

DOI:10.16136/j.joel.2023.02.0144

一种改进级联 U-Net 网络的结肠息肉分割算法

王龙业¹, 张凯信^{1*}, 曾晓莉², 肖 舒¹, 肖 越¹, 敬 梁¹

(1. 西南石油大学 电气信息学院, 四川 成都 610500; 2. 西藏大学 信息科学技术学院, 西藏 拉萨 850000)

摘要: 结肠镜图像中息肉的精确分割是诊断结肠癌的关键环节, 针对目前结肠息肉分割算法存在孔洞、分割粗糙以及分割不完全的问题, 提出了一种改进级联 U-Net 结构的结肠息肉分割算法。运用特征融合思想, 设计了多尺度语义嵌入模块和残差模块, 充分利用深、浅层特征的语义信息。引入注意力机制, 在模型的级联处构建了改进空洞卷积模块, 扩大卷积感受野并增强特征捕获能力。改进了卷积层模块和分割损失函数, 提升模型的泛化性和鲁棒性。在 Kvasir-SEG 数据集上进行实验分析, 相似系数、平均交并比、召回率和准确率分别达到了 90.39%、88.34%、83.62% 和 95.12%。实验结果表明, 该文所提算法改善了分割图像内部孔洞、边缘粗糙及分割不完全的问题, 优于其他息肉分割算法。

关键词: 结肠息肉; 图像分割; 空洞卷积; 级联 U-Net; 分割损失函数

中图分类号: TP391 文献标识码: A 文章编号: 1005-0086(2023)02-0214-11

A colon polyp segmentation algorithm based on improved cascaded U-Net network

WANG Longye¹, ZHANG Kaixin^{1*}, ZENG Xiaoli², XIAO Shu¹, XIAO Yue¹, JING Liang¹

(1. College of Electrical Information, Southwest Petroleum University, Chengdu, Sichuan 610500, China; 2. College of Information Science and Technology, Tibet University, Lhasa, Xizang 850000, China)

Abstract: Accurate segmentation of polyps in colonoscopy images has become a key aspect in the diagnosis of colon cancer. A colon polyp segmentation algorithm with improved cascade U-Net structure is proposed to address the problems of holes, rough segmentation and incomplete segmentation in the current colon polyp segmentation algorithm. Using the idea of feature fusion, a multi-scale semantic embedding module and a residual module are designed to make full use of the semantic information of deep and shallow features. An attention mechanism is introduced and an improved null convolution module is built at the cascade of the model to expand the convolutional field of perception and enhance feature capture. The convolutional layer module and segmentation loss function are improved to promote the generalization and robustness of the model. The experimental analysis on the Kvasir-SEG dataset achieves 90.39%, 88.34%, 83.62% and 95.12% for similarity coefficient, average intersection ratio, recall and accuracy, respectively. The experimental results show that the proposed algorithm improves the problems of internal holes, rough edges and incomplete segmentation of segmented images and outperforms other polyp segmentation algorithms.

Key words: colon polyps; image segmentation; atrous convolution; cascaded U-Net; segmentation loss function

0 引 言

结直肠癌是临床上死亡率最高的癌症之

一^[1], 其病变过程较长, 最初是结肠息肉引起的良性息肉, 但如果检查和治疗不及时, 随着时间的发展, 良性息肉会逐步发展成腺瘤, 而腺瘤则是癌前

* E-mail: 2839966954@qq.com

收稿日期: 2022-03-09 修订日期: 2022-04-06

基金项目: 国家自然科学基金(61561045)和四川省科技计划项目(2019JDRC0012)资助项目

病变^[2]。因此,早筛查、早诊治结肠息肉是预防结肠癌发生的有效手段。由于传统息肉检测主要依赖临床医师的手动操作,长时间工作会造成极高的漏检率^[3],故结合人工智能的辅助诊断手段已经被广泛关注和研究。

2009年,GROSS等^[4]将多尺度滤波用于息肉分割,虽然可以实现息肉的简单分割,但分割区域粗糙。2012年,BERNAL等^[5]提出了一种扇形堆积深度(sector accumulation depth of valleys accumulation,SA-DOVA)能量图方法,可以实现对息肉区域定位,但误定位率较高。2015年,BERNAL等^[6]为解决误定位问题,采用窗谷堆积中值深度(window median depth of valleys accumulation,WM-DOVA)能量图检测肠道息肉,从而获取息肉边界。

以上方法虽然可以实现息肉的简单分割,但均只考虑部分特征信息,没有综合考虑息肉的纹理形状和背景信息等特征,致使分割结果存在边缘粗糙、空洞等问题。近年随着深度学习技术的发展,结肠息肉分割算法被深入研究。2018年,WANG等^[7]使用SegNet实现息肉的像素级分割,虽分割速度较快,但精度不高。TASHK等^[8]基于U-Net网络进行息肉分割,虽取得较好的分割效果,但网络的特征融合方式过于简单,信息损失量过大。

为解决以上问题,ZHOU等^[9]以U-Net的编码器和解码器为基础,提出了U-Net++网络,增加了跳跃连接数量。U-Net++虽然能提取不同层次的特征并相互整合,减小低级和高级特征之间的差距,但其卷积感受野固定,只能提取较少多尺度信息。2019年,JHA等^[10]提出了ResUNet++网络,虽然通过添加压缩激励模块(squeeze-and-excitation,SE)^[11]和空间金字塔池化(atrous spatial pyramid pooling,ASPP)^[12]实现多尺度的特征融合和特征通道信息的提取,但并未充分利用多尺度原始特征,丢失了特征的空间信息。HUANG等^[13]在U-Net和U-Net++的基础上提出了U-Net3+,它利用了全尺度的跳跃连接和深度监督,将不同尺度的低级和高级语义结合,从多尺度聚合特征中学习深层表示,然而模型存在泛化性不足的问题。

2020年,JHA等^[14]打破了传统单一编码器解码器结构的常见设计思路,基于级联模式的双编

码器和双解码器结构,提出了级联形式的双U型结构模型DoubleU-Net。虽然该网络采取了迁移学习的策略,并使用VGG-19^[15]预训练网络作为编码器提升模型的分割精度和泛化性,但提取上下文空间特征信息较少。2021年,LIU等^[16]提出了改进的MADoubleU-Net,利用空间注意力模块,克服了DoubleU-Net不能对空间维度信息进行相关性计算的问题,并且使用的传统空洞卷积丢失了大量细节。为解决上述问题,2022年,SHI等^[17]提出了利用Kronecker空洞卷积的阶梯型U-Net,弥补了传统空洞卷积丢失的细节特征,但仍存在低级和高级特征融合效率低的问题。

