

文章编号 1004-924X(2024)18-2846-15

采用级联策略融合边界特征的多尺度 息肉分割网络

易见兵^{1,2*}, 万建辉^{1,2}, 曹 锋^{1,2}, 李 俊^{1,2}, 陈 鑫^{1,2}

(1. 江西理工大学 信息工程学院, 江西 赣州 341000;
2. 多维智能感知与控制江西省重点实验室, 江西 赣州 341000)

摘要:结直肠息肉分割能有效辅助医生筛查大肠腺瘤,但息肉分割存在噪声较多、边界区分度不够等问题。针对以上问题,本文设计了一种采用级联策略融合边界特征的多尺度息肉分割网络。首先,本文提出了一种改进的通道分组空间增强模块,以增强骨干网络提取的图像特征,从而提高通道和空间位置的相关性。其次,考虑到边界区分度不够,设计了一个级联特征融合网络,以更好地保留边界信息并提高边界区分度,从而提高分割精度。最后,引入了一种双分支混合上采样模块来获取更多的特征细节信息,以实现特征的互补以及捕获更完整有效的特征。在 CVC-ClinicDB 和 Kvasir 数据集上进行测试,本文算法的平均 Dice 系数分别为 0.944, 0.920, 平均交并比分别为 0.900, 0.869;而 M²SNet 算法的平均 Dice 系数分别为 0.922, 0.912, 平均交并比分别为 0.880, 0.861。在 ETIS-LaribPolypDB, CVC-300 和 CVC-ColonDB 数据集上进行测试,本文算法的平均 Dice 系数分别为 0.776, 0.915, 0.782;而 M²SNet 算法的平均 Dice 系数分别为 0.749, 0.903, 0.758。实验结果表明本文算法的分割精度较高,泛化能力较强。

关键词:多尺度息肉分割;通道分组空间增强;边界特征增强;级联特征融合;双分支上采样

中图分类号:TP391 **文献标识码:**A **doi:**10.37188/OPE.20243218.2846

Multi-scale polyp segmentation network employing cascaded strategy to fuse boundary features

YI Jianbing^{1,2*}, WAN Jianhui^{1,2}, CAO Feng^{1,2}, LI Jun^{1,2}, CHEN Xin^{1,2}

(1. College of Information Engineering, Jiangxi University of Science and Technology,
Ganzhou 341000, China;

2. Jiangxi Province Key Laboratory of Multidimensional Intelligent Perception and Control,
Ganzhou 341000, China)

* Corresponding author, E-mail: yijianbing8@jxust.edu.cn

Abstract: Colorectal polyp segmentation can effectively assist doctors in screening for colorectal adenomas, but polyp segmentation has problems such as more noise and insufficient boundary distinguishability. In response to these issues, this paper designed a multi-scale polyp segmentation network that adopts cascaded strategy to fuse boundary features. Firstly, this paper proposed an improved channel grouping spatial enhancement module to enhance the image features extracted by the backbone network, thereby im-

收稿日期:2024-04-21;修订日期:2024-06-25.

基金项目:国家自然科学基金资助项目(No. 62066018, No. 62366017);江西省自然科学基金资助项目(No. 20181BAB202004);江西省研究生创新专项资金资助(No. YC2023-S662)

proving the correlation between channels and spatial positions. Secondly, considering the insufficient boundary distinction, a cascaded feature fusion network was designed to better retain boundary information and improve boundary distinction, so as to improve the segmentation accuracy. Finally, a dual-branch hybrid upsampling module was introduced to obtain more detail feature information, so as to realize the complementarity of features and capture more complete and effective features. Tested on the CVC-ClinicDB and Kvasir datasets, our algorithm achieves mean Dice coefficients of 0.944 and 0.920, and mean Intersection over Union of 0.900 and 0.869 respectively, compared to the M²SNet algorithm with average Dice coefficients of 0.922 and 0.912, and mean IoU of 0.880 and 0.861 respectively. Tested on the ETIS-Larib-PolypDB, CVC-300, and CVC-ColonDB datasets, our algorithm achieves mean Dice coefficients of 0.776, 0.915, and 0.782 respectively, while the M²SNet algorithm achieves mean Dice coefficients of 0.749, 0.903, and 0.758 respectively. Experimental results show that the proposed algorithm has high segmentation accuracy and strong generalization ability.

Key words: multi-scale polyp segmentation; channel group spatial enhancement; boundary feature enhancement; cascade feature fusion; dual-branch upsampling

1 引 言

结直肠癌(Colorectal Cancer, CRC)是全球第三大常见癌症^[1],通过筛查和切除肿瘤前病变(大肠腺瘤)来预防CRC非常关键,定期进行结肠镜检查可以减少30%的结直肠癌发病率^[2]。在结直肠图像中,同一类型的息肉外观(如大小、颜色和质地)往往各不相同^[3];息肉与其周围黏膜之间的边界也比较模糊^[4],缺乏分割方法所需的强烈对比度,容易导致息肉分割不准确和息肉漏检^[5]等问题。因此,研究一种精度较高的息肉分割算法来预防CRC具有重要意义^[6]。

结直肠息肉分割方法^[7]可以分为基于传统的机器学习方法和基于深度学习的方法。传统的基于机器学习的息肉分割方法能够充分地利用提取到的颜色、纹理、形态外观等特征,但需要大量的先验知识提取手工特征,且设计的手工特征只对一些特定的数据集上的分割任务有效,分割性能不稳定^[8],模型的泛化能力较差,实时性和精度都难以满足临床要求。随着深度学习^[9]技术的大力发展,其被广泛应用于医学图像分割等领域。相对于传统的息肉分割算法,深度学习模型在分割精度、泛化能力等方面都具有一定的优势,可以应对不同采样环境下的图像及息肉形状多样性,其鲁棒性较强。CNN网络^[10]在图像级别的分类和回归任务中表现良好,全卷积神经网络(Fully Convolutional Networks, FCN)^[11]继承

CNN的思想,将CNN的全连接层替换成为卷积层,采用反卷积层对最后一个卷积层的特征图进行上采样,在上采样的特征图中逐像素进行分类,达到了在语义级别上分割息肉的效果。2015年U-Net^[12]的提出为医学图像分割带来了革命性的改变,极大地推动了深度学习技术在医学图像分割中的应用。U-Net的编码器和解码器分别用于捕获上下文和恢复细节特征,同时利用跳跃连接增强浅层和深层特征的融合,模型可扩展性强。文献[13]提出改进的DoubleUnet网络,在两个子网络的编解码器中都引入SE注意力模块,使得网络为每个模块赋予不同的权重,从而增强特征学习的表征能力。Zhou等人^[14]提出了一种U-Net++网络,使用嵌套密集跳跃连接,以减少编码器和解码器之间的语义差别,并设计了一个对网络进行剪枝的方案来提高U-Net++的性能。Fan等人^[15]提出了一种平行反向注意力网络PraNet用于息肉的精确分割,但其对小息肉的分割精度仍然偏低。Liu等人^[16]提出了一种多视角方向注意力网络MVOANet,其结合方向和点感知,可有效整合息肉的几何方位信息和整个结直肠区域的点信息。Huang等人^[17]提出了一种MMFIL-Net网络,通过对编码器提取的多级特征图进行同步处理和整合,解决了不同阶段息肉的尺度和尺寸变化问题。Jain等人^[18]提出了一种息肉分割模型CoInNet,其采用了一种新颖的特征提取机制,通过统计特征注意力单元考虑不同

