

DOI:10.16136/j.joel.2022.04.0510

基于双支路特征融合的 MRI 颅脑肿瘤图像分割研究

熊 炜^{1,2*}, 周 蕾¹, 乐 玲¹, 张 开¹, 李利荣¹

(1. 湖北工业大学 电气与电子工程学院, 湖北 武汉 430068; 2. 美国南卡罗来纳大学计算机科学与工程系, 南卡罗来纳 29201)

摘要: 针对磁共振成像 (magnetic resonance imaging, MRI) 颅脑肿瘤区域误识别与分割网络空间信息丢失问题, 提出一种基于双支路特征融合的 MRI 脑肿瘤图像分割方法。首先通过主支路的重构 VGG 与注意力模型 (re-parameterization visual geometry group and attention model, RVAM) 提取网络的上下文信息, 然后使用可变形卷积与金字塔池化模型 (deformable convolution and pyramid pooling model, DCPM) 在副支路获取丰富的空间信息, 之后使用特征融合模块对两支路的特征信息进行融合。最后引入注意力模型, 在上采样过程中加强分割目标在解码时的权重。提出的方法在 Kaggle_3m 数据集和 BraTS2019 数据集上进行了实验验证, 实验结果表明该方法具有良好的脑肿瘤分割性能, 其中在 Kaggle_3m 上, Dice 相似系数、杰卡德系数分别达到了 91.45% 和 85.19%。

关键词: 磁共振成像 (magnetic resonance imaging, MRI) 颅脑肿瘤图像分割; 双支路特征融合; 重构 VGG 与注意力模型 (re-parameterization visual geometry group and attention model, RVAM); 可变形卷积与金字塔池化模型 (deformable convolution and pyramid pooling model, DCPM)

中图分类号: TP391 文献标识码: A 文章编号: 1005-0086(2022)04-0383-10

Research on MRI brain tumor image segmentation based on dual-branch feature fusion

XIONG Wei^{1,2*}, ZHOU Lei¹, YUE Ling¹, ZHANG Kai¹, LI Lirong¹

(1. School of Electrical & Electronic Engineering, Hubei University of Technology, Wuhan, Hubei 430068, China; 2. Department of Computer Science & Engineering, University of South Carolina, Columbia, SC 29201, USA)

Abstract: To address the problem of MRI brain tumor region misidentification and spatial information loss of segmentation network, an MRI brain tumor image segmentation method based on dual-branch feature fusion is proposed. First, the contextual information of the network is extracted by structurally the re-parameterization visual geometry group and attention model (RVAM) in the primary branch, and then the rich spatial information is obtained in the secondary branch using deformable convolution and pyramid pooling model (DCPM), after which the feature fusion module is used to fuse the feature information of the two branches. Finally, the attention model is introduced to strengthen the weight of segmented targets in the up-sampling process at decoding. The proposed method has been experimentally validated on the Kaggle_3m and BraTS2019 datasets, and the experimental results show that our method has good brain tumor segmentation performance, where the Dice similarity coefficient and Jaccard coefficient reach 91.45% and 85.19% on Kaggle_3m, respectively.

Key words: magnetic resonance imaging (MRI) brain tumor image segmentation; dual-branch feature fusion; re-parameterization VGG and attention model (RVAM); deformable convolution and pyramid pooling model (DCPM)

* E-mail: xw@mail.hbut.edu.cn

收稿日期: 2021-07-22 修订日期: 2021-08-27

基金项目: 国家自然科学基金 (61571182, 61601177)、国家留学基金 (201808420418)、湖北省自然科学基金 (2019CFB530) 和湖北省科技厅重大专项 (2019ZYYD020) 资助项目

1 引言

脑部肿瘤又称颅内肿瘤,分为良性与恶性2种类型,常用的检测方式为磁共振成像(magnetic resonance imaging, MRI)^[1],并且主要的治疗方法为手术治疗,对肿瘤进行切除。但仅依靠基于人工的传统方法对切片进行检测与术前对肿瘤的形态分析,效率较低。随着计算机视觉领域的发展,计算机辅助诊断系统(computer aided diagnosis system, CADs)^[2]逐渐完善,不仅降低了医生的阅片负担,同时也提高了肿瘤的诊断效率与准确度。但由于检测个体自身的差异性、脑瘤形状复杂多变、边缘界线模糊和肿瘤出现位置不定等困难,脑部肿瘤自动分割任务难度依旧很大^[3]。

常用的图像分割结构为编码器-解码器结构。经典的分割网络如 U-Net^[4],其通过添加跳跃连接,将低层卷积获取的前景边缘信息辅助解码器进行上采样,从而进行图像分辨率恢复。Seg-Net^[5],其通过复用编码器中的最大池化索引,辅助非线性上采样操作。为在语义分割任务中,加速算法的计算速率并同时重视空间信息,YU等^[6]设计了 BiSeNet 网络,通过使用小步长的空间路径保留空间位置信息,并通过拥有快速下采样率的语义路径获取高级语义信息,其在自然场景下的语义分割任务表现良好。JIN等^[7]提出的 DU-Net 在 U-Net 的基础上引入了可变形卷积^[8]模块,在模块中增加额外偏移量以提高网络的特征提取能力,在视网膜血管的分割任务中,增强了空间信息的提取能力。LALONDE等^[9]提出的卷积-反卷积胶囊网络 SegCaps,通过引入局部约束和变换矩阵权值共享,在肺部 CT 分割中,取得了 SOTA 的分割效果。为获取更多的空间特征,MENG等^[10]

将卷积神经网络(convolution neural network, CNN)、注意力优化模型(attention refinement module, ARM)和图卷积网络(graph convolution network, GCN)^[11]进行组合,注重目标的轮廓信息,从而在图像中获取丰富的低层特征与高层语义信息,在超声胎儿头部(fetal head, FH)与彩色眼底图像分割中,取得了最优的效果。

由于 BiSeNet 网络能够高速率和高精度地处理自然场景中的分割任务,故将其应用在颅脑图像分割中,但网络在提取脑部肿瘤时,仍存在空间信息丢失与误识别问题,为此,本文根据 BiSeNetv1 网络进行改进,采取双支路进行目标特征提取,其中,重构 VGG 与注意力模型(re-parameterization visual geometry group and attention model, RVAM)为主支路通过 RepVGG(re-parameterization VGG)^[12]与 ARM 模型提取高级语义信息,可变形卷积与金字塔池化模型(deformable convolution and pyramid pooling model, DCPM)为副支路通过可变形卷积^[13]与空洞空间卷积金字塔池化(atrous spatial pyramid pooling, ASPP)模型^[14]获取丰富的空间信息,并使用特征融合模块(feature fusion model, FFM)^[6]对上述信息进行特征融合。最后设计了一个注意力模型作为辅助网络,在上采样过程中加强分割目标在解码时的权重,提高脑部肿瘤的分割精度。

