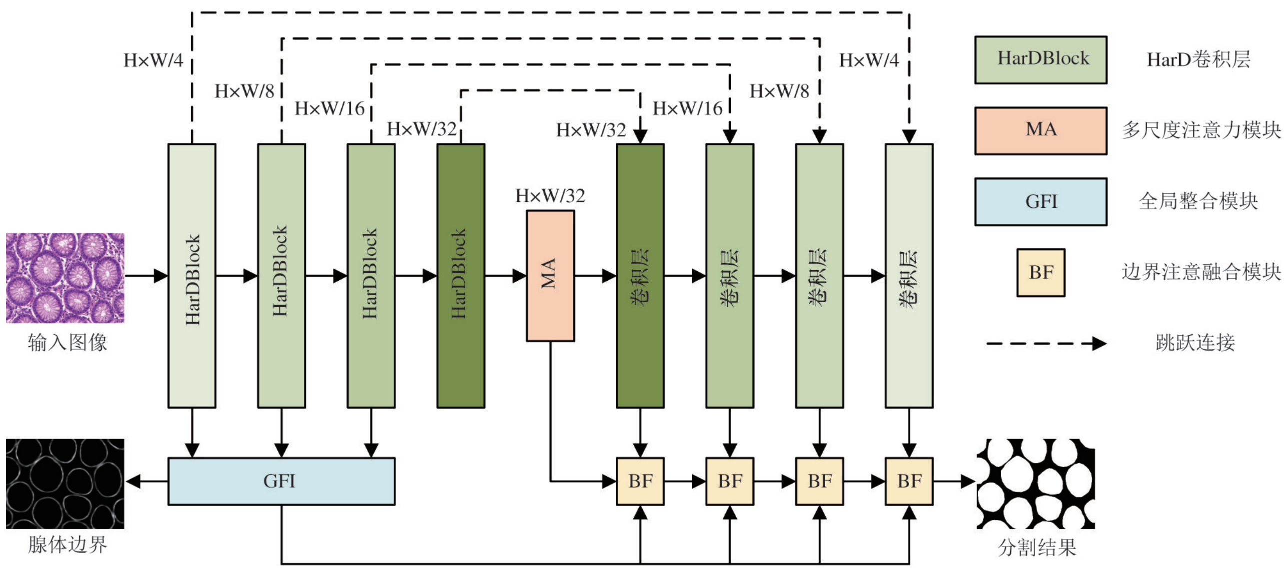


图 1 网络整体框架

操作后最终得到与输入图像一样尺寸的预测图。以下将通过主干部分、GFI 模块、MA 模块以及 BF 模块 4 个部分对所提出方法展开介绍。

1.2 预训练主干网络

为了减少训练时间以及降低过拟合的可能性,本文采用了迁移学习策略,使用预训练的 HarD-

Net68[19]作为主干网络。网络在 DenseNet[20]的基础上作出改进,移除了全局密集连接,采用如图 2 所示的连接方式,推理速度更快且在 ImageNet 上达到了更高的准确率。此外,HarDNet68 在 stride-8 上拥有最多的层来改进局部特征学习,从而更加有利于小规模目标检测,能够更好地检测腺体狭小的边界像素点。

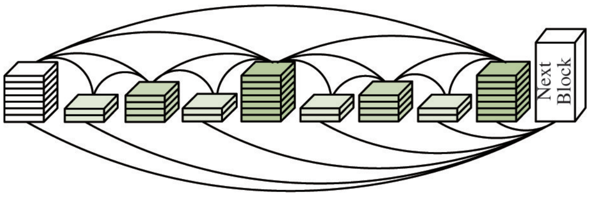


图 2　卷积层之间采用稀疏连接的 HarDBlock

1.3　边界提取分支

在腺体分割任务中,通过边界提取分支强调腺体边界的重要性已证实能有效细化分割结果,并且提供给网络准确的腺体边界对整体网络的性能会有更好的指导作用。由于腺体与周围环境的低对比度,边界预测通常与背景相混淆。此外,腺体之间狭小的间隙也使得边界特征难以被模型捕获,相邻很近的腺体边界可能被误分割在一起。因此,有必要进一步提高腺体边界特征的重要性,这不仅需要边界细节信息,还需要语义信息。为此本研究在网络浅层引入了全局特征整合(GFI)模块,使用真实的边界标签对模块的输出进行监督,利用 3 个不同尺度的特征提取边界,保留重要的细节信息的同时引入语义信息。