特征图之间的关系,学习如何突出图像中的息肉区域。Sharma 等人^[19]提出了一种基于稳定扩散模型的新型框架 ControlPolypNet 网络,使用新颖的自定义屏蔽输入控制来控制息肉生成过程,从而生成保留重要信息的图像,该图像用于分割任务的训练数据,有助于息肉分割性能的提升。特征金字塔网络 (Feature Pyramid Network, FPN)^[20]是息肉分割常用的一种方法,但其使用极其简单的元素求和方法进行特征融合,导致特征交互效率较低,因此, Su 等人^[21]在 FPN 中引入特征二次融合模块 (Feature Secondary Fusion Module, FSFM) 来提高息肉的分割性能。Zhao 等人^[22]重新设计跳跃连接部分,对跨层级的差异性特征进行聚合,提出了一种具有单尺度减法单元的网络结构 MSNet,其设计的减法单元 (Subtraction Unit, SU) 对相邻层级的特征图进行逐像素减法运算,用于捕获息肉浅层特征和深层特征的互补信息并突出其差异,将多个 SU 金字塔式地连接起来构成多尺度减法模块。多尺度减法模块中使用减法运算降低了输入到解码器的特征在不同层次之间的冗余度,但对特征的重用仍然导致高计算资源需求。在 MSNet 单尺度减法单元结构的基础上, Zhao 等人^[23]提出了一个简单但通用的多尺度减法网络 M²SNet,其设计了一种更强大的层内多尺度减法单元,为解码器提供像素级和结构级的差异信息,然后,金字塔式地为不同层级的多尺度减法单元配备不同的感受野,从而实现层间多尺度特征的聚合并获得更丰富的多尺度差异信息。

现有息肉分割算法已取得较好的分割效果,但对边界与背景之间的区分度不够,为了更好地利用边界特征信息,本文在 M²SNet 算法^[23]的基础上,设计了一种采用级联策略融合边界特征的多尺度息肉分割网络。本文算法主要工作如下:

(1) 针对骨干网络提取的图像特征在空间上容易受到相似形态和噪声背景的干扰,导致算法的错误定位和识别^[24]问题,本文在编码阶段,使用改进的通道分组空间增强模块 (Channel Grouping Spatial Enhancement, CGSE) 对骨干网络提取的图像特征进行增强,以提高通道和空间位置的相关性;

(2) 针对结直肠息肉与背景之间边界区分度

不足问题,本文在多尺度减法模块中设计了一个级联特征融合网络 (Cascade Feature Fusion Network, CFFN),在该网络中首先使用融合模块融合两个低层次的特征,接着通过级联卷积模块处理融合后的特征,以更好地保留边界信息并提高边界区分度,最后聚合处理后的低层特征和增强后的骨干网络特征,以提高网络特征的表达能力,提高分割的准确度;

(3) 针对单一上采样方法容易导致信息丢失与图像模糊化问题,本文在多尺度减法模块中引入双分支混合上采样模块,帮助分割网络有效地处理边界和纹理特征信息,以实现特征的互补和捕获更加完整有效的特征。

2 网络结构

本文提出的算法网络由三个部分组成:编码器、多尺度减法模块、解码器,网络结构如图 1 所示。将经过数据预处理的结直肠息肉图像输入到编码器,其中输入图像的特征维度为 $C_0 \times H \times W$, C_0 表示输入图像的通道数,在本文中其值为 3, H 和 W 分别表示输入图像的高和宽。在编码器中,本文首先采用骨干网络 Res2Net-50^[25]提取五个层次的息肉图像特征 (E^1, E^2, E^3, E^4, E^5),接着采用改进的通道分组空间增强模块来增强特征 (E^2, E^3, E^4, E^5),对增强后的每个层次特征分别采用 3×3 卷积处理,将通道减少到 64 个,这样可以减少后续操作的参数量,最后将处理好的骨干特征输入到多尺度减法模块中。多尺度减法单元^[23]作为多尺度减法模块中的一部分,其利用大小为 $1 \times 1, 3 \times 3, 5 \times 5$ 的固定权重的多尺度卷积滤波器来计算细节和结构差值,从而生成特征。在多尺度减法模块中,本文设计了一种级联特征融合网络 (CFFN),该网络使用特征 (E^1, E^2) 以及经过 3×3 卷积操作增强特征作为输入,以 CFFN₀² 模块所在分支为例,输入特征经过 CFFN₀² 模块处理后,其得到的输出特征将输入到多尺度减法单元中,对于多尺度减法单元,其输入分别为 CFFN₀² 模块输出特征和经过双分支混合上采样模块处理后的 CFFN₀³ 模块输出特征,经过多尺度减法单元输出的特征作为 CFFN₁² 模块的一个输入特征,接着 CFFN₁² 模块的输出特征与经过双分支混合上采样模块处理

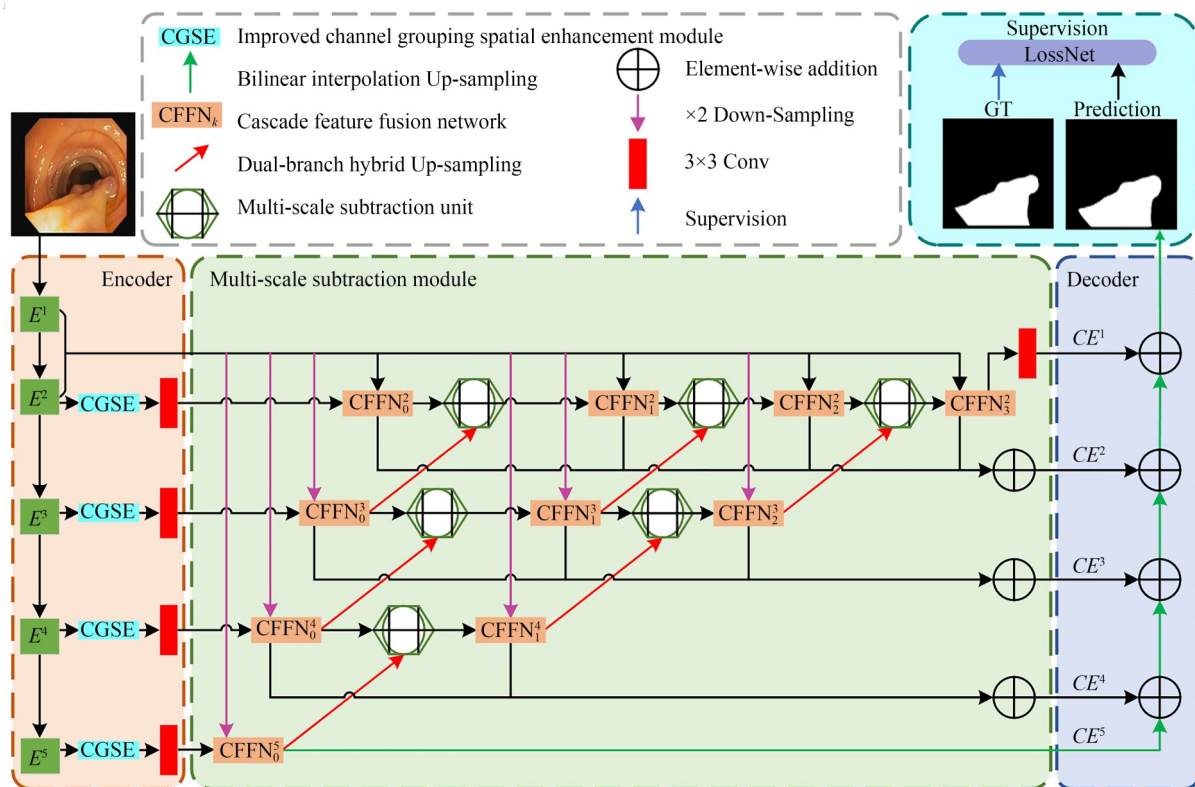


图1 本文算法的网络结构

Fig. 1 Network architecture of our algorithm

后的 $CFFN_1^3$ 模块输出特征共同输入到多尺度减法单元,生成输出特征,以此类推。同一分支的各 CFFN 模块的输出特征汇总,生成三个互补特征 (CE^4 , CE^3 , CE^2)。最后, $CFFN_0^5$ 输出特征 CE^5 、三个互补特征 (CE^4 , CE^3 , CE^2)、经过 $CFFN_3^2$ 和一次 3×3 卷积处理过的特征 E^1 和 E^2 (CE^1) 都会逐步输入到解码器中,其中, CE^5 的特征维度为 $64 \times H/32 \times W/32$, CE^4 的特征维度为 $64 \times H/16 \times W/16$, CE^3 的特征维度为 $64 \times H/8 \times W/8$, CE^2 的特征维度为 $64 \times H/4 \times W/4$, CE^1 的特征维度为 $64 \times H/4 \times W/4$, 得到最后的输出特征并生成最终的预测结果,输出特征的维度为 $1 \times H \times W$ 。在训练阶段,预测结果和 Ground truth 都会输入到 LossNet 中以计算损失。