2 本文方法介绍

2.1 总体框架

本文分割颅内肿瘤的总框架如图1所示,首先在主支路使用 RepVGG 作为 backbone 进行特征提取,并将其输出的最后两层分别送入 ARM 中,使用注意力机制对其进行目标权重优化,完善高级语

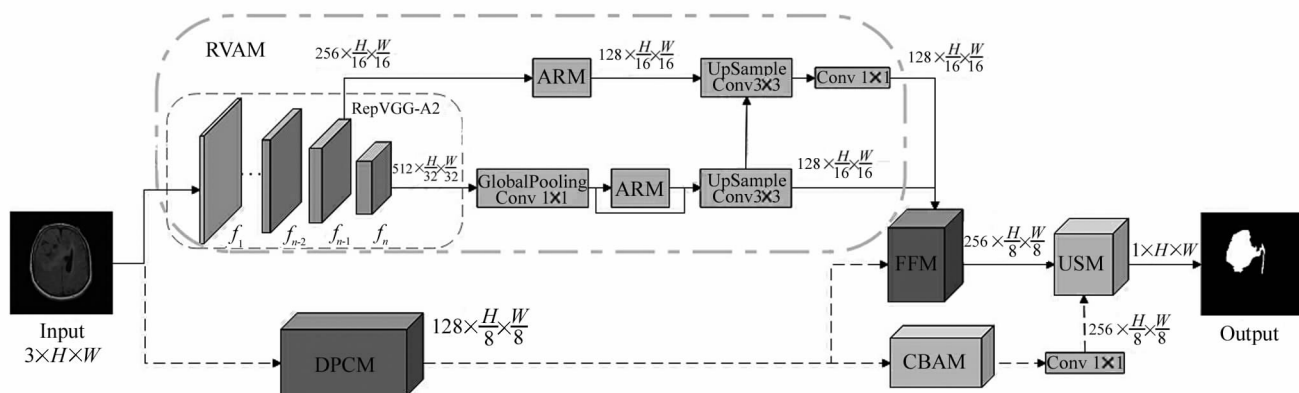


图1 本文总体框架

Fig. 1 Overall framework

义特征的提取。然后在副支路使用提出的 DCPM, 丰富网络模型对图像空间信息的提取, 解决边缘和轮廓等细节信息丢失问题。然后将两者输出送入特征融合模块, 融合两个支路提取到的语义特征。最后使用解码模型(upsample module, USM)将通过神经网络提取的分割特征图分辨率恢复为原始图像大小, 完成对脑部肿瘤的分割任务。同时, 本文利用卷积注意力机制(convolutional block attention module, CBAM)^[15], 将 DCPM 提取的空间信息送入上采样模型中, 加强肿瘤这一目标特征在解码网络中的权重, 进一步提升训练模型的效果, 提高分割的准确率。

2.2 主编码器结构

最近, RepVGG 网络结构仅仅依靠在 VGG 结构上重参数化, 就在图像分类和语义分割领域取得了 SOTA 效果^[12]。故本文拟使用该网络作为肿瘤分割的 backbone, 并将其最后 2 层的输出分别送入 ARM, 提取高级语义信息。RepVGG 网络结构如图 2 所示(图中只详细画出第三组内的网络结构), 由

22 或 28 个 RepVGG Blocks 组成, 以下以 22 个卷积块进行模型描述。模型由 5 组 Blocks 组成, 5 组内的 Layers(Blocks)依次为[1, 2, 4, 14, 1]。设置改变通道数的宽度乘数 a 和 b , 并且为在最后一组内的 layer 中获取丰富的特征信息, 设置 $>$ 。并通过均匀缩放通道数[64, 128, 256, 512]来构成每层的输入、输出通道。其中, 前 4 组 Blocks 使用参数 a , 第 5 组使用参数 b , 若直接在分辨率大的原始特征图上进行大幅卷积(即: 当 $a > 1$ 时, 通道为 3 的特征图直接输出通道为 $64a$ 的特征图), 会造成图像空间信息损失过大等问题^[12], 故设置第一组的输出通道为 $\min(64, 64a)$, 其余 4 组设置为 $[64a, 128a, 256a, 512b]$ 。在每个 layer 中, 设 x 为输入, $f(x)$ 为 Conv3×3 支路的输出, $g(x)$ 为 1×1 卷积实现的 shortcut, 则输出可表示为:

$$y = x + g(x) + f(x)。$$
 (1)

特别的是, 当 x 与 $f(x)$ 的特征维度不匹配时, $y = g(x) + f(x)$ 。故 RepVGG Block 在每个卷积层都加上一个或两个并行分支, 使得多路分支并行融合,

