

多模态磁共振脑肿瘤图像自动分割算法研究

何承恩*, 徐慧君**, 王忠***, 马丽萍

四川大学电气工程学院, 四川 成都 610065

摘要 脑肿瘤图像自动分割的难点在于肿瘤形态各异,且类别不平衡情况比较严重,常规的卷积神经网络难以预测出高精度分割图像。针对以上问题,在原始 3D-UNet 的基础上提出一种改进模型,以混合膨胀卷积模块代替常规卷积模块,指数级地增大神经元的感受野,同时减小网络深度,避免上采样时无法恢复小目标的情况。同时以混合损失函数代替原来的 Dice 损失函数,加强稀疏类分类错误时对模型的惩罚,迫使模型更好地学习分类错误的样本。实验结果表明,混合膨胀卷积模块与混合损失函数能分别提高整个肿瘤区域和肿瘤核心区域的预测精度,提出的 3D-HDC-UNet 模型改善了脑肿瘤自动分割的多项性能参数。

关键词 图像处理; 脑肿瘤自动分割; 卷积神经网络; 混合膨胀卷积; 混合损失函数; 磁共振图像

中图分类号 TP391

文献标志码 A

doi: 10.3788/AOS202040.0610001

Automatic Segmentation Algorithm for Multimodal Magnetic Resonance-Based Brain Tumor Images

He Cheng'en*, Xu Huijun**, Wang Zhong***, Ma Liping

College of Electrical Engineering, Sichuan University, Chengdu, Sichuan 610065, China

Abstract Automatic segmentation of brain tumor images is difficult to achieve owing to the diversity in tumor shapes and severe imbalance in the segmentation categories. Conventional convolutional neural network can hardly predict high precision segmentation images. To solve the abovementioned problems, an improved model based on the original three-dimensional (3D)-UNet was proposed, which replaced the conventional convolution module with a hybrid dilated convolution module to exponentially increase the receptive field of neurons, reducing the network depth and avoiding scenarios wherein small targets could not be recovered during up-sampling. Furthermore, the hybrid loss function was used to replace the original Dice loss function to increase the penalty faced by the model when classification errors of sparse classes occurred, forcing the model to learn the features of these classes better. Experiment results showed that the hybrid dilated convolution module and the hybrid loss function could respectively improve the prediction accuracy of the whole tumor region and the core tumor region. Multiple performance parameters of brain tumor automatic segmentation were improved using this 3D-HDC-UNet model.

Key words image processing; brain tumor automatic segmentation; convolutional neural network; hybrid dilated convolution; hybrid loss function; magnetic resonance image

OCIS codes 100.4996; 170.3880; 100.3008

1 引 言

从医学图像中准确分割脑肿瘤区域对于疾病诊断、制定放射治疗计划等意义重大。磁共振成像(MRI)因其具有亚毫米级的空间分辨率,常用于肿瘤检测^[1]。MRI 有 4 种成像模态:T1、T1c、T2 和 Flair,不同的模态用于突出显示不同的肿瘤区域^[2]。手动分割肿瘤区域是一项耗时且高度依赖医生个人

经验的工作。其难点在于肿瘤形态各异,且可能分布于大脑的任意区域,快速、准确、自动地描绘出患者的肿瘤区域难度较大。

医学图像分割的根本问题是提取像素或体素的特征,然后对其分类。技术发展路线可分为手动分割、基于强度的分割、基于地图集的分割以及深度学习的方法。

手动分割是最原始的方法,耗时、繁琐,且存在

收稿日期: 2019-08-19; 修回日期: 2019-11-20; 录用日期: 2019-11-28

基金项目: 四川省教育厅 2018 自然科学重点科研项目(18ZA0307, 18ZA0308)

* E-mail: 530208058@qq.com; ** E-mail: 839000429@qq.com; *** E-mail: 1648277629@qq.com

人工差异。基于强度的分割主要是将像素的灰度值与某一个或多个阈值相比较来区分不同的组织^[3]。常用的方法有区域分割法和聚类分割法,区域分割法具有较低的计算负荷,但对噪声很敏感,聚类分割法对初始化参数敏感,且难以考虑空间信息或像素间的依赖关系^[4]。

“地图集”就是人工标记完备的数据库,基于地图集的分割就是将测试图像与当前数据库的数据进行非刚性配准,然后用标签传播的方法将地图集的标签映射到测试图像中。但非刚性配准的准确度不高,且计算复杂度高^[5]。

2012年以来,以神经网络为代表的深度学习学习方法性能大幅超过传统算法的性能。深度学习方法不依赖手工制作的特征,可以直接从数据中自动地学习特征的各个层次结构^[6]。Shelhamer等^[7]提出全卷积网络(FCN),将原始卷积神经网络中最后用于分类的全连接层换成卷积层,以处理任意输入尺寸的图像。Ronneberger等^[8]在FCN的基础上设计了2D-Unet网络,该网络由对称的“收缩路径(Contracting Path)”和“扩展路径(Expanding Path)”组成,收缩路径用于捕捉上下文信息,扩展路径用于定位感兴趣的结构,且在同一深度的收缩层和扩展层之间添加跳跃连接。此模型达到了2015年2D医学图像分割的最先进水平。然而大部分医学图像是3D的,2D切片会丢失相邻切片之间的三维空间信息,降低分割精度,且单纯地将2D分割结果拼接成3D分割结果,可能会出现锯齿、断层等问题^[9]。Çiçek等^[10]利用3D卷积核将其拓展为3D-Unet网络。Milletari等^[11]在3D-Unet的基础上加入残差模块,缓解了网络过深时梯度消失的问题。Kayalibay等^[12]在最先进的3D-Unet架构下采用Jaccard损失函数,改善了类别不平衡问题,取得了BraTS 2015数据集的最好成绩。

本文以BraTS 2017脑肿瘤数据集作为实验数据,在3D-Unet架构基础上,构建4通道全卷积神经网络。在收缩路径中加入混合膨胀卷积模块(HDC)以扩大每一层的感受野。在扩展路径中,将上采样得到的特征映射(Feature Maps)与本层的特征映射进行拼接,生成不同分辨率下的预测图,然后通过逐元素相加得到最终预测图。损失函数方面,受文献^[13]的启发,将多类别Dice损失函数和Focal损失函数^[14]相结合,得到混合损失函数(HLF),加强稀疏类分类错误对模型的惩罚,有助于改善类别不平衡问题。

2 自动分割方法

2.1 数据集与预处理

实验数据选取BraTS 2017数据集^[15],其中包含210位高级别胶质瘤(HGG)和75位低级别胶质瘤(LGG)患者的4种MRI模态图像,原始尺寸为 $240\text{ mm} \times 240\text{ mm} \times 155\text{ mm}$,分别来自于19个机构的临床数据。每位患者的肿瘤图像都由4位放射科医生手动标注出4类标签,分别代表3种肿瘤亚型与背景:1)标签1代表坏死的(Necrotic, NCR)和非增强肿瘤(Non-Enhancing Tumor, NET)核心;2)标签2代表瘤周水肿(Peritumoral Edema, ED);3)标签4代表增强肿瘤(Enhancing Tumor, ET);4)标签0代表除肿瘤以外的背景。不同模态的图像可以突显不同的肿瘤亚型:Flair图像突显整个肿瘤(Whole Tumor, WT)区域;T2图像突显由NCR和NET组成的肿瘤核心(Tumor Core, TC)区域;T1c图像突显增强肿瘤(Enhancing Tumor, ET)区域。

BraTS 2017数据集图像已经剔除了头骨部分并配准到相同的解剖模板,分辨率为 $1\text{ mm} \times 1\text{ mm} \times 1\text{ mm}$ 。经过观察,将图像裁剪为 $128\text{ mm} \times 128\text{ mm} \times 128\text{ mm}$,进一步减少了无关信息但并未减少有用的肿瘤区域信息。为了避免数据不足可能导致的过拟合风险,在每次训练开始之前应用随机旋转、缩放、弹性变形、镜像等数据增强技术对数据集进行扩充。

2.2 3D-HDC-Unet 模型架构

3D-HDC-Unet模型如图1所示。C-M代表卷积模块,HDC代表混合膨胀卷积模块,S-C代表跨步卷积模块,UP代表上采样模块,L-M代表定位模块,S-L代表分割层。 \oplus 代表矩阵相加, \odot 代表矩阵拼接,下方数字表示数据经过当前层之后的尺寸,例如 16×128^3 代表输入数据经过16个尺寸为 $128\text{ mm} \times 128\text{ mm} \times 128\text{ mm}$ 的3D卷积核处理后,输入数据尺寸变为16层 $128\text{ mm} \times 128\text{ mm} \times 128\text{ mm}$ 的图像。

模型的左半部分为收缩路径。第一个C-M模块是由16个尺寸为 $3 \times 3 \times 3$ 的卷积层、实例标准化(Instance Normalization, IN)层和PReLU层组成的卷积模块。卷积层将输入的 $(4, 128, 128, 128)$ 体素块转换为16个的 $128\text{ mm} \times 128\text{ mm} \times 128\text{ mm}$ 的体素块,相比起全连接网络,卷积操作的稀疏交互和参数共享两个特性,使得需要优化的参数数量大大缩减,每一层的神经元只响应前一层局部区域范

围内的神经元,提高了模型的训练效率,提升了模型的可扩展性。普通 3D 卷积运算如图 2(a)所示,普

通的 3D 卷积核的膨胀率 r 可视为 1,该卷积核的感受野,即能覆盖的区域为 $3 \text{ mm} \times 3 \text{ mm} \times 3 \text{ mm}$ 。