2.1 通道分组空间增强模块

结直肠息肉图像通常具有复杂的形态、纹理和颜色特征,结直肠息肉图像分割中骨干网络提取的图像特征在空间上容易受到相似形态和噪声背景的影响。通道分组是一种有效提取特征

的方法,通过将通道划分为多个组别,可以使模型更好地理解图像中的不同特征。本文通过在网络结构中引入通道分组,以提高模型对息肉特征的提取能力,从而改善分割的准确性和稳定性。但是全局平均池化操作会导致一部分的细节信息丢失,而在全局平均池化之后再进行点积运算,可能会加重细节信息的丢失甚至增加模型的计算复杂度,因此,本文提出了一种改进的通道分组空间增强模块,在全局平均池化操作之后不再进行点积运算,以保存更多的细节信息,进而增强特征并提高通道及空间位置的相关性,改进后的网络结构如图 2 所示。

本文采用改进的通道分组空间增强模块 (CGSE) 分别对特征 E^2 , E^3 , E^4 , E^5 进行增强,其维度为 $256 \times H/4 \times W/4$, $512 \times H/8 \times W/8$, $1024 \times H/16 \times W/16$, $2048 \times H/32 \times W/32$ 。每个独立的分组空间位置的特征向量用 x_i 表示, $X = \{x_1, \dots, x_m\}$, $x_i \in R^{\frac{C}{G}}$, $m = h \times w$ 。其中, C 表示此模块输入的特征通道数; h 和 w 分别表示输入特征图的高和宽; G 表示分组数。根据实验设置

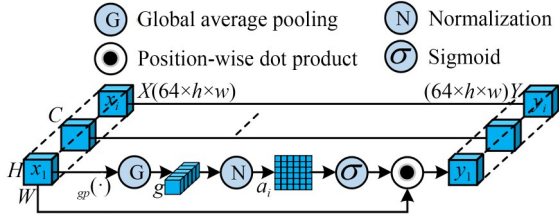


图2 改进的通道分组空间增强模块

Fig. 2 Improved channel grouping spatial enhancement module

的 Batch-size, 结合输入的特征通道数, 并且为防止过小的分组造成特征细节缺失, 因此特征通道数分别设置为 256, 512, 1024, 2 048, 分组数 G 分别设置为 16, 32, 64, 128。首先, 分别对每个分组进行全局平均池化, 以此逼近每个分组学习到的特征语义向量, 并抑制可能存在的噪声, 得到全局特征 g :

$$g = F_{sp}(X) = \frac{1}{m} \sum_{i=1}^m X_i. \quad (1)$$

接下来, 对这一全局特征进行归一化处理, 得到特征 a_i , 以抑制特征之间的偏移, 使得不同通道的特征能够更好地进行信息交互, 增强模型的感知能力。最后, 将 a_i 经过 sigmoid 函数后与局部特征 x_i 进行点积运算, 得到增强的特征分组 y_i , 所有增强后的特征构成特征组 $Y = \{y_1, \dots, y_m\}$, $y_i \in \mathbb{R}^{\frac{C}{G}}$, $m = h \times w$, 增强后的特征维度与输入时的特征维度一致。

2.2 级联特征融合网络

息肉在视觉上是嵌入在背景中的, 且息肉与周围背景之间的边界并不清晰, 有效地聚合边界

特征有助于更好地捕获和整合细节信息, 从而提高息肉分割的准确率。要充分利用边界特征, 需要考虑两个关键问题: 一是要有效地将边界特征与增强后的骨干特征聚合, 二是要减少边界特征中的冗余和噪声。文献[26-27]表明, 只有低层次特征能够保留足够的边界信息, 但是这些边界信息中可能会存在噪声。在分割任务中, 使用门控融合模块[28]能够有效地整合多尺度特征信息, 从而提高分割的准确性, 并且门控融合模块可以帮助网络对特征信息进行选择性的过滤, 从而减少噪声的干扰, 提高模型对真实目标的感知。因此, 本文设计了一种级联特征融合网络(CFFN), 在整合低层次特征的同时有效减少了提取边界信息中的冗余和噪声, 并有效地保留了特征细节信息, 提高了边界区分度, 接着充分地将边界特征和增强后的骨干网络特征聚合[29], 以提高分割性能。如图3所示, 级联特征融合网络由门控融合模块(Gated Fusion Module, GFM)、级联卷积模块(Cascade Conv Module, CCM)和特征聚合模块(Feature Aggregation Module, FAM)组成。

本文首先使用门控融合模块(GFM)融合两个低层次特征 E^1 和 E^2 , 其特征维度分别为 $64 \times H/4 \times W/4$ 和 $256 \times H/4 \times W/4$, 得到融合后的特征 F_0^{edge} , 将其输入到级联卷积模块(CCM)中。在 CCM 中, 依据级联策略思想, 即级联 k 个卷积操作, $k \in (0, 1, 2, 3)$, 当 $k=0$ 时, 表示融合后的特征不进行 3×3 卷积, 直接将 F_0^{edge} 输入到特征聚合模块(FAM)中; 当 $k=1$ 时, 表示融合后的

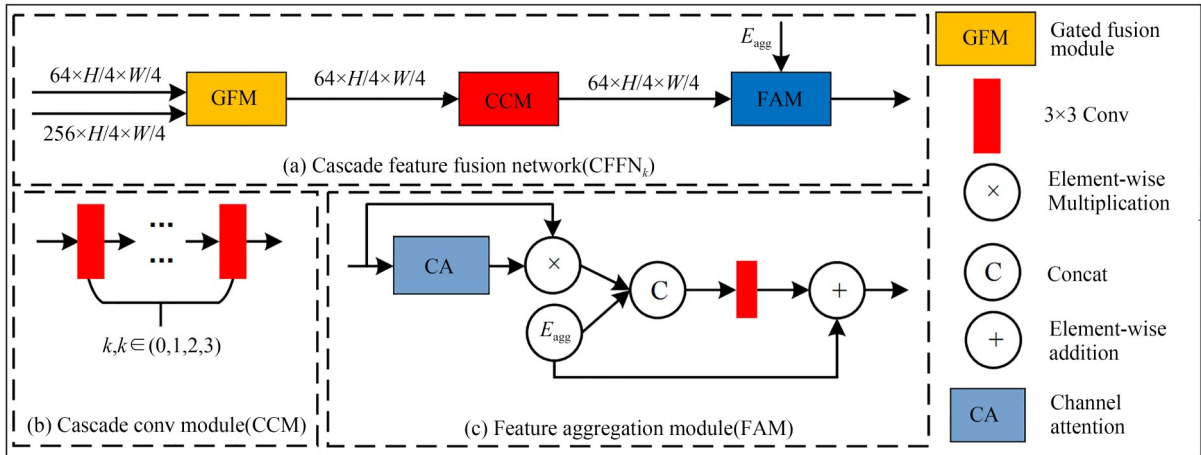


图3 级联特征融合网络(CFFN)

Fig. 3 Cascade feature fusion network(CFFN)