GFI 模块如图 3 所示,首先本方法将 3 个不同尺度的输入特征分别通过 RFB 模块[21]。降低通道维度后得到通道数统一的 $f_1(H \times W/4)$、$f_2(H \times W/8)$、$f_3(H \times W/16)$,再将 f_3 通过 3 个不同的上采样和 3×3 卷积层得到 $f_{34}(H \times W/4)$、$f_{32}(H \times W/8)$、$f_{32}'(H \times W/8)$,其中,H 表示特征图的高度,W 表示特征图的宽度。同样,对 f_2 二倍上采样得到 f_{22}、f_{22} 再与同尺寸的 f_{34} 和 f_1 元素级相乘得到 f_1'。将 f_{32} 和 f_2 元素级相乘后得到 f_2',再利用通道级连接整合 f_2'、f_{32}' 以及最大池化后的 f_1',对整合后的特征进行二倍上采样和卷积后得到 f_2''。最后再将 f_1' 和 f_2'' 通道级连接通过一系列卷积后输出最终的边界预测图。计算公式如式(1)、(2)所示。

$$f_2'' = \mathrm{up}_2(\mathrm{concat}(f_2 \odot \mathrm{up}_2(f_3),\ \mathrm{pool}(f_1'),\ \mathrm{up}_2(f_3))) \quad (1)$$

$$f_e = \mathrm{concat}(f_1 \odot \mathrm{up}_2(f_2),\ \mathrm{up}_4(f_3),\ f_2'' \quad (2)$$

其中,⊙表示点积运算,up_2 和 up_4 分别表示 2 倍和 4 倍上采样,concat 表示通道级连接,pool 表示平均池化,公式中上采样和通道级连接之后默认使用卷积核为 3×3 大小尺寸的卷积操作,每个卷积操作后默认进行 BN 操作和 ReLU 操作。

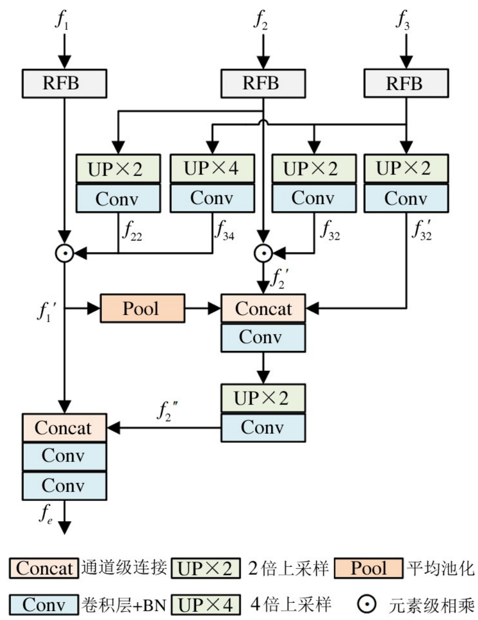


图 3　全局特征整合模块

1.4　多尺度注意力模块

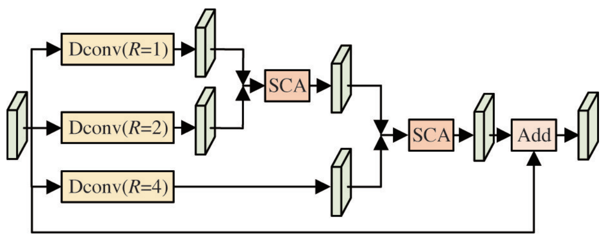
多尺度上下文信息可以提高语义分割任务的性能[21],然而,有效地获取并整合这些信息的方法仍然值得探索。膨胀卷积扩大了卷积层的感受野,实现上下文语义信息联系,在图像分割领域得到广泛应用。受到上下文金字塔融合网络(context pyramid fusion network, CPF-Net)[22]的启发,本文采用具有尺度感知能力的多尺度注意力(MA)模块代替传统编解码网络结构的最后一层操作,MA 结构如图 4(a)所示,通过 3 个具有不同扩张率 $R(1,2,4)$ 的膨胀卷积捕获多尺度上下文信息,其中使用了 2 个尺度感知模块(scale-aware module, SCA)整合不同尺度的特征。SCA 本质上是一种注意力机制,可以动态选择合适的尺度特征并通过自学习将其融合,SCA 模块结构如图 4(b)所示。具体来讲,首先