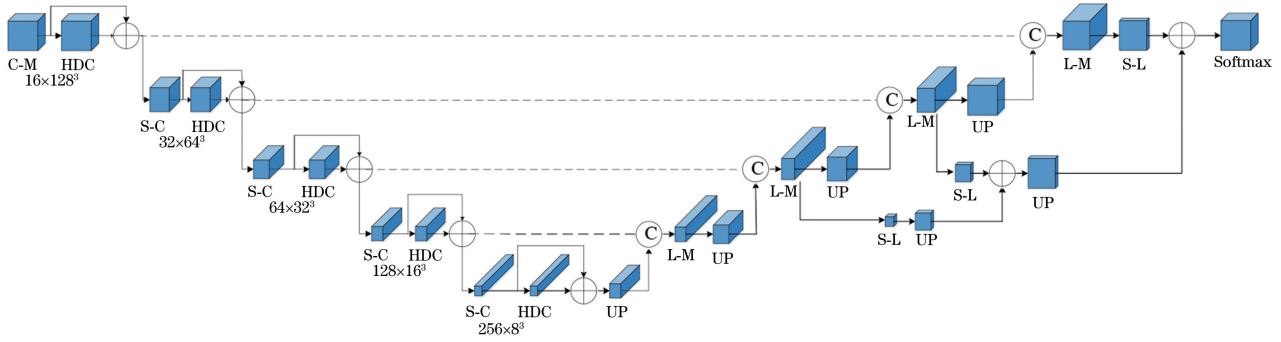


图 1 3D-HDC-Unet 模型结构示意图

Fig. 1 Schematic diagram of 3D-HDC-Unet model structure

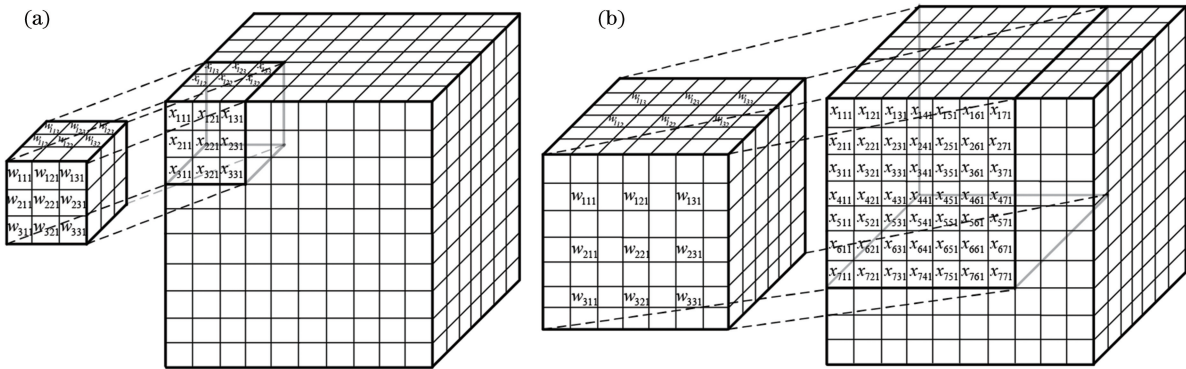


图 2 3D 卷积与输入体素块的点乘运算示意图。(a)普通 3D 卷积运算;(b)膨胀率 $r=2$ 的 3D 膨胀卷积运算

Fig. 2 Dot product operation diagram of 3D convolution with input voxel block. (a) Original 3D convolution operation; (b) 2-dilated 3D convolution operation

IN 层是针对整幅输入图像的所有像素值的灰度值作归一化处理,使所有体素块的灰度值服从均值为 0、方差为 1 的正态分布,避免网络中每一层的输入的数据分布差异过大。

PReLU 层实质为带参数的线性整流单元 (Parametric Rectified Linear Unit),定义了来自上一层神经网络的某通道输入变量 X_m 的非线性输出结果,即

$$f(x) = \begin{cases} X_m, & \text{if } X_m > 0 \\ a_m X_m, & \text{if } X_m \leq 0 \end{cases}, \quad (1)$$

式中: a_m 是可学习的参数,采用带动量的更新方式使其在反向传播中逐渐确定^[16]。然后数据被送入混合膨胀残差模块,进行特征提取。前向传播时第 l 层的 3D 卷积核的权重 W_l 与第 $l-1$ 层体素块 X_{l-1} 卷积之后,加上偏置项 b 后经过 PReLU 函数 f 处理,即可得到第 l 层的特征映射 F_l ^[17],即

$$F_l = f\left(\sum_l W_l * X_{l-1} + b\right), \quad (2)$$

式中: b 为偏置项 b 组成的矩阵。

混合膨胀残差模块如图 3 所示,由 IN 层、PReLU 层和 3 组膨胀率分别为 1、2、3 的膨胀卷积层共同构成。

首先是在常规卷积的基础上引入变化的膨胀率 (Dilation)。膨胀卷积 (Dilated Convolution) 可以增加单个神经元的感受野 (Receptive Field),如图 2(b) 所示。膨胀卷积通过在普通卷积核的相邻元素之间插入零值来扩大卷积核的尺寸,相邻元素间隔的步长即膨胀率 r 。大小为 $k \times k \times k$ 的 3D 卷积核,得到的膨胀卷积核大小为 $k_d \times k_d \times k_d$,其中 $k_d = k + (k-1) \cdot (r-1)$,对应第 l 层的感受野变为 $(2^{l+1}-1) \times (2^{l+1}-1)$,卷积核的参数量不变,但感受野呈指数增大。若膨胀率不变,神经元只能以棋盘的方式查看信息,这种方式会产生棋盘效应,不利于图像分割^[18]。

本研究采用膨胀率变化的混合膨胀卷积核来减弱棋盘效应。例如对于原来 $r=2$ 的层,将 3 个级联的膨胀卷积层作为一个组,分别将其膨胀率设置为 1、2、3,这样顶层可以访问更大范围的体素信息。这一操作对于模型识别较大的对象非常重要。混合膨

胀卷积能在模型参数个数保持不变、下采样次数较少的情况下,大大增加三维卷积层的感受野^[19],如图 4 所示。图中的正方形区域为混合膨胀卷积的感

受野。在混合膨胀卷积模块的首尾两端引入残差连接^[20],使网络更容易优化,避免了网络过深后的梯度消失问题。

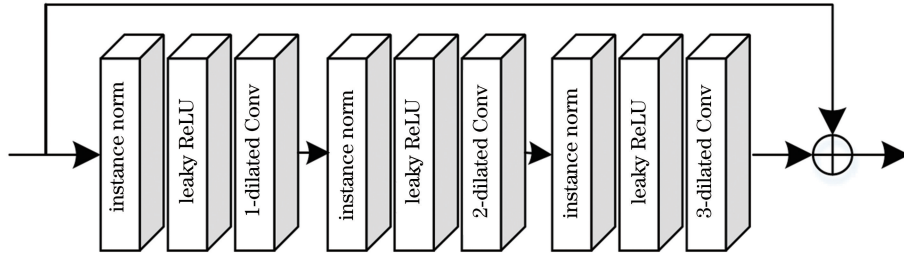
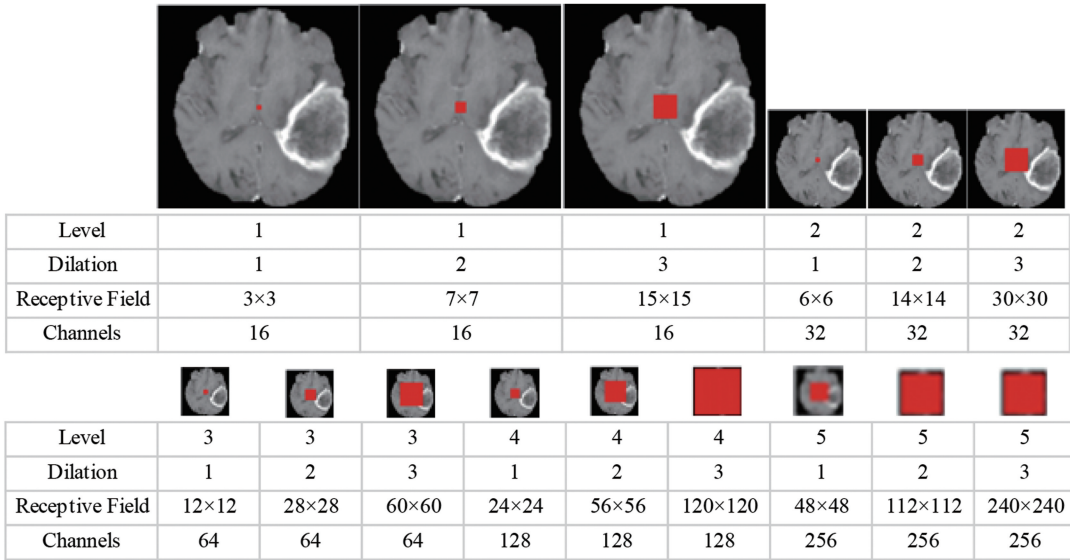