虽然U-Net、U-Net++、ResUNet++、U-Net3+和DoubleU-Net等基础编解码结构的分割方法,可以对高光区域较小、形状规则、边缘清晰的息肉实现较好的分割,但仍存在肠道高光^[18]区域影响、特征融合效率低、模型泛化性不足、信息损失量大、息肉漏分割、分割图像粗糙问题。针对上述问题,本文提出一种改进级联结构U-Net的结肠息肉分割算法,与现有的结肠息肉分割算法相比,主要在以下3个方面进行了创新改进:

1) 数据预处理:利用模糊逻辑和对抗生成网络对息肉图像进行去高光处理,来减弱结肠息肉图像中高光区域对息肉分割结果的影响,并通过数据增强的方式增加数据集的数量,增强网络对细腻特征的学习。

2) 网络模型结构:在编解码部分为了提升网络的表征能力及空间和通道特征提取效率,加入了卷积层重构模块(rep convolution block,Rep-Conv-block)和注意力模块(convolutional block attention module,CBAM)^[19];在多尺度特征融合部分为了减小高级特征与低级特征之间语义的差距,设计了多尺度语义嵌入模块(multi-scale resolution semantic embedding branch block,M-block)和瓶颈残差模块(bottleneck residuals path,Res-Path);在改进空洞卷积部分为了扩大卷积感受野和增强特征捕获能力,设计了多尺度池化空洞卷积模块(multi-scale pooling atrous convolution modules,A-block)。

3) 损失函数:针对结肠息肉数据集中存在数据不平衡问题,改进了分割损失函数,提出了交叉熵损失和对数余弦 L_γ 损失结合的非对称损失函数,提升了模型的泛化能力和鲁棒性。

1 级联结构的 U-Net 图像分割算法

1.1 MACU-Net 网络总体结构

为了解决结肠息肉分割图像中存在的图像内部空洞、边缘粗糙和分割不完全等问题,本文提出了一种多尺度空洞注意力级联网络(multiscale atrous attention cascade u-network, MACU-Net),其主要包括编解码部分、多尺度特征融合部分和改进空洞卷积部分,整体结构如图1所示。

MACU-Net 网络结构主要特点如下:1) 在级联网络的两个 U 型结构桥接处加入 CBAM,提升网络空间和通道特征提取效率;2) 在编码器和解码器中加入了改进的 RepConv-block,增强了网络的表征能力;3) 在网络两个 U 型结构的跳跃连接处分别加入了 M-block 和 ResPath,减小了高级特征与低级特征之间语义的差距,提升了多尺度特征融合性能;4) 在网络的级联处加入了 A-block,扩大了卷积感受野并增强了特征捕获能力。

1.2 编解码部分

在第一个 U 型结构的编码部分,输入了不同尺度的图像来避免网络深层纹理细节丢失的问题;在第二个 U 型结构的解码部分,通过转置卷积将每层的特征图进行上采样,并进行拼接操作,实现了特征图的多尺度融合输出。在两个 U 型结构编码到解码

的桥接处加入了 CBAM 模块,增强了对空间和通道特征提取效率。在 A-block 模块的输出和输入图像做乘积操作(Multiply),丰富了语义信息。整个网络的编解码部分中,除第五层使用 1 组 3×3 卷积操作(conv 3×3)外,其余各层均采用本文提出的 Rep-Conv-block。在网络编码部分的特征图进行了最大池化层(MaxPooling)的下采样操作,并在多尺度特征融合部分和解码部分进行特征融合,然后使用转置卷积进行上采样操作(Upsampling),实现特征图的尺度变换。

虽然分割网络性能会随着深度增加而增加,但是也伴随着网络退化问题,直到残差结构的出现^[20]在一定程度上缓解了网络的退化。借鉴 RepVGG 网络的相关思想^[21],本文提出了 RepConv-block,如图2所示,每个重构模块中采用两组 3×3 卷积操作(convolution, Conv)和残差结构,加上批归一化(batch normalization, BN)、ReLU 激活函数和 Drop-Block。

在模型的初始阶段,使用 1×1 卷积残差结构,如图2(a)和图2(b)所示。为了使网络在深层获得鲁棒性更好的特征表示,在图2(c)和图2(d)中使用了同时含有 1×1 卷积和 Identity 的复杂残差结构。由于残差结构有多个分支,给网络增加了多条梯度流动路径,进一步提升了网络的表征能力。