对不同尺度的输入特征图 F_p 和 F_q 进行通道级拼接以及一系列卷积操作后得到 2 个尺度特征图 P 和 Q。然后 P、Q 分别经过 softmax 分类器得到 2 个像素级注意力映射 p、q, 接着与原先特征图进行点积得到对应感受野信息的特征图,最后相加得到融合后的特征图 F, 计算公式如式(3)、(4)。

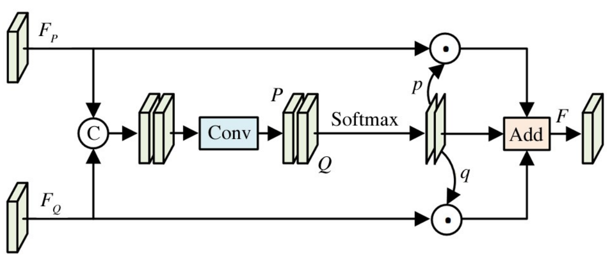
$$p_i = \frac{e^{P_i}}{e^{Q_i} + e^{P_i}},\ q_i = \frac{e^{P_i}}{e^{Q_i} + e^{P_i}},\quad i = [1,2,3,\cdots,H \times W] \tag{3}$$

$$F = p \odot F_p + q \odot F_q \tag{4}$$

其中,$\odot$ 表示点积运算,H 表示特征图的高度,W 表示特征图的宽度。



(a) MA



(b) SCA

图 4　多尺度注意力模块

1.5　边界注意融合模块

传统的实例-边界双分支网络往往在网络的最后一层通过元素级相加或通道级连接实现 2 个分支的信息融合。但这样可能存在不足,仅使用传统方法并不能充分进行 2 个分支之间的交互,也使得网络难以自适应地调整权重。为解决这个问题,本文使用所提出的边界注意融合(BF)模块自底而上的融合来自主干网的多层特征,并逐步将边缘信息嵌入到高级特征图中。BF 模块结构如图 5 所示,通过多尺度注意力模块的提取后,解码器能够提供丰富的细节和其他有用的信息,但这些信息仍然有一定的冗余性。因此,本文引入空间注意力 SA[23] 从空间维度调整特征,重点突出关键的空间位置。具体来说,本方法首先将上一层的特征与对应解码器特征进行通道级连接后作为输入特征,再沿着通道轴将平均池操作和最大池操作应用于输入特征,接着将 2 个的输出结果通道级连接,生成一个有效的特征描述符。最后,利用一个卷积核为 7×7 的卷积层和 sigmoid 算子生成所需的空间注意力图。计算公式如下:

$$A = \sigma(\mathrm{conv}_{7\times7}(\mathrm{concat}(\mathrm{avgpool}(f^i),\ \mathrm{maxpool}(f^i)))) \tag{5}$$

其中,σ 表示 sigmoid 算子,avgpool 和 maxpool 分别表示沿通道轴的平均池化和最大池化。通过将输入特征与空间注意力图元素级相乘突出感兴趣区域,再结合残差连接,将解码器初始特征 f^i 细化为空间增强特征 f_s^i。具体公式如下:

$$f_s^i = A_s^i \odot f^i + f^i \tag{6}$$

其中,$\odot$ 表示元素级相乘。空间增强特征在腺体内部区域响应高,在腺体背景区域响应低,从而实现了突出重要空间位置和抑制干扰的效果。如前所述,腺体模糊的边界易产生误分割,为了在每个解码阶段保持边界的完整度和清晰度,本文引入 GFI 模块提取的边界图 S_b 来强调重要的边界细节。同时,为了在不丢失其他重要信息的情况下突出边界位置,本文仍然使用残差连接进行集成,并在最后通过自乘实现边界注意。具体公式如下:

