

中图分类号: TP391.7 文献标识码: A 文章编号: 1006-8961(2020)10-2159-12

论文引用格式: Zhao Y M, Li Q and Guan X. 2020. Lightweight brain tumor segmentation algorithm based on a group convolutional neural network. Journal of Image and Graphics 25(10): 2159-2170(赵奕名, 李强, 关欣. 2020. 组卷积轻量级脑肿瘤分割网络. 中国图象图形学报, 25(10): 2159-2170 [DOI: 10.11834/jig.200247])

组卷积轻量级脑肿瘤分割网络

赵奕名, 李强, 关欣

天津大学微电子学院, 天津 300072

摘要: 目的 脑肿瘤是一种严重威胁人类健康的疾病。利用计算机辅助诊断进行脑肿瘤分割对于患者的预后和治疗具有重要的临床意义。3D 卷积神经网络因具有空间特征提取充分、分割效果好等优点, 广泛应用于脑肿瘤分割领域。但由于其存在显存占用量巨大、对硬件资源要求较高等问题, 通常需要在网络结构中做出折衷, 以牺牲精度或训练速度的方式来适应给定的内存预算。基于以上问题, 提出一种轻量级分割算法。方法 使用组卷积来代替常规卷积以显著降低显存占用, 并通过多纤单元与通道混合单元增强各组间信息交流。为充分利用多显卡协同计算的优势, 使用跨卡同步批量归一化以缓解 3D 卷积神经网络因批量值过小所导致的训练效果差等问题。最后提出一种加权混合损失函数, 提高分割准确性的同时加快模型收敛速度。结果 使用脑肿瘤公开数据集 BraTS2018 进行测试, 本文算法在肿瘤整体区、肿瘤核心区和肿瘤增强区的平均 Dice 值分别可达 90.67%、85.06% 和 80.41%, 参数量和计算量分别为 3.2 M 和 20.51 G, 与当前脑肿瘤分割最优算法相比, 其精度分别仅相差 0.01%、0.96% 和 1.32%, 但在参数量和计算量方面分别降低至对比算法的 1/12 和 1/73。结论 本文算法通过加权混合损失函数来提高稀疏分类错误对模型的惩罚, 有效平衡不同分割难度类别的训练强度, 本文算法可在保持较高精度的同时显著降低计算消耗, 为临床医师进行脑肿瘤分割提供有力参考。

关键词: 核磁共振成像; 脑肿瘤分割; 深度学习; 组卷积; 加权混合损失函数

Lightweight brain tumor segmentation algorithm based on a group convolutional neural network

Zhao Yiming, Li Qiang, Guan Xin

School of Microelectronics, Tianjin University, Tianjin 300072, China

Abstract: **Objective** Brain tumor is a serious threat to human health. The invasive growth of brain tumor, when it occupies a certain space in the skull, will lead to increased intracranial pressure and compression of brain tissue, which will damage the central nerve and even threaten the patient's life. Therefore, effective brain tumor diagnosis and timely treatment are of great significance to improving the patient's quality of life and prolonging the patient's life. Computer-assisted segmentation of brain tumor is necessary for the prognosis and treatment of patients. However, although brain-related research has made great progress, automatic identification of the contour information of tumor and effective segmentation of each sub-region in MRI (magnetic resonance imaging) remain difficult due to the highly heterogeneous appearance, random location,

收稿日期: 2020-06-01; 修回日期: 2020-07-24; 预印本日期: 2020-07-31

基金项目: 国家自然科学基金项目(61471263); 天津市自然科学基金项目(16JCZDJC31100)

Supported by: National Natural Science Foundation of China(61471263); Tianjin Municipal Natural Science Foundation(16JCZDJC31100)

and large difference in the number of voxels in each subregion of the tumor and the high degree of gray-scale similarity between the tumor tissue and neