

基于双边融合的医学图像分割算法

梁礼明*, 尹江, 吴媛媛, 冯骏

江西理工大学电气工程与自动化学院, 江西 赣州 341000

摘要 针对小目标识别与分割问题,提出了一种基于双边融合的网络模型BFNet(Bilateral Fusion Network),该模型具有双分支结构。一支为具有较窄的通道和较浅的结构层,用于关注相邻像素点之间的联系。另一支引入RFB(Receptive Field Block)和DFB(Dense Fusion Block)两个模块,其具有较宽的通道和较深的结构层,可以获得高级语义的上下文信息。随后由一个引导聚合层融合两分支的特征表示。以息肉和皮肤病变区域为应用对象,使用三个公开的医学分割数据集来评估所提算法的性能。实验结果表明,在息肉和皮肤病变分割任务中,所提算法优于现有的医学图像分割算法。特别是在自动息肉检测Kvasir-SEG数据集中,所提算法的平均Dice和平均交并比分别达到了92.3%和86.2%,均已高于已有算法。

关键词 医用光学; 图像处理; 息肉分割; 皮损分割; 医学图像分割; 双边融合

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202259.0817003

Medical Image Segmentation Algorithm Based on Bilateral Fusion

Liang Liming*, Yin Jiang, Wu Yuanyuan, Feng Jun

School of Electrical Engineering and Automation, Jiangxi University of Science and Technology,
Ganzhou, Jiangxi 341000, China

Abstract Aiming at the problem of small target recognition and segmentation, a network model based on bilateral fusion (BFNet) is proposed, which has a dual-branch structure. One has a narrow channel and a shallower structural layer, which focuses on the connections between adjacent pixels. The other introduces two modules, such as receptive field block (RFB) and dense fusion block (DFB), which have wider channels and deeper structural layers and can obtain high-level semantic context information. It is then represented by a guide aggregation layer that fuses the features of the two branches. Three open medical segmentation datasets are used to evaluate the performance of the proposed algorithm. The experimental results show that the proposed algorithm is superior to existing medical image segmentation algorithms in the segmentation task of polyps and skin lesions. Especially, in the automatic polyp detection Kvasir-SEG dataset, the average Dice and average cross ratio of the proposed algorithm reached 92.3% and 86.2% respectively, which are both higher than the existing algorithms.

Key words medical optics; image processing; polyp segmentation; skin lesion segmentation; medical image segmentation; bilateral fusion

收稿日期: 2021-07-20; 修回日期: 2021-08-16; 录用日期: 2021-08-25

基金项目: 国家自然科学基金(51365017, 61463018)、江西省自然科学基金面上项目(20192BAB205084)、江西省教育厅科学技术研究重点项目(GJJ170491)

通信作者: lianglm67@163.com

1 引言

医学图像分割是一种将图像分割成易于分析和解释的感兴趣区域(ROI)的技术,通常是临床应用中的一项关键任务,常用于计算机辅助诊断(CADx)和治疗计划的指导^[1],有助于临床医生对疾病进行更精准的诊断。

目前,临床中的图像分割任务主要依靠有经验的医生对病灶区域和息肉进行精准的手工分割。但是,手工分割往往是一项重复且费时费力的工作,同时受制于医生主观认识上的差异,造成分割结果不一致,因此需要一种自动和高效的图像分割算法。近些年,深度学习在图像分割领域发展迅速且取得了众多成果。然而,基于医学图像的自动分割算法仍然是一个新兴领域,其中一个重大挑战是缺乏测试数据集。此外,公开的医学数据集通常是不平衡的。在某种程度上,许多分割算法的性能在自然图像上是可以接受的,但医学图像所需的分割精度较高,同时也需要很高的稳定性,不稳定或不正确的分割将会直接影响后续的治疗计划,从而失去分割的意义^[2]。

近年来,卷积神经网络(CNN)在医学图像分割方面取得了显著的成功。在不同的CNN体系结构中,FCN(Fully Convolutional Networks)^[3]及其扩展U-Net^[4]编解码器网络在医学图像分割中得到了广泛的应用。Zhang等^[5]设计了一种深度残差U型网络,利用残差连接可以更好地输出分割图。Ibtehaz等^[6]对U-Net进行了改进,提出了一种新的MultiResU-Net(Multimodal Deep Residual U-Network)医学图像分割架构,引入MultiRes模块可以更好地处理边界模糊问题。Oktay等^[7]将AGs(Attention Gates)与U-Net结合,有效捕获了目标位置先验信息,其作用是抑制输入图像中不相关的区域,着重关注目标显著特征。文献[8]通过Attention U-Net^[7]来定位目标,提出了一种改进的DenseUnet^[9]分割方法。Jha等^[10]构建的ResUNet++是标准ResUNet的增强版本,通过在网络中添加多个模块,如ASPP(Atrous Spatial Pyramid Pooling)和注意块加强了上下文联系。文献[11]提出了一种用于语义和实例分割任务的神经网络体系结构,即UNet++模型,并通过减少未知的网络深度、重新设计跳过连接以及为架构设计剪枝方案来提高UNet++模型的性能。