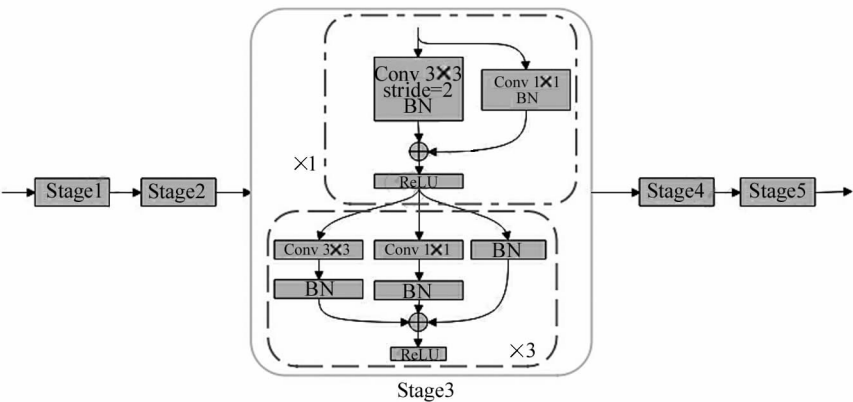


图 2 RepVGG 网络结构
Fig. 2 RepVGG network structure

提升模型的泛化性能和表征能力。

分别将通过 RepVGG 模型训练出的最后 2 层特征张量输入 ARM 中, 使用高级语义信息来引导前馈

网络, 进一步提升目标特征在编码过程中的权重。注意力模型如图 3 所示, 首先将输入特征通过全局平均池化捕获全局上下文信息, 并使用 Conv1×1、

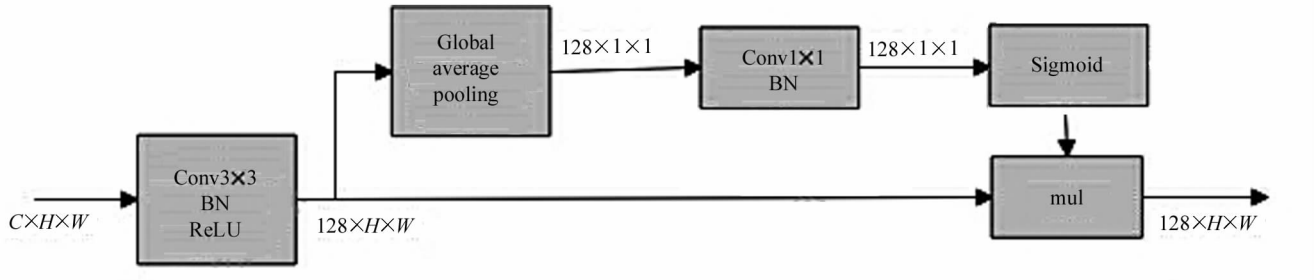


图 3 注意力模型
Fig. 3 Attention model

BN、Sigmoid 与 mul(两矩阵对应位相乘)共同计算注意力向量来指导特征学习,获取更加丰富的全局上下文信息。

2.3 副编码器结构

本文提出的 DCPM 模型如图 4 所示,由于本支路为获取丰富的边缘和轮廓等细节的空间信息,故网络的设计不用过深,只需要在不破坏空间信息的前提下,对空间信息编码。本文设计的副编码器下采样倍数为 8 倍,与 BiSeNetv1 相同。首先将原始图像送入网络,经过可变形卷积和普通卷积为一个卷积 block 的 3 层网络模块(DCN-CNN),最后通过 ASPP 网络,多尺度地捕捉图像更丰富的空间信息。

在同一层 Conv 中,激活单元的感受野是一样的,但是由于脑部肿瘤的位置多变、形状复杂等难点,若仅使用 CNNs,无法有效地应对几何图形(目标)的变化,故采用自适应的尺度和感受野,获取精确的定位,即在卷积模块中学习额外偏移量(Offsets)来获取空间中的肿瘤位置。但若仅适应物体的形变,可变形卷积网络可能会获取感兴趣区域(肿瘤区域)之外的不相关特征信息,从而影响网络的性能,故本文使用 v2 版本的可变形卷积网络(DCNv2)^[13],在 v1 的基础之上增加控制输入强度的标量,通过更有效的建模能力和训练使网络更关注肿瘤区域,提升特征提取能力。在标准卷积中,给定一个有 k 个采样位置的卷积核,则 w_k 为第 k 个位置的权重, p_k 为预先设定的偏移,令 $y(p)$ 为输出特征图 y 上位置 p 的特征,其表达式为:

$$y(p) = \sum_{p_k \in R} w(p_k) \cdot x(p + p_k), \quad (2)$$

式中, $w(p_k)$ 为位置 p 对应卷积核中第 k 个位置的权重(卷积核内的权重参数), $x(p)$ 为输入特征图上位置 p 的特征。 R 为卷积核内的参数集合, p_k 为 R 中所在位置的枚举。

故设 Δp_k 为第 k 个位置上可学习的偏移量,可变形卷积 v1 计算式表示为:

$$y(p) = \sum_{p_k \in R} w(p_k) \cdot x(p + p_k + \Delta p_k). \quad (3)$$

在 v1 基础上加入可学习的控制输入强度的标量 $\Delta m_k \in [0, 1]$, 可调节的变形卷积表达式为:

$$y(p) = \sum_{k=1}^K w_k \cdot x(p + p_k + \Delta p_k) \cdot \Delta m_k, \quad (4)$$

式中, Δp_k 同 v1 相同, 为第 k 个位置上可学习的偏移量。 Δm_k 为第 k 个位置上调节参数, 可学习采样点的权重。

将经过 3 个 DCN-CNN 组成的解码网络后的输出特征图, 送入 ASPP 模块, 模块中共有 5 路分支, 其中一路使用自适应平均池化和 Conv1×1, 将特征图的通道数减小为 128, 并使用双线性插值法恢复图像尺寸, 即生成尺寸为 $128 \times \frac{H}{8} \times \frac{W}{8}$ 的特征图。一路应用 Conv1×1, 在不改变特征图分辨率的前提下改变图像深度, 获取 $128 \times \frac{H}{8} \times \frac{W}{8}$ 的特征图。其余 3 路分别使用膨胀率为 6, 12, 18 的空洞卷积, 并在卷积操作中, 设置相应填充(Padding)的 0 元素个数与膨胀率相同, 使得各路获取的输出特征图大小为 $128 \times \frac{H}{8} \times \frac{W}{8}$ 。最后将 5 路的输出特征图按通道数相加, 并使用 Conv1×1, 改变通道数为 128, 生成 $128 \times \frac{H}{8} \times \frac{W}{8}$ 的特征图。在进行 ASPP 时, 特征图中的神经元包含多个接受域大小, 可多尺度地对目标信息进行编码, 从而可以多个比例捕捉图像的上下文信息, 丰富高级语义信息, 提高特征提取性能。

在完成主编码器支路、副编码器支路的特征提取后, 将两者的输出特征通过一个特征融合模块, 提取更为丰富的上下文语义信息与空间细节信息。特征融合模块如图 5 所示, 首先使用级联(Concatenate), 按通道数将两者输出拼接为 $256 \times \frac{H}{8} \times \frac{W}{8}$, 再通过 BN 与 ReLU 操作, 得到批归一化和激活后

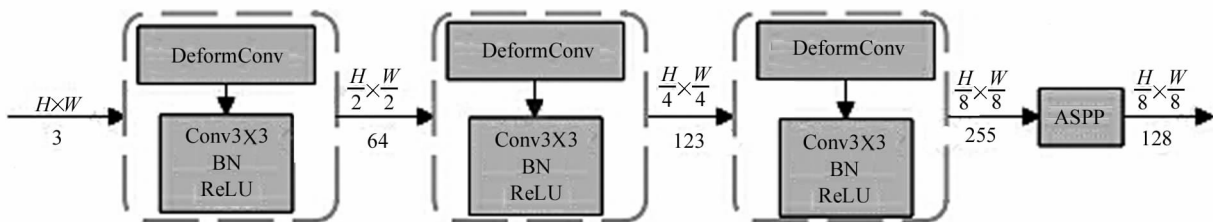


图 4 DCPM 模型

Fig. 4 DCPM model

的特征输出 $256 \times \frac{H}{8} \times \frac{W}{8}$ 。然后将其进行全局平均池化处理,继续通过 1×1 的卷积后,特征图大小为 $64 \times 1 \times 1$,接着再通过 1×1 的卷积与激活函数为

Sigmoid 的网络结构,输出 $256 \times 1 \times 1$ 的特征图。最后使用 mul 和 add 操作,得到特征大小为 $256 \times \frac{H}{8} \times \frac{W}{8}$ 的特征向量。

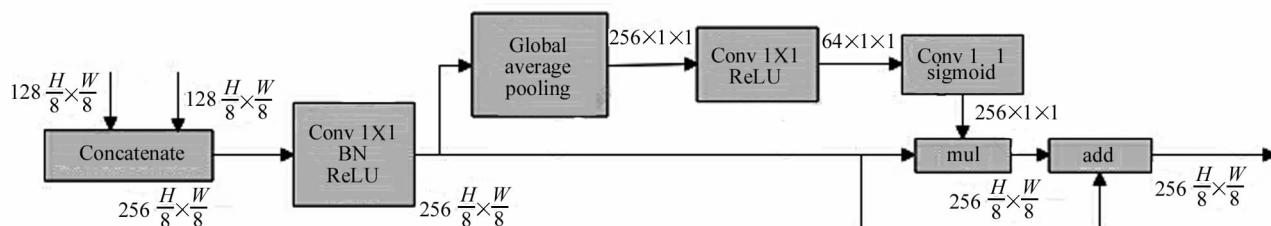


图5 特征融合模块

Fig. 5 Feature fusion module

2.4 解码器结构

在解码结构中,加入 CBAM 模块,利用副支路获得的空间信息辅助分割模型对特征融合网络模块的输出进行上采样。注意力机制模块如图 6 所示,首先将输入大小为 $128 \times \frac{H}{8} \times \frac{W}{8}$ 的特征图 F 输入注意力模块(channel attention module, CAM)中,分别经过基于 W 和 H 的全局最大池化和全局平均池化得到两个 $128 \times 1 \times 1$ 的特征图,然后送入多层感知器(multilayer perceptron, MLP),此处设计为含 2 层的神经网络,第 1 层神经元个数(输出通道数)设置为 8,激活函数为 ReLU,第 2 层神经元个数设置为 128。接着将 MLP 的输出进行逐像素(element-wise)的加操作与 Sigmoid 激活操作。最后,将通道注意力特征与输入特征 F 进行 element-wise 的乘操作,输出大小为 $128 \times \frac{H}{8} \times \frac{W}{8}$ 特征图。

将 CAM 的输出特征图作为空间注意力机制(spatial attention module, SAM)的输入,之后分别经过基于通道的全局最大池化与全局平均池化。然后将两者结果进行 concatenate,接着经过 1 个

Conv 7×7 的卷积操作,降维为 1 通道,再通过 Sigmoid 进行激活,得到空间注意力特征。最后将其与输入特征 F 进行 element-wise 的乘操作,输出大小为 $128 \times \frac{H}{8} \times \frac{W}{8}$ 特征图。

本文设计了一个简单的上采样结构进行解码,解码结构如图 7 所示,采用 3×3 卷积、 1×1 的卷积、BN、ReLU 的组合,最后通过 PixelShuffle 操作将图像尺寸变为 $1 \times 256 \times 256$ 。第一个卷积保持通道数不变化,输出大小为 $256 \times \frac{H}{8} \times \frac{W}{8}$ 的特征图,接着通过批归一化和 ReLU 激活函数,继续通过第二个卷积 Conv 1×1 ,输出 $64 \times \frac{H}{8} \times \frac{W}{8}$,最后通过像素重组(PixelShuffle)输出与原始图像相同分辨率的分割结果图,设输入尺寸为 $(N, C \cdot f_{up}^2, H_{in}, W_{in})$,则输出尺寸表示为:

$$(N, C, H_{in} \cdot f_{up}, W_{in} \cdot f_{up}), \quad (5)$$

式中, f_{up} 为上采样因子,即图像分辨率的扩大倍数, $H_{out} = H_{in} \cdot f_{up}$, $W_{out} = W_{in} \cdot f_{up}$ 。所以,经过解码网络结构后,输出大小为 $1 \times H \times W$ 的分割概率图。