图 3 混合膨胀残差模块

Fig. 3 Hybrid dilated convolutional residual module



Note: In layer 3 4 5, the size of images is enlarged to scale for better demonstration.

图 4 收缩路径中输入图片降采样与混合膨胀卷积模块的感受野示意图

Fig. 4 Sketch map of input images sub-sampling and HDC receptive field

最后再将数据送入下一层步长为 2 的 $3 \times 3 \times 3$ 的跨步卷积层进行下采样,将体素块 3 个维度的尺寸减半,特征映射层数加倍,一方面可以提升卷积层的感受野,另一方面可以降低存储空间需求,最终实现效率和精度的平衡。随着网络的加深,输入图像块的分辨率越低,特征的抽象程度越高。抽象程度低的特征映射用于捕捉边缘、转角等特征,抽象程度高的特征映射用于捕捉相邻体素的非线性强度关系。

扩展路径中,首先利用双线性插值法对低分辨率的特征映射进行上采样,将 3 个维度的尺寸扩大一倍,然后利用 $3 \times 3 \times 3$ 的卷积将其特征映射的数量减半。同时,将此特征映射与同一层级的收缩路径中的特征映射在通道维度上进行拼接,再通过定位模块将其重新组合,以低空间分辨率编码上下文信息,并将该信息传送到更高空间分辨率的特征中。

定位模块由两组重复的 C-M 模块组成。

在扩展路径的第 3、2、1 层分别进行不同分辨率的预分割。分割层首先由卷积层将特征映射的通道维数降至 4,然后再接 IN 层和 PReLU 层。经过上采样后分别与上一层的特征映射进行矩阵相加。通过整合网络不同层级上的分割层,生成最终的预测图,最后利用 Softmax 层进行体素级预测。Softmax 函数也称为归一化指数函数,用于计算输入张量 \mathbf{X} 属于某类标签 \mathbf{Y} 的概率 $P(y = \mathbf{Y} | \mathbf{X})$,定义为

$$P(y = \mathbf{Y} | \mathbf{X}) = \frac{\exp(\mathbf{W}_Y^T \mathbf{X})}{\sum_{k=1}^K \exp(\mathbf{W}_k^T \mathbf{X})}, \quad (3)$$

式中: $\exp(\mathbf{W}_Y^T \mathbf{X})$ 表示标签为 \mathbf{Y} 的体素张量 (Tensor) \mathbf{X} 经过神经网络各层加权之后的值, $K = 4$

代表类别个数,输出为一个4维张量 (Y_1, Y_2, Y_3, Y_4) , $Y_i \in (0, 1)$,对应该体素属于各类别标签的概率,概率值最高的项对应的标签为该体素的预测标签。

2.3 混合损失函数

训练数据中的类别不平衡是阻碍模型精度提高的一大障碍,表1列出了BraTS 2017数据集中各类别体素占总体素的百分比^[12]。可以看到标签0的体素是标签1和标签4体素的百倍之多,这导致稀疏类预测错误对于模型整体损失来说影响甚小,不利于稀疏类预测精度的提高。

表1 BraTS 2017数据集中各类别体素占总体素的比值
Table 1 Ratio of different types of voxels to total voxel in the BraTS 2017 dataset

Label	Label 0	Label 1	Label 2	Label 4
Percentage /%	96	0.8	2.4	0.8

为此,本研究构造了一种结合Dice损失和Focal损失的混合损失函数来克服这个问题。Dice相似系数(DSC)是一种集合相似度的度量函数,用于计算预测区域与真实分割区域的重叠程度。Dice损失函数为负的DSC,定义为

$$L_{Dice} = -\frac{2P_{TP,x}}{2P_{TP,x} + P_{FN,x} + P_{FP,x}}, \quad (4)$$

式中: $P_{TP,x}(y)$ 、 $P_{FN,x}(y)$ 和 $P_{FP,x}(y)$ 分别是将体素 x 预测为类别 y 的真阳、假阴和假阳的概率,定义为

$$P_{TP,x} = \sum_{x=1}^X p_x(y)t_x(y), \quad (5)$$

$$P_{FN,x} = \sum_{x=1}^X [1 - p_x(y)]t_x(y), \quad (6)$$

$$P_{FP,x} = \sum_{x=1}^X p_x(y)[1 - t_x(y)], \quad (7)$$

其中 $p_x(y)$ 是将体素 x 预测为类别 y 的预测概率, $t_x(y)$ 是体素 x 对应的真实标签的one-hot编码值, X 为输入MRI图像的体素总数, K 为标签类别总数,本研究中取4。多类别分类时的Dice损失函数为

$$L_{Dice} = K - \sum_{y=0}^{K-1} \frac{2P_{TP,x}}{2P_{TP,x} + P_{FN,x} + P_{FP,x}}。 \quad (8)$$

Focal Loss是通过降低易分类样本的损失权重,提高难分类样本(比如稀疏样本)的权重,来迫使模型更好地学习分类较差的样本,定义为

$$L_{Focal} = \frac{1}{X} \sum_{y=0,1,2,4} \sum_{x=1}^X [1 - p_x(y)]^\gamma t_x(y) \log[p_x(y)]. \quad (9)$$