从上述相关工作可以看出,用于医学图像分割的深层CNN体系结构已产出大量的研究成果,并且研究成果主要集中于所构建的医学图像分割模型的泛化性能^[6,10-11]。然而,在医学影像中仍有许多具有挑战性的图像分割任务,如结肠镜检查中的扁平息肉,这些扁平息肉在结肠镜检查中往往会被遗漏,如果不及早发现,就可能发展为癌症,为此需要一种更精准的分割模型来应对具有挑战性的医学图像问题。

本文提出一种双边融合的网络模型,在网络的双分支部分使用改进的HRNet(High-Resolution Network)^[12]体系结构。将HRNet作为主干网络的主要原因如下:1)与其他预训练的模型相比,HRNet是并行连接的卷积流,而不是串联的卷积流(并行连接的卷积流在特征提取的过程中能够始终保持高分辨率,无需从低分辨率恢复到高分辨率,这可以使空间上的特征表示更精准);2)HRNet能跨分辨率重复交换信息,所产生的特征表示在语义上更丰富;3)HRNet允许更深的网络产生更好的输出分割掩码,并通过融合此架构来提高网络的整体分割性能。

本文主要贡献如下。1)提出了一种新的医学图像分割模型BFNet(Bilateral Fusion Network),该网络结构为两个不同分支提供并行的方式以同时对图像进行特征提取,然后由一个聚合层进行特征融合。2)设计了一个深层语义融合分支,引入了一个基于RFB(Receptive Field Block)模块和DFB(Dense Fusion Block)的轻量级网络,以增强感受野并捕获丰富的上下文信息。3)多个数据集的对比实验中体现了所提算法的性能优于其他算法。本文在三个不同的医学图像数据集上进行了相关实验,其中有两个数据集来自皮肤镜检查,一个来自结肠镜检查。与现有的算法相比,BFNet在Kvasir-SEG息肉分割数据集^[13]以及ISIC2016^[14]和ISIC2017^[15]皮肤病变分割数据集上显示出更优的分割性能。4)引入增强损失的训练方法,以进一步提高算法的分割性能而不会增加训练成本。

2 本文算法

本节首先介绍图像分割网络BFNet的整体结构,分别阐述所设计的各个模块。两个分支网络都是以HRNet为骨干网络,分支网络分别为深层语义融合分支和浅层细节融合分支,最后由聚合层融合

双分支的语义信息。

2.1 网络模型的整体结构

图 1 显示了网络模型的整体结构,其主要有浅层细节分支(HR1)、深层语义分支(HR2)和双边聚合(Agg)层三个组成部分。浅层语义分支主要负责空间细节,通过融合低分辨率的特征来丰富空间细节信息。同时,浅层语义分支着重关注相邻像素点之间的联系,对小物体的识别有一定的

帮助。深层语义分支用来捕获高级语义信息和整个图像的上下文联系,有助于模糊目标的精确定位及大目标的识别。聚合层部分采用的是双边聚合层,其中 down 表示下采样操作,up 表示上采样操作,⊗表示卷积操作。此外,在损失函数部分设计了一些辅助的分割头,在不增加网络推理的代价下提高网络的分割性能。