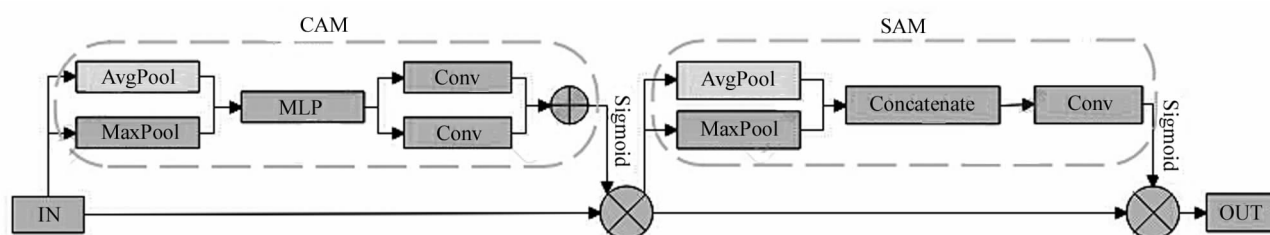


图6 CBAM 模块

Fig. 6 CBAM module

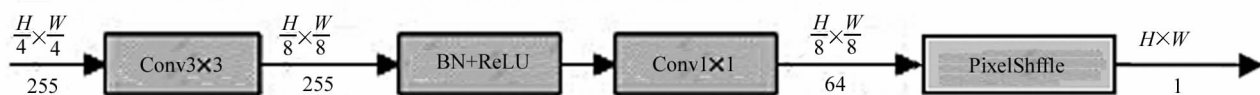


图7 解码结构

Fig. 7 Decoding structure

3 实验过程与分析

3.1 数据集

本文所提出的双支路特征融合网络模型分别在 kaggle_3m 数据集和 BraTS2019 数据集上进行了实验,并与近 3 年的处理方法进行了大量的对比实验。

kaggle_3m 数据集来源于 kaggle 脑肿瘤分割比赛。数据集中含有 110 例患者,对应于 110 个序列肿瘤图片组,每个序列内患者脑切片数量从 20—80 不等。肿瘤类型为低级神经胶质瘤(low grade gliomas, LGG),序列类型为磁共振成像液体衰减反转恢复序列(fluid attenuated inversion recovery, FLAIR),能使肿瘤区域呈现高信号,而其余脑部组织呈现相对低信号。每张图片大小为 256×256 ,且对图片进行了颅骨分离,使得网络在训练的时候,能更集中训练大脑内部区域。

BraTS2019 数据集来源于脑肿瘤分割比赛。数据来源于多机构的术前 MRI 扫描,其中含有高级神经胶质瘤(high grade gliomas, HGG)(WHO3—4 级)的病人有 259 例,含有 LGG(WHO1—2 级)的病人有 76 例。每个病例中包含 4 种模态的 MRI 序列和一个相应的分割真值(seg)文件,每个序列(单个 3 维图像)的尺寸为 $(155 \times 240 \times 240)$,本文使用 FLAIR 模态的 MRI 进行实验。并将每个 3 维图像切片为 2 维序列,即生成 155 张图片大小为 240×240 的脑部图像。

3.2 实验内容

3.2.1 评价指标

本次实验进行定量评测时,采用医学影像分割领域常用的 4 个评价指标。并且在测试结果分析时,保留了序列的完整性,以 1 个 2 维序列切片(1 个 3 维图像)为单元进行,即根据某 1 组 2 维序列(某一患者)进行指标计算。具体评价指标含义式如下。

a) Dice 相似系数(dice similarity coefficient, DSC),一种基于样本集合的相似度进行度量的指标,用于计算 2 个样本的相似度,表达式为:

$$DSC(Pre, G) = \frac{2 |Pre \cap G|}{|Pre| + |G|}, \quad (6)$$

式中, Pre 为样本预测值(预测区域所包含的点集), G 为样本真实值(真实区域所包含的点集), DSC 函数值范围为 $[0-1]$, 指标越高表示方法性能越优。

b) 杰卡德系数(jaccard index),表达式为:

$$Jaccard\ index = \frac{(Pre \cap G)}{Pre \cup G}, \quad (7)$$

与 DSC 公式不同,该公式无法衡量个体间差异具体值的大小,只能衡量 2 个切片内某一位置像素(预测图片与真值图像)相似与否。故该系数只计算个体间具有的特征是否一致,数值越大表示分割性能越优。

c) 欠分割率,又称漏报率(false negative rate, FNR),其计算式为:

$$FNR = \frac{FN}{Pre \cup G}, \quad (8)$$

式中, FN 为假阴性指标,即在所有的正类中,预测为反类的值。 FNR 数值越小表示分割方法性能越优。

d) $F1$ 分数($F1$),又称为平衡 F 分数(balanced F score),是精确率和召回率的调和平均数,其表达式如式(9)所示,数值越大,分割性能越优。

$$F1 = \frac{2 \times P \times R}{(P + R)}, \quad (9)$$

式中,精确率 $P = \frac{TP}{TP + FP}$,召回率 $R = \frac{TP}{TP + FN}$, TP 为真阳性, FP 为假阳性, FN 为假阴性。

3.2.2 实验过程

本文实验平台的软硬件配置如下:操作系统为 64 位 Ubuntu18.04 LTS, CPU 为 AMD Ryzen 5 3500x(3.6 GHz),显卡型号为 NVIDIA RTX 2060S (显存为 8 GB),内存为 16 GB,所使用的深度学习环境为 Python 3.6 和 Pytorch 1.3.1。

在脑部 MRI 切片中,属于肿瘤的区域非常小,背景区域远远大于目标区域,导致在网络训练时,模型更可能偏向于学习非目标区域的特征信息,导致分割精度降低。故在原损失函数 Dice 基础上将预测图像的权重加大,即将预测值取平方,进而在训练逻辑

回归模型的时候,使得模型更加侧重于学习颅内肿瘤特征,让误差函数降到最低,从而获取最优的模型,本文采取改进的Dice损失函数,其表达式为^[16]:

$$L_{AdaDice} = 1 - \frac{2 \sum_i p_i^2 g_i}{\sum_i p_i^2 + \sum_i g_i^2}, \quad (10)$$

式中, $p_i \in (0, 1)$ 为分割模型输出中属于第 i 个像素点的样本预测概率值,即样本预测值, $g_i \in \{0, 1\}$ 为真实标签图像中属于第 i 个像素点的样本真实概率值,即样本真实值。

实验训练时,将输入图像大小调整为 256×256 , 预处理阶段的图像剪裁尺寸为 5%, 旋转角度为 15° , 分割网络设置迭代次数为 200, batchsize 为 16, 初始学习率为 0.0001, 学习率衰减为 0.9, 权重衰减为 $5e-4$, 优化器设置为 Adam。

为验证本文 backbone 的有效性以及选取最合适的 RepVGG 参数, 设计多组对比实验, 在 BiSeNetv1 网络的基础上改变主支路的 backbone, 使用 4 种不同的网络模型在 kaggle_3m 颅脑肿瘤数据集上进行对比, 对比实验结果如表 1 所示。第 1 列为对比方法, 第 2 至 4 列为性能指标, 最后 1 列为参数对比。表中前 3 种实验所使用的 backbone 为近年来较为高效的分割网络主干。而最后 1 种为本文所采用的 RepVGG, 其中 RepVGG-A2 与 RepVGG-B1g4 在参数 blocks、 a 和 b 上不同, 且 g4 代表 4 分组卷积。在 RepVGG-A2 中, $\text{blocks} = [2, 4, 14, 1]$, $a =$