特征进行一次 3×3 卷积,得到输出特征 F_1^{edge} ;当 $k=2$ 时,表示融合后的特征进行两次 3×3 卷积,得到输出特征 F_2^{edge} ;当 $k=3$ 时,表示融合后的特征进行三次 3×3 卷积,得到输出特征 F_3^{edge} 。输出特征 F_3^{edge} 会经过两步操作,首先特征 F_3^{edge} 会再进行一次 3×3 卷积操作,生成边界预测特征 F_4^{edge} ,其特征维度为 $64 \times H/4 \times W/4$,用于解码操作。其次,特征 F_3^{edge} 会与特征 $F_0^{\text{edge}}, F_1^{\text{edge}}, F_2^{\text{edge}}$ 分别作为输入特征输入到 FAM 中,其特征维度均为 $64 \times H/4 \times W/4$ 。在 FAM 中,首先将输入特征输入到通道注意力(Channel Attention, CA)中,以帮助模型学习不同通道之间的相关性和重要性,提高模型的分割准确性和泛化能力。然后,将经过 CA 处理的输入特征与未经过 CA 处理的输入特征进行元素乘法运算,得到特征 F_{att} ,再将特征 F_{att} 和输入特征 E_{agg} 使用 Concat 操作进行合并。合并后的特征(即 $F_{\text{cat}} = \text{CAT}(F_{\text{att}}, E_{\text{agg}})$)会再进行一次 3×3 卷积操作。为了提高网络效率并保留原始信息,使用残差连接来处理合并后的特征和来自编码器中增强后的骨干特征。该过程可以表示为:

$$F_{\text{agg}} = F_{\text{cat}} \oplus E_{\text{agg}}, \quad (2)$$

其中:在 CFFN_k 中,当 $k=0$ 时, E_{agg} 表示编码器中增强后的骨干特征,其特征维度在不同层级的 CFFN_0 中依次可以表示为 $256 \times H/4 \times W/4$, $512 \times H/8 \times W/8$, $1024 \times H/16 \times W/16$, $2048 \times H/32 \times W/32$;当 $k=1, 2, 3$ 时, E_{agg} 表示多尺度减法单元输出的特征,其特征维度与当 $k=0$ 时的输入特征的特征维度一致。 F_{agg} 表示聚合后的特征,特征维度与输入特征 E_{agg} 的特征维度一致, $\oplus$ 表示元素加法操作。聚合后的特征将作为多尺度减法单元的输入。值得注意的是,聚合后的边界特征可以很好地被纳入分割网络中,因此有效地提高了算法的分割性能。

2.3 双分支混合上采样模块

特征聚合模块输出的特征通常包含较多边界信息和纹理信息,而结直肠息肉图像的分割需要同时处理边界信息和纹理信息,本文通过引入双分支混合上采样模块^[30],可以分别处理边界特征和纹理特征,并在上采样过程中进行融合,从而有效地提高特征图的分辨率,有助于保留息肉图像中的细节和提高模型对结直肠息肉边界和

内部结构的识别能力,使得分割结果更加精确。

双分支混合上采样模块的结构如图 4 所示。输入特征 $X_u \in \mathbb{R}^{64 \times h \times w}$ 被分为上分支和下分支,其中 h 和 w 分别表示输入特征的高和宽,上分支通过步长为 2 的反卷积运算进行重构,得到特征 $X_u^1 \in \mathbb{R}^{64 \times 2h \times 2w}$ 。下分支通过双线性插值对图像和纹理进行重采样,得到特征 $X_u^2 \in \mathbb{R}^{64 \times 2h \times 2w}$ 。将两种不同的上采样方法得到的特征图串联起来,得到一个融合特征 X_u^f ,实现两种特征的互补,从而捕捉到更加完整有效的特征。最后,经过一个 1×1 卷积变化进行通道调整,得到最终的特征 $X_{\text{out}} \in \mathbb{R}^{64 \times 2h \times 2w}$ 。

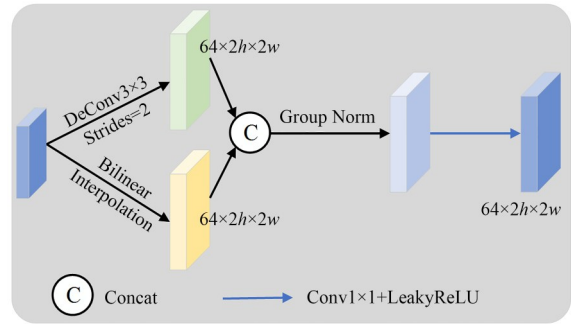


图 4 双分支混合上采样模块

Fig. 4 Dual-branch hybrid upsampling module

2.4 损失函数

在本文中,损失函数如式(3)所示。

$$L_{\text{total}} = L_{\text{IoU}}^w + L_{\text{BCE}}^w + L_f, \quad (3)$$

其中: L_{IoU}^w 表示基于全局约束的加权 IoU 损失,与标准 IoU 损失不同的是,加权 IoU 损失通过增加样本像素的权重来突出其重要性; L_{BCE}^w 表示基于局部(像素级)约束的加权二元交叉熵(Binary Cross Entropy, BCE)损失,与标准的 BCE 损失相比,加权二元交叉熵损失更注重样本像素,而不是对所有像素进行等权重分配。使用的定义与文献[15]和文献[31]中的定义相同,其有效性也在该文献中得到验证。

L_f 表示 LossNet 损失^[23],其被设计为一个感知损失网络,利用 ImageNet 预训练的分类网络 VGG-16 来分别提取预测结果和 Ground truth 的多尺度特征,该多尺度特征用来计算预测结果和 Ground truth 之间的感知损失。这种感知损失被用于训练网络模型,通过 SGD 优化器来更新网络模型的参数,使网络模型能够生成更高质量的

图像。LossNet 网络本身的参数是被冻结的,其通过计算感知损失来间接地指导本文分割网络参数的更新。

损失 L_f 的计算如下:

$$L_f = l_f^1 + l_f^2 + l_f^3 + l_f^4, \quad (4)$$

$$l_f^i = \|F_P^i - F_G^i\|_2, \quad (5)$$

其中: F_P^i 和 F_G^i 表示分别从预测结果和 Ground Truth 中提取的第 i 层特征映射, l_f^i 表示 L2 损失。

L_{IoU}^w 和 L_{BCE}^w 的定义如下:

$$L_{IoU}^w = 1 - \frac{\sum_{i=1}^W \sum_{j=1}^H (g_{i,j} * p_{i,j}) * (1 + 5\mu_{i,j})}{\sum_{i=1}^W \sum_{j=1}^H (g_{i,j} + p_{i,j} - g_{i,j} * p_{i,j}) * (1 + 5\mu_{i,j})}, \quad (6)$$

$$L_{BCE}^w = \frac{\sum_{i=1}^W \sum_{j=1}^H (1 + 5\mu_{i,j}) \sum_{l=0}^l I(g_{i,j} = l) \log \Pr(p_{i,j} = l | \Phi)}{\sum_{i=1}^W \sum_{j=1}^H 5\mu_{i,j}}, \quad (7)$$

其中: $g_{i,j}$ 和 $p_{i,j}$ 分别表示 Ground truth 和预测结果中像素 (i, j) 的值, $I(\cdot)$ 表示指示函数, 符号 $l \in \{0, 1\}$ 表示两种标签, Φ 表示模型的所有参数, $\Pr(p_{i,j} = l | \Phi)$ 表示预测概率, $\mu_{i,j}$ 表示像素重要性的指标, 根据中心像素与其周围的插值计算, 对于所有像素, $\mu_{i,j} \in (0, 1)$, 如果 $\mu_{i,j}$ 很大, 则认为它是一个重要的像素 (例如边缘或孔), 值得更多关注, 相反, 如果 $\mu_{i,j}$ 很小, 则认为它是一个普通像素, 不值得关注。

3 实验及结果分析

3.1 实验数据及预处理

本文采用息肉分割数据集 CVC-ClinicDB^[32], Kvasir^[33], CVC-300^[34], CVC-ColonDB^[35], ETIS-LaribPolypDB^[36] 验证算法的有效性、泛化性和普适性。数据集 CVC-ClinicDB 由医学图像计算与计算机辅助干预国际会议 (Medical Image Computing and Computer-Assisted Intervention, MICCAI) 于 2015 年发布, 该数据集包括 612 张息肉图片, 图片分辨率为 384×288 pixels; 数据集 Kvasir 由挪威奥斯陆大