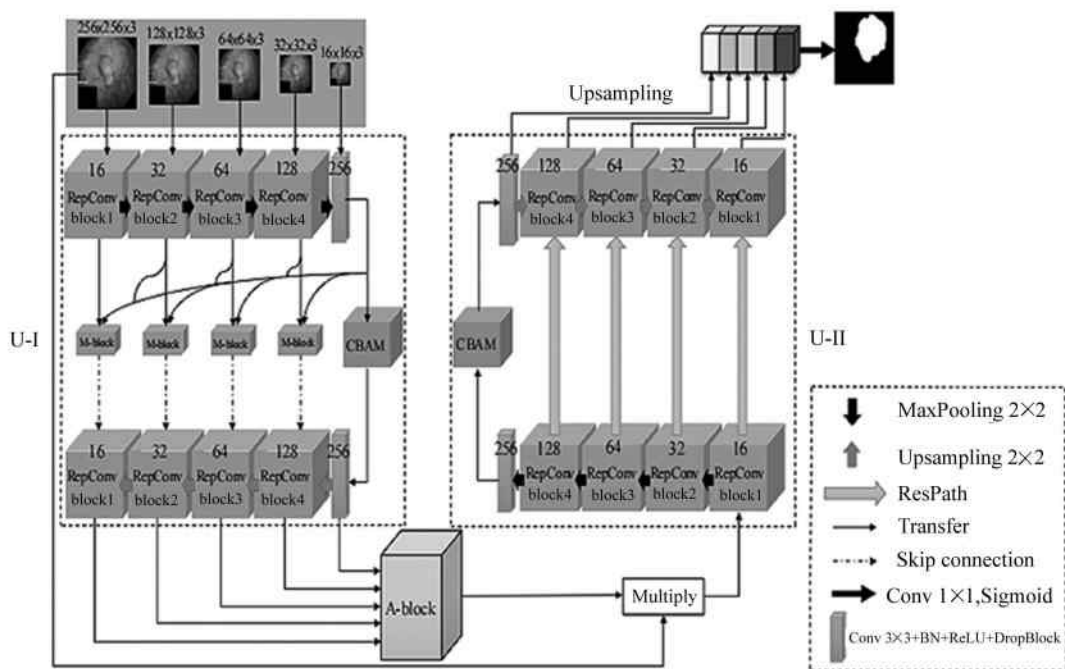


图1 MACU-Net 网络整体结构

Fig. 1 Overall structure of MACU-Net network

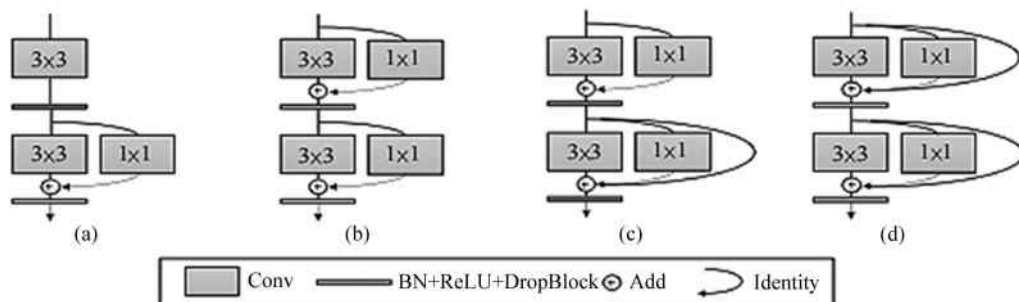


图2 卷积层重构模块结构:(a) 模块1; (b) 模块2; (c) 模块3; (d) 模块4

Fig. 2 Structure of RepConv-block: (a) RepConv-block1; (b) RepConv-block2; (c) RepConv-block3; (d) RepConv-block4

1.3 多尺度特征融合部分

1.3.1 M-block

图像分割网络中融合不同尺度的特征表示是提高网络性能的重要课题^[22]。本文提出了一种 M-block 模块。

M-block 模块结构如图 3(a) 所示, 其将具有高级语义信息的特征图进行一次 3×3 卷积后, 再通过 SE^[11] 模块进行处理得到特征图之间的相关性, 经上采样后与低语义特征图进行相加操作 (addition), 最后得到融合后的特征。

M-block 模块在 MACU-Net 模型中的位置如图 3(b) 所示, 编码部分的特征图经过 M-block 和 CBAM 模块后与解码部分进行拼接操作。多尺度语义信息的特征融合过程表示如式(1)所示:

$$Y_L = y_l + \sum_{i=l}^{L-1} \text{Upsample}(y_{i+1} \cdot SE_{i+1}) \quad (1)$$

式中, $\text{Upsample}(y_{i+1} \cdot SE_{i+1})$ 表示将高级语义特征经过 SE 模块后再进行上采样, Y_L 表示语义融合后

的特征, y_l 表示语义融合前的特征, L 表示参与融合的层数, $L \in [1, 5]$, l 表示具体层数, $l \in [1, 5]$ 。特别注意, Y_1 是经过 CBAM 模块后的特征, $Y_5 - Y_2$ 均为经过 M-block 模块的融合特征。

1.3.2 ResPath

U-Net 系列网络中的跳跃连接将池化操作期间丢失的特征, 由编码器传递到解码器之中, 但这样的连接会造成语义上的差距^[23], 本文提出了一种 ResPath 模块, 如图 4 所示。

在图 4 模块中, 跳跃连接结构上合并了卷积层, 可以减小语义差距。瓶颈部分使用 3×3 卷积, 升维和降维使用 1×1 卷积, 并使用 1×1 卷积进行残差连接。随着网络的变深, 在编码和解码特征之间语义差异逐渐减小的同时, 也会逐渐减少跳跃连接上的瓶颈模块数量。在网络结构中, 沿着第二个 U 型结构的 4 条跳跃连接路径分别使用 4, 3, 2, 1 个瓶颈模块。

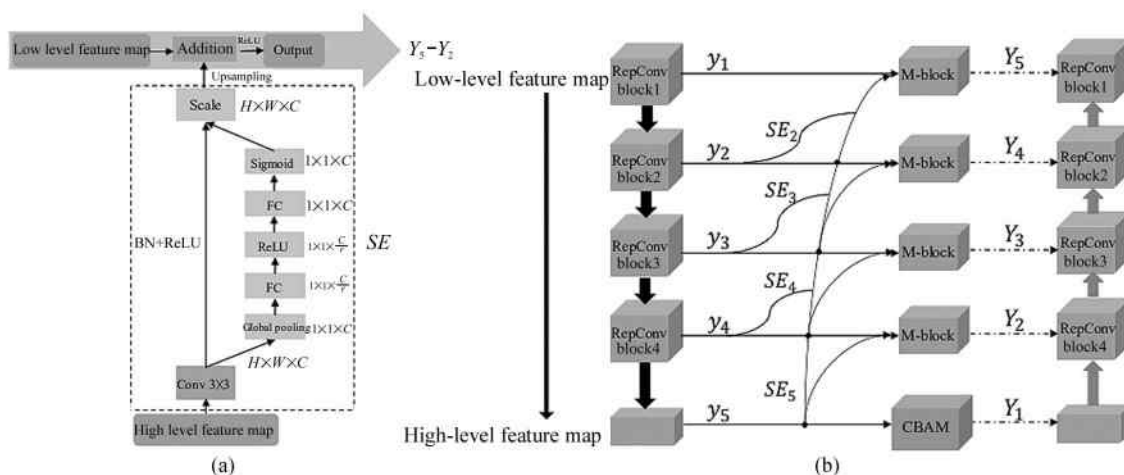


图3 M-block 位置结构图:(a) 结构图; (b) 位置图

Fig. 3 Location structure map of M-block module: (a) Structure map; (b) Location map

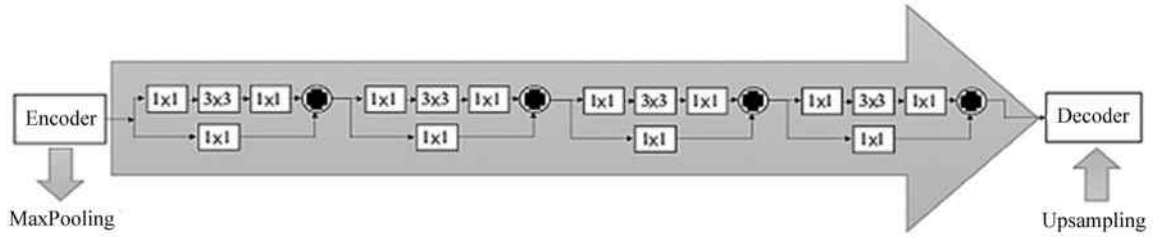


图4 ResPath 模块结构

Fig. 4 Structure of ResPath module

1.4 改进空洞卷积部分

采取密集方式连接空洞卷积(dense atrous spatial pyramid pooling, DenseASPP)^[24],在获取大感受野的同时,多次使用池化层,降低了特征的分辨率,丢失了大量的细节信息。本文提出了一种 A-block 模块,如图 5 所示。

