

一种基于级联卷积网络的三维脑肿瘤精细分割

褚晶辉, 李晓川, 张佳祺, 吕卫*

天津大学电气自动化与信息工程学院, 天津 300072

摘要 提出了一种基于级联的 2.5 维(2.5D)卷积神经网络。将该任务拆分为脑肿瘤整体分割、肿瘤核分割、增强肿瘤分割三个子任务,并将三个结果合并生成最终结果。在每个子任务中,对三维(3D)图像进行轴向、矢向和冠向的裁取,生成 2.5D 图像;将 2.5D 图像输入至所提 2.5D V-Net 中进行训练;将 2.5D 分割结果拼接成 3D 结果,生成不同子任务的分割结果图。结果表明,所提方法对肿瘤整体、肿瘤核和增强肿瘤分割的平均 dice 值分别可达 0.9071、0.8542 和 0.8140,基本满足临床需要。

关键词 图像处理; 三维图像; 图像分割; 卷积神经网络; 磁共振图像

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.101001

Fine-Granted Segmentation Method for Three-Dimensional Brain Tumors Using Cascaded Convolutional Network

Chu Jinghui, Li Xiaochuan, Zhang Jiaqi, Lü Wei*

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

Abstract A cascaded 2.5-dimensional (2.5D) convolutional neural network is proposed. The task is divided into three sub-tasks of whole tumor segmentation, tumor core segmentation and enhancing tumor segmentation, and the results are combined to generate the final result. In each sub-task, the three-dimensional (3D) images are horizontally, coronally and sagittally cropped to generate 2.5D images. The 2.5D images are fed into the proposed 2.5D V-Net for training. The 2.5D segmentation results are concatenated as the 3D results to generate the segmentation results of different sub-tasks. The results show that the average dice scores for the segmentation of whole tumor, tumor core and enhancing tumor by the proposed method are 0.9071, 0.8542, and 0.8140, respectively, which basically meet the clinic need.

Key words image processing; three-dimensional image; image segmentation; convolutional neural network; magnetic resonance image

OCIS codes 100.2000; 100.4996; 100.6890

1 引言

脑肿瘤是一种致死率极高的肿瘤,死亡率占人体肿瘤发病比例的 2.4%,脑肿瘤中脑胶质瘤是发病率最高、预后最差的原发性肿瘤。胶质瘤包括胶质母细胞瘤(高级别细胞瘤, HGG)与低级别细胞瘤(LGG)两种,具有不同程度的侵略性及预后。一般来讲,脑胶质瘤包含水肿区域、坏死核心与非增强肿瘤核心区域、增强肿瘤核心区域。准确识别肿瘤及

其各子区域的位置、轮廓信息是医生制定精确诊断和治疗方案(如预测病患存活期、放射治疗等)的前提。临床上,医生个人医疗知识的积累、经验水平的差异以及视觉疲劳等不确定因素,都会影响对图像结果的分析。因此,构建脑肿瘤精细分割的系统模型是极其重要的。

通常,医生通过脑部多模态三维(3D)磁共振成像(MRI)对患者进行诊断。MRI 技术是一种通过产生、采集磁共振信号并对其进行空间编码与傅里

收稿日期: 2018-09-25; 修回日期: 2018-12-03; 录用日期: 2018-12-13

基金项目: 国家自然科学基金(61271069)

* E-mail: luwei@tju.edu.cn

叶采样等操作实现的医学病灶成像技术^[1]。3D MRI 技术即将顺序排列的二维 MRI 图像序列重建为 3D 图像,是脑肿瘤检测任务中最主要的成像方法。与 X 射线平片、血管造影扫描等二维医学影像相比,3D MRI 可为医生提供空间中任意坐标点的影像信息,方便工作人员进行定量与定性的数学、医学分析。在 3D MRI 过程中,由于造影剂等辅助条件的差异可扫描出不同模态的 MRI 图像。对于脑部 MRI 而言,常见的模态有 4 种:磁共振成像液体衰减反转恢复序列 (FLAIR) 图像、T1 加权像 (T1 weighted image)、T1ce 图像 (post-contrast T1-weighted image) 与 T2 加权像 (T2 weighted image),不同模态的 MRI 图像对肿瘤或其内部不同子区域具有不同的成像效果。例如:T1 加权像可较准确地识别肿瘤核心的位置和轮廓;T2 加权像对水肿区域具有较高识别度;T1ce 可较准确检测肿瘤核心中是否存在增强肿瘤;FLAIR 可较清晰地扫描出肿瘤整体的轮廓。可见,根据多模态 3D MRI,医生可以多角度观察脑肿瘤,也能够获取更充足的脑部信息,避免了信息的遗漏,提高了诊断的准确性。

由于肿瘤微环境的异质性以及脑肿瘤不同子区域的外观不同,从多模态 MRI 图像中进行精细分割是很复杂的。传统分割方法如基于图论与最大流的 graph cut、grab cut、denseCRF 等机器学习算法^[2-3]可实现多种前景与背景的分割,在自然图像上取得了较精准的分割效果。由于这些方法大多通过利用图像的颜色信息与边界反差信息来实现分割,而医学影像的特征不如自然图像清晰,且轮廓欠鲜明,所以分割效果并不理想^[4-8]。

自 2012 年起,随着深度学习的普及以及全卷积网络 (FCN)^[9]的提出,基于深度卷积网络的图像分割方法得到了发展和应用,取得了出色的效果。DeepLab 网络^[10-12]通过引入空洞卷积^[13],在不加降采样层的前提下扩大感受野以获取全局信息,提高分割精度。金字塔场景分解网络 (PSPNet)^[14]在空间金字塔池化 (SPP)^[15]思想的基础上,将不同尺度降采样的特征进行融合,获得更全面的深层信息。类似地,基于卷积神经网络的分割算法在二维自然图像上均取得了较好的分割效果^[16-17]。

对于医学图像而言,V-Net^[18]的提出解决了 3D 医学影像的分割问题,提高了三维图像的分割准确率。V-Net 是一个用 3D 卷积搭建的类似二维 U-Net^[19-20]的网络结构。此外,V-Net 在卷积中加入了残差学习^[21-22],并采用 Dice Loss 函数^[18]的损失函