$1.5, b = 2.75$ 。在 RepVGG-B1g4 中, $\text{blocks} = [4, 6, 16, 1]$, $a = 2, b = 4$ 。从表中得出在参数总量相差不大的情况下, RepVGG 可取得比前 3 组更高的分割精度。结合实验中 RepVGG 其他版本参数的设计与参数数量, RepVGG-A2 模型性能相对最优。故选择 RepVGG-A2 为本文具体的 backbone。

为了在 RepVGG-A2 上对比加入不同网络模块的有效性, 设计了 4 组对比实验, 数据集使用 kaggle_3m, 对比实验结果如表 2 所示。其中 RepVGG-A2 为第 1 组基础对比, 第 2、3 组在此基础上分别添加 DCPM、CBAM。第 4 组为同时添加 DCPM、CBAM 后的网络模型。从表得出, 单独加入 DCPM 结构, 模型在 DSC、Jaccard 和 FNR 3 个指标上均有提升。单独加入 CBAM 模块, 模型在 3 个指标上均有提升。同时加入 2 个模块后, DSC 提升了 0.2%, Jaccard 提升了 1.23%, FNR 提升了 0.09%。

表 1 不同基础网络对比实验

Tab. 1 Comparative experiment of different basic networks

Methods	DSC/%	Jaccard/%	FNR/%	Params/M
ResNet50	90.77	83.17	0.28	26.90
ResNext50	90.74	83.08	0.14	26.37
RegNetY-6.4GF	90.28	81.58	0.23	25.43
RepVGG-A2	91.25	83.96	0.19	29.70
RepVGG-B1g4	90.23	83.46	0.12	41.17

表 2 不同网络模型对比实验

Tab. 2 Comparative experiment of different network models

Methods	DSC/%	Jaccard/%	FNR/%	Params/M
RepVGG-A2	91.25	83.96	0.19	29.70
RepVGG-A2+DCPM	91.27	84.01	0.21	31.39
RepVGG-A2+CBAM	91.33	84.36	0.15	29.87
RepVGG-A2+DCPM+CBAM	91.45	85.19	0.10	31.56

3.2.3 实验结果

本文提出的网络模型与近年来的医学分割网络模型进行评价指标比较, 对比实验结果分别如表 3、4、5 所示, 所有对比实验数据均由本地设备复现。

如表 3 所示, 在 kaggle_3m 数据集中, DSC 较原 BiSeNetV1 网络提升了 2.57%, Jaccard index 提升了 3.75%, F1 分数提升了 1.69%, 欠分割率提升了 0.1%。并且, 本文方法相较于其他文献方法, 在 DSC、杰卡德系数、F1 分数和欠分割率上均取得了

最优值。

如表 4 所示, 在 BraTS2019_HGG 数据集中, DSC 较原 BiSeNetV1 网络提升了 1.49%, Jaccard index 提升了 0.86%, F1 分数提升了 0.57%, 欠分割率提升了 4.98%。并且, 本文方法相较于其他文献方法, 在 DSC、杰卡德系数和 F1 分数上取得了最优值, 在欠分割率指标上比 Multi_Scale_Attention 方法略低。

如表 5 所示, 在 BraTS2019_LGG 数据集中,

DSC 较原 BiSeNetV1 网络提升了 1.07%, *Jaccard index* 提升了 1.96%, *F1* 分数提升了 1.34%, 欠分割率提升了 1.24%。并且, 本文方法相较于其他文献方法, 在 *DSC*、杰卡德系数和 *F1* 分数上取得了最优值, 在欠分割率指标上比 Multi_Scale_Attention 方法略低。综上所述, 本文提出的颅脑肿瘤分割方法对以上 3 个数据集均有较好的效果, 可证明本文方法对于颅脑肿瘤分割的有效性。

表 3 kaggle_3m 上与主流算法评价指标比较 %

Tab.3 Comparison with mainstream algorithm evaluation indicators on Kaggle_3m %

Methods	DSC	Jaccard	F1	FNR
BiSeNetV1 ^[6]	88.88	81.44	89.62	0.20
BiSeNetV2 ^[17]	90.25	83.47	90.85	5.48
DANet ^[18]	83.73	72.25	83.73	7.10
CENet ^[19]	91.16	83.65	90.95	0.26
Multi_Scale_Attention ^[20]	90.57	82.31	90.16	4.23
Ours	91.45	85.19	91.31	0.10

表 4 BraTS2019_HGG 上与主流算法评价指标比较 %

Tab.4 Comparison with mainstream algorithm evaluation indicators on BraTS2019_HGG %

Methods	DSC	Jaccard	F1	FNR
BiSeNetV1 ^[6]	88.12	78.71	87.96	10.02
BiSeNetV2 ^[17]	86.99	76.86	86.76	13.51
DANet ^[18]	80.48	67.69	80.44	14.30
CENet ^[19]	88.45	78.92	88.06	10.29
Multi_Scale_Attention ^[20]	87.04	74.67	85.30	1.69
Ours	89.61	79.57	88.53	5.04

表 5 BraTS2019_LGG 上与主流算法评价指标比较 %

Tab.5 Comparison with mainstream algorithm evaluation indicators on BraTS2019_LGG %

Methods	DSC	Jaccard	F1	FNR
BiSeNetV1 ^[6]	86.54	76.17	86.07	3.51
BiSeNetV2 ^[17]	84.18	72.68	83.74	5.71
DANet ^[18]	82.54	70.30	82.19	3.28
CENet ^[19]	87.37	77.66	87.16	6.71
Multi_Scale_Attention ^[20]	84.29	70.32	81.78	1.47
Ours	87.61	78.13	87.41	2.27