解码部分的输出特征在经过池化整合模块与级联的空洞卷积后,再次与输出特征进行相加操作(Add),将相加后的特征图通过 Upsampling 进行拼接操作(Concatenation),再利用 1×1 卷积和 BN 进行特征降维,最终得到输出特征图(output)。

空洞卷积能够增加感受野,感受野如式(2)所示,两个空洞卷积的叠加会得到更大的感受野,卷积核大小分别为 K_1 和 K_2 ,新的感受野如式(3)所示:

$$R = (d-1) \times (K-1) + K, \quad (2)$$

$$K = K_1 + K_2 - 1. \quad (3)$$

ASPP(3,6,12,18,24)中最大感受野如式(4)所示,本文 A-block(3,6,12,18,24)上的感受野如式(5)所示:

$$R_{\max} = \max[R_{3,3}, R_{3,6}, R_{3,12}, R_{3,18}, R_{3,24}] = R_{3,24} = 51, \quad (4)$$

$$R_{\max} = R_{3,3} + R_{3,6} + R_{3,12} + R_{3,18} + R_{3,24} - 3 = 128, \quad (5)$$

式中, $R_{\max}$ 表示金字塔结构的最大感受野, $R_{k,d}$ 表示卷积核大小 K 和空洞率 d 的感受野,可以看出使用级联方法叠加的空洞卷积具有更大的感受野。

A-block中的多尺度池化模块如图6所示,其中

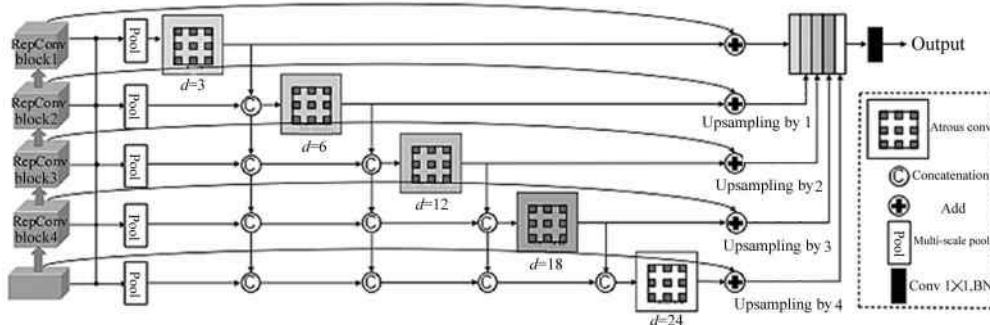


图5 A-block 的结构

Fig. 5 Structure of A-block

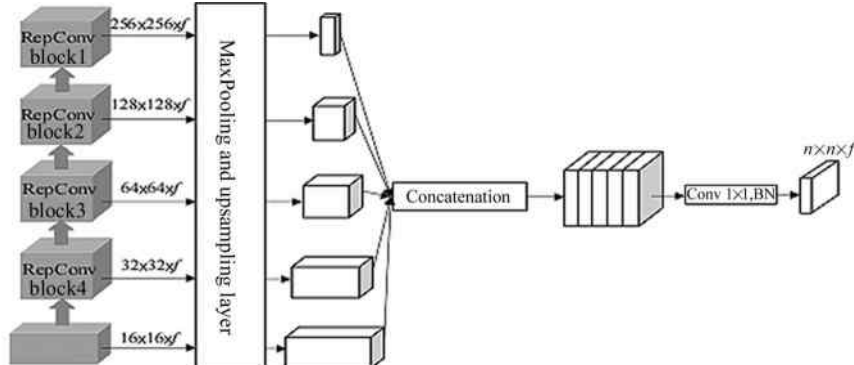


图6 多尺度池化模块结构

Fig. 6 Structure of multi-scale pool module

f 为特征数, n 为输出特征图的尺寸。该模块依靠解码部分的不同尺度特征, 首先进行 Upsampling 和 MaxPooling 对特征信息进行尺度变化, 然后将特征图进行 Concatenation, 最后引入 1×1 卷积和 BN 进行特征降维, 最后得到输出特征图。

2 损失函数

在传统图像分割任务中, 常使用交叉熵损失函数, 虽利用了正确标签预测的信息, 却忽略了非正确预测的信息, 交叉熵损失函数定义如式(6)所示:

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N g_i \log(p_i) + \bar{g}_i \log(\bar{p}_i), \quad (6)$$

式中, N 代表了所有像素值, $p_i \in [0, 1]$ 表示的是像素 i 属于息肉部分的预测概率, $\bar{p}_i$ 表示的是像素 i 属于背景的预测概率, $g_i \in \{0, 1\}$ 表示的是像素 i 属于息肉部分的真实标签值, $\bar{g}_i$ 表示的像素 i 属于背景的真实标签值。

在医学图像分割任务中, 数据的不均衡问题具有一定挑战性, 在结肠息肉的数据集中, 息肉区域往往小于背景区域, 目前的许多工作是通过分配更多权重来惩罚错误预测, 在 Tversky 损失函数^[25]中, 通过分配不同的权重, 来对假阳性和假阴性进行更多惩罚, Tversky 损失函数定义如式(7)所示:

$$L_{Tversky} = 1 - \frac{\sum_{i=1}^N p_i g_i}{\sum_{i=1}^N p_i g_i + \alpha \sum_{i=1}^N \bar{p}_i g_i + \beta \sum_{i=1}^N p_i \bar{g}_i}. \quad (7)$$

在 Tversky 损失函数中, L_γ 损失函数^[26]是一个特例, 表达式如式(8)所示:

$$L_\gamma = 1 - \frac{\gamma \sum_{i=1}^N p_i g_i}{\gamma \sum_{i=1}^N p_i g_i + (\gamma - 1) \sum_{i=1}^N \bar{p}_i g_i + \sum_{i=1}^N p_i \bar{g}_i}, \quad (8)$$

式中, $\gamma = 1 + \beta^2$, 仅具有超参数 β 的非对称 L_γ 损失函数可以推广为相似性系数和交并比指标。为了防止分母为零, 在 L_γ 第二项分子分母同时加上了一个极小的常数 S , 将 S 设置为 10^{-8} 。改进后的损失函数 L_S 如式(9)所示:

$$L_S = 1 -$$

$$\frac{\sum_{i=1}^N p_i g_i + S}{\sum_{i=1}^N p_i g_i + \frac{(\gamma - 1)}{\gamma} \sum_{i=1}^N \bar{p}_i g_i + \frac{1}{\gamma} \sum_{i=1}^N p_i \bar{g}_i + S}, \quad (9)$$