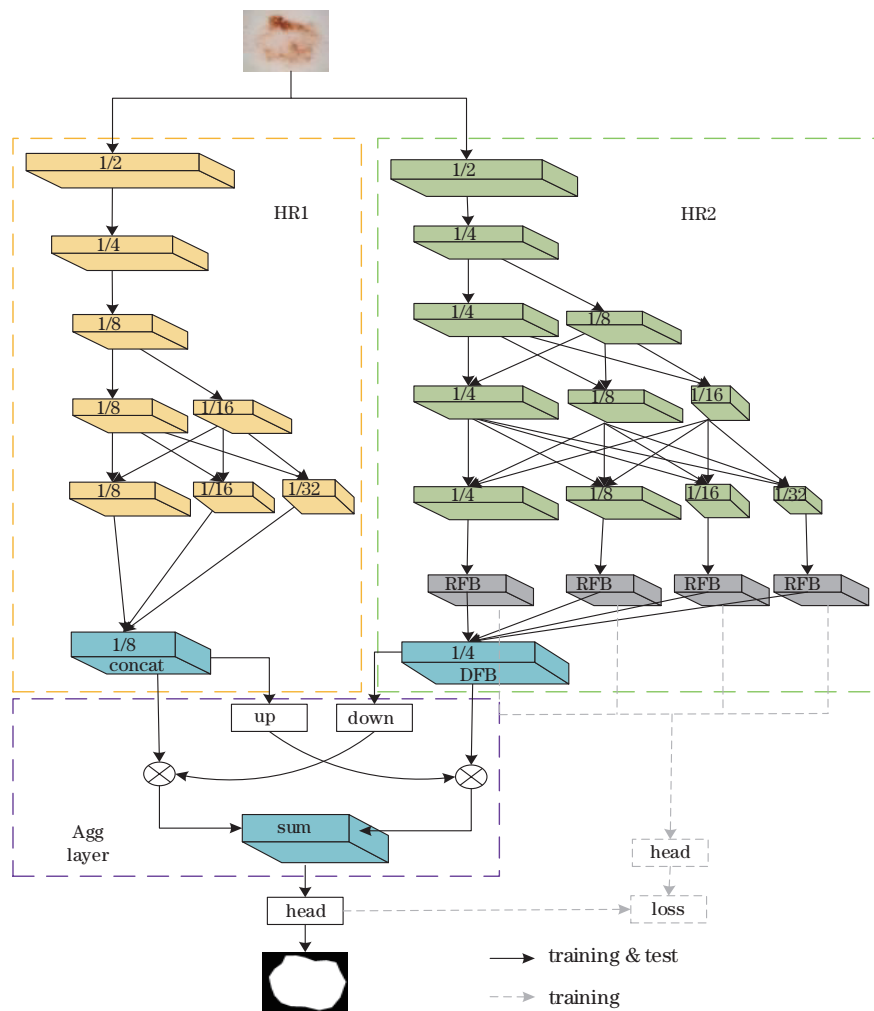


图 1 基于双边融合的医学图像分割网络模型 BFNet 的整体结构

Fig. 1 Overall structure of medical image segmentation network model BFNet based on bilateral fusion

2.2 深层语义融合分支

在皮肤镜图像中,当病变区域与相邻区域高度相似时,算法很难分割出这个病变区域。此外,受光线强弱的影响,病变区域的边界轮廓容易与背景融为一体,导致算法难以精确分割出病变区域的边缘。深度信息不会随光线强弱的变化而改变,因此可以根据深度值对病变区域进行分割。深层语义融合分支将 HRNet 作为主干网络,加入了 RFB^[16] 和 DFB。

深层语义融合分支网络的主要结构如图 2 所示,其中 Conv 表示卷积运算,concat 表示特征融合操作。

2.2.1 主干网络

主干网络由一个头部网络(图 2 线条块)和并行网络(图 2 阴影块)组成。首先将图像输入到头部网络中,头部网络由两个卷积核大小均为 3×3、步长均为 2 的卷积层构成,此时将图像的分辨率降低到输入图像的 1/4。随后将低分辨率图像输入

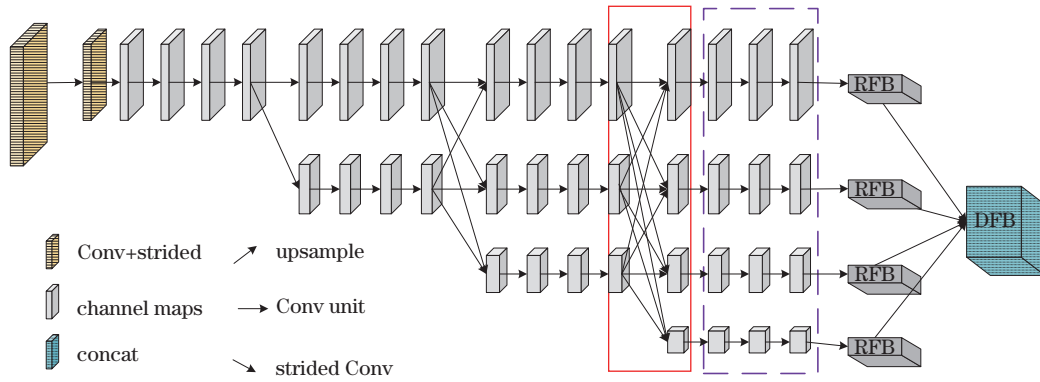


图 2 深层语义融合网络的结构

Fig. 2 Structure of deep semantic fusion network

进并行网络(图 2 阴影块)中,并行网络由并行多分辨率卷积层(图 2 点框)和重复多分辨率融合层(图 2 实线框)组成。并行多分辨率卷积层从高分辨率卷积流开始作为第一阶段,逐步增加从高分辨率到低分辨率的流以形成新的阶段,同时并行连接多个分辨率流。因此,后一阶段并行流的分辨率是由前一阶段的分辨率和一个额外更低的分辨率组成。重复多分辨率融合层通过双线性上采样,在不改变通道数的情况下重新缩放分辨率表示,并将 4 个表示连接起来,然后进行 1×1 的卷积以混合这 4 个表示。

2.2.2 RFB

图 3 显示的是 RFB^[16],该模块主要是使用不同内核大小的卷积层和空洞卷积层,在保持参数尽可能少的同时,在更大的区域和更多的上下文中捕获信息。最后用 1×1 的卷积来融合这些特征。图 3 中 BN 表示归一化,ReLU 表示激活函数, 1×1 、 1×3 和 3×1 等表示卷积核的大小。受文献[17]的启发,本文将 RFB 添加到不同尺度的特征图(图 2 阴影块)中,以增加网络对不同尺度的感受野。

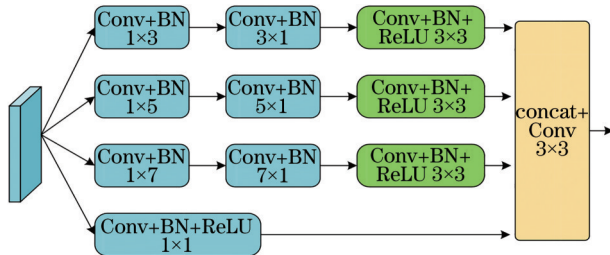


图 3 RFB 的结构

Fig. 3 Structure of RFB

2.2.3 DFB

CNN 的高层特征中每个通道都有特定类别的

语义信息,为了充分利用不同分辨率的特征的语义信息,受现有文献[18]密度融合思想的启发,将多尺度输入的图像特征逐个上采样并与上一个特征相乘进行融合。通过整合相同尺度的空间特征通道和深度特征通道,使得编码相同的语义信息互相增强。图 4 展示的是 DFB,其中 V_{jk} 表示子流, j 表示子流是处于第几阶段, k 表示子流的分辨率系数。当 k 值相同时,子流的分辨率大小相同,即每层的分辨率是相同的。