(9)式相当于在标准交叉熵损失前添加一个调节因

子 $[1 - p_x(y)]^\gamma$, $\gamma \in N$ 为聚焦参数,其功能在于: 1)当样本分类错误, $p_x(y)$ 值较小时,调节因子接近于1,该体素对损失函数贡献正常的惩罚值;当样本分类容易, $p_x(y)$ 较大时,调节因子接近于0,该体素对损失函数的贡献降低。2)聚焦参数 γ 可以调整易分类体素对损失函数权重降低的速率。当 $\gamma=0$ 时, L_{Focal} 相当于标准交叉熵损失, γ 越大,易分类体素对损失函数的贡献越小。例如 $\gamma=1$, $p_x(y)=0.8$ 时易分类样本与标准交叉熵损失相比损失值降低为1/5,这也变相地增加了纠正错误分类样本的重要性^[14]。 γ 的取值对损失函数 f 的影响如图5所示, γ 一定时,模型对某体素预测正确的概率越高,经过衰减后其对损失函数的贡献越低; γ 越大,这种趋势越明显。

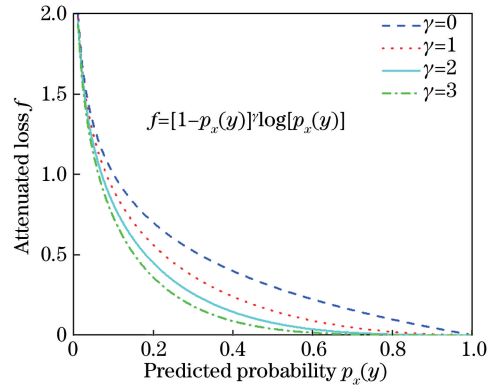


图5 γ 的取值对损失函数的影响

Fig. 5 Effect of γ on attenuated loss

本文的损失函数由 L_{Dice} 和 L_{Focal} 混合而成,即

$$L = L_{Dice} + \lambda L_{Focal} = K - \sum_{y=0,1,2,4} \frac{2P_{TP,x}}{2P_{TP,x} + P_{FN,x} + P_{FP,x}} - \lambda \frac{1}{X} \sum_{y=0,1,2,4} \sum_{x=1}^X [1 - p_x(y)]^\gamma t_x(y) \log[p_x(y)], \quad (10)$$

混合比例由 λ 调节。由实验可得,当 $\lambda=0.5$, $\gamma=1$ 时,模型的综合性能最佳。

3 仿真实验

3.1 评价指标

为了定量评估此模型的性能,选取医学图像分割中常用的4种评价指标: Dice系数(D_{sc})、灵敏度(S_{en})^[21]、特异度(S_{pe})、Hausdorff距离(D_{HD})。 D_{sc} 表示预测肿瘤区域与真实肿瘤区域的重叠程度;灵敏度又称为真阳性率、召回率,表示真实肿瘤区域中被预测为肿瘤的比例;特异度又称为真阴性率,表示

健康组织区域中被预测为健康的比例; Hausdorff 距离用于度量两个点集的最大不匹配程度,对分割出的边界比较敏感。其定义为

$$D_{sc} = \frac{2P_{TP,x}}{2P_{TP,x} + P_{FN,x} + P_{FP,x}}, \quad (11)$$

$$S_{en} = \frac{P_{TP,x}}{P_{TP,x} + P_{FN,x}}, \quad (12)$$

$$S_{pe} = \frac{P_{TN,x}}{P_{TN,x} + P_{FP,x}}, \quad (13)$$

$$D_{HD} = d_H(P, T) = \max\{\sup_{p \in P} \inf_{t \in T} d(p, t), \sup_{t \in T} \inf_{p \in P} d(p, t)\}, \quad (14)$$

(14)式中的 P 和 T 分别表示模型预测的肿瘤区域和真实肿瘤区域在二维图像上的像素集, p 和 t 分别表示各自像素集中的像素点。

3.2 实验结果

从 BraTS 2017 数据集的 285 名患者中随机选出 5 名作为测试集,将剩余 280 名患者均分成 5 组,每组 56 人,在 BraTS 2017 数据集上进行 5 次交叉验证(5-fold Cross-validation),如图 6 所示。每一次都将其中一组作为验证集,其余 4 组作为训练集,如此训练出 5 个模型,将这 5 个模型在验证集上算出的评价指标的平均数作为最终模型的性能参数。

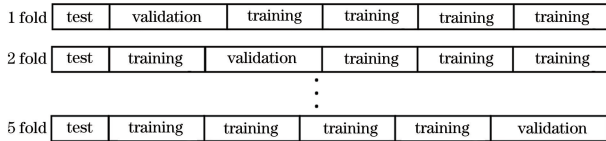


图 6 5 次交叉验证

Fig. 6 5-fold cross-validation

每次训练一位患者的 MRI 图像,总共训练 300 轮,每一轮训练 224 幅图像。优化器选用 Adam,在反向传播时,首先利用

$$g_l \leftarrow \frac{1}{X} \nabla_w \sum_i L(f(x_i; W_{l-1}), y_i), \quad (15)$$

计算第 l 层的损失函数关于前一层 $l-1$ 的梯度 g_l , 然后根据

$$m_l \leftarrow \beta_1 m_{l-1} + (1 - \beta_1) g_l, \quad (16)$$

$$v_l \leftarrow \beta_2 v_{l-1} + (1 - \beta_2) g_l^2, \quad (17)$$

定义的一阶动量 m_l 和二阶动量 v_l , 分别初始化为 $m_1 = g_1, v_1 = g_1^2, \beta_1$ 和 β_2 分别采用建议值 0.9 和 0.999, 利用

$$\hat{m}_l = \frac{m_l}{(1 - \beta_1^l)}, \hat{v}_l = \frac{v_l}{(1 - \beta_2^l)} \quad (18)$$

对其进行偏差纠正后,再根据

$$W_l = W_{l-1} - \alpha \frac{\hat{m}_l}{\sqrt{\hat{v}_l} + \epsilon} \quad (19)$$