数,在二类分割问题上取得了很好的效果。由于大多数医学影像的分割问题均属于二分割,V-Net 在医学图像分割领域非常实用。但在脑肿瘤分割任务中,由于需要对肿瘤内部子区域进行精细分割,经实验发现,仅将 V-Net 修改为多类分割并不能取得较好的效果。此外,脑肿瘤分割任务中,如何处理多模态的 3D 图像同样是亟待解决的难点之一。目前,对 3D 图像的处理方式主要有两种。一种是使用 3D 卷积网络进行图像处理,由于 3D 卷积核的空间特点,其提取的特征会更加充分,包含数据各个维度的信息。Kamnitsas 等^[23]使用 3D 卷积神经网络提取脑肿瘤样本的特征,并结合全连接条件随机场等后处理手段完成了多模态脑肿瘤 MRI 图像的分割任务。Chen 等^[24]将残差学习的思想扩展到 3D 卷积,设计了 Voxresnet 的网络结构,同样完成了 3D 图像的分割任务。但是,3D 图像通常数据量较少,且在实际训练过程中 3D 图像所占的图形处理器 (GPU) 显存过大,不利于网络的加深,所以 3D 卷积方法的分割精度难以提高。另一种是使用二维 (2D) 卷积网络进行图像处理。Pereira 等^[25]将 3D 图像裁剪为若干二维图像序列,并使用 2D 卷积核对每张二维图像进行特征提取,获得了较出色的分割效果。Havaei 等^[26]将级联思想运用到分割中来,并通过 2D 卷积进行特征的提取,实现了全自动的脑肿瘤分割。由于每个 3D 图像可裁剪为若干张二维图像,所以 2D 卷积网络的数据量比 3D 卷积网络更加充分。对于深度学习这类数据驱动算法来说,2D 卷积网络的训练效果通常更好。但是,2D 卷积核不能提取空间特征,单纯将 2D 的分割结果拼接成 3D 分割结果,可能会出现锯齿、断层等问题,影响分割的精度。

为解决上述问题,本文提出了一种基于级联的多帧二维卷积神经网络,即级联 2.5 维 (2.5D) 卷积网络,并搭建了一个支持多模态的脑肿瘤精细分割系统。实验证明,所搭建的网络结构在脑肿瘤内各部分均可获得较高的分割精度。在临床诊断中,所提脑肿瘤精细分割系统可以较好地辅助医生诊断。例如,为医生提供辅助信息,医生可根据计算机标明的位置进一步判断肿瘤的性质、判断患者预后,或者辅助医生制定手术方案,根据肿瘤位置确定更准确的开刀位置等。

2 算法原理

2.1 总体框架

在脑肿瘤分割的任务中,若直接对脑部的 3D

MRI 图像进行分割,会造成不同子区域之间的误判以及模糊的边界,且易在脑后、脑干等非肿瘤区域误检出肿瘤。为提高脑肿瘤整体、肿瘤核和增强肿瘤的分割精度、减少假阳性,提出了一种基于级联的多帧二维卷积神经网络,即级联 2.5D 卷积网络。如图 1 所示,级联 2.5D 卷积网络将肿瘤及其子区域分割任务逐级细分为对肿瘤整体(WT,包括肿瘤内各部分)的分割,对肿瘤核(TC,包括增强肿瘤、非增强肿瘤及肿瘤坏死区域)的分割和对增强肿瘤(ET)的分割,这种级联的网络结构提高了脑肿瘤精细分割的准确度。

如图 2 所示,所搭建的级联 2.5D 卷积网络包括

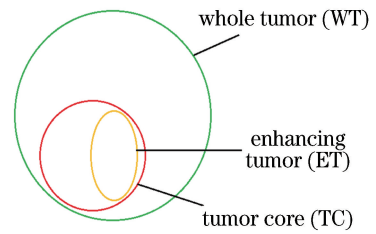


图 1 脑肿瘤整体、肿瘤核和增强肿瘤位置关系图
Fig. 1 Position relationship among whole tumor, tumor core and enhancing tumor

图像截取模块、肿瘤整体分割模块、肿瘤核分割模块和增强肿瘤分割模块。

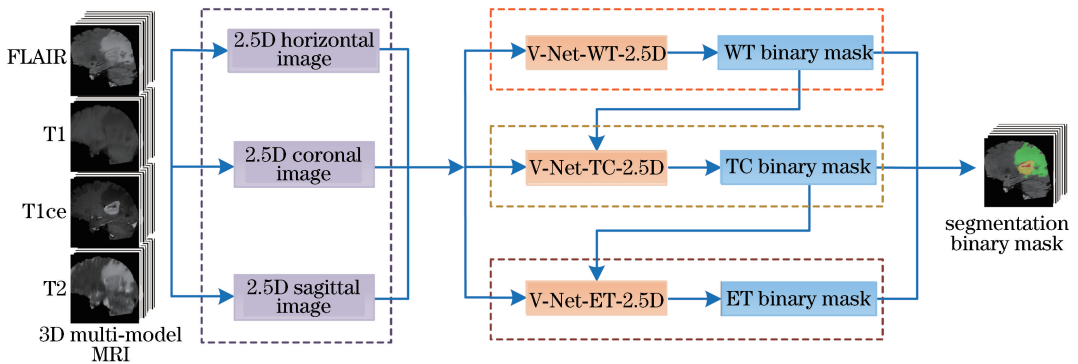


图 2 基于级联 2.5D 卷积神经网络的分割系统

Fig. 2 Segmentation system based on cascaded 2.5D convolutional neural network

1) 对四模态的 3D MRI 进行图像裁取得到 2.5D 图像。2) 将 2.5D 图像输入到肿瘤整体分割模块中进行处理,得到肿瘤整体分割图像。3) 将肿瘤整体的分割图像和 2.5D 图像一起输入到肿瘤核分割模块中,得到肿瘤核的分割图像。4) 将肿瘤核的分割图像与 2.5D 图像一起输入到增强肿瘤分割模块中,得到增强肿瘤的分割图像。5) 将肿瘤整体分割图像、肿瘤核分割图像和增强肿瘤分割图像进行合并,得到最终的肿瘤分割图像。

2.2 图像截取模块

在卷积神经网络的实现过程中,由于 3D 图像占用显存过大,提出了一种多帧的二维图像格式。为和单帧的二维图像区分,称之为 2.5D 图像。假设一个四模态的 3D 图像大小为 $4 \times 128 \text{ pixel} \times 128 \text{ pixel} \times 128 \text{ pixel}$ (4 表示图像的四个模态),则将其看作 128 张大小为 $128 \text{ pixel} \times 128 \text{ pixel}$ 的四模态 2D 图像序列,其中任取连续的 N 张图像即可组成 2.5D 图像。假设 $N = 11$,2.5D 图像大小为 $4 \times 11 \times 128 \text{ pixel} \times 128 \text{ pixel}$,可将其理解为 11 层连续的四模态 2D 图像。在截取过程中,128 张序列图像的任意一张均可作为 11 层连续 2D 图像的中