式中, $\frac{(\gamma - 1)}{\gamma}$ 和 $\frac{1}{\gamma}$ 分别用于控制假阳性和假阴性的权重。

由于 Tversky 系数的非凸性, 会导致一定程度的次优梯度, 所以本文在 L_γ 损失函数前加上了对数双曲余弦函数, 利用对数双曲余弦函数充当了凸代理, 在一定程度上减少了非凸性, 使曲线变得更加平滑, 从而优化了梯度。改进的 $L_{\logcosh}$ 损失函数如式(10)所示:

$$L_{\logcosh} = \log(\cosh(L_S)). \quad (10)$$

本文提出了一种非对称的混合损失函数, 利用了非对称对数余弦 Tversky 系数来处理不均衡问题, 同时考虑了交叉熵的优点, 将不对称对数余弦 $L_{\logcosh}$ 损失函数和交叉熵相结合, 如式(11)所示:

$$L_M = \alpha L_{CE} + (1 - \alpha)(L_{\logcosh}), \quad (11)$$

式中, α 和 $1 - \alpha$ 分别用于控制交叉熵损失和 $L_{\logcosh}$ 损失对混合损失函数的贡献程度。

3 数据集预处理

结肠镜检查过程中, 人体腹腔存在大量的黏液会造成图像的反光区域。在模型的训练过程中, 反光区域会影响对息肉分割的精度, 因此在训练之前对图像进行去反光处理是有必要的。传统去高光区域的方法是将超过某一阈值的像素点直接划分成高光区域, 但由于高光区域是渐变的, 阈值去高光方法会造成截断式划分的问题。

本文选取了模糊数学^[27]中的隶属度函数来划分高光区域, 在完成高光区域的划分之后, 将生成的高光区域划分及对应的二值掩码送入生成对抗模型进行去高光处理。本文引用的模糊逻辑方法和其他滤波方法相比, 能更好地保留图像细节, 更有利于分割模型的训练, 预处理前后的图像如图 7 所示。

为了减少因数据集数量不足导致的过拟合现象, 使模型学习到更多的特征信息。本文将训练集部分进行去高光处理之后, 再将图像和对应的掩模执行相同的数据增强, 增强之后的数据集为 8000 张。

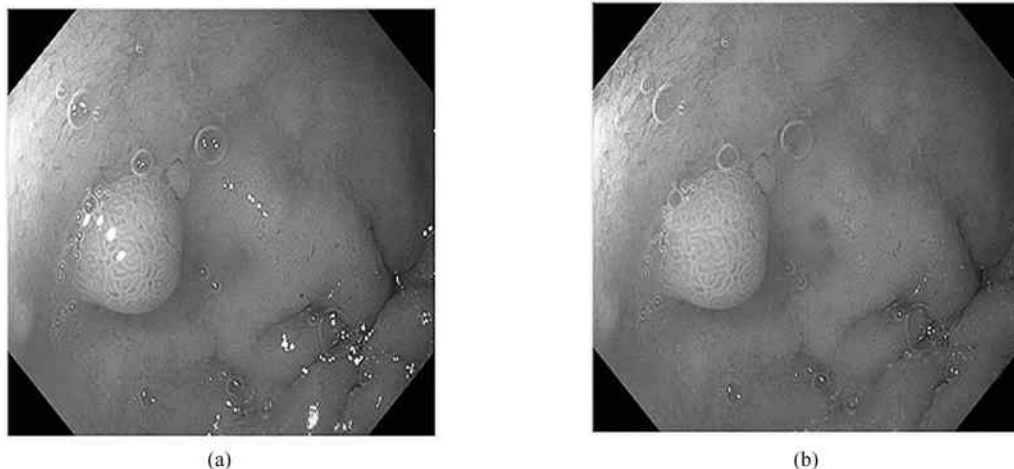


图 7 图像去高光对比:(a) 原始图像;(b) 处理后图像

Fig. 7 Comparison of image removed highlight:(a) Original image; (b) Processed image

4 实验结果与分析

4.1 数据集划分

为了综合评估分割模型的性能,本实验使用了 Kvasir-SEG^[28] 数据集来进行训练和测试,将

MACU - Net 模型在 CVC - ClinicDB^[29]、CVC - ColonDB^[29] 和 ETIS-LaribPolypDB 数据集上进行泛化实验来测试模型的通用性及鲁棒性,表 1 展示了以上数据集的具体信息。将 Kvasir-SEG 数据集中的图像按照 8 : 1 : 1 的比例划分为训练集、验证集

表 1 数据集基本信息

Tab. 1 Basic information about the dataset

Modality	Dataset	No. of images	Original resolution	Input resolution
Endoscopy	Kvasir-SEG	1 000	Variable	256 × 256
Endoscopy	CVC-ClinicDB	612	384 × 288	256 × 256
Endoscopy	CVC-ColonDB	380	574 × 500	256 × 256
Endoscopy	ETIS-LaribPolypDB	196	1 225 × 966	256 × 256

和测试集 3 部分。

4.2 参数设置和评价指标

实验部分,所有的模型都是在 Tensorflow-gpu-2.4.1 框架,AMD Ryzen 5 3600 处理器,内存 16 GB,NVIDIA GeForce RTX 3060,显存 12 GB 的设备上训练了 100 个 epoch。在训练过程中,模型使用了 Adam 优化器优化模型参数,损失函数选择了本文提出的非对称混合损失函数,初始学习率设为 1×10^{-3} ,batch size 设置为 8。

为了验证所提分割模型的有效性,实验使用了在医学图像分割中常用的评价指标,精确率(Precision)、召回率(Recall)、相似系数(dice coefficient, Dice)、前景交并比(IoU)、背景交并比(IoU_B)、平均交并比(mIoU)和准确率(Accuracy),具体定义如式(12)–(18)所示:

$$Precision = \frac{TP}{TP + FP}, \quad (12)$$

$$Recall = \frac{TP}{TP + FN}, \quad (13)$$

$$Dice = \frac{2TP}{FP + 2TP + FN}, \quad (14)$$

$$IoU = \frac{TP}{FP + TP + FN}, \quad (15)$$

$$IoU_B = \frac{TN}{FP + TN + FN}, \quad (16)$$

$$mIoU = \frac{IoU + IoU_B}{2}, \quad (17)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}, \quad (18)$$