$$f_b^i = (S_b \odot f_s^i + f_s^i) \odot f_s^i \tag{7}$$

其中,f_b^i 表示边界增强特征,S_b 为在边界提取分支中 GFI 模块生成的边界图。本文通过 4 个 BF 模块

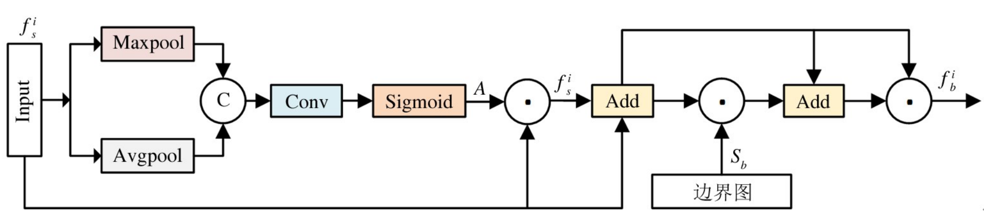


图 5　边界融合模块

自下而上增强预测结果,同时引入空间注意和边界注意,在减小编码阶段信息冗余的同时增强网络对边界细节的捕获能力,最终实现两分支信息的充分有效融合。

2 实验结果及分析

2.1 实验数据集

本文使用公开可用的组织学图像数据集 GlaS[24] 对所提出的 ABG-Net 进行评估。该数据集来自 2015 年的 MICCAI 腺体分割挑战赛,包含 165 张苏木精和伊红(H&E)染色玻片的图像。整个数据集被分为训练集(85 个样本)和测试集(80 个样本),其中测试集分为测试集 A(60 个样本)和测试集 B(20 个样本),每张图像都提供了癌症分级(即良性或恶性)的像素级和图像级注释,表 2 显示了该数据集的具体组成。

表 1 GlaS 数据集

	良性	恶性	总数
训练集	37	48	85
测试集 A	33	27	60
测试集 B	4	16	20

本文对初始的训练集进行了数据增强,首先将输入的图像和相应的真实标注图片统一调整为 512 ×512 像素,再对调整过的图片使用旋转操作(0 °,90 °,180 °,270 °)、镜像反转、亮度增强和随机裁剪进行数据扩充,最终训练集被扩充为 2 040 张,其中随机划分出 320 张图片用作验证集,并通过对训练集中的真实标签做形态学操作得到对应腺体图片的真实边界标签。

2.2 评价指标

为保证定量分析的有效性,本文采用 MICCAI2015 腺体分割挑战中使用的 3 个评估标准来评估所提出方法的性能,各指标定义如下:

(1)F1 分数(F1 score)用于测量单个腺体对象的检测精度,定义为精确度和召回率的调和平均值,F1 分数越高说明预测结果越好。具体公式如下:

$$F1 = 2 \times \frac{pre \times rec}{pre + rec} \tag{8}$$

$$pre = TP/(TP + FP),\ rec = TP/(TP + FN) \tag{9}$$

其中,TP、FP、FN 分别表示真阳性、假阳性、假阴性的数量。各指标以分割结果中腺体对象与真实标签中腺体对象的最大重叠面积作为衡量标准,分割结果中的腺体对象如果至少与它的真实标签相交 50% 以上,则被视为真阳性,否则被视为假阳性;一个真实腺体个体若没有对应的分割对象,或与其相应的分割对象重叠的面积小于 50% 将被视为假阴性。

(2)对象级 Dice 系数(Object Dice)测量 2 组样本之间的相似性,对象级 Dice 系数越高说明预测结果越好。具体公式如下:

$$D = \frac{2 \mid G \cap P \mid}{\mid G \mid + \mid P \mid} \tag{10}$$

$$D_{\text{object}}(G,S) = \frac{1}{2}\Big[\sum_{i=1}^{n_s} \omega_i D(G_i, S_i) + \sum_{j=1}^{n_G} \tilde{\omega}_j D(\tilde{G}_j, \tilde{S}_j)\Big] \tag{11}$$

其中,G 和 P 分别表示真实标签和预测结果。S_i 表示第 i 个分割对象,G_i 表示与 S_i 最大程度重叠的真实标签,$\tilde{G}_j$ 表示第 j 个地面真实样本对象,$\tilde{S}_j$ 表示与 $\tilde{G}_j$ 最大程度重叠的分割对象,其中 n_S 和 n_G 分别是分割对象和真实标签对象的总数,$\omega_i = \mid S_i \mid / \sum_{i=1}^{n_s} \mid S_j \mid$,$\tilde{\omega}_j = \mid \tilde{G}_j \mid / \sum_{j=1}^{n_s} \mid \tilde{G}_j \mid$。

(3)对象级豪斯道夫距离(Object Hausdorff)用于评估预测结果与真实标签的形状相似度,侧重衡量边界的分割精度。与前 2 个指标不同,豪斯道夫距离的数值越小,说明分割的效果越好。具体公式如下:

$$H(G,S) = \max\{\sup_{x \in G} \inf_{y \in S} \| x - y \|, \sup_{y \in S} \inf_{x \in G} \| x - y \|\} \tag{12}$$

$$H_{\text{object}}(G,S) = \frac{1}{2}\Big[\sum_{i=1}^{n_s} \omega_i H(G_i, S_i) + \sum_{i=1}^{n_G} \tilde{\omega}_i H(\tilde{G}_i, \tilde{S}_i)\Big] \tag{13}$$

其中各字母表示意义与式(11)一致。

2.3 损失函数

本文使用深度监督,在解码器多个阶段使用真

实标签进行监督。总的损失函数被定义为

$$L_{total} = L_{BCE} G_e, S_e + \sum_{i=2}^{i=5} \varphi_i L_{hybrid}(G, S_i) \quad (14)$$

其中, S_e 是表示边界分支的输出, S_i 表示腺体分割分支的 4 个输出, G 表示腺体分割的真实标签, G_e 表示真实标签 G 导出的真实边缘映射, φ_i 是每个损失的权重。具体而言,本文使用 2 种不同的损失分别训练边界分支和腺体分割分支。L_{BCE} 是标准的二进制交叉熵(BCE)损失函数,被用于测量相关的边缘映射 S_e 和真实边缘映射 G_e 之间的损失。本文使用混合损失 L_{hybrid} [25] 训练腺体分割, L_{hybrid} 定义如式(15)所示,其中 L_{iou} 表示加权的联合区间(intersection over union, IoU)损失,用于测量 2 个集合的相似性。L_{ssim} 表示结构化损失,它能够捕获图像中的结构信息。L_{BCE}、L_{ssim} 和 L_{iou} 的定义与文献[25]中的定义相同。

$$L_{hybrid} = L_{BCE} + L_{ssim} + L_{iou} \quad (15)$$

2.4 实验细节

本研究的实验环境是 Python3.6 和 PyTorch1.9.1,所有的模型都在 NVIDIA Quadro RTX A6000 GPU 上进行训练。输入的腺体图像大小尺寸为 512 × 512 像素,以端到端的方式训练,并采用多尺度的训练策略(参数为 0.75,1.00,1.25)。本文采用 Adam 作为优化器,默认值分别为 $\beta 1 = 0.9$, $\beta 2 = 0.999$, 批大小设置为 16,最大训练轮次(epoch)为 100,初始学习率为 0.000 1,在每 20 个轮次之后下降学习率。

2.5 实验结果及分析

2.5.1 总体方法实验结果

定量分析:为了验证所提出的 ABG-Net 的有效性,本文使用上述的训练方法与另外 10 种方法在公共数据集(GlaS)上进行了对比实验,在计算每个模型的性能指标之后进行排名并求和,各指标得分排名之和越低,模型的综合性能越好。为保证实验的公平性,实验中对所有的模型均采用和 ABG-Net 相同的训练方式,各模型的定量分割结果如表 2 所示。可以看出,本文提出的方法在 11 种方法中排名的和最小,证明了该模型的有效性。其中 ABG-Net 在测试集 B 上各项指标均高于其他方法,如表 1 所示,相比测试集 A,测试集 B 中结构严重变异的恶性腺体占比要大得多,网络更难提取这些腺体图片的结构特征;另外,前后景像素差异小也导致上采样时容易丢失边界像素点。本文所提出的 ABG-Net 通过提取多尺度上下文信息能更为准确地把握腺体的结构信息,定位腺体的空间位置,并通过边界分支作为辅助,补充恢复图像尺寸时丢失的边缘信息,能够更好地分割恶性腺体。