心,因此,每个 $4 \times 128 \text{ pixel} \times 128 \text{ pixel} \times 128 \text{ pixel}$ 大小的 3D 图像可生成 128 个 $4 \times 11 \times 128 \text{ pixel} \times 128 \text{ pixel}$ 的 2.5D 图像(若中心层在 3D 图像两端则通过向外填充补充 11 层)。此外,由于截取角度的不同,可按照空间中不同坐标轴的指向重复上述截取过程,称之为轴向截取、矢向截取和冠向截取。因此,通过这个过程,每个 $4 \times 128 \text{ pixel} \times 128 \text{ pixel} \times 128 \text{ pixel}$ 大小的 3D 图像可被截取成 $128 \times 3 = 384$ 个 2.5D 图像,用于网络训练。

在训练过程中,2.5D 图像的截取包括图像截取与标签截取两部分。图像截取过程如前文所述,标签截取过程指将 2.5D 图像的中心 3 层的标签图截取出来。例如,前文获得了大小为 $4 \times 11 \times 128 \text{ pixel} \times 128 \text{ pixel}$ 的 2.5D 图像,将其中间 3 层,即大小为 $4 \times 3 \text{ pixel} \times 128 \text{ pixel} \times 128 \text{ pixel}$ 的图像的分割标签图称为 2.5D 标签。

一般来讲,基于卷积网络的分割算法^[23-26]在训练过程中,图像与标签的大小是相同的,但如上文所述,在所提算法中二者的大小并不相同,这是因为所设计的 2.5D V-Net 具有不同大小的输入和输出。本文将在 2.3 小节详细论述 2.5D V-Net 网络的机理。

2.3 2.5D 卷积网络

对经典网络结构 V-Net 进行了改进,搭建了 2.5D V-Net 网络结构。如图 3 所示,网络结构主要

采用层内卷积与层间卷积两种卷积操作实现,由编码器、解码器和跳层连接三个部分组成。

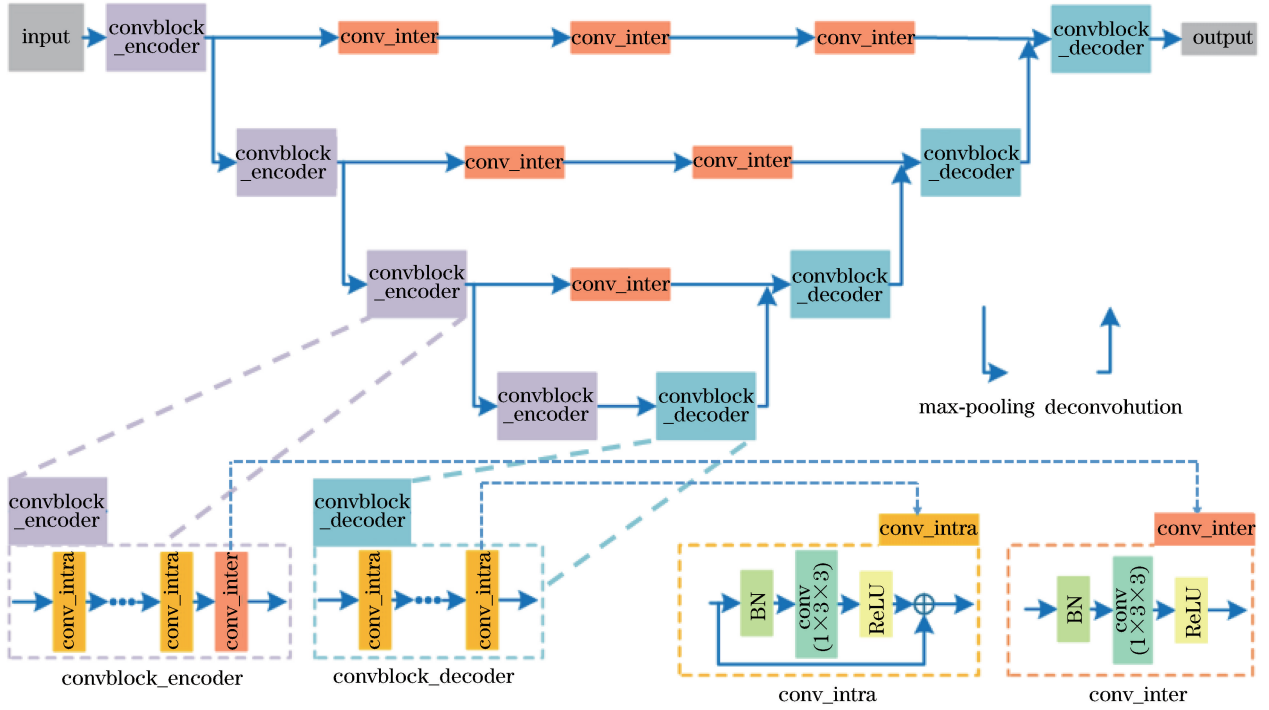


图 3 所提 2.5D V-Net 网络结构图

Fig. 3 Structural diagram of proposed 2.5D V-Net

所采用的层内卷积和层间卷积均通过改变 3D 卷积层的卷积核大小来实现。层内卷积层将 3D 卷积层的卷积核大小设置为 $1 \times 3 \times 3$,使得卷积层只提取层内信息,而不提取层间信息。因此,层内卷积层的实质是同时对多张二维图像进行 2D 卷积,即可用来提取 2.5D 数据中每层平面图像的特征,层内卷积层通过对输入数据进行边缘填零来保持数据输入与输出大小的一致,假设某层内卷积层包含 64 个卷积核,若其输入大小为 $16 \times 11 \times 128 \text{ pixel} \times 128 \text{ pixel}$,则卷积层的输出为 $64 \times 11 \times 128 \text{ pixel} \times 128 \text{ pixel}$ 的特征。层间卷积层通过将 3D 卷积层的卷积核设置为 $3 \times 1 \times 1$,使卷积层只提取层间信息,而不提取层内信息。另外,层间卷积层采用非填零卷积操作,因此卷积本身会改变数据的大小,例如,某层间卷积层包含 64 个卷积核,其输入数据大小为 $64 \times 11 \times 128 \text{ pixel} \times 128 \text{ pixel}$,则其输出为 $64 \times 9 \times 128 \text{ pixel} \times 128 \text{ pixel}$ 的特征。