学医院的内窥镜专家采集并标注, 该数据集包含 1 000 张内窥镜息肉图片, 数据集中图片分辨率从 332×487 到 1920×1070 不等; 数据集 CVC-300 包含 60 张息肉图片, 来自 EndoScene 的测试集, EndoScene 包含 36 个患者的 912 张内窥镜下的息肉图片; 数据集 CVC-ColonDB 包含 380 张图片; 数据集 ETIS-LaribPolypDB 包含 196 张分辨率为 1255×996 pixels 的息肉图片, 图片中的息肉均较小, 且与周围组织高度相似, 具有较高的挑战性。本文采用数据扩增的方法来增加训练集样本数量, 包括中心裁剪、随机旋转、高斯模糊、弹性变换、RGB 平移等。

3.2 评价指标

为了对实验结果进行定量分析, 本文采用了六种评价指标对息肉分割的性能进行评估, 分别是平均 Dice 系数 (mDice)、平均交并比 (mean Intersection over Union, mIoU)、平均绝对误差 (Mean Absolute Error, MAE)、加权 F-measure (F_β^w)、S-measure (S_a) 和 E-measure ($E_\phi^{\max}$)。mDice 是一种度量相似度的标准, 为每类置信度 (Dice) 累加平均。mIoU 是每一类 IoU 结果的累加平均。平均绝对误差 (MAE) 是指模型预测值 X 与样本真实值 Y 之间距离的平均值, 用于评估像素级别的精度。F-measure (F_β^w) 是精度和召回率的调和平均。在本文的实验中, 使用了 F_β^w 的改进版本, 即加权 F-measure (F_β^w)。其计算公式为:

$$F_\beta^w = (1 + \beta^2) \frac{P^w \times R^w}{\beta^2 P^w + R^w}, \quad (8)$$

其中, β 是权衡精度和召回率的参数。

S-measure (S_a) 是结合图像结构相似性度量 (S_0) 与区域结构相似性度量 (S_r) 的分割评价指标, 其计算公式为:

$$S_a = \alpha \times S_0 + (1 - \alpha) \times S_r, \quad (9)$$

其中, α 是一个权衡参数, 默认设置为 0.5。

E-measure ($E_\phi^{\max}$) 是一种增强对齐度量, 用于评估像素级和全局级的相似性, 其计算公式为:

$$E_\phi^{\max} = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H \phi_{FM}(i, j), \quad (10)$$

其中: ϕ_{FM} 表示增强对齐矩阵。在上述评价指标中, MAE 值越小表示算法性能越好, 其他评价指标值越大表示算法性能越好。

3.3 参数设置

本文的硬件实验环境:GPU 为 NVIDIA Ge-Force RTX 2080Ti 12 GB 显卡;软件实验环境为:Ubuntu 操作系统,Pytorch 深度学习框架。本文的输入图片大小为 352×352 ,并采用多尺度训练策略^{[15][22][37]},即采用随机水平翻转和旋转的数据增强方式避免算法训练过拟合。本文采用随机梯度下降法(SGD),动量和权重衰减分别设置为 0.9 和 0.000 5,初始学习率设为 0.005,使用 Warm-up 和线性衰减策略来调整学习率,Batch-size 大小设置为 16,Epoch 为 50。

3.4 实验结果分析

3.4.1 不同息肉分割算法性能对比

为了验证所提出的息肉分割算法的性能,本文在 CVC-ClinicDB 和 Kvasir 数据集上进行了相关实验。在实验中,本文随机从 CVC-ClinicDB 和 Kvasir 数据集中分别提取 550 和 900 张图片,并对提取的共 1 450 张图片进行数据扩增,将其作为训练集,数据集中未提取的剩余图片作为测试集,对网络进行训练得到相应模型。将本文算法与其他经典算法进行对比,包括 U-Net^[12],U-Net++^[14],SFA^[38],PraNet^[15],CaraNet^[39],FS-FM^[21],BLE-Net^[40],SAM-CLNet^[41],MSNet^[22],M²SNet^[23]等算法。为保证实验的公平性,本文算法采用的训练集和测试集与其他算法相同,表 1 和表 2 中其他算法的分割结果均引用自原论

文。表 1 和表 2 分别表示在 CVC-ClinicDB 和 Kvasir 数据集上各算法性能对比的结果。从表 1 的实验结果中可以发现,本文算法相较于基准算法 M²SNet,mDice 系数、mIoU, F_{β}^w , S_a , $E_{\phi}^{\max}$ 分别提高了 0.022,0.020,0.021,0.006,0.017,MAE 减少了 0.003。从表 2 的实验结果中可以发现,本文算法相较于基准算法 M²SNet,mDice 系数、mIoU, F_{β}^w , S_a , $E_{\phi}^{\max}$ 分别提高了 0.008,0.008,0.009,0.002,0.008,MAE 减少了 0.003。从上面的分析中可以得到,相较于基准算法 M²SNet,本文算法在六个指标上的性能都有一定程度的提升。此外,相较于 CaraNet 算法,本文算法在多个指标上是表现较好的,表 1 数据显示,对于 mDice, mIoU 和 F_{β}^w 指标,本文算法分别比 CaraNet 算法高出 0.008,0.013 和 0.007,对于 MAE 指标,本文算法比 CaraNet 算法减少了 0.001;表 2 数据显示,对于 mDice,mIoU 和 F_{β}^w 指标,本文算法分别比 CaraNet 算法高出 0.002,0.004 和 0.001,对于 MAE 指标,本文算法比 CaraNet 算法减少了 0.001。但对于 S_a 和 $E_{\phi}^{\max}$ 指标,本文算法略逊于 CaraNet 算法,这可能是由于本文算法在结构信息提取方面相对于 CaraNet 算法有所不足。此外,本文还将提出的算法与基于大模型的分割算法 SAM-CLNet 进行了对比,从表 1 和 2 的结果可以看出,本文算法在各项指标上的性能均优于 SAM-CLNet 算法。

表 1 各模型在 CVC-ClinicDB 数据集的性能比较
Tab. 1 Performance comparison of each model on the CVC-ClinicDB dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
U-Net	0.824	0.767	0.811	0.889	0.917	0.019
U-Net++	0.797	0.741	0.785	0.872	0.898	0.022
SFA	0.698	0.615	0.647	0.793	0.816	0.042
PraNet	0.902	0.858	0.896	0.935	0.958	0.009
CaraNet	0.936	0.887	0.931	0.954	0.991	0.007
MSNet	0.921	0.879	0.914	0.941	0.972	0.008
FSFM	0.934	0.884	0.931	0.953	0.982	0.010
BLE-Net	0.905	0.878	0.922	0.950	0.981	0.008
SAM-CLNet	0.853	0.783	0.845	0.897	0.953	0.014
M ² SNet	0.922	0.880	0.917	0.942	0.970	0.009
Ours	0.944	0.900	0.938	0.948	0.987	0.006

表 2 各模型在 Kvasir 数据集的性能比较

Tab. 2 Performance comparison of each model on the Kvasir dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
U-Net	0.821	0.756	0.794	0.858	0.901	0.055
U-Net++	0.824	0.753	0.808	0.862	0.907	0.048
SFA	0.725	0.619	0.670	0.782	0.828	0.075
PraNet	0.901	0.848	0.885	0.915	0.943	0.030
CaraNet	0.918	0.865	0.909	0.929	0.968	0.023
FSFM	0.913	0.861	0.906	0.922	0.952	0.027
BLE-Net	0.905	0.854	0.897	0.919	0.955	0.025
SAM-CLNet	0.857	0.788	0.854	0.882	0.929	0.042
MSNet	0.907	0.862	0.893	0.922	0.944	0.028
M ² SNet	0.912	0.861	0.901	0.922	0.953	0.025
Ours	0.920	0.869	0.910	0.924	0.961	0.022

Note: ↑ indicates that the higher the metric, the better the model performance, and ↓ indicates that the lower the indicator, the better the model performance. Bold data indicates the best performance.