表 2 不同方法在公共数据集 GlaS 上的比较结果

网络	ObjectF1				Object Dice				Object Hausdorff				合计
	TestA		TestB		TestA		TestB		TestA		TestB		
	得分	排名	得分	排名	得分	排名	得分	排名	得分	排名	得分	排名	
FCN[11]	0.783	11	0.692	11	0.795	11	0.767	11	105.04	11	147.28	10	65
Unet[6]	0.870	8	0.695	10	0.876	7	0.786	9	57.09	8	148.46	11	53
DeepLab[7]	0.862	9	0.764	5	0.859	9	0.804	8	65.72	10	124.97	8	49
Segnet[8]	0.858	10	0.753	7	0.864	8	0.807	7	62.62	9	118.51	7	48
customized Unet[18]	0.872	7	0.752	8	0.832	10	0.817	6	55.51	7	111.17	5	43
MIMONet[15]	0.913	4	0.724	9	0.906	5	0.785	10	49.15	5	133.98	9	42
文献[16]	0.877	6	0.757	6	0.886	6	0.84	1	55.45	6	107.60	4	29
DCAN[10]	0.893	5	0.843	3	0.908	4	0.833	5	44.13	4	116.82	6	27
MildNet[14]	0.914	3	0.844	2	0.913	2	0.836	3	41.54	3	105.89	3	16
GCSBA-Net[17]	0.916	1	0.832	4	0.914	1	0.834	4	41.49	2	102.88	2	14
本文方法	0.916	1	0.847	1	0.909	3	0.840	1	37.65	1	102.24	1	8

定性分析:定性实验中用到的对比网络为 FCN、Segnet、Unet 和 DeepLab,图 6 展示了 ABG-Net 和对比网络模型在公共数据集(GlaS)上的分割结果。本文从数据集中选择了 4 张具有代表性的腺体图片进行展示,包括 2 张间隙狭小的良性腺体图片以及 2 张分化程度较差的恶性腺体图片。定性实验结果表明,其他对比方法在处理密集的良性腺体时,难以分割出每个独立的腺体实例;在分割边界模糊的恶性腺体时容易误分割前景和背景,预测图与标签有很大差别。而本文提出的 ABG-Net 可以有效地解决这 2 个问题,取得了明显更为准确的分割结果。主要原因是相比其他对比网络,ABG-Net 能够通过边界分支捕获更为完整的边界信息,并在解码器的每个阶段补充由于上采样丢失的细节信息,从而细化分割结果。这在良性腺体的分割案例中表现为给予边界更多的关注,原本粘在一起的边界间隙在预测结果中更为清楚;而在恶性腺体的分割案例中表现为能够准确分割出腺体的模糊轮廓。

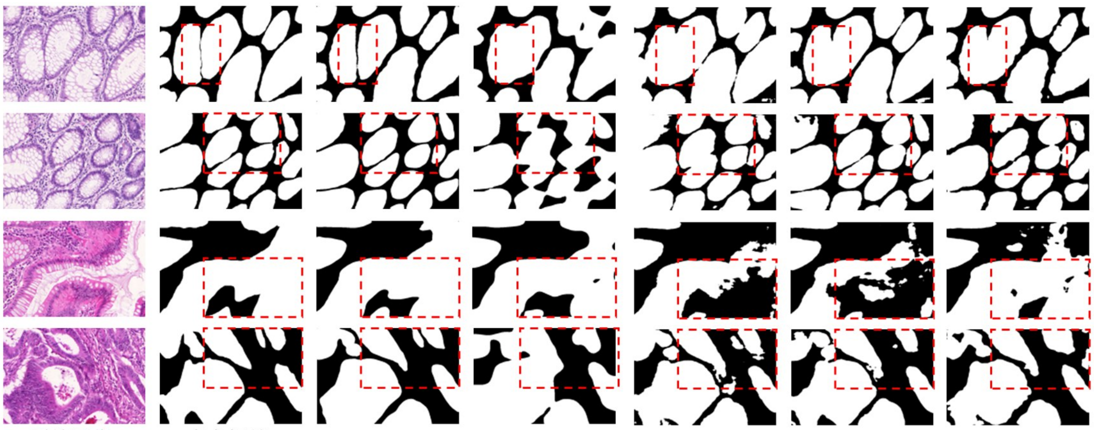


图 6 各方法腺体分割效果对比

2.5.2 消融实验

这个部分将通过消融实验分别验证本文所引入的 GFI 模块、MA 模块以及 BF 模块的有效性。消融实验同样使用上述数据集进行训练和测试,其余实验细节均保持一致。本文分为 2 个阶段进行实验,首先是验证边界提取分支中使用的 GFI 模块的有效性,然后在此基础上测试语义分割分支中使用的 MA 模块以及 BF 模块对网络性能的影响。