如图 3 所示,编码器由 4 个编码卷积块组成,用于提取特征与特征降维。其中,每个编码器卷积块

通过多个层内卷积、一个层间卷积、一个降采样层的级联来实现。这里,降采样层采用滑窗大小为 $1 \times 2 \times 2$ 的最大池化层,类似于层内卷积层,降采样层仅对特征进行层内池化,例如某降采样层的输入大小为 $64 \times 9 \times 128 \text{ pixel} \times 128 \text{ pixel}$,则其输出为 $64 \times 9 \times 64 \text{ pixel} \times 64 \text{ pixel}$ 的特征。

解码器由 4 个解码卷积块组成,用于图像重建与尺度恢复。解码卷积块通过一个反卷积层、多个层内卷积层的级联来实现。其中,反卷积的卷积核大小为 $1 \times 2 \times 2$,可以看作是降采样层的反变换,例如某反卷积层的输入大小为 $256 \times 3 \times 32 \text{ pixel} \times 32 \text{ pixel}$,则其输出为 $256 \times 3 \times 64 \text{ pixel} \times 64 \text{ pixel}$ 的特征。

跳层连接部分用于深层特征与浅层特征的融合。由于在每条跳层连接之前,浅层特征与深层特征大小并不相等,因此,本文在不同跳层中加入数量不等的层间卷积层,使两组特征具有空间上相等的大小,方便融合。

通过在 V-Net 的基础上进行改进,实现了其 2.5D 版本。如图 3 所示,在 2.5D V-Net 结构中,每

条支路上的特征均进行了4次层间卷积,因此,假设输入大小为 $4 \times 11 \times 128 \text{ pixel} \times 128 \text{ pixel}$,则网络输出大小为 $c \times 3 \times 128 \text{ pixel} \times 128 \text{ pixel}$,即输入的2.5D图像中间部分的分割结果,其中 c 表示分割的种类数。例如,若分割种类包括肿瘤与背景两类,

则 $c=2$ 。

2.4 肿瘤整体分割模块

为降低脑后、脑干等区域的假阳性,提高肿瘤整体的分割精度,设计了一个肿瘤整体分割模块,得到肿瘤整体的分割图像,如图4所示。

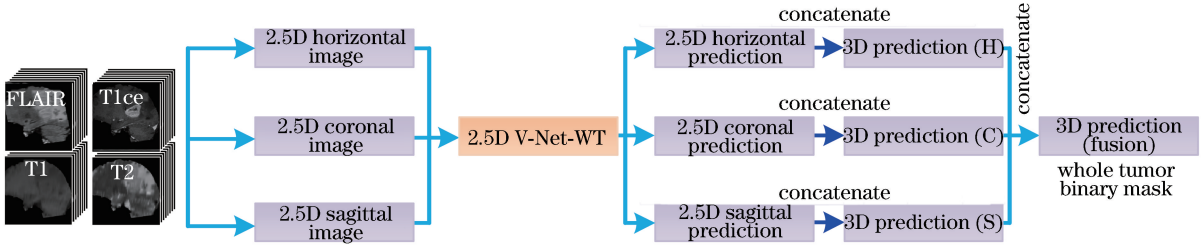


图4 肿瘤整体分割模块结构图

Fig. 4 Structural diagram of whole tumor segmentation module

首先,对3D MRI进行裁切,由于脑肿瘤仅存在于脑内区域,而脑内区域仅占3D MRI的中心部分,因此在训练网络之前,对原始MRI进行裁切,取出其中心的 $4 \times 128 \text{ pixel} \times 128 \text{ pixel} \times 128 \text{ pixel}$ 以进行2.2小节所述的2.5D图像裁取工作,可获得128个轴向2.5D图像、128个矢向2.5D图像和128个冠向2.5D图像。将这3组图像输入到2.3小节所述2.5D V-Net中,可得到3组2.5D分割图像。通过拼接每组2.5D分割图像,可以获得3个3D分割

图像,之后,对3个分割图像进行融合,可获得最终的肿瘤整体分割图像。

2.5 肿瘤核分割模块

由于在进行肿瘤核分割时,部分样本易造成边界预测不准确,假阳性显著的问题。为获得准确的肿瘤核边界,采用一个肿瘤核分割模块,将肿瘤整体分割图像作为辅助信息,同4个模态的2.5D图像共同作为分割网络的输入,以此来修正肿瘤核的分割边界,提高分割精度,分割流程图如图5所示。

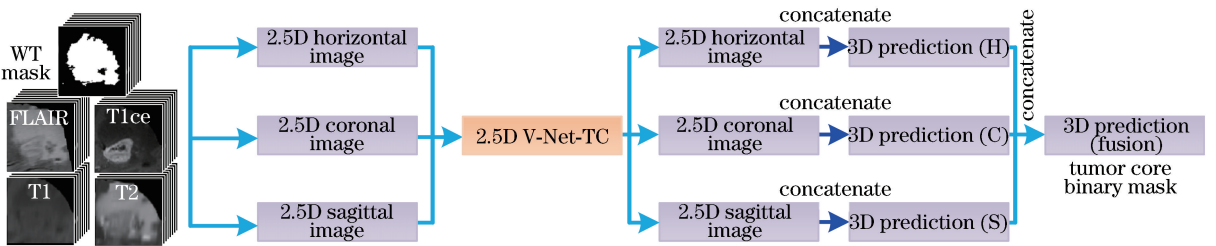


图5 肿瘤核分割模块结构图

Fig. 5 Structural diagram of tumor core segmentation module

首先,将肿瘤整体的分割结果与前一步大小为 $4 \times 128 \text{ pixel} \times 128 \text{ pixel} \times 128 \text{ pixel}$ 的3D MRI进行合并,得到大小为 $5 \times 128 \text{ pixel} \times 128 \text{ pixel} \times 128 \text{ pixel}$ 的3D图像,其中5表示4个MRI模态与1个肿瘤整体分割结果的叠加。之后,利用肿瘤整体分割结果对肿瘤进行定位,将肿瘤所在体素块裁切出来,并对其进行尺度归一化处理,生成大小为 $5 \times 96 \text{ pixel} \times 96 \text{ pixel} \times 96 \text{ pixel}$ 的3D MRI以进行2.5D图像裁取。类似肿瘤整体分割过程,将轴向、矢向、冠向2.5D数据输入网络,获得3组肿瘤核的2.5D分割图像。通过对每组2.5D分割图像进行拼接,可以获得3个3D分割图像,之后,通过对3个