对权重值进行更新,步长 $\alpha = 0.001$ 。

利用 Keras + TensorFlow 深度学习框架搭建此模型,在 Nvidia GTX 1070 GPU 上进行约 36 h 的训练,训练好的模型大约需要 5 s 来完成一名患者的 MRI 图像的分割。

为了探究 HDC 模块和 HLF 各自对脑肿瘤分割性能的提升程度,在原始 3D-Unet 模型基础上分别添加 HDC 模块和 HLF,训练出两个模型:3D-Unet + HDC 和 3D-Unet + HDC + HLF,连同使用普通 Dice 损失函数的双通道 3D CNN 网络,普通 3D-Unet 网络, K 均值聚类(K-means Clustering), 混合水平集(Hybrid Level), 结合了 FCM、自动阈值法和混合水平集的混合算法(Hybrid Algorithm)、随机森林(Random Forests),记录他们在测试集上的 4 种参数性能,分割的区域包括 WT(标签 1、2、4)、TC(标签 1、4)和 ET(标签 4)。结果记录于表 2。由于传统算法 FCM、混合水平集、混合算法等近年来很少使用,在以 BraTS 数据集为测试集的分割结果中,只有 WT 的参数,并未区分不同肿瘤区域。由表 2 可知,随着 HDC 和 HLF 的加入,肿瘤分割的各项参数均有不同程度的提升,尤其以 Hausdorff 距离提升最为明显,这是因为它对边缘的匹配程度最为敏感。3D-HDC-Unet 的测试集 DSC 箱型图如图 7 所示。

表 2 8 种模型的评价指标数据对比

Table 2 Comparison of evaluation index data of eight models

Model	DSC			Sensitivity			Specificity			Hausdorff distance		
	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET
Two-path 3D CNN ^[21]	0.85	0.78	0.73	0.80	0.76	0.75	0.97	0.97	0.96	10.6	15.4	7.70
3D-Unet ^[22]	0.88	0.76	0.72	0.90	0.78	0.75	0.98	0.98	0.98	13.6	22.3	13.8
3D-Unet+HDC	0.90	0.75	0.71	0.92	0.78	0.73	0.99	0.99	0.97	6.06	7.85	5.33
K-means ^[23]	0.79	—	—	0.94	—	—	0.98	—	—	—	—	—
Hybrid level set ^[23]	0.80	—	—	0.77	—	—	0.98	—	—	—	—	—
Hybrid algorithm ^[24]	0.90	—	—	0.89	—	—	0.98	—	—	—	—	—
Random forests ^[24]	0.86	0.78	0.66	0.83	0.72	0.57	0.99	0.99	0.99	7.61	8.70	3.76
Proposed method	0.90	0.80	0.77	0.92	0.84	0.81	0.99	0.99	0.97	6.06	7.69	5.20

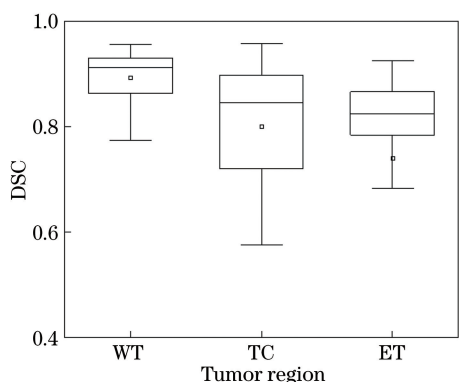


图 7 3D-HDC-Unet 测试集 DSC 箱型图

Fig. 7 3D-HDC-Unet test set DSC box diagram

本文提出的模型算法流程图如图 8 所示。以某位患者 4 种模态的多切片数据作为原始数据,经过裁剪、随即旋转、缩放、扭曲等操作得到增强后的数据,然后送入收缩路径,经过 4 次下采样得到 4 种不同尺度下的特征映射,然后再经过扩展路径的 4 次上采样,并将第 3、4 次的上采样结果作预分割,最终合成出与原始输入图片同尺寸的肿瘤区域分割图。

以患者 Brats17_TCIA_167_1 为例,分割结果

如图 9 所示,(a)~(c)为患者的 MRI 三种模态的图像,Flair 和 T2 能明显看出 WT 的轮廓,T1c 能明显看出 TC 和 ET 的轮廓。(d)~(f)、(g)~(i)和(j)~(l)分别为 42.531,69.094,90.344 mm 处的轴向分割图,从左到右依次为真实分割图、3D-Unet 分割图和 3D-HDC-Unet 分割图,可以看到本文提出的网络能够准确地分割出整个肿瘤的区域,能够比较准确地判断肿瘤核心和增强肿瘤的区域。相比原始的 3D-Unet,3D-HDC-Unet 在目标边缘处、ET 和 TC 的部分与真实分割图更接近。

结合表 2 的统计数据可以看出,HDC 模块有助于提升模型识别大区域目标的准确性,在 BraTS 2017 训练数据集上能达到 0.90 的 DSC。而 HLF 有助于提升模型识别小区域目标的准确性,对较难分割的 TC 和 ET 部分的 DSC 提升较大。SEN 的数值与 DSC 相似,模型对健康组织(背景)几乎没有误判,一方面是因为健康组织体素的灰度值差异较小,另一方面是因为背景所占的体素很多,可以使模型充分学习到其特征。