3.4.2 不同息肉分割算法的泛化性能对比

为了验证本文算法模型的泛化性能,采用论文 3.4.1 部分的训练集作为模型的训练集,并以 CVC-ColonDB、CVC-300 和 ETIS-LaribPolypDB 数据集作为模型的测试集。将本文算法与其他经典算法进行对比,包括 U-Net^[12], U-Net++^[14], SFA^[38], PraNet^[15], CaraNet^[39], FSFM^[21], BLE-Net^[40], SAM-CLNet^[41], MSNet^[22], M²SNet^[23]等算法。为保证实验的公平性,本文算法采用的训练集和测试集与其他算法相同,表 3~5 中其他算法的分割结果均引用自原文。表 3~表 5 分别展示了在 CVC-ColonDB、CVC-

300 和 ETIS-LaribPolypDB 三个数据集上各算法的泛化性能对比结果。从表 3~表 5 结果可以得知,本文算法在三个数据集上的 mDice 系数相较于基准算法 M²SNet 分别实现了 0.024、0.012 和 0.027 的提升;同时,在 mIoU 指标上,本文算法也较 M²SNet 算法分别提高 0.029、0.010、0.016。以上述结果表明,本文算法在分割精度上较基准算法 M²SNet 有一定的提升。进一步地与其他经典算法相比,本文算法在保持较高精度的同时,还展现出了较强的泛化能力。尤为值得注意的是,与基于大模型的分割算法 SAM-CLNet 相比,本文算法在多个指标上均展现出较大优势,对于

表 3 CVC-ColonDB 数据集上不同算法分割结果对比

Tab. 3 Comparison of segmentation results of different algorithms on the CVC-ColonDB dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
U-Net	0.519	0.449	0.498	0.711	0.763	0.061
U-Net++	0.490	0.413	0.467	0.691	0.762	0.064
SFA	0.467	0.351	0.379	0.634	0.648	0.094
PraNet	0.716	0.645	0.699	0.820	0.847	0.043
CaraNet	0.773	0.689	0.729	0.853	0.902	0.042
FSFM	0.781	0.709	0.775	0.852	0.898	0.032
BLE-Net	0.731	0.658	0.710	0.832	0.859	0.044
SAM-CLNet	0.711	0.627	0.694	0.815	0.862	0.040
MSNet	0.755	0.678	0.737	0.836	0.883	0.041
M ² SNet	0.758	0.685	0.737	0.842	0.869	0.038
Ours	0.782	0.714	0.765	0.855	0.893	0.034

表 4 CVC-300数据集上不同算法分割结果对比

Tab. 4 Comparison of segmentation results of different algorithms on the CVC-300 dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
U-Net	0.717	0.639	0.684	0.842	0.867	0.022
U-Net++	0.714	0.636	0.687	0.838	0.884	0.018
SFA	0.465	0.332	0.341	0.640	0.604	0.065
PraNet	0.873	0.804	0.843	0.924	0.938	0.010
CaraNet	0.903	0.838	0.887	0.940	0.989	0.007
FSFM	0.910	0.846	0.893	0.942	0.978	0.006
BLE-Net	0.879	0.805	0.850	0.928	0.973	0.009
SAM-CLNet	0.876	0.800	0.851	0.922	0.975	0.008
MSNet	0.869	0.807	0.849	0.925	0.943	0.010
M ² SNet	0.903	0.842	0.881	0.939	0.965	0.009
Ours	0.915	0.852	0.896	0.938	0.979	0.006

表 5 ETIS-LaribPolypDB数据集上不同算法分割结果对比

Tab. 5 Comparison of segmentation results of different algorithms on the ETIS-LaribPolypDB dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
U-Net	0.406	0.343	0.366	0.682	0.645	0.036
U-Net++	0.413	0.342	0.390	0.681	0.704	0.035
SFA	0.297	0.219	0.231	0.557	0.515	0.109
PraNet	0.630	0.576	0.600	0.791	0.792	0.031
CaraNet	0.747	0.672	0.709	0.868	0.894	0.017
FSFM	0.775	0.702	0.733	0.850	0.883	0.013
BLE-Net	0.673	0.594	0.622	0.810	0.837	0.032
SAM-CLNet	0.665	0.586	0.635	0.810	0.859	0.019
MSNet	0.719	0.664	0.678	0.840	0.830	0.020
M ² SNet	0.749	0.678	0.712	0.846	0.872	0.017
Ours	0.776	0.694	0.737	0.850	0.887	0.019

Note: ↑ indicates that the higher the metric, the better the model performance, and ↓ indicates that the lower the indicator, the better the model performance. Bold data indicates the best performance.

CVC-ColonDB数据集,在mDice、mIoU指标上,本文算法分别比SAM-CLNet算法高出0.071,0.087,同时在MAE指标上减少了0.006,这充分证明了本文算法在泛化性方面相较于分割大模型类算法SAM-CLNet具有一定优势。

3.4.3 可视化对比

图5是本文算法在CVC-ClinicDB、Kvasir、CVC-300、CVC-ColonDB和ETIS-LaribPolypDB五个数据集上的部分息肉图像分割的可视化结果。其中第一行到第五行分别为CVC-ClinicDB、Kvasir、CVC-300、CVC-ColonDB、ETIS-LaribPolypDB数据集的分割结果。通过

对比前两行的可视化结果,能够直观地看出本文算法在精准描绘息肉轮廓方面的较强能力,边界清晰且准确,有力地验证了算法在核心分割任务中的较高精确度。随后三行则是对本文算法泛化性实验的可视化结果展示,尽管这些数据集与训练数据来自不同渠道,且部分图像具有边缘模糊、形状不规则等特点,导致分割结果在某些细节上(如边缘平滑度)相较于前两行结果略有不足,但本文算法仍能在保持息肉主体形状完整性的同时,相较于同等条件下的其他算法,展现出较为良好的分割形态保持能力。以上结果不仅体现了算法在应对不同数据集时

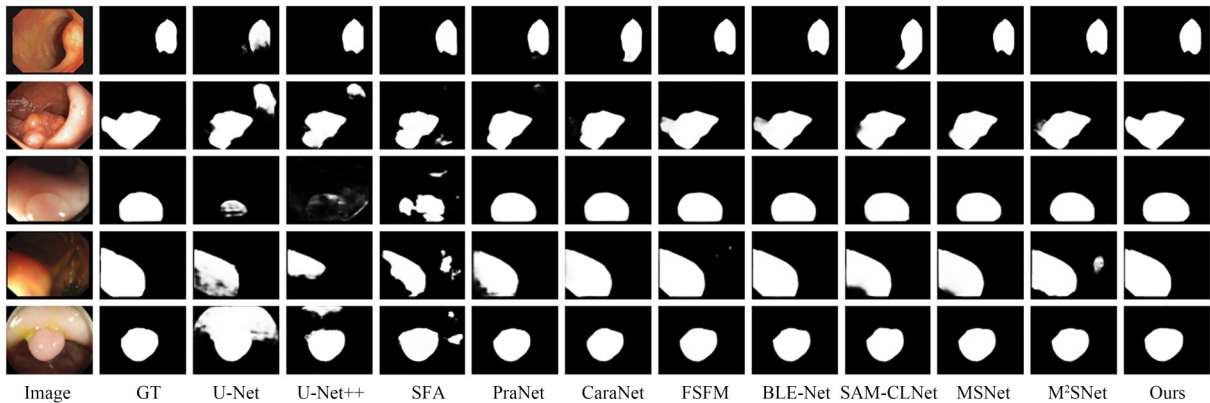


图5 各算法在五个数据集上的部分分割结果

Fig. 5 Partial segmentation results of each algorithm on five datasets

的一致性和稳定性,也进一步证实了其较好的泛化能力。

3. 4. 4 消融实验

为了验证本文算法各改进模块的有效性,本文在5个数据集上进行了消融实验,实验的训练集和测试集划分与论文3.4.1和3.4.2部分相同。Method1表示本文采用的基准网络M²SNet; Method2表示在Method1的基础上,将双分支混合上采样模块应用于所提出的网络中,以探索其在提升息肉特征图分辨率同时保持边界纹理信息方面的潜力; Method3表示在Method2的基础上加入改进的通道分组空间增强模块,旨在通过增强通道间及空间位置的相关性来抑制噪声干扰,提升分割精度; Method4即本文所提出的方法,表示在Method3的基础上添加设计的级联特征融合网络,以强化结直肠息肉与背景之间的边界区分能力。