分割图像进行融合,可获得最终的肿瘤核分割图像。

2.6 增强肿瘤分割模块

利用一个增强肿瘤分割模块,将肿瘤核分割图像作为辅助信息,同4个模态的2.5D图像共同作为分割网络的输入,以此来获得较准确的增强肿瘤分割结果,分割流程图如图6所示。

与2.5小节类似,将肿瘤核的分割结果与前一步大小为 $4 \times 96 \text{ pixel} \times 96 \text{ pixel} \times 96 \text{ pixel}$ 的3D MRI图像进行合并,生成大小为 $5 \times 96 \text{ pixel} \times 96 \text{ pixel} \times 96 \text{ pixel}$ 的3D图像,其中5表示4个MRI模态与1个肿瘤核分割结果的叠加。之后,利用肿瘤核分割结果对肿瘤核进行定位,将肿瘤核所

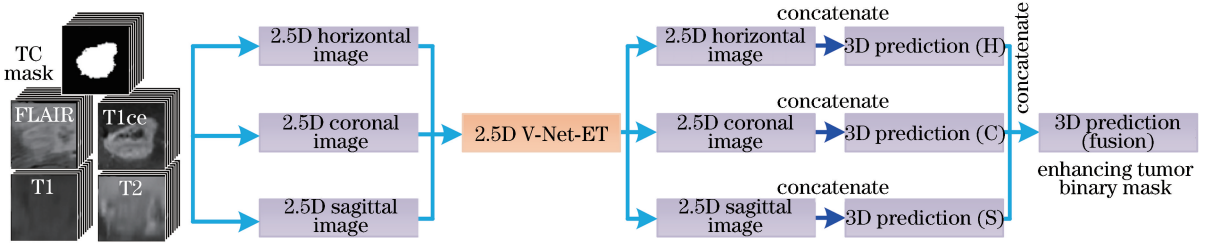


图6 增强肿瘤分割模块结构图

Fig. 6 Structural diagram of enhancing tumor segmentation module

在体素块裁切出来,并对其进行尺度归一化处理,生成大小为 $5 \times 64 \text{ pixel} \times 64 \text{ pixel} \times 64 \text{ pixel}$ 的 3D MRI 以进行 2.5D 图像裁取。类似肿瘤核分割过程,将轴向、矢向、冠向 2.5D 数据输入网络,获得 3 组增强肿瘤的 2.5D 分割图像。通过对每组 2.5D 分割图像进行拼接,可以获得 3 个 3D 分割图像,之后,通过对 3 个分割图像进行融合,可获得最终的增强肿瘤分割图像。

2.7 损失函数

Dice Loss 是一种常用于二类分割问题的损失函数,其表达式为

$$D = 1 - \frac{2 \times \sum_{i \in V} p_i \cdot l_i}{\sum_{i \in V} p_i^2 + \sum_{i \in V} l_i^2}, \quad (1)$$

式中: V 为 3D 图像中的所有体素点; p_i 为体素点 i 被预测为目标类的概率; l_i 为该体素点的实际标签。类似地,Jaccard Loss 也是可用作二类分割问题的损失函数,其表达式为

$$J = 1 - \frac{\sum_{i \in V} p_i \cdot l_i}{\sum_{i \in V} p_i^2 + \sum_{i \in V} l_i^2 - \sum_{i \in V} p_i \cdot l_i}. \quad (2)$$

Dice Loss^[18] 具有收敛快、精度高的优点,但在训练后期随着分母数值的增加,收敛速度骤降,不利于分割精度进一步提高。Jaccard Loss^[27] 通过在分母中减去 $\sum_{i \in V} p_i \cdot l_i$,有效避免了该问题,但其在训练前期收敛较慢。此外,由于该损失函数对假阳性体素的惩罚力度较小,容易造成分割结果较大、边界不准。因此,提出一种加权的损失函数,其表达式为

$$L = \alpha \sum_k D_k + (1 - \alpha) \sum_k J_k, \quad (3)$$

式中: D_k 和 J_k 分别为目标类为第 k 类的 Dice Loss 与 Jaccard Loss; α 为二者的权重关系。训练过程中,

$$\alpha = 0.99t_{ep}, \quad (4)$$

式中: t_{ep} 为该训练过程所处的训练周期次数; α 随 t_{ep} 的增加而减小。

3 实验及结果

3.1 数据集与评价指标

本文所用数据集来源为 BraTS2018^[28-30],数据包括 210 例 HGG 数据和 75 例 LGG 样本,每例数据包括配准后的 FLAIR、T1、T1ce、T2 四种模态的 3D MRI 影像,不同模态图像的相同坐标体素值对应患者脑中的同一位置,图像大小为 $155 \text{ pixel} \times 240 \text{ pixel} \times 240 \text{ pixel}$ 。如图 7 所示,胶质瘤可分为多种异质性子区域,如水肿区域(绿色)、坏死核心与非增强肿瘤核心区域(红色)、增强肿瘤核心区域(黄色)。本文将 285 例样本随机分为数量相当的 10 份,以进行十折交叉验证。

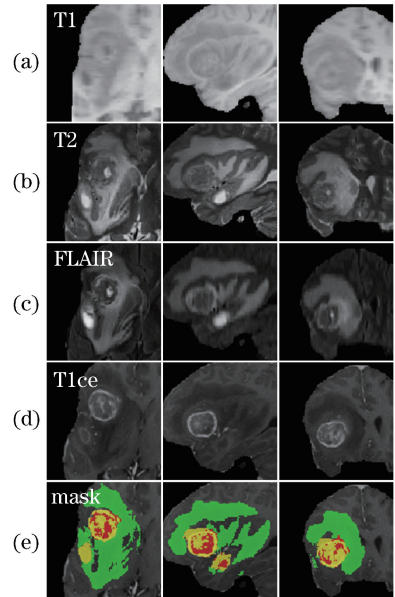


图7 不同模态的 MRI 在三个截面上的图像。(a) T1 图像;(b) T2 图像;(c) FLAIR 图像;(d) T1ce 图像;(e)精细分割标签图

Fig. 7 MRI images in three sections for different modes. (a) T1 images; (b) T2 images; (c) FLAIR images; (d) T1ce images; (e) label mask for fine segmentation