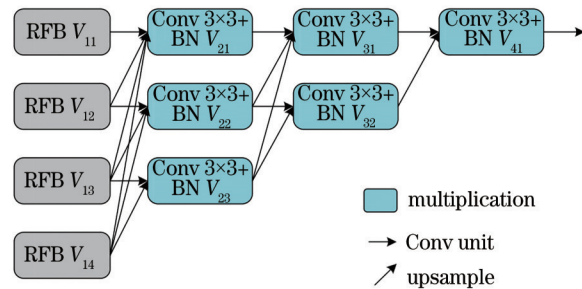


图 4 DFB 的结构

Fig. 4 Structure of DFB

2.3 浅层细节融合分支

在医学图像中,小目标病变区域尺寸较小、边缘较模糊,相比于大目标病变区域更难被精准识别,导致整体的病变区域分割精度较低。针对此问题,本文设计了浅层细节分支,该分支包含了一个小跨度的浅层结构,主要通过融合小尺度的特征来丰富空间细节信息。浅层语义融合分支网络的结构如图 5 所示,其主要由一个头部网络(图 5 线条块)和并行网络(图 5 阴影块)组成。头部网络由三个卷积核均为 3×3 、步长均为 2 的卷积层构成,并行网络由三行并联的卷积层组成。并行网络同样由并行多分辨率卷积层(图 5 虚线框)和重复多分辨率

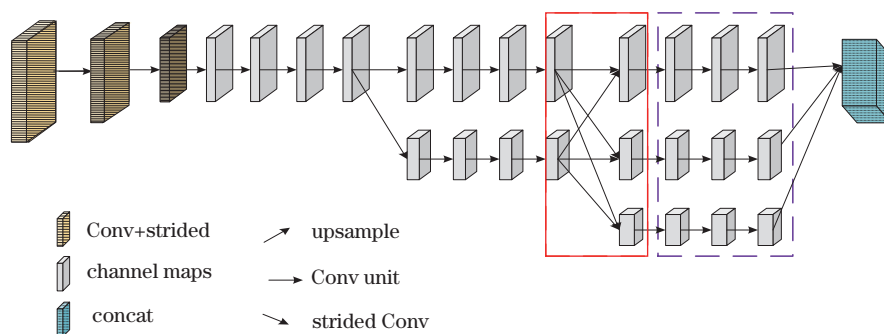


图 5 浅层细节融合网络的结构

Fig. 5 Structure of shallow detail fusion network

融合层(图 5 实线框)组成,与 2.2.1 节的主干网络结构类似。

2.4 双边融合模块

深层语义分支和浅层语义分支的特征表示是互补的,双边融合模块的设计是为了融合两个分支的特征。如图 1 的 Agg 层所示,该模块利用不同尺度特征的融合来引导深层语义分支和浅层语义分支的特征表示,对两个分支分别进行上采样与下采样,随后将相同尺度的特征进行融合,从而获得不同尺度的特征表示,这些特征表示进一步编码了多尺度信息,这种引导方法能够使两个分支之间的特征信息进行高效融合。

2.5 增强损失的训练方法

为了进一步提高分割精度,本文采用增强损失的训练方式,灵感来自文献[19]。增强损失的训练方式在训练阶段可以增强网络对目标的特征表示,在预测推理阶段不参与计算,因此在推理阶段几乎不增加计算复杂度。图 1 中虚线连接的损失只参见训练阶段。图 6 显示了分割头(head)的细节,上采样采用的是双线性插值,同时 1×1 与 3×3 表示核大小, H 、 W 和 C 分别表示高度、宽度和通道维度, S 表示上采样的尺度比, N 为最终输出维度。通过控制通道维度 C 来调整辅助分割头和主分割头的计算复杂度,同时通过调整系数 S 使输出图像的高和宽与输入图像相同,最终对分割头输出的特征进行损失计算。

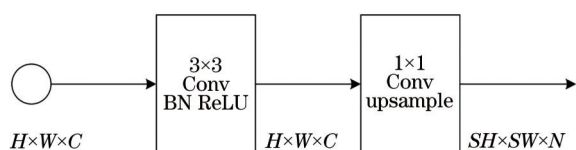


图 6 分割头的结构

Fig. 6 Structure of partition head

2.6 与现有方法的联系和区别

本文所提出的双边融合网络主要针对医学图像数据量少、小目标以及边缘难以分割的情况。为了解决这一问题,设计了深层语义融合分支和浅层细节融合分支,灵感主要源于文献[19-21]。相比于文献[20],本文设计的深层分支在 HRNet 的基础上引入了 RFB 和 DFB,对于医学图像边缘难以分割的情况起到一定的改善作用。同时,提出了一个浅层分支,在 HRNet 的基础上进行了简化,其主要目的是通过融合小尺度的特征来丰富空间细节信息,改善了整体网络对医学图像中小目标的识别。相比于文献[21],本文借鉴了文献[21]的双边融合思想。但在医学图像分割方面,文献[20-21]都达不到令人满意的结果,这些算法主要是为语义分割设计,在 Cityscapes 公开数据集上均表现得较为优秀,但在医学图像分割的任务中达不到想要的效果。为此,本文借鉴文献[19-21]中的优秀思想并针对医学图像的特点对网络进行改进。为了验证本文所构建的双边融合分割模型的性能,将本网络在三个数据集(Kvasir-SEG、ISIC2016 和 ISIC2017)上进行验证。