消融实验量化结果如表6~表10所示,图6是在5个数据集上进行消融实验的部分息肉图像可视化结果,其中第一行到第五行分别为CVC-

表6 在CVC-ClinicDB数据集上的消融实验结果

Tab.6 Ablation experimental results on the CVC-ClinicDB dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
Method1	0.922	0.880	0.917	0.942	0.970	0.009
Method2	0.934	0.887	0.926	0.944	0.979	0.006
Method3	0.935	0.892	0.929	0.946	0.979	0.006
Method4	0.944	0.900	0.938	0.948	0.987	0.006

表7 在Kvasir数据集上的消融实验结果

Tab.7 Ablation experimental results on the Kvasir dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
Method1	0.912	0.861	0.901	0.922	0.953	0.025
Method2	0.914	0.867	0.906	0.920	0.952	0.024
Method3	0.917	0.867	0.908	0.921	0.956	0.022
Method4	0.920	0.869	0.910	0.924	0.961	0.022

表8 在CVC-300数据集上的消融实验结果

Tab.8 Ablation experimental results on the CVC-300 dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
Method1	0.903	0.842	0.881	0.939	0.965	0.009
Method2	0.903	0.844	0.882	0.943	0.960	0.006
Method3	0.908	0.848	0.890	0.937	0.968	0.007
Method4	0.915	0.852	0.896	0.938	0.979	0.006

表9 在ETIS-LaribPolypDB数据集上的消融实验结果

Tab.9 Ablation experimental results on the ETIS-Larib-PolypDB dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
Method1	0.749	0.678	0.712	0.846	0.872	0.017
Method2	0.752	0.678	0.714	0.844	0.873	0.018
Method3	0.761	0.692	0.726	0.851	0.865	0.015
Method4	0.776	0.694	0.737	0.850	0.887	0.019

表 10 在 CVC-ColonDB 数据集上的消融实验结果
Tab.10 Ablation experimental results on the CVC-ColonDB dataset

Method	mDice ↑	mIoU ↑	F_{β}^w ↑	S_a ↑	$E_{\phi}^{\max}$ ↑	MAE ↓
Method1	0.758	0.685	0.737	0.842	0.869	0.038
Method2	0.765	0.690	0.748	0.842	0.891	0.035
Method3	0.776	0.697	0.762	0.843	0.886	0.033
Method4	0.782	0.714	0.765	0.855	0.893	0.034

Note: ↑ indicates that the higher the metric, the better the model performance, and ↓ indicates that the lower the indicator, the better the model performance. Bold data indicates the best performance.

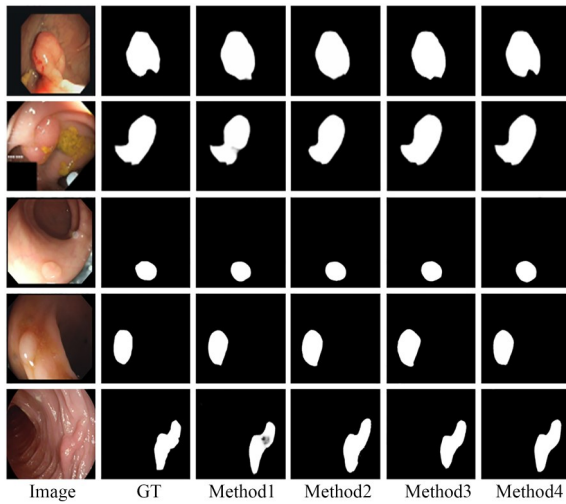


图 6 消融实验可视化结果

Fig. 6 Visual results of ablation experiments

ClinicDB, K-vasir, CVC-300, CVC-ColonDB, ETIS-LaribPolyp-DB 数据集的分割结果。Method2 在 5 个数据集上普遍展现出较好的性能,例如在 CVC-ClinicDB 数据集上,其 mDice 和 mIoU 分别较 Method1 提升了 0.012 和 0.007,同时 MAE 比降低了 0.003,这有力证明了双分支混合上采样模块在保留息肉精细边界与纹理信息、防止边界信息丢失方面的较好作用。Method3 的

各项评价指标进一步提升,例如在 CVC-ColonDB 数据集上 mDice 和 mIoU 分别较 Method2 高出 0.011 和 0.007,这表明改进的通道分组空间增强模块通过增强通道及空间位置相关性,有效降低了噪声影响,提升了分割精度。尤为值得一提的是,Method4 作为集成所有改进模块的最终方案,在 5 个数据集均实现了性能的较大提升,例如在 ETIS-LaribPolypDB 数据集上,其 mDice 较 Method3 高出 0.015,这直接体现了级联特征融合网络在提高息肉与背景边界区分度、进而优化整体分割性能方面的较好效果。从以上分析结果可以得到本文设计的各模块对网络性能都有一定程度的提升。

4 结 论

本文提出了一种采用级联策略融合边界特征的多尺度分割网络。首先,利用改进的通道分组空间增强模块对骨干网络特征进行增强,提高了通道和空间位置的相关性;其次,通过一个级联特征融合网络,使分割网络能够在减少冗余噪声的同时提高边界区分度;最后,为了帮助网络有效地处理边界信息和纹理信息,本文在多尺度减法模块中引入双分支混合上采样模块以提升网络的分割精度。在 CVC-ClinicDB 和 Kvasir 数据集上进行测试,本文算法的平均 Dice 系数分别为 0.944, 0.920, 平均交并比分别为 0.900, 0.869; 在 ETIS-LaribPolypDB, CVC-300 和 CVC-ColonDB 数据集上进行测试,本文算法的平均 Dice 系数分别为 0.776, 0.915, 0.782。实验结果表明,本文提出的算法分割精度较高、泛化性较好,有效解决了结直肠息肉图像分割时存在背景干扰、噪声较多、区分度不够等问题,为辅助医生诊断提供了有效的帮助。

未来将在跨域分割等领域进行研究,以扩大算法的应用场景。

参考文献:

- [1] 贾立新, 胡奕标, 金燕, 等. 融合多种注意力机制的结直肠息肉分割神经网络[J]. 计算机辅助设计与图形学学报, 2023, 35(3): 463-473.
- JIA L X, HU Y B, JIN Y, *et al.* Polyp segmenta-

- tion network combined with multi-attention mechanism [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2023, 35 (3): 463-473. (in Chinese)
- [2] XIA P, ZHANG G Y, LEI B J, *et al.* Polyp image