采用相似系数 Dice 作为评价指标用于评估分割结果,通过比较脑肿瘤及其各部分的 Dice 值来评估精细分割的效果,Dice 值越高表示分割精度越高。Dice 的计算公式为

$$R_{\text{Dice}} = \frac{2 \times |V_{\text{Pred}} \cap V_{\text{GT}}|}{|V_{\text{Pred}}| + |V_{\text{GT}}|}, \quad (5)$$

式中: V_{Pred} 为网络输出的分割结果二值图; V_{GT} 为数据给定的分割标签二值图; \cap 为两张二值图对应体素点之间的与运算; $|\cdot|$ 为二值图的代数求和运算。

3.2 参数设置

类似其他 3D 医学图像的预处理步骤,对多模态 MRI 图像的预处理包括归一化、裁取及数据增

强^[23-26]。裁取并归一化后的图像是原图中心的 128 pixel×224 pixel×224 pixel 的 3D 体素块;本文采用的数据增强方法包括图像翻转、图像旋转和图像弹性形变。本文网络结构基于 PyTorch^[31] 框架实现,所用 GPU 为 Nvidia GTX 1080Ti。训练过程中采用 ADAM^[32] 优化器,优化器中的相关参数设置为 $\alpha=0.1, \beta=0.001$,学习率初始化为 5×10^{-3} 。

3.3 结果分析

分析讨论了级联 V-Net(Cascaded V-Net)与传统端到端分割卷积神经网络的效果。如表 1 所示,实验对比了传统多类分割 V-Net 与 3D、2D、2.5D 级联分割 V-Net 的效果,表中数据为各算法在训练集上的交叉验证结果。

表 1 十折交叉验证的各类算法比较

Table 1 Comparison of various algorithms for ten-fold cross-validation

Method	End-to-end-V-Net	Cascaded V-Net-3D	Cascaded V-Net-2D	Cascaded V-Net-2.5D
Dice WT	0.8714	0.8832	0.8986	0.9071
Dice TC	0.7562	0.7924	0.8413	0.8542
Dice ET	0.6524	0.7122	0.7919	0.8140

表 1 中,Dice WT、Dice TC、Dice ET 分别表示肿瘤整体(包含肿瘤各部分)、肿瘤核(包含坏死核心区域及增强/非增强肿瘤核心区域)、增强肿瘤的分割相似系数。End-to-end-V-Net 所表示的网络结构是与 V-Net-WT 网络结构相似的 3D 卷积分割网络。在训练过程中,网络输出为 4 通道的热力图,分别代表各体素点属于水肿、增强肿瘤、非增强肿瘤及坏死、背景的概率。

由表 1 可知,Cascaded V-Net-3D 网络相比同等规模的 End-to-end-V-Net 网络表现出更加突出的网络性能,这说明对于脑瘤精细分割的任务,将多类分割转换为多个二类分割网络的级联有助于降低网络的训练难度,并提高脑肿瘤的分割精度。

表 1 中还对比了基于 V-Net 的 3D、2D、2.5D 网络对分割效果的影响。经实验发现,2D 网络相比 3D 网络具有更好的网络性能。一种合理的解释为,2D 网络具有更充分的数据量,假设输入样本大小为 $z \times h \times w$,2D 网络的数据量为 3D 网络的 $z+h+w$ 倍,因此 2D 网络的分割结果更加准确。对于增强肿瘤区域,2D 网络的 Dice-ET 比 3D 网络提高了 0.0797,假阳性样本点明显减少。该现象说明,3D 网络结构由于样本稀缺,易预测较多的假阳性样本点,对于增强肿瘤分割,部分 LGG 样本不存在或仅存在少量该区域,因此 3D 网络的假阳性缺陷更明显。此外,与 2D 网络相比,2.5D 网络进一步提高了

分割精度。因为 2.5D 网络提取层间信息,对于每个体素点,可以提取到更加丰富的特征。而且 2.5D 网络具有与 2D 网络规模相同的训练样本数量,可集成 2D 与 3D 网络的优势,获得较高的性能。表 1 中的各网络分割效果如图 8 所示。

图 8 和图 9 所用样例 1、2 来源均为 BraTS2018 数据集的验证集部分,由于该部分数据标签未公开,未列出所示样例的金标准。2.5D V-Net 有效减少了各个分割模块中的假阳性体素点,且可预测出更加准确的轮廓。

如表 2 所示,实验分析了 2.5D V-Net 对轴向、矢向及冠向的分割效果及融合之后的分割效果,证明了多方向融合的分割精度高于单方向。从表 2 可以看出,三个方向的分割结果有差异,一种合理的解释是,由于不同方向上的图像标注本身存在优劣差异,且这在图像标注工作中不可避免,因此经训练的网络模型会体现出不同方向的预测差异。如表 2 所示,冠向预测结果通常低于另外两个方向。轴向、矢向及冠向融合前后的预测效果如图 10 所示。由图 10 可见,融合可解决网络的假阳性问题,同时提高分割的边界精度。

最后,对采用不同损失函数的分割效果进行了对比。如表 3 所示,Dice Loss 比 Cross Entropy Loss 略有提高。一种合理的解释为,本文第 2 节中所提的三个卷积网络,分割任务类别较少,且不同

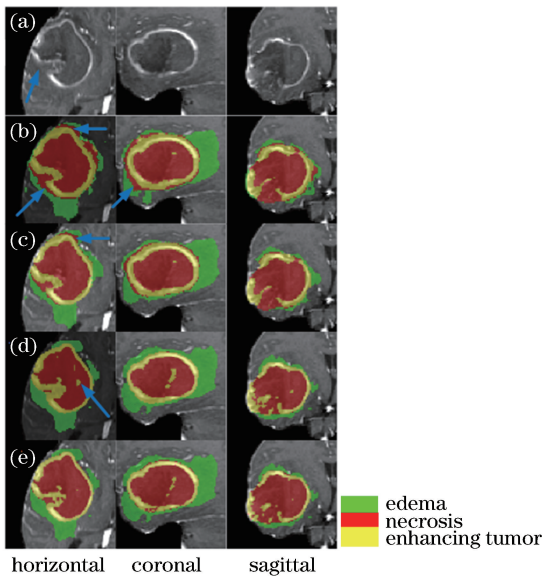


图 8 不同网络结构在三个方向对样例 1 的分割结果。(a) T1ce 图像;(b) end-to-end V-Net;(c) cascaded V-Net-3D;(d) cascaded V-Net-2D;(e) cascaded V-Net-2.5D