3 实验

本节首先介绍实验所使用的数据集、实验环境和评价指标,然后与现有的医学图像分割算法进行对比,最后对本文网络模型内的两条分支网络和各个模块的作用进行实验分析。

3.1 实验数据集与实验环境设置

3.1.1 数据集

本节描述了三个公共的医学图像数据集,三个数据集分别为 ISIC2016 和 ISIC2017 皮肤病变分割数据集以及 Kvasir-SEG 息肉分割数据集。ISIC2016 与 ISIC2017 数据集是常用的皮肤病变分

割数据集。ISIC2016 数据集包含了 900 组训练数据图像和 379 组测试数据图像, ISIC2017 数据集包含了 2000 组训练数据图像和 600 组测试数据图像, 每组图像都包括一张原始图像和一张金标准图像。其中原始图像是 JPEG 格式的皮肤镜图像, 金标准图像是专家手工标注、PNG 格式的二进制蒙版图像。所有图像都被精细标注, 每个像素均被分配一个类别标签。Kvasir-SEG 数据集是一个具有挑战性的息肉分割数据集, 它是从 Kvasir 数据集的一个子类(息肉类)中选择的 1000 幅图像集。本文采用与文献[22]相同的数据划分方法, 在实验评估中使用 880 个实例进行训练, 120 个实例进行测试。

3.1.2 实验环境设置

本文实验是基于百度飞桨的 paddlepaddle2.1 实现的, 采用了由百度提供的 PaddleSeg 图像分割开发套件。实验是在百度的云平台(AI Studio)上进行的。使用的机器是 v100 型 GPU 加速图像处理器, 内存大小为 32 G, 显存大小为 16 G。在实验训练过程中, 对输入数据进行数据增强处理, 这个数据增强处理过程是与训练过程同步进行的。实验主要采用尺寸随机按比例缩放的图像, 最大尺寸为原始图像的 2 倍, 最小尺寸为原始图像的 1/2。然后将所有输入图像的大小调整为 256 pixel×256 pixel 并作为网络输入。训练过程中, batch_size 为 16, epoch 为 120(将训练数据全部训练一次为一个 epoch)。优化算法采用的是随机梯度下降(SGD)算法, 学习率为 0.05, 同时加入一个动量衰减值 y , 值为 0.9, 而衰减的速率为 0.0001。为了缓解 Kvasir-SEG 数据集中背景和目标之间的不平衡, 训练过程中在交叉熵损失函数上为不同类别添加一个权重系数。

3.2 损失函数与评估指标

本文实验采用了加权交叉熵损失函数, 表达式为

$$L = -\frac{1}{n} \sum_{i=1}^n [w_0 y_i \ln(p_i) + w_1 (1 - y_i) \ln(1 - p_i)], \quad (1)$$

式中: w_0 和 w_1 分别表示背景和目标的损失函数权重, 由实验可知, 当 $w_0=1$ 和 $w_1=1.2$ 时, 网络的分割精度最高; n 表示类别的数量; y_i 表示样本 i 的类别, 正类为 1, 负类为 0; p_i 表示样本 i 预测为正类的概率。

从有效性和复杂性两个方面对算法进行评估,

分别使用平均交并比(mIoU)、平均 Dice 系数、准确率、召回率、像素精度和 Kappa 值 6 个指标进行衡量。上述指标的计算公式分别为

$$x_{mIoU} = \frac{1}{k+1} \left(\frac{x_{TP}}{x_{TP} + x_{FP} + x_{FN}} + \frac{x_{TN}}{x_{TN} + x_{FP} + x_{FN}} \right), \quad (2)$$

$$x_{mDice} = \frac{1}{k+1} \left(\frac{2x_{TP}}{2x_{TP} + x_{FN} + x_{FP}} + \frac{2x_{TN}}{2x_{TN} + x_{FN} + x_{FP}} \right), \quad (3)$$

$$P = \frac{x_{TP}}{x_{TP} + x_{FP}}, \quad (4)$$

$$R = \frac{x_{TN}}{x_{TN} + x_{FP}}, \quad (5)$$

$$x_{Acc} = \frac{x_{TP} + x_{TN}}{x_{TP} + x_{TN} + x_{FP} + x_{FN}}, \quad (6)$$

$$x_{Pe} = \frac{(x_{TP} + x_{FP}) \times (x_{TP} + x_{FN})}{(x_{TP} + x_{FP} + x_{FN} + x_{TN})^2}, \quad (7)$$

$$x_{Kappa} = \frac{x_{Acc} - x_{Pe}}{1 - x_{Pe}}, \quad (8)$$

式中: $k+1$ 表示包含背景和非病变区域的总类别数; x_{TP} 表示真阳性像素点的个数; x_{TN} 表示真阴性像素点的个数; x_{FP} 表示假阳性像素点的个数; x_{FN} 表示假阴性像素点的个数。