本文方法与主流算法对颅脑肿瘤图像分割结果可视化对比如图 8、图 9、图 10 所示, 绿色轮廓线对应

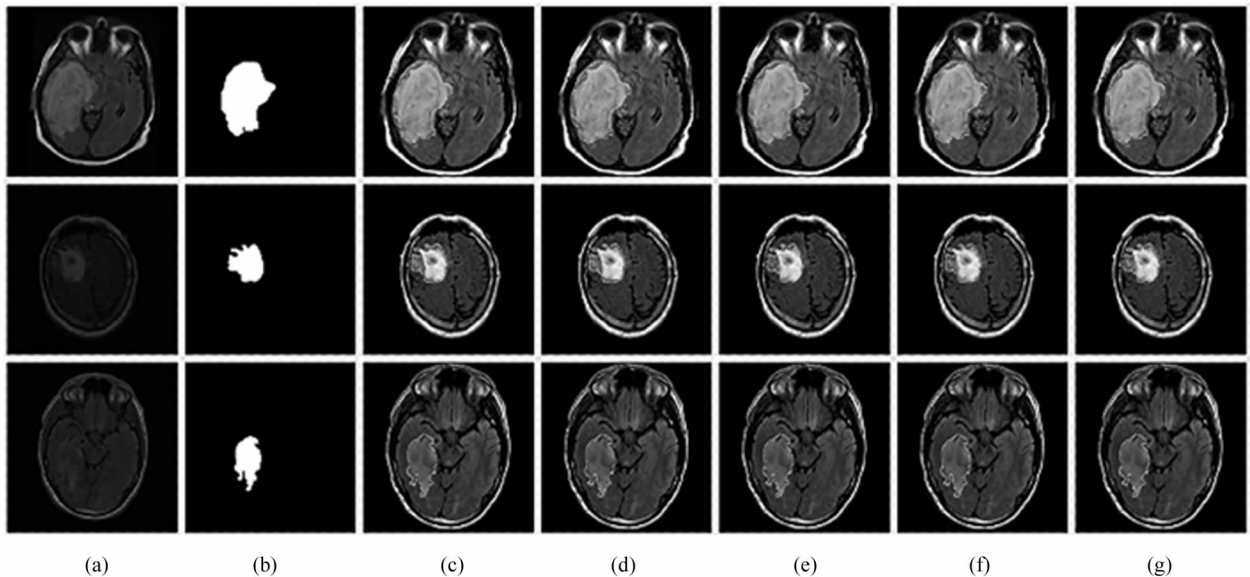


图 8 Kaggle_3m 上与主流方法分割结果可视化对比: (a) 原图; (b) GT 图像; (c) BiSeNetV1; (d) BiSeNetV2; (e) CENet; (f) Multi_Scale_Attention; (g) Ours

Fig.8 Visual comparison of segmentation results with mainstream methods on Kaggle_3m;

(a) Original image; (b) Ground truth; (c) BiSeNetV1; (d) BiSeNetV2; (e) CENet; (f) Multi_Scale_Attention; (g) Ours

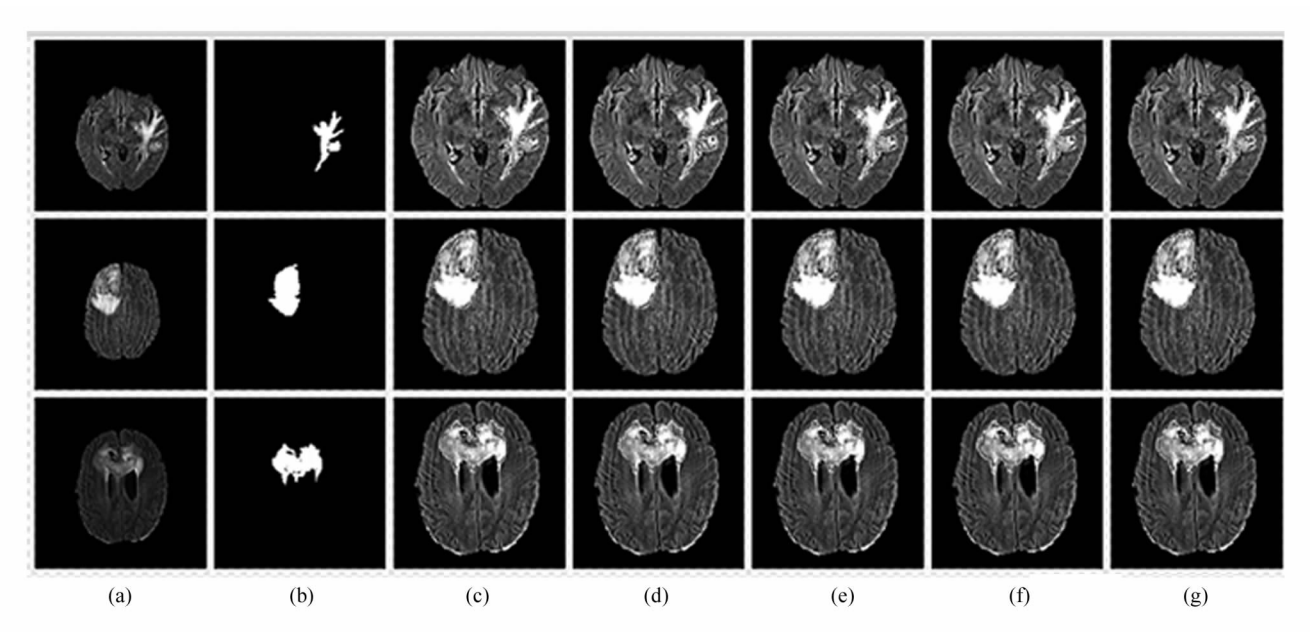


图 9 BraTS2019_HGG 上与主流方法分割结果可视化对比:(a) 原图;(b) GT 图像;(c) BiSeNetV1;
(d) BiSeNetV2;(e) CENet;(f) Multi_Scale_Attention;(g) Ours

Fig. 9 Visual comparison of segmentation results with mainstream methods on BraTS2019_HGG;
(a) Original image;(b) Ground truth;(c) BiSeNetV1;(d) BiSeNetV2;(e) CENet;(f) Multi_Scale_Attention;(g) Ours