Fig. 8 Segmentation results of different networks in three directions for sample 1. (a) T1ce images; (b) end-to-end V-Net; (c) cascaded V-Net-3D; (d) cascaded V-Net-2D; (e) cascaded V-Net-2.5D

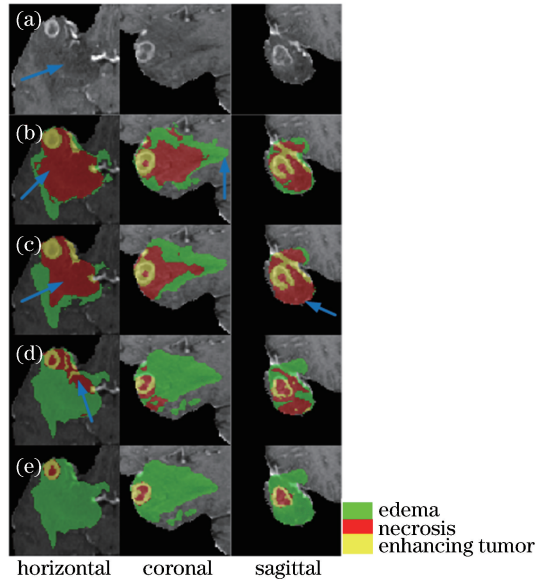


图 9 不同网络结构在三个方向对样例 2 的分割结果。(a) T1ce 图像;(b) end-to-end V-Net;(c) cascaded V-Net-3D;(d) cascaded V-Net-2D;(e) cascaded V-Net-2.5D

Fig. 9 Segmentation results of different networks in three directions for sample 2. (a) T1ce images; (b) end-to-end V-Net; (c) cascaded V-Net-3D; (d) cascaded V-Net-2D; (e) cascaded V-Net-2.5D

表 2 输入不同方向图片时 2.5D V-Net 分割效果比较

Table 2 Segmentation performance comparison of 2.5D V-Net fed by images in different directions

Direction	Horizontal prediction	Coronal prediction	Sagittal prediction	Fused prediction
Dice WT	0.8989	0.8808	0.8912	0.9071
Dice TC	0.8457	0.8367	0.8401	0.8542
Dice ET	0.8081	0.7922	0.8001	0.8140

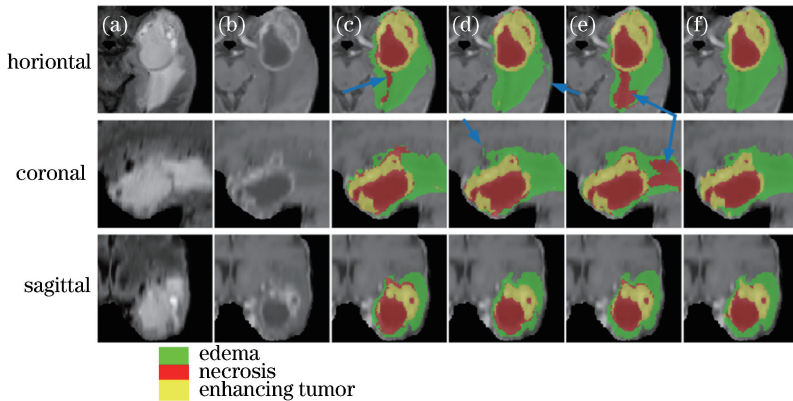


图 10 融合前后分割结果的定性对比。(a) FLAIR 图像;(b) T1ce 图像;(c)水平向;(d)冠向;(e)矢向;(e)融合结果
Fig. 10 Qualitative comparison of segmentation results before and after fusion. (a) FLAIR images; (b) T1ce images; (c) horizontal prediction; (d) coronal prediction; (e) sagittal prediction; (f) fused prediction

类别样本点差异较大,因此 Cross Entropy Loss 难以体现其优势。另外,当训练进行到一定阶段,Dice Loss 会逐渐降低,对于仍然误分割的点惩罚减弱,

因此文中所提 Dice Loss 与 Jaccard Loss 的加权损失函数,会进一步优化分割效果。

表3 基于不同损失函数的分割效果对比

Table 3 Segmentation performance comparison based on different loss functions

Loss function	Cross entropy	Dice loss	Dice loss+ Jaccard loss
Dice WT	0.9042	0.9051	0.9071
Dice TC	0.8411	0.8501	0.8542
Dice ET	0.7908	0.8089	0.8140

4 结 论

提出了一种基于级联卷积神经网络的多模态脑肿瘤精细分割,通过实验对比,论证了级联的二类分割网络结构与端到端的卷积神经网络结构相比,对脑肿瘤的精细分割准确率更高。此外,还对比分析了在卷积神经网络结构相同的前提下,采用3D、2D和2.5D网络性能的优劣。经实验证实,2.5D网络在3D图像精细分割中最具优势。因为与3D网络相比,2.5D网络具有更大的样本规模,而且可以通过多方向的融合提高分割精度;与2D网络相比,2.5D网络可以提取更丰富的图像特征用以预测。最后,提出了加权损失函数Dice Loss与Jaccard Loss,可进一步提升网络分割效果。最终实现的级联卷积网络可实现脑肿瘤(包括高级别细胞瘤与低级别胶质瘤)的高精度精细分割,且具有较高稳健性。

参 考 文 献

- [1] Mayer G S, Vrscay E R. Self-similarity of Fourier domain MRI data[J]. *Nonlinear Analysis: Theory, Methods & Applications*, 2009, 71(12): e855-e864.
- [2] Wolf D, Prankl J, Vincze M. Fast semantic segmentation of 3D point clouds using a dense CRF with learned parameters [C] // *International Conference on Robotics and Automation (ICRA)*, May 26-30, 2015, Seattle, WA, USA. New York: IEEE, 2015: 4867-4873.
- [3] Freedman D, Zhang T. Interactive graph cut based segmentation with shape priors[C] // *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, June 20-25, 2005, San Diego, CA, USA. New York: IEEE, 2005: 755-762.
- [4] Ren L, Li Q, Guan X, *et al.* 3D segmentation of brain tumors in MRI based on improved continuous max-flow[J]. *Laser & Optoelectronics Progress*, 2018, 55(11): 111011.
任璐, 李镝, 关欣, 等. 改进的连续型最大流算法的脑肿瘤MRI三维分割[J]. *激光与光电子学进展*,

2018, 55(11): 111011.