3.3 实验结果分析

实验结果如表 1~4 所示。除本文算法以外, 表 1 的对比实验结果均来自文献[20,22], 不同的实验环境和实验参数可能会对各项评价指标的结果产生影响。表 2~4 的对比实验均是在相同的实验环境和参数设置下进行的。

3.3.1 与其他算法的对比结果

1) 在 Kvasir-SEG 数据集上的对比结果

本文首先将所提算法与现有算法在 Kvasir-SEG 数据集上进行对比实验。为了验证本文提出的双边融合网络对于医学图像中息肉自动分割任务的优越性, 将所得结果与文献[20,22]进行对比, 实验结果如表 1 所示。

从表 1 可以看到, 本文算法在平均交并比和平均 Dice 上均优于现有算法。具体来说, 本文设计的双边融合网络在 Kvasir-SEG 测试集上的平均交并比、平均 Dice、准确率、召回率和像素精度分别为 86.2%、92.3%、89.1%、86.7% 和 96.1%; 相比于文献[12], 本文算法的平均交并比、平均 Dice、准确率分别提高了 10.3 个百分点、7.8 个百分点和 1.3

表 1 所提算法与其他算法在 Kvasir-SEG 数据集上的比较结果

Table 1 Comparison results of proposed algorithm and other algorithms on Kvasir-SEG dataset

Algorithm	mIoU / %	mDice / %	Precision / %	Recall / %	Overall accuracy / %
Ref. [4]	47.1	59.7	67.2	61.7	89.4
Ref. [5]	57.2	69.0	74.5	72.5	91.7
Ref. [10]	61.3	71.4	78.4	74.2	91.7
Ref. [3]	73.7	83.1	88.2	83.5	95.2
Ref. [12]	75.9	84.5	87.8	85.9	95.2
Ref. [23]	73.3	81.3	86.1	84.0	94.9
Ref. [24]	74.4	84.1	89.0	83.6	95.3
Ref. [25]	77.6	85.7	89.1	86.2	96.1
Ref. [25]	78.6	86.4	90.6	85.9	96.1
Ref. [26]	81.0	87.6	94.4	86.0	96.8
Ref. [20]	84.8	90.4	90.7	92.3	96.9
BFNet (ours)	86.2	92.3	89.1	86.7	96.1

个百分点,这证实了本文的双边融合网络在息肉分割的任务上可以取得更优的分割结果。

本文进一步验证将 HRNet 作为骨干网络的性能,结果如表 1 所示。使用已有的优秀骨干网络进一步提高了本文网络的分割性能,相比于文献[25] (使用 ResNet50 和 ResNet101 作为编码器的 DeepLabv3+ 算法),本文算法在 Kvasir-SEG 测试集的 5 种评价指标均优于文献[25]。

2) 在 ISIC2016 和 ISIC2017 数据集上的对比结果

为了更清楚地体现本文提出的双边融合网络模型在医学图像分割上的泛化性,在 ISIC2016 及 ISIC2017 数据集上将本文算法与医学图像分割经典算法进行对比,主要评价指标为平均交并比、像素精度和 Kappa 值,实验结果如表 2 和表 3 所示。

表 2 所提算法与其他算法在 ISIC2016 数据集上的比较结果
Table 2 Comparison results of proposed algorithm and other algorithms on ISIC2016 dataset

Algorithm	mIoU / %	Overall accuracy / %	Kappa / %
Ref. [4]	89.11	93.32	87.34
Ref. [7]	89.19	95.33	88.46
Ref. [11]	89.53	95.48	88.46
Ref. [27]	89.64	95.50	89.96
BFNet (ours)	89.66	95.54	88.90

本文算法在 ISIC2016 数据集上的平均交并比达到了 89.66%,像素精度达到了 95.54%;在 ISIC2017 数据集上的平均交并比达到了 83.41%,像素精度达到了 93.39%。从表 2 和表 3 可以看出,在平均交并比与像素精度两项主要评价指标上,本

表 3 所提算法与其他算法在 ISIC2017 数据集上的比较结果
Table 3 Comparison results of proposed algorithm and other algorithms on ISIC2017 dataset

Algorithm	mIoU / %	Overall accuracy / %	Kappa / %
Ref. [4]	78.34	91.57	74.87
Ref. [7]	80.32	92.45	77.49
Ref. [11]	80.16	92.38	77.28
Ref. [27]	82.23	93.20	80.51
BFNet (ours)	83.41	93.39	81.45

文算法均高于其他分割算法。

相比于现有医学图像分割算法,本文算法不仅在息肉分割任务中能取得优异的结果,在皮肤镜图像分割任务中也能表现出较强的优势。这表明本文提出的双边融合网络在不同数据集上均可以学习到具有判别性的特征表示,能够有效适应医学图像数据中类别和背景的多样性。

3.3.2 消融实验与分析

为了进一步验证本文算法的有效性,将双边融合网络分成三个基本模块,分别为浅层融合网络、深层融合网络以及双边融合模块,单独考虑三个模块对网络性能的影响,并在 ISIC2017 数据集上进行消融实验,实验结果如表 4 所示。

首先探讨两个分支对模型的影响,表 4 的第一组实验是仅使用一条分支的分割精度。浅层融合分支缺乏足够的高级语义特征信息,而深层融合分支缺乏低级空间细节,这会导致分割结果达不到令人满意的效果。表 4 的第二组实验表明,两个不同分支的不同组合的分割结果都优于仅一条分支,这