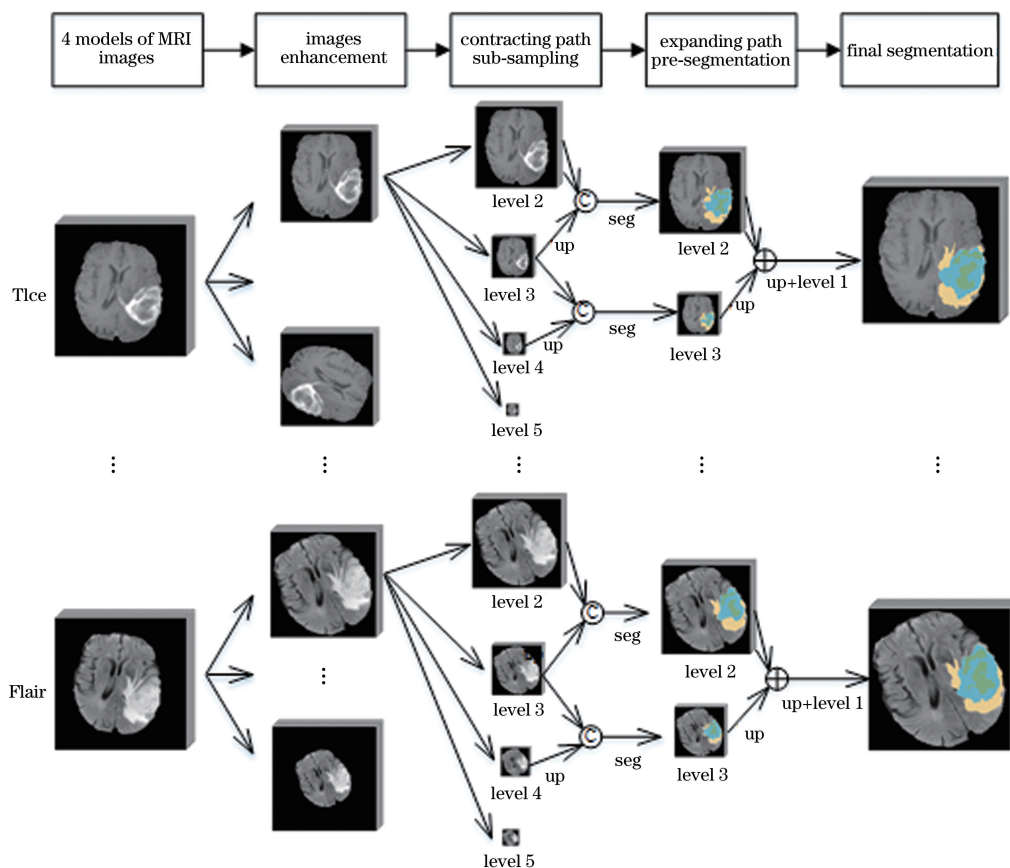


图 8 算法流程图与各阶段图像处理结果

Fig. 8 Algorithm flowchart and image processing results of each stage

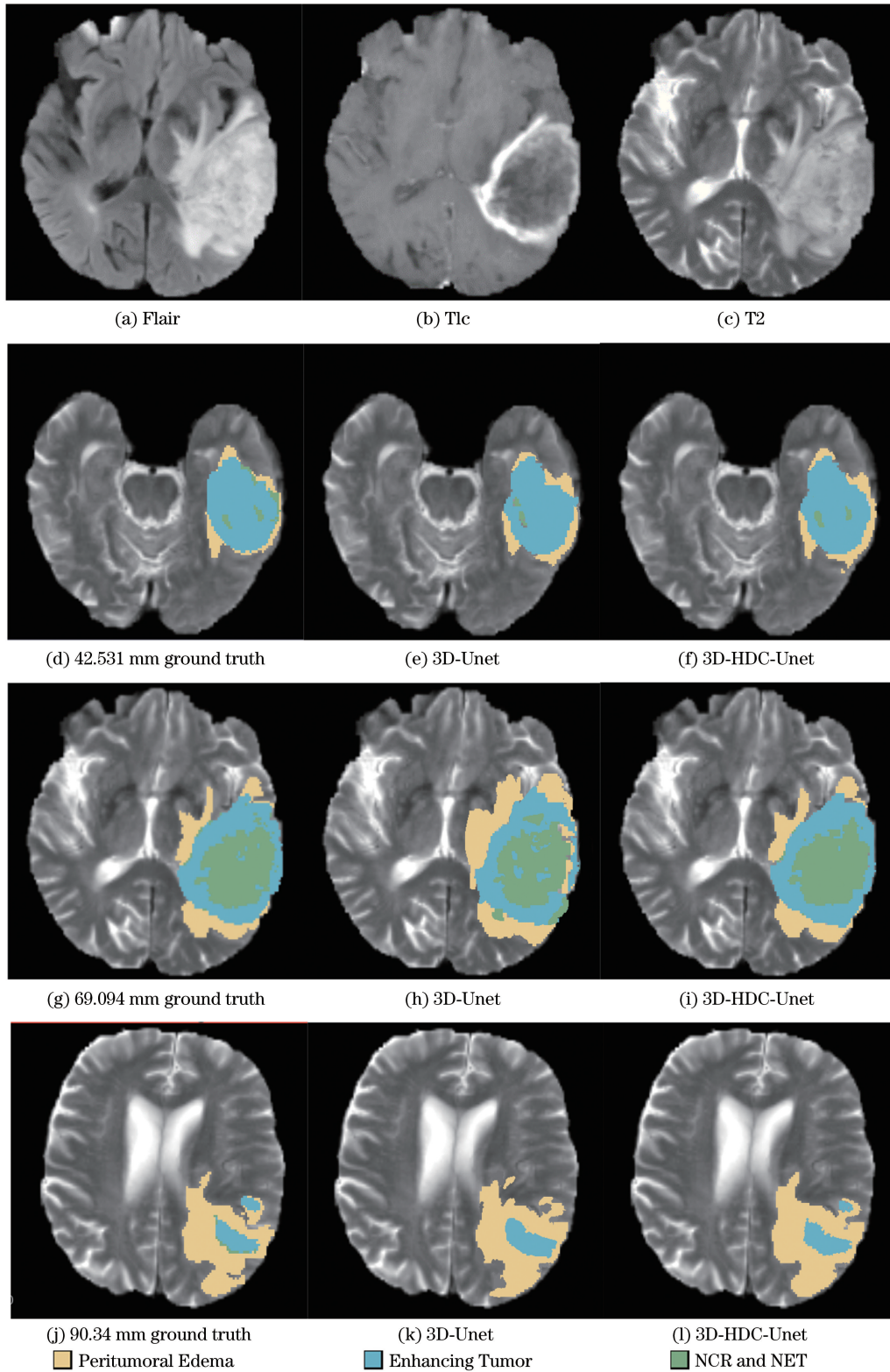


图 9 分割结果图

Fig. 9 Segmentation result diagram

由于 BraTS 2017 的在线评估端口已经关闭,无法测试此模型在官方测试集上的具体性能结果,根据文献[22]中的数据,可知模型在官方未公开发表的测试集上的性能相比从训练数据中分割出的测试