- [5] Li R Z, Liu Y Y, Yang M, *et al.* Three-dimensional point cloud segmentation algorithm based on improved region growing [J]. *Laser & Optoelectronics Progress*, 2018, 55(5): 051502.
李仁忠, 刘阳阳, 杨曼, 等. 基于改进的区域生长三维点云分割[J]. *激光与光电子学进展*, 2018, 55(5): 051502.
- [6] Xie Z N, Zheng D, Chen J Y, *et al.* A tumor segmentation method of improved Chan-Vese model for liver cancer ablation computed tomography image [J]. *Laser & Optoelectronics Progress*, 2017, 54(2): 021702.
谢志南, 郑东, 陈嘉耀, 等. 改进Chan-Vese模型的肝癌消融CT图像肿块分割方法[J]. *激光与光电子学进展*, 2017, 54(2): 021702.
- [7] Chu J H, Wang X Y, Lü W. 3D segmentation of breast MRI based on inter-frame correlations [J]. *Journal of Tianjin University*, 2017, 50(8): 835-842.
褚晶辉, 王星宇, 吕卫. 基于帧间相关性的乳腺MRI三维分割[J]. *天津大学学报*, 2017, 50(8): 835-842.
- [8] Yao H B, Bian J W, Cong J W, *et al.* Medical image segmentation model based on local sparse shape representation [J]. *Laser & Optoelectronics Progress*, 2018, 55(5): 051011.
姚红兵, 卞锦文, 丛嘉伟, 等. 基于局部稀疏形状表示的医学图像分割模型[J]. *激光与光电子学进展*, 2018, 55(5): 051011.
- [9] 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.
- [10] Chen L C, Papandreou G, Kokkinos I, *et al.* DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [11] Chen L C, Papandreou G, Kokkinos I, *et al.* Semantic image segmentation with deep convolutional nets and fully connected CRFs [EB/OL]. (2016-06-07) [2018-06-30]. <https://arxiv.org/abs/1412.7062>.
- [12] Chen L C, Papandreou G, Schroff F, *et al.* Rethinking atrous convolution for semantic image segmentation [EB/OL]. (2017-12-05) [2018-06-30]. <https://arxiv.org/abs/1706.05587>.

- [13] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions [EB/OL]. (2016-04-30) [2018-06-30]. <https://arxiv.org/abs/1511.07122>.
- [14] Zhao H S, Shi J P, Qi X J, *et al.* Pyramid scene parsing network[C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 3690-3695.
- [15] He K M, Zhang X Y, Ren S Q, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition[M] // He K M, Zhang X Y, Ren S Q, *et al.* Computer Vision - ECCV 2014. Cham: Springer International Publishing, 2014: 346-361.
- [16] Guo C C, Yu F Q, Chen Y. Image semantic segmentation based on convolutional neural network feature and improved superpixel matching[J]. Laser & Optoelectronics Progress, 2018, 55(8): 081005.
郭呈呈, 于凤芹, 陈莹. 基于卷积神经网络特征和改进超像素匹配的图像语义分割[J]. 激光与光电子学进展, 2018, 55(8): 081005.
- [17] An Z, Xu X P, Yang J H, *et al.* Design of augmented reality head-up display system based on image semantic segmentation[J]. Acta Optica Sinica, 2018, 38(7): 0710004.
安喆, 徐熙平, 杨进华, 等. 结合图像语义分割的增强现实型平视显示系统设计与研究[J]. 光学学报, 2018, 38(7): 0710004.
- [18] Milletari F, Navab N, Ahmadi S A. V-Net: fully convolutional neural networks for volumetric medical image segmentation [C] // Fourth International Conference on 3D Vision (3DV), October 25-28, 2016, Stanford, CA, USA. New York: IEEE, 2016: 565-571.
- [19] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M] // Ronneberger O, Fischer P, Brox T. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241.
- [20] Çiçek Ö, Abdulkadir A, Lienkamp S S, *et al.* 3D U-Net: learning dense volumetric segmentation from sparse annotation [M] // Çiçek Ö, Abdulkadir A, Lienkamp S S, *et al.* Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016. Cham: Springer International Publishing, 2016: 424-432.
- [21] Szegedy C, Ioffe S, Vanhoucke V, *et al.* Inception-v4, inception-resnet and the impact of residual connections on learning [EB/OL]. (2016-08-23) [2018-06-30]. <https://arxiv.org/abs/1602.07261>.
- [22] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition[C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [23] Kamnitsas K, Ledig C, Newcombe V F J, *et al.* Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation [J]. Medical Image Analysis, 2017, 36: 61-78.
- [24] Chen H, Dou Q, Yu L, *et al.* VoxResNet: deep voxelwise residual networks for volumetric brain segmentation[EB/OL]. (2016-04-21) [2018-06-30]. <https://arxiv.org/abs/1608.05895>.
- [25] 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.
- [26] Havaei M, Davy A, Warde-Farley D, *et al.* Brain tumor segmentation with deep neural networks[J]. Medical Image Analysis, 2017, 35: 18-31.
- [27] Cai J, Lu L, Xie Y, *et al.* Improving deep pancreas segmentation in CT and MRI images via recurrent neural contextual learning and direct loss function [EB/OL]. (2017-07-18) [2018-06-30]. <https://arxiv.org/abs/1707.04912>.
- [28] Menze B H, Jakab A, Bauer S, *et al.* The multimodal brain tumor image segmentation benchmark (BRATS) [C] // IEEE Transactions on Medical Imaging, December 4, 2014. New York: IEEE, 2015: 1993-2024.
- [29] Bakas S, Akbari H, Sotiras A, *et al.* Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features[J]. Scientific Data, 2017, 4: 170117.
- [30] Bakas S, Akbari H, Sotiras A, *et al.* Segmentation labels and radiomic features for the pre-operative scans of the TCGA-GBM collection[EB/OL]. [2018-08-30]. <https://doi.org/10.7937/K9/TCIA.2017.KLXWJJ1Q>.
- [31] Paszke A, Gross S, Chintala S, *et al.* Automatic differentiation in PyTorch[C/OL] // 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA. [2018-08-30]. <https://openreview.net/pdf?id=BJJsrmiCZ>.
- [32] Kingma D P, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2017-01-30) [2018-08-30]. <https://arxiv.org/abs/1412.6980>.