表 4 不同模块对网络性能的影响
Table 4 Influence of different modules on network performance

Shallow network	Deep network			Agg	EL	mIoU / %	mDice / %	Overall accuracy / %
	Backbone	RFB	DFB					
✓						79.81	88.30	92.00
	✓	✓	✓			80.11	88.59	92.14
✓	✓			✓	✓	83.00	90.48	93.11
✓	✓	✓		✓	✓	83.34	90.68	93.30
✓	✓	✓	✓	✓		82.69	89.88	93.00
✓	✓	✓	✓	✓	✓	83.41	90.73	93.39

说明两条分支可以提供互补的特征表示,从而使模型有更好的分割性能。表 4 的第三组实验表明,在加入增强损失(EL)方法后,网络模型的自动分割性能得到了较大提升。

对于 ISIC2017 数据集,在双边分支融合网络的基础上依次在深层网络中增加 RFB 和 DFB,可以看到 mIoU 由 83.00% 分别增加至 83.34% 和 83.41%,平均 Dice 由 90.48% 分别增加至 90.68% 和 90.73%,像素精度由 93.11% 分别增加至 93.30% 和 93.39%,这证明了 RFB 及 DFB 对提高皮肤镜图像分割精度的有效性。

3.4 结果可视化

图 7 展示了所提算法在 ISIC2017 数据集上的可视化对比结果,第 3 列为本文网络 BFNet 的分割结果。从图 7 可以看出:所提算法在各种复杂的病变区域中都取得了精确的分割结果;当病变区域的边缘与背景相似时,现有算法的分割结果很难令人满意;如图 7 第 3 行所示,所提算法可以完整地分割边界模糊的病变区域,而其他算法对边界模糊的分割目标存在一些遗漏;如图 7 最后一行所示,所提算法能很好地对小目标进行较为精确的分割,优于其他算法。从可视化结果可以得出,所提算法能过处理皮肤镜图像的各种情况并能分割出高精度的结果。

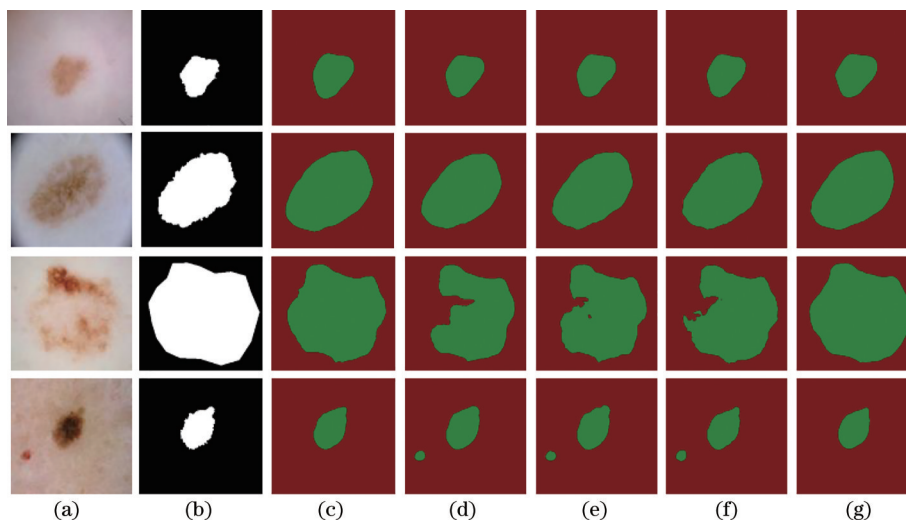


图 7 不同算法在 ISIC2017 数据集上的可视化结果对比。(a)原图;(b)专家手工标注图;(c)所提算法;(d)文献[4];(e)文献[7];(f)文献[11];(g)文献[27]

Fig. 7 Comparison of visualization results of different algorithms on ISIC2017 dataset. (a) Original images; (b) manual labeling images by experts; (c) proposed algorithm; (d) Ref. [4]; (e) Ref. [7]; (f) Ref. [11]; (g) Ref. [27]

4 结 论

本文提出了一种基于双边融合的医学图像分割模型 BFNet,其主要有三个模块,即两个分支模块和一个引导聚合模块。两个分支模块分别为深

层语义融合模块和浅层细节融合模块,本文用一个引导聚合层将它们集成在一个网络中,有效融合了多层次的空间特征和深度特征。从消融实验结果可知,两个不同分支的不同组合的分割结果都优于仅一条分支,这说明两条分支可以获得更多的互补

信息,增强了小目标与目标边缘的完整性。在三个公开医学图像数据集上对不同算法进行对比实验。结果表明,本文算法能够有效处理各种复杂性的医学图像,在多个具有挑战性的医学图像分割数据集上的分割性能均优于现有算法。所提算法可以应用于其他语义分割的计算机视觉任务,例如遥感图像描述生成、场景识别和城市景观分割等。

接下来的工作将为不同的医学图像分割任务进行针对性的研究,并着重于简化架构,同时保留其产生高分割掩模的能力。与HRNet相比,双分支融合网络使用了更多参数,这会导致训练时间增加。未来的研究方向为更多地集中在设计参数更少的简化架构上,同时保持其分割性能。