式中,TP 表示标签图像和预测图像的息肉区域的像素数量,FP 表示在分割预测区域内而在标签分割区域外的像素数量,TN 表示标签图像和预测图像的非

息肉公共区域的像素数量, FN 表示在分割预测区域外在标签分割区域内的像素数量。

4.3 结果与分析

4.3.1 损失函数对实验结果影响

由统计可知, 本文用于训练和测试的结肠息肉数据集的正负样本, 存在严重失衡的问题, 非对称混合损失函数对于 FN 和 FP 的检测权重是不同的, 训练时可以根据训练数据的实际情况改变对应的权重, 从而减少数据不均衡给分割结果带来的不良影响。本文测试并比较了交叉熵损失函数(见式(6))、Tversky 损失函数(见式(7))、logcosh 损失函数(见式(10))和非对称混合损失函数(见式(11))的性能。

本文的 MACU-Net 模型在 4 种损失函数上的验证损失曲线如图 8 所示。从图中可以看出: 非对称混合损失函数和其他 3 种损失函数相比, 损失收敛速度更快, 训练过程更加稳定且损失值最小, 迭代收敛到接近 0.08。

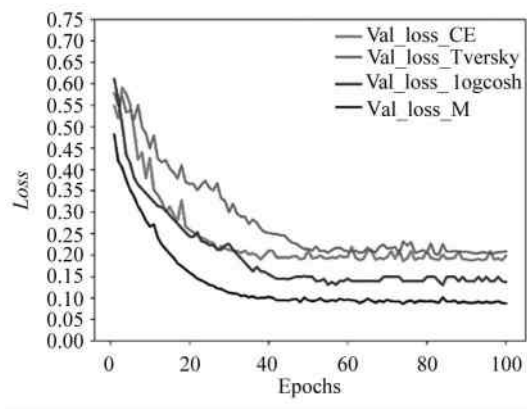


图 8 损失曲线

Fig. 8 Curve of loss

4 种损失函数对分割结果的性能比较如表 2 所示, 从表中可看出, 本文的非对称混合损失函数, 对分割模型的性能有一定提升, 其中和 Tversky 损失函数相比提升最为明显, $mIoU$ 、 $Dice$ 和 $Precision$ 分

表 2 损失函数性能对比

Tab. 2 Comparison of loss function performance

Loss function	$Dice$	IoU	IoU_B	$mIoU$	$Recall$	$Precision$
L_{CE}	0.8998	0.807 1	0.950 3	0.878 7	0.832 8	0.940 8
$L_{Tversky}$	0.881 3	0.797 3	0.941 2	0.869 3	0.837 3	0.932 5
$L_{logcosh}$	0.887 1	0.801 2	0.948 6	0.874 9	0.834 6	0.938 1
L_M (ours)	0.903 9	0.810 8	0.956 0	0.883 4	0.836 2	0.941 7

别上升了 1.41%、2.26% 和 0.92%。

4.3.2 性能比较与分析

MACU-Net 模型的准确率在训练和验证集上的表现如图 9 所示, 训练和验证精度在训练开始一段时间后便迅速改善, 且随着迭代次数的增加, 改善的速度逐渐减慢。经过 100 次迭代后, 模型性能趋于稳定, 最终准确率达到 95.12%。

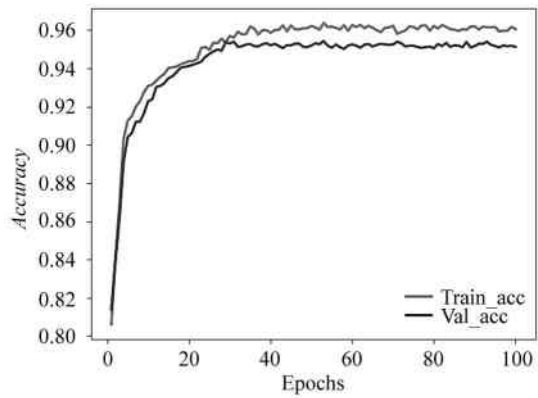


图 9 准确率曲线

Fig. 9 Curve of accuracy

为了排除模型中各个单独组件的相互影响, 通过消融实验对 CBAM、M-block、A-block、ResPath 和 RepConv-block 进行了测试。消融实验从只含有两个基础 U 型结构开始, 逐渐过渡到 MACU-Net 模型, 如表 3 所示。

表 3 消融实验

Tab. 3 Ablation experiments

Number	Experimental description
1	Basic cascade model
2	1+ De-reflection dataset
3	2+CBAM model
4	3+M-block model
5	4+ResPath model
6	5+A-block model
7	6+RepConv-block model
8	MACU-Net

根据表 3 中的设置, 分别构建出相应的分割模型, 所有实验均在同一参数设置下对模型进行训练, 消融实验在 Kvasir-SEG 数据集上的结果如表 4 所示。

表 4 消融实验结果对比

Tab. 4 Comparison of ablation experimental results

Number	<i>Dice</i>	<i>IoU</i>	<i>IoU_B</i>	<i>mIoU</i>	<i>Recall</i>	<i>Precision</i>
1	0.868 8	0.780 1	0.939 3	0.859 7	0.818 8	0.900 8
2	0.869 1	0.780 5	0.939 8	0.860 1	0.819 2	0.901 2
3	0.873 2	0.788 3	0.942 4	0.865 3	0.820 8	0.921 3
4	0.886 8	0.799 2	0.943 8	0.871 5	0.823 4	0.928 7
5	0.890 3	0.801 1	0.949 7	0.875 4	0.830 7	0.908 5
6	0.893 6	0.805 4	0.952 1	0.878 7	0.829 6	0.931 5
7	0.899 4	0.808 2	0.950 5	0.879 3	0.831 1	0.937 8
8	0.903 9	0.810 8	0.956 0	0.883 4	0.836 2	0.941 7

分析消融实验结果可知,实验 4 在实验 3 的基础上加入了 M-block, *Dice* 和 *IoU* 指标分别上升了 1.36% 和 1.09%; 实验 5 在实验 4 的基础上增加了 ResPath, *Dice* 和 *IoU* 指标分别上升了 0.35% 和 0.19%; 实验 6 在实验 5 的基础上加入了 A-block, *Dice* 和 *IoU* 指标分别上升了 0.33% 和 0.43%; 实验 7 在实验 6 的基础上加入 RepConv-block, 使 *Dice* 和 *IoU* 指标分别上升了 0.58% 和 0.28%。

根据上述参数的设置,分别与 FCN、U-Net、U-Net++、ResU-Net++、DeepLabv3+、DoubleU-Net、MACU-Net(ours) 网络进行了对比测试,实验结果见表 5。显然,MACU-Net 模型在 *Dice*、*mIoU* 和 *Recall* 指标上分别达到了 90.39%、88.34% 和 83.62%, 与基础 U-Net 网络相比分别提升了 34.28%、31.27% 和 21.26%, 与 DoubleU-Net 网络相比分别提升了 1.71%、0.77% 和 1.45%。