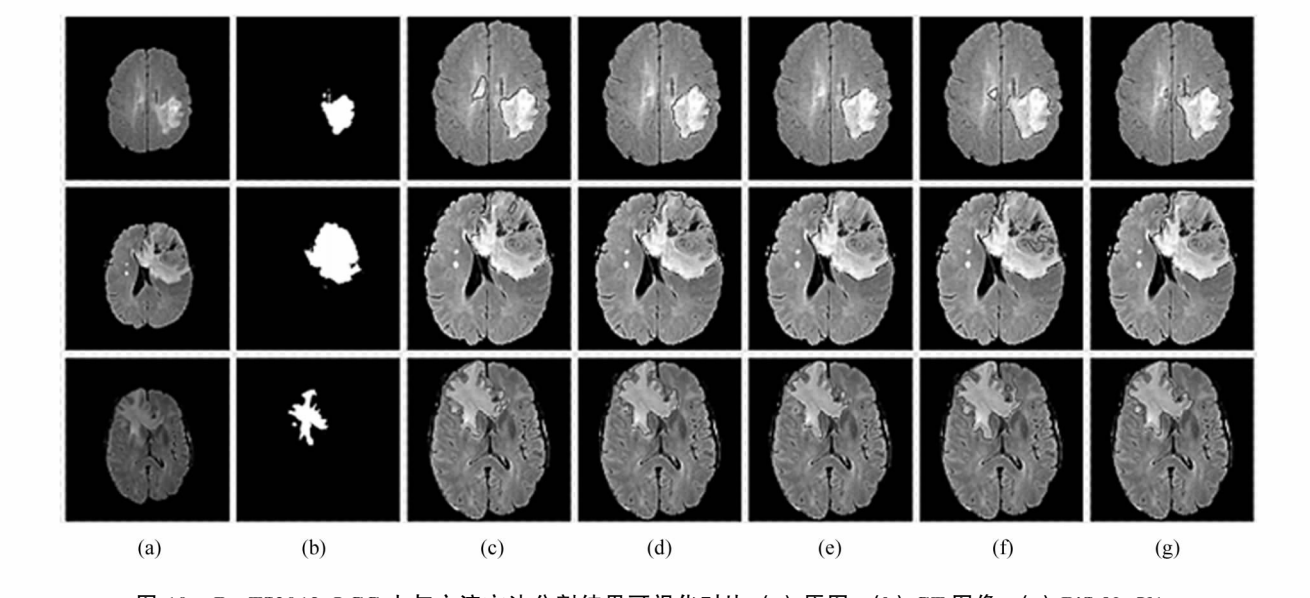


图 10 BraTS2019_LGG 上与主流方法分割结果可视化对比:(a) 原图;(b) GT 图像;(c) BiSeNetV1;
(d) BiSeNetV2;(e) CENet;(f) Multi_Scale_Attention;(g) Ours

Fig. 10 Visual comparison of segmentation results with mainstream methods on BraTS2019_LGG;
(a) Original image;(b) Ground truth;(c) BiSeNetV1;(d) BiSeNetV2;(e) CENet;(f) Multi_Scale_Attention;(g) Ours

真值区域,红色轮廓线对应网络预测区域。其中,(a)为原始颅脑切片图,(b)为对应的 GT 图像,(c)——(f)为对比方法,(g)为本文方法。从可视化图像可见,主流方法能够大致分割肿瘤区域,但是对于某些不规则的肿瘤形状,仍会出现过分割和欠分割现象。而本文方法相较于其他对比方法,分割区域较为精

确,误分割也相对较少。综上所述,本文方法较其他方法分割效果表现更好,证明本文提出的分割方法的有效性。

4 结 论

本文提出了一种采用双支路进行特征融合的

MRI 颅脑肿瘤图像分割方法。首先通过设计 RVAM 主支路结构,提取更加有效的高级肿瘤特征信息,减小肿瘤区域误识别的可能性。然后设计了 DCPM 副支路模型,将获取的空间信息融合到主支路中,减少分割网络空间信息丢失问题。最后使用注意力机制在上采样中加强肿瘤特征的权重。实验证明,本文提出的方法降低了肿瘤误识别率,同时减少了分割网络的特征信息丢失。与其他主流方法相比,本文提出的方法分割精度更高,性能更优,但仍存在部分肿瘤区域难以提取。未来的工作是在简化网络模型复杂度的前提下,设计更优的颅脑肿瘤分割图像方法。

参考文献:

- [1] GAO J H, LEI H, CHEN Q, et al. Magnetic resonance imaging: progresses and perspective[J]. Scientia Sinica Vitae, 2020, 50(11): 1285-1295.
- [2] ZHENG G Y, LIU X B, HAN G H. Survey on medical image computer aided detection and diagnosis systems[J]. Ruan Jian Xue Bao/Journal of Software, 2018, 29(5): 1471-1514.
- [3] AL-SAFFAR Z A, YILDIRIM T. A hybrid approach based on multiple Eigenvalues selection (MES) for the automated grading of a brain tumor using MRI[J]. Computer Methods and Programs in Biomedicine, 2021, 201: 105945.
- [4] 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, October 5-9, 2015, Munich, Germany. Berlin: Springer, 2015: 234-241.
- [5] 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.
- [6] YU C, WANG J, PENG C, et al. Bisenet: Bilateral segmentation network for real-time semantic segmentation[C]//The 15th European Conference on Computer Vision, September 8-14, 2018, Munich, Germany. Berlin: Springer, 2018: 325-341.
- [7] JIN Q, MENG Z, PHAM T D, et al. DUNet: a deformable network for retinal vessel segmentation[J]. Knowledge-Based Systems, 2019, 178: 149-162.
- [8] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks[C]//IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 764-773.
- [9] LALONDE R, XU Z, IRMAKCI I, et al. Capsules for biomedical image segmentation[J]. Medical Image Analysis, 2021, 68: 101889.
- [10] Meng Y, Wei M, Gao D, et al. CNN-GCN aggregation enabled boundary regression for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention, October 4-8, 2020, Lima, Peru. Berlin: Springer, 2020: 352-362.
- [11] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[J]. ArXiv Preprint ArXiv:1609.02907, 2016.
- [12] DING X, ZHANG X, MA N, et al. RepVGG: making VGG-style ConvNets great again[EB/OL]. (2021-03-29) [2021-07-22]. <https://arxiv.org/abs/2101.03697>.
- [13] ZHU X, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 9308-9316.
- [14] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-12-05) [2021-07-22]. <https://arxiv.org/abs/1706.05587>.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision, September 8-14, 2018, Munich, Germany. Berlin: Springer, 2018: 3-19.
- [16] YANG H G. Research on medical image segmentation method based on deep learning[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2019.
杨洪光. 基于深度学习的医学图像分割方法研究[D]. 重庆: 重庆邮电大学, 2019.
- [17] Yu C, Gao C, Wang J, et al. Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation[J]. International Journal of Computer Vision, 2021, 129: 3051-3068.
- [18] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 3146-3154.
- [19] GU Z, CHENG J, FU H, et al. Ce-net: context encoder network for 2d medical image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 38(10): 2281-2292.
- [20] SINHA A, DOLZ J. Multi-scale self-guided attention for medical image segmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 25(1): 121-130.

作者简介:

熊 炜 (1976—), 男, 博士, 副教授, 硕士生导师, 主要从事文档图像分析与理解、计算机视觉与模式识别等方面的研究。