参 考 文 献

- [1] Lê M, Unkelbach J, Ayache N, et al. GPSSI: Gaussian process for sampling segmentations of images[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 38-46.
- [2] Song J, Xiao L, Lian Z C, et al. Overview and prospect of deep learning for image segmentation in digital pathology[J]. Journal of Software, 2021, 32(5): 1427-1460.
宋杰,肖亮,练智超,等.基于深度学习的数字病理图像分割综述与展望[J].软件学报,2021,32(5): 1427-1460.
- [3] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 3431-3440.
- [4] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [5] Zhang Z X, Liu Q J, Wang Y H. Road extraction by deep residual U-Net[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(5): 749-753.
- [6] Ibtehaz N, Rahman M S. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation[J]. Neural Networks, 2020, 121: 74-87.
- [7] Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: learning where to look for the pancreas[EB/OL]. (2018-04-11)[2021-05-04]. <https://arxiv.org/abs/1804.03999>.
- [8] Tian F Y, Zhou M Q, Yan F, et al. Spinal CT segmentation based on AttentionNet and DenseUnet [J]. Laser & Optoelectronics Progress, 2020, 57(20): 201008.
田丰源,周明全,闫峰,等.基于AttentionNet和DenseUnet的脊椎CT分割[J].激光与光电子学进展,2020,57(20):201008.
- [9] Sekuboyina A, Kukačka J, Kirschke J S, et al. Attention-driven deep learning for pathological spine segmentation[M]//Glocker B, Yao J H, Vrtovec T, et al. Computational methods and clinical applications in musculoskeletal imaging. Lecture notes in computer science. Cham: Springer, 2018, 10734: 108-119.
- [10] Jha D, Smedsrud P H, Riegler M A, et al. ResUNet++: an advanced architecture for medical image segmentation[C]//2019 IEEE International Symposium on Multimedia (ISM), December 9-11, 2019, San Diego, CA, USA. New York: IEEE Press, 2019: 225-230.
- [11] 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, 2020, 39(6): 1856-1867.
- [12] Wang J D, Sun K, Cheng T, et al. Deep high-resolution representation learning for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(10): 3349-3364.
- [13] Jha D, Smedsrud P H, Riegler M A, et al. Kvasir-SEG: a segmented polyp dataset[M]//Ro Y M, Cheng W H, Kim J, et al. MultiMedia modeling. Lecture notes in computer science. Cham: Springer, 2019, 11962: 451-462.
- [14] Gutman D, Codella N C F, Celebi M E, et al. Skin lesion analysis toward melanoma detection: a challenge at the international symposium on biomedical imaging (ISBI)2016, hosted by the international skin imaging collaboration (ISIC)[EB/OL]. (2016-05-04)[2021-05-04]. <https://arxiv.org/abs/1605.01397>.
- [15] Codella N C F, Gutman D, Celebi M E, et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging

- collaboration (ISIC)[C]//2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), April 4-7, 2018, Washington, DC, USA. New York: IEEE Press, 2018: 168-172.
- [16] Liu S T, Huang D, Wang Y H. Receptive field block net for accurate and fast object detection[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11215: 404-419.
- [17] Wu Z, Su L, Huang Q M. Cascaded partial decoder for fast and accurate salient object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 3902-3911.
- [18] Yun D, Ryu J, Lim J. Dual aggregated feature pyramid network for multi label classification[J]. Pattern Recognition Letters, 2021, 144: 75-81.
- [19] Yu C Q, Gao C X, Wang J B, et al. BiSeNet V2: bilateral network with guided aggregation for real-time semantic segmentation[EB/OL]. (2020-05-05)[2021-05-04]. <https://arxiv.org/abs/2004.02147v1>.
- [20] Huang C H, Wu H Y, Lin Y L, et al. HarDNet-MSEG: a simple encoder-decoder polyp segmentation neural network that achieves over 0.9 dicemean and 86 fps[EB/OL]. (2021-01-18)[2021-05-04]. <https://arxiv.org/abs/2101.07172>.
- [21] Zhang H L, Li Q, Guan X. An improved three-dimensional dual-path brain tumor image segmentation network[J]. Acta Optica Sinica, 2021, 41(3): 0310002.
张恒良, 李镛, 关欣. 一种改进的三维双路径脑肿瘤图像分割网络[J]. 光学学报, 2021, 41(3): 0310002.
- [22] Jha D, Ali S, Tomar N K, et al. Real-time polyp detection, localization and segmentation in colonoscopy using deep learning[EB/OL]. (2020-11-15)[2021-05-04]. <https://arxiv.org/abs/2011.07631>.
- [23] 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), July 28-30, 2020, Rochester, MN, USA. New York: IEEE Press, 2020: 558-564.
- [24] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6230-6239.
- [25] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 833-851.
- [26] Zhang Q, Cui Z P, Niu X G, et al. Image segmentation with pyramid dilated convolution based on ResNet and U-Net[M]//Liu D R, Xie S L, Li Y Q, et al. Neural information processing. Lecture notes in computer science. Cham: Springer, 2017, 10635: 364-372.
- [27] Qin X B, Zhang Z C, Huang C Y, et al. U2-Net: going deeper with nested U-structure for salient object detection[J]. Pattern Recognition, 2020, 106: 107404.