表 5 不同分割模型在 Kvasir-SEG 数据集上的结果对比

Tab. 5 Comparison of segmentation results of different models in Kvasir-SEG dataset

Method	<i>Dice</i>	<i>IoU</i>	<i>IoU_B</i>	<i>mIoU</i>	<i>Recall</i>	<i>Precision</i>
FCN	0.554 1	0.394 6	0.613 5	0.504 0	0.591 4	0.729 2
U-Net	0.561 1	0.433 4	0.708 1	0.570 7	0.623 6	0.921 2
U-Net++	0.667 3	0.640 3	0.937 2	0.788 7	0.690 1	0.871 6
ResU-Net++	0.838 8	0.574 5	0.931 9	0.753 2	0.706 2	0.877 3
DeepLabv3+	0.690 9	0.751 1	0.942 5	0.846 8	0.835 8	0.917 5
DoubleU-Net	0.886 8	0.800 8	0.950 7	0.875 7	0.821 7	0.944 8
MACU-Net	0.903 9	0.810 8	0.956 0	0.883 4	0.836 2	0.941 7

4.3.3 分割结果与分析

本文提出的 MACU-Net 模型与其他分割模型

在 Kvasir-SEG 数据集上分割结果的对比如图 10 所示。从示例 3 中可以看出模型学习到更高级的更加

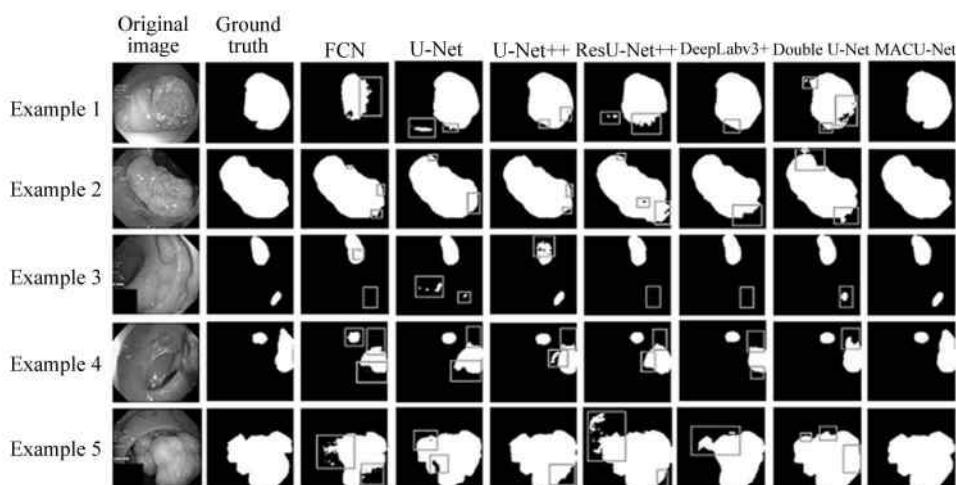


图 10 MACU-Net 与其他模型分割对比图

Fig. 10 Comparison chart of MACU-Net and other models segmentation

细腻的特征,示例5显示了 MACU-Net 强大的分割效果,在边界部分和标签更加接近,没有存在分割图像内部空洞的现象,细节部分也更加细腻。

此外,通过对 Kvasir-SEG、CVC-ClinicDB、CVC-ColonDB 和 ETIS-LaribPolypDB 4 个数据集进行分

割对比实验来测试 MACU-Net 模型的泛化能力,分割结果如图 11 所示。由图可见分割的图像边界清晰连贯,证明了 MACU-Net 模型具有良好的泛化性与鲁棒性,在未参与训练的数据集上同样表现出了优异的性能。

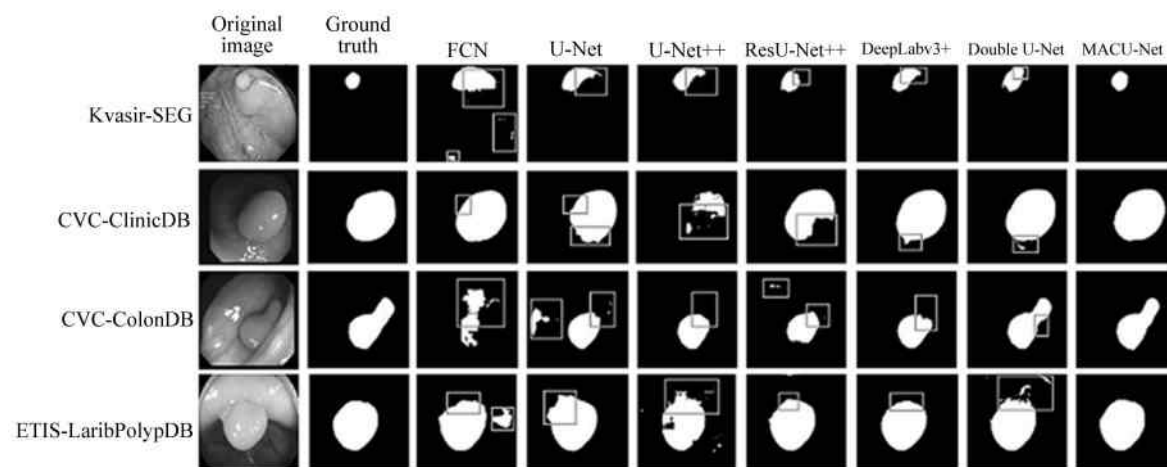


图 11 MACU-Net 在不同数据集上分割对比图

Fig. 11 Comparison of MACU-Net segmentation on different dataset

5 结 论

本文提出一种双编解码模式的结肠息肉分割算法。为了使模型训练更加稳定,在编码器和解码器中加入 RepConv-block 并设计非对称混合损失函数。为减小特征之间语义的差距,基于特征融合思想,在跳跃连接处加入 M-block 和残差跳跃连接模块。为增强模型对特征的提取能力,在 U 型结构之间加入 A-block。实验结果表明,无论从主观的视觉评价还是客观的指标评价,该算法能实现结肠息肉的有效分割,且分割性能优于其他同类算法。然而,本文算法仍存在网络参数过多、训练时间较长等缺陷,需继续优化,减少模型参数,提升其实用性。

参考文献:

- [1] TORRE L A, BRAY F, SIEGEL R L, et al. Global cancer statistics, 2012[J]. CA: A Cancer Journal for Clinicians, 2015, 65(2): 87-108.
- [2] GSCHWANTLER M, KRIWANNEK S, LANGNER E, et al. High-grade dysplasia and invasive carcinoma in colorectal adenomas: a multivariate analysis of the impact of adenoma and patient characteristics[J]. European Journal of Gastroenterology & Hepatology, 2002, 14(2): 183-188.
- [3] LEUFKENS A M, VAN OIJEN M G H, VLEGGAAR F P, et al. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study[J]. Endoscopy: Journal for Clinical Use Biopsy and Technique, 2012, 44(5): 470-475.
- [4] GROSS S, KENNEL M, STEHLE T, et al. Polyp segmentation in NBI colonoscopy [C]//Bildverarbeitung für die medizin, March 22-25, 2009, Heidelberg, Germany. Berlin: Springer, 2009: 252-256.
- [5] BERNAL J, J SÁNCHEZ, VILARIO F. Towards automatic polyp detection with a polyp appearance model[J]. Pattern Recognition, 2012, 45(9): 3166-3182.
- [6] 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.
- [7] WANG L, QIAN Y, HU Y. IDDF2018-ABS-0259 segmentation of intestinal polyps via a deep learning algorithm [C]//International Digestive Disease Forum (IDDF), June 9-10, 2018, Hong Kong, China. London: BMJ Publishing Group, 2018:83-84.
- [8] TASHK A, HERP J, NADIMI E. Fully automatic polyp detection based on a novel U-Net architecture and morphological post-process[C]//2019 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO), December 8-10, 2019, Athens, Greece. New York: IEEE, 2019:37-41.
- [9] ZHOU Z W, RAHMAN SIDDIQUEE M M, 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 Deci-

- sion Support, September 20-22, 2018, Granada, Spain. Berlin: Springer, 2018;3-11.
- [10] JHA D, SMEDSRUD P H, RIEGLER M A, et al. ResUNet++: An advanced architecture for medical image segmentation[C]//2019 International Symposium on Multimedia (ISM), December 9-11, 2019, San Diego, CA, USA. New York: IEEE, 2019;225-231.
- [11] JIE H, LI S, SAMUEL A, et al. Squeeze-and-excitation networks. [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8):2011-2023.
- [12] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-06-17) [2022-03-09]. <https://arxiv.org/abs/1706.05587>.
- [13] HUANG H, LIN L, TONG R, et al. UNet 3+: A full-scale connected unet for medical image segmentation[C]//2020 International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 4-8, 2020, Barcelona, Spain. New York: IEEE, 2020;1055-1059.
- [14] JHA D, RIEGLER M A, JOHANSEN D, et al. DoubleUNet: A deep convolutional neural network for medical image segmentation[C]//33rd International symposium on computer-based medical systems (CBMS), July 28-30, 2020, Rochester, MN, USA. New York: IEEE, 2020; 558-564.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04) [2022-03-09]. <https://arxiv.org/abs/1409.1556>.
- [16] LIU J W, LIU Q H, LI X O, et al. Improved colonic polyp segmentation method based on double U-shaped network [J]. Acta Optica Sinica, 2021, 41(18):1810001.
刘佳伟, 刘巧红, 李晓欧, 等. 一种改进的双 U 型网络的结肠息肉分割方法[J]. 光学学报, 2021, 41(18): 1810001.
- [17] SHI Y G, LI Y, ZHOU Z G, et al. Polyp segmentation using stair-structured U-Net[J]. Journal of Electronics and Information Technology, 2022, 44(1): 39-47.
时永刚, 李伟, 周治国, 等. 基于阶梯结构的 U-Net 结肠息肉分割算法[J]. 电子与信息学报, 2022, 44(1): 39-47.
- [18] SASMAL P, IWAHORI Y, BHUYAN M K, et al. Active contour segmentation of polyps in capsule endoscopic images[C]//2018 International Conference on Signals and Systems (ICSigSys), May 1-3, 2018, Bali, Indonesia. New York: IEEE, 2018; 201-204.
- [19] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the 15th European Conference on Computer Vision (ECCV), September 8-14, 2018, Munich, Germany. Berlin: Springer-Verlag, 2018;3-19.
- [20] SHANG W, SOHN K, ALMEIDA D, et al. Understanding and improving convolutional neural networks via concatenated rectified linear units[C]//33rd International Conference on Machine Learning, June 20-22, 2016, New York City, USA. San Diego, CA: PMLR, 2016: 2217-2225.
- [21] DING X, ZHANG X, MA N, et al. RepVGG: Making VGG-style ConvNets great again[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 20-25, 2021, Nashville, TN, USA. New York: IEEE, 2021: 13733-13742.
- [22] ZHANG Z, ZHANG X, PENG C, et al. ExFuse: Enhancing feature fusion for semantic segmentation[C]//Proceedings of the 15th European conference on computer vision (ECCV), September 8-14, 2018, Munich, Germany. Berlin: Springer-Verlag, 2018; 269-284.
- [23] IBTEHAZ N, RAHMAN M S. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation[J]. Neural Networks, 2020, 121: 74-87.
- [24] YANG M, YU K, ZHANG C, et al. DenseASPP for semantic segmentation in street scenes[C]//Proceedings of 2018 IEEE Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: 2018: 3684-3692.
- [25] SALEHI S S M, ERDOGMUS D, GHOLIPOUR A. Tversky loss function for image segmentation using 3D fully convolutional deep networks[C]//8th International Workshop on Machine Learning in Medical Imaging, September 10-14, 2017, Quebec City, Canada. Berlin: Springer, 2017;379-387.
- [26] HASHEMI S R, SALEHI S S M, ERDOGMUS D, et al. Asymmetric loss functions and deep densely-connected networks for highly-imbalanced medical image segmentation: Application to multiple sclerosis lesion detection [J]. IEEE Access, 2018, 7: 1721-1735.
- [27] GUO J F, LI X, PANG Z Q, et al. Research on custom fuzzy logic and generative adversarial networks in image highlight processing [J]. Journal of Chinese Computer Systems, 2021, 42(8): 1715-1719.
郭继峰, 李星, 庞志奇, 等. 自定义模糊逻辑与 GAN 在图像高光处理中的研究[J]. 小型微型计算机系统, 2021, 42(8):1715-1719.
- [28] JHA D, SMEDSRUD P H, RIEGLER M A, et al. Kvasirseg: A segmented polyp dataset[C]//26th International Conference on Multimedia Modeling, January 5-8, 2020, Daejeon, Korea (South). Berlin: Springer, 2020; 451-462.
- [29] VÁZQUEZ D, BERNAL J, SÁNCHEZ F J, et al. A benchmark for endoluminal scene segmentation of colonoscopy images[J]. Journal of Healthcare Engineering, 2017, 2017:4037190.

作者简介:

王龙业 (1976—), 男, 博士, 教授, 硕士生导师, 主要从事计算机视觉、通信信号设计等研究。