边界提取分支消融实验:如前文所述,对于双分支网络来说,为了更好地辅助语义分割分支,边界分支需要能够提供一个更加准确的边界图。于是本文首先验证了所提出 GFI 模块对腺体边界的提取效果。实验采用预训练的 HarDNet 作为主干网络,保留原始的腺体分割分支,并在网络浅层引出边界分支通过真实的边界标签进行监督。本文把文献[26]中所使用的边界分支提取方法作为对照实验,即只在主干网络的第 2 层引出一条卷积支路作为边界提取分支。图 7 展示了 2 种方法对边界模糊的恶性腺体的边界提取效果。可以看出,由于缺乏高级语义信息,在分割前后背景差距小的恶性腺体时,仅通过在网络浅层添加一条卷积分支并不能捕获到完整的腺体边界,并且分割结果中存在过多噪声干扰,这对后续的腺体分割是不利的。结果显示 GFI 模块对腺体边界的提取更为完整,并且没有多余的噪声。

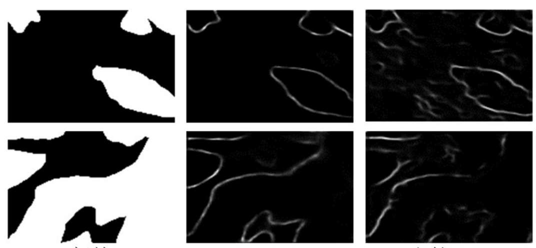


图 7 不同边界分割方法效果对比

语义分割分支消融实验:在这个阶段,本文验证了语义分割分支中所使用的 2 个模块的有效性。实验将预训练的 HarDNet 定义为主干网络,并添加 GFI 模块后作为基准线进行对比实验(此时网络并

没有用到边界分支);在编码器末端,实验测试所引入的 MA 模块对实验的影响,此时网络仅使用通道级连接以及 2 层卷积层进行边界融合,实验结果表示为 GFI + MA;再在解码阶段测试所提出 BF 模块对性能的影响,实验结果表示为 GFI + BF;表 3 中给出了对比实验的定量结果,可以看出,与基准线 HarDNet 相比,引入的模块能显著优化网络性能,这验证了增强边界分支的边界提取能力对腺体语义分割有更好的辅助作用,以及解码器阶段 2 个模块自适应融合边界信息的有效性。最终 ABG-Net 集成了上述模块,得到了最优的性能结果。

表 3 不同模块对性能的影响

方法	Object Hausdorff		Object Dice	
	testA	testB	testA	testB
HarDNet	69.30	124.98	0.858	0.814
GFI + MA	55.19	127.02	0.873	0.825
GFI + BF	42.48	110.94	0.895	0.830
ABG-Net	37.65	102.24	0.909	0.840

3 结 论

本文提出一种新的基于腺体边界的分布注意引导的结直肠腺体分割网络(ABG-Net),旨在解决结直肠腺体分割中存在的腺体结构变异、边界模糊以及腺体间隙狭小等问题。模型引入多尺度注意力模块捕获不同尺度的上下文特征,来弥补网络在上采样中丢失的细节信息;提出全局特征整合模块聚合底层特征中的全局空间信息,捕获完整的腺体边界;提出边界注意融合模块结合空间注意力和边界信息,减少冗余信息的同时强化边界特征,进一步细化腺体分割结果。在公开数据集上进行的实验证明了 ABG-Net 能够准确分割出腺体的模糊边界,并且对比其他方法取得了综合第一的定量指标结果。同时,消融实验验证了所使用模块的有效性。在未来的工作中,希望将本文提出的方法应用于更多的医学图像分割场景。

参考文献

[1] BRAY F, LAVERSANNE M, WEIDERPASS E, et al. The ever-increasing importance of cancer as a leading cause of premature death worldwide[J]. Cancer, 2021, 127(16):3029-3030.

[2] COMPTON C C. Updated protocol for the examination of specimens from patients with carcinomas of the colon and rectum, excluding carcinoid tumors, lymphomas, sarcomas, and tumors of the vermiform appendix: a basis for checklists[J]. Archives of Pathology & Laboratory Medicine, 2000,124(7):1016-1025.

[3] 赵宝奇, 尉飞, 孙军梅, 等. 结合密集连接块和自注意力机制的腺体细胞分割方法[J]. 计算机辅助设计与图形学学报, 2021,33(7):991-999.

[4] RATHORE S, HUSSAIN M, ALI A, et al. A recent survey on colon cancer detection techniques[J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2013,10(3):545-563.

[5] 金旭, 文可, 吕国锋, 等. 深度学习在组织病理学中的应用综述[J]. 中国图象图形学报, 2020,25(10):1982-1993.

[6] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//The 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: IEEE, 2015:234-241.