- segmentation based on multi-scale ResNeSt-50 aggregation network and message passing [J]. *Opt. Precision Eng.*, 2023, 31(18): 2765-2780. (in Chinese)
- [3] LIANG L M, HE A J, LI R J, *et al.* Cross-scale and cross-dimensional adaptive transformer network for colorectal polyp segmentation [J]. *Opt. Precision Eng.*, 2023, 31(18): 2700-2712. (in Chinese)
- [4] HU K L, ZHAO L P, FENG S, *et al.* Colorectal polyp region extraction using saliency detection network with neutrosophic enhancement [J]. *Computers in Biology and Medicine*, 2022, 147: 105760.
- [5] LI W S, HUANG Z P, LI F Y, *et al.* CIFG-Net: cross-level information fusion and guidance network for Polyp Segmentation [J]. *Computers in Biology and Medicine*, 2024, 169: 107931.
- [6] MA H C, XU C, NIE C, *et al.* DBE-net: dual boundary-guided attention exploration network for polyp segmentation [J]. *Diagnostics*, 2023, 13(5): 896.
- [7] 梁礼明, 周琬颂, 冯骏, 等. 基于高分辨率复合网络的皮肤病变分割 [J]. *光学精密工程*, 2022, 30(16): 2021-2038.
- LIANG L M, ZHOU L S, FENG J, *et al.* Skin lesion segmentation based on high-resolution composite network [J]. *Opt. Precision Eng.*, 2022, 30(16): 2021-2038. (in Chinese)
- [8] 王雪, 李占山, 吕颖达. 基于多尺度感知和语义适配的医学图像分割算法 [J]. *吉林大学学报(工学版)*, 2022, 52(3): 640-647.
- WANG X, LI ZH SH, LÜ Y D. Medical image segmentation based on multi-scale context-aware and semantic adaptor [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2022, 52(3): 640-647. (in Chinese)
- [9] 支佩佩, 邓健志, 钟震霄. 基于卷积和注意力机制的医学细胞核图像分割网络 [J]. *生物医学工程杂志*, 2022, 39(4): 730-739.
- ZHI P P, DENG J Z, ZHONG Z X. Medical nucleus image segmentation network based on convolution and attention mechanism [J]. *Journal of Biomedical Engineering*, 2022, 39(4): 730-739. (in Chinese)
- [10] LECUN Y, BOSER B, DENKER J S, *et al.* Backpropagation applied to handwritten zip code recognition [J]. *Neural Computation*, 1989, 1(4): 541-551.
- [11] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.
- [12] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2015: 234-241.
- [13] 徐昌佳, 易见兵, 曹锋, 等. 采用 DoubleUNet 网络的结直肠息肉分割算法 [J]. *光学精密工程*, 2022, 30(8): 970-983.
- XU C J, YI J B, CAO F, *et al.* Colorectal polyp segmentation algorithm using DoubleUNet network [J]. *Opt. Precision Eng.*, 2022, 30(8): 970-983. (in Chinese)
- [14] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHS 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, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, S, 2018*, 11045: 3-11.
- [15] FAN D P, JI G P, ZHOU T, *et al.* PraNet: Parallel reverse attention network for polyp segmentation [C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2020: 263-273.
- [16] LIU Y, YANG Y, JIANG Y Q, *et al.* Multi-view orientational attention network combining point-based affinity for polyp segmentation [J]. *Expert Systems with Applications*, 2024, 249: 123663.
- [17] MUHAMMAD Z U D, MUHAMMAD U, HUANG Z J, *et al.* MMFIL-Net: multi-level and multi-source feature interactive lightweight network for polyp segmentation [J]. *Displays*, 2024, 81: 102600.
- [18] JAIN S, ATAIE R, GUPTA A, *et al.* CoIn-Net: a convolution-involution network with a novel statistical attention for automatic polyp segmentation [J]. *IEEE Transactions on Medical Imaging*, 2023, 42(12): 3987-4000.
- [19] SHARMA V, KUMAR A, JHA D, *et al.* ControlPolypNet: towards controlled colon polyp syn-

- thesis for improved polyp segmentation[C]. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Seattle, WA, USA. IEEE, 2024: 2325-2334.
- [20] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. IEEE, 2017: 936-944.
- [21] SU Y Z, XIE Q S, YE J, *et al.* An accurate polyp segmentation framework via feature secondary fusion[C]. 2023 *IEEE 20th International Symposium on Biomedical Imaging (ISBI)*. Cartagena, Colombia. IEEE, 2023: 1-5.
- [22] ZHAO X Q, ZHANG L H, LU H C. Automatic polyp segmentation via multi-scale subtraction network[C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2021: 120-130.
- [23] ZHAO X Q, JIA H P, PANG Y W, *et al.* M²SNet: Multi-Scale in multi-scale subtraction network for medical image segmentation[EB/OL]. 2023: arXiv: 2303.10894. <https://arxiv.org/pdf/2303.10894.pdf>.
- [24] LI Y X, LI X, YANG J. Spatial Group-Wise Enhance: enhancing semantic feature learning in CNN[C]. *Asian Conference on Computer Vision*. Cham: Springer, 2023: 316-332.
- [25] 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.
- [26] ZHANG Z J, FU H Z, DAI H, *et al.* ET-net: A generic edge-attention guidance network for medical image segmentation[C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2019: 442-450.
- [27] 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.
- [28] CHEN X K, LIN K Y, WANG J B, *et al.* Bi-directional Cross-Modality Feature Propagation with Separation-and-Aggregation Gate for RGB-D Semantic Segmentation[C]. *European Conference on Computer Vision*. Cham: Springer, 2020: 561-577.
- [29] ZHOU T, ZHOU Y, HE K L, *et al.* Cross-level feature aggregation network for polyp segmentation[J]. *Pattern Recognition*, 2023, 140: 109555.
- [30] YANG L, ZHAI C X, LIU Y H, *et al.* CFHA-Net: a polyp segmentation method with cross-scale fusion strategy and hybrid attention[J]. *Computers in Biology and Medicine*, 2023, 164: 107301.
- [31] WEI J, WANG S H, HUANG Q M. F³Net: fusion, feedback and focus for salient object detection[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 12321-12328.
- [32] BERNAL J, SÁNCHEZ F J, FERNÁNDEZ-ESPARRACH G, *et al.* WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physicians[J]. *Computerized Medical Imaging and Graphics*, 2015, 43: 99-111.
- [33] JHA D, SMEDSRUD P H, RIEGLER M A, *et al.* Kvasir-SEG: a segmented polyp dataset[C]. *International Conference on Multimedia Modeling*. Cham: Springer, 2020: 451-462.
- [34] VÁZQUEZ D, BERNAL J, SÁNCHEZ F J, *et al.* A benchmark for endoluminal scene segmentation of colonoscopy images[J]. *Journal of Healthcare Engineering*, 2017, 2017(1): 4037190.
- [35] TAJBAKSH N, GURUDU S R, LIANG J M. Automated polyp detection in colonoscopy videos using shape and context information[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(2): 630-644.
- [36] SILVA J, HISTACE A, ROMAIN O, *et al.* Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2014, 9(2): 283-293.
- [37] LV Y Q, ZHANG J, DAI Y C, *et al.* Simultaneously localize, segment and rank the camouflaged objects[C]. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, TN, USA. IEEE, 2021: 11586-11596.
- [38] FANG Y Q, CHEN C, YUAN Y X, *et al.* Selective feature aggregation network with area-boundary constraints for polyp segmentation[C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2019: 302-310.

- [39] LOU A, GUAN S, LOEW M. CaraNet: context axial reverse attention network for segmentation of small medical objects[J]. *Journal of Medical Imaging (Bellingham, Wash)*, 2023, 10 (1): 014005.
- [40] TAN, CHEN H P, LYU Y D, *et al.* BLE-Net: boundary learning and enhancement network for polyp segmentation [J]. *Multimedia Systems*, 2023, 29(5): 3041-3054.
- [41] ZHAO Y M, ZHOU T, GU Y Q, *et al.* Segment anything model-guided collaborative learning network for scribble-supervised polyp segmentation [EB/OL]. 2023: arXiv: 2312.00312. <https://arxiv.org/pdf/2312.00312.pdf>.

作者简介:



易见兵(1980—),男,江西宜春人,博士,副教授,硕士生导师,2017年于深圳大学获得博士学位,主要从事计算机视觉、图像配准、高性能计算等方面的研究。E-mail: yijianbing8@jxust.edu.cn



万建辉(1999—),男,江西抚州人,硕士研究生,2022年于江西理工大学应用科学学院获得学士学位,主要从事计算机视觉、图像分割等方面的研究。E-mail: wan15079426139@163.com