集的性能来说,会有 5% 以内的下降。但至少从训练数据中得到的性能参数可以看出,3D-HDC-Unet 仍然取得了领先成绩。

4 结 论

在原始 3D-UNet 的基础上,本研究加入 HDC 模块以增加神经元的感受野,可在较浅的网络中实现多尺度特征提取,并大大降低模型参数规模。用混合损失函数替换单一的 Dice 损失函数,提高了稀疏类对损失函数的贡献,迫使模型更好地学习不易分类的样本,从而提升了 3D-HDC-UNet 在 BraTS 2017 数据集上的分割性能。经过 36 h 的训练,模型在训练数据分割出的验证集上的 WT、TC、ET 的平均 Dice 系数为 0.90、0.80、0.77,均优于原始 3D-UNet 模型。

参 考 文 献

- [1] Niessen W J. MR brain image analysis in dementia: from quantitative imaging biomarkers to ageing brain models and imaging genetics [J]. *Medical Image Analysis*, 2016, 33: 107-113.
- [2] Ren L, Li Q, Guan X, et al. Three-dimensional segmentation of brain tumors in magnetic resonance imaging based on improved continuous max-flow[J]. *Laser & Optoelectronics Progress*, 2018, 55(11): 111011.
任璐, 李镛, 关欣, 等. 改进的连续型最大流算法脑肿瘤磁共振成像三维分割[J]. *激光与光电子学进展*, 2018, 55(11): 111011.
- [3] Shanthi K J, Kumar M S. Skull stripping and automatic segmentation of brain MRI using seed growth and threshold techniques[C]//2007 International Conference on Intelligent and Advanced Systems, November 25-28, 2007, Kuala Lumpur, Malaysia. New York: IEEE, 2007: 422-426.
- [4] Saman S, Narayanan S J. Survey on brain tumor segmentation and feature extraction of MR images [J]. *International Journal of Multimedia Information Retrieval*, 2019, 8(2): 79-99.
- [5] Bach Cuadra M, de Craene M, Duay V, et al. Dense deformation field estimation for atlas-based segmentation of pathological MR brain images [J]. *Computer Methods and Programs in Biomedicine*, 2006, 84(2/3): 66-75.
- [6] Pereira S, Pinto A, Alves V, et al. Brain tumor segmentation using convolutional neural networks in MRI images [J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1240-1251.
- [7] 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.
- [8] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9351: 234-241.
- [9] Chu J H, Li X C, Zhang J Q, et al. Fine-grained segmentation method for three-dimensional brain tumors using cascaded convolutional network [J]. *Laser & Optoelectronics Progress*, 2019, 56(10): 101001.
褚晶辉, 李晓川, 张佳祺, 等. 一种基于级联卷积网络的三维脑肿瘤精细分割[J]. *激光与光电子学进展*, 2019, 56(10): 101001.
- [10] Çiçek Ö, Abdulkadir A, Lienkamp S S, et al. 3D U-net: learning dense volumetric segmentation from sparse annotation[M]//Ourselin S, Joskowicz L, Sabuncu M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9901: 424-432.
- [11] Milletari F, Navab N, Ahmadi S A. V-net: fully convolutional neural networks for volumetric medical image segmentation [C]//2016 Fourth International Conference on 3D Vision (3DV), October 25-28, 2016, Stanford, CA, USA. New York: IEEE, 2016: 565-571.
- [12] Kayalibay B, Jensen G, van der Smagt P. CNN-based segmentation of medical imaging data[J/OL]. (2017-07-25) [2019-08-18]. <https://arxiv.org/abs/1701.03056>.
- [13] Zhu W, Huang Y, Tang H, et al. Anatomynet: deep 3D squeeze-and-excitation U-Nets for fast and fully automated whole-volume anatomical segmentation[J]. *bioRxiv*, 2018: 392969.
- [14] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018: 1.
- [15] Menze B H, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS) [J]. *IEEE transactions on medical imaging*, 2014, 34(10): 1993-2024.
- [16] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1026-1034.
- [17] Liu Y, Stojadinovic S, Hryciushko B, et al. A deep convolutional neural network-based automatic

- delineation strategy for multiple brain metastases stereotactic radiosurgery [J]. *PLoS One*, 2017, 12 (10): e0185844.
- [18] Yu F, Koltun V, Funkhouser T. Dilated residual networks[C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 636-644.
- [19] Wang P Q, Chen P F, Yuan Y, et al. Understanding convolution for semantic segmentation [C] // 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), March 12-15, 2018, Lake Tahoe, NV, USA. New York: IEEE, 2018: 1451-1460.
- [20] He K M, Zhang X Y, Ren S Q, et al. Identity mappings in deep residual networks[M] // Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9908: 630-645.
- [21] Havaei M, Davy A, Warde-Farley D, et al. Brain tumor segmentation with deep neural networks [J]. *Medical Image Analysis*, 2017, 35: 18-31.
- [22] Menze B H, Jakab A, Bauer S, et al. Proceedings of the MICCAI challenge on multimodal brain tumor image segmentation (BRATS) 2017 [J]. *MICCAI Challenge on Multimodal Brain Tumor Image Segmentation*, 2017, 10: 1-347.
- [23] Tong Y F, Li Q, Guan X. An improved multi-modal brain tumor segmentation hybrid algorithm [J]. *Journal of Signal Processing*, 2018, 34(3): 340-346. 童云飞, 李镛, 关欣. 改进的多模式脑肿瘤图像混合分割算法[J]. *信号处理*, 2018, 34(3): 340-346.
- [24] Zhang C, Shen X J, Cheng H, et al. Brain tumor segmentation based on hybrid clustering and morphological operations [J]. *International Journal of Biomedical Imaging*, 2019, 2019: 7305832.