[7] 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. Munich, Germany: ECCV, 2018:801-818.

[8] 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.

[9] CHEN H, QI X, YU L, et al. DCAN: deep contour-aware networks for accurate gland segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016:2487-2496.

[10] CHEN H, QI X, YU L, et al. DCAN: deep contour-aware networks for object instance segmentation from histology images[J]. Medical Image Analysis, 2017,36:135-146.

[11] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015:3431-3440.

[12] XU Y, LI Y, LIU M, et al. Gland instance segmentation by deep multichannel side supervision[C]// The 19th International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens, Greece: MICCAI, 2016:496-504.

[13] XU Y, LI Y, WANG Y, et al. Gland instance segmenta-

tion using deep multichannel neural networks[J]. IEEE Transactions on Biomedical Engineering, 2017,64(12): 2901-2912.

[14] GRAHAM S, CHEN H, GAMPER J, et al. MILD-Net: minimal information loss dilated network for gland instance segmentation in colon histology images[J]. Medical Image Analysis, 2019,52:199-211.

[15] RAZA S E A, CHEUNG L, EPSTEIN D, et al. Mimo-Net: a multi-input multi-output convolutional neural network for cell segmentation in fluorescence microscopy images[C]//2017 IEEE 14th International Symposium on Biomedical Imaging. Melbourne, Australia: IEEE, 2017:337-340.

[16] 王红玉, 张堧琦, 卜起荣, 等. 基于双路径特征融合的结肠组织病理腺体分割方法[J]. 西北大学学报(自然科学版), 2021,51(4):577-586.

[17] WEN Z, FENG R, LIU J, et al. GCSBA-Net: Gabor-based and cascade squeeze bi-attention network for gland segmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2020,25(4):1185-1196.

[18] RASTOGI P, KHANNA K, SINGH V. Gland segmentation in colorectal cancer histopathological images using U-net inspired convolutional network[J]. Neural Computing and Applications, 2022,34(7):5383-5395.

[19] HUANG C H, WU H Y, LIN Y L. HarDnet-MSEG: a simple encoder-decoder polyp segmentation neural network that achieves over 0.9 mean dice and 86 fps[EB/OL]. (2021-01-20)[2022-11-18]. https://arxiv.org/pdf/2101.07172.pdf.

[20] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017:4700-4708.

[21] LIU S, HUANG D. Receptive field block net for accurate and fast object detection[C]//Proceedings of the European Conference on Computer Vision. Munich, Germany: ECCV, 2018:385-400.

[22] 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.

[23] CHEN L C, YANG Y, WANG J, et al. Attention to scale: scale-aware semantic image segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 3640-3649.

[24] SIRINUKUNWATTANA K, PLUIM J P W, CHEN H, et al. Gland segmentation in colon histology images: the glas challenge contest[J]. Medical Image Analysis, 2017, 35:489-502.

[25] QIN X, ZHANG Z, HUANG C, et al. BASNet: boundary-aware salient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019: 7479-7489.

[26] FAN D P, ZHOU T, JI G P, et al. INF-Net: automatic COVID-19 lung infection segmentation from CT images[J]. IEEE Transactions on Medical Imaging, 2020,39(8):2626-2637.

Attention and boundary guided network for colorectal gland segmentation

FAN Zhengbang*, SHI Shuling*, MA Yue**, LI Sheng*
(*College of Information Engineering, Zhejiang University of Technology, Zhejiang 310023)
(**Zhejiang Aida Technology Co., Ltd., Zhejiang 310012)

Abstract

Automatic and accurate segmentation of gland contour from colorectal histopathology is of great help to colorectal pathological diagnosis. However, due to the narrow gaps between glands and the morphological variability of different grades of glands, accurately segmenting each gland instance is a significant challenge. In this paper, an attention and boundary guided network is proposed for gland segmentation. The boundary branch is supervised by the ideal boundary map, and the global feature integration module is employed to extract the gland boundary, which is then input to the subsequent decoding stage to assist gland segmentation. A multi-scale attention module is introduced to extract multi-scale context information and increase the model's receptive field. Finally, the boundary-attention fusion module is proposed to supplement boundary details, further refine the segmentation results, and obtain the final gland segmentation map. The effectiveness of the proposed model is validated on a publicly available colorectal adenocarcinoma dataset GlaS, achieving superior performance over other networks.

Key words: gland segmentation, colorectal adenocarcinoma, histological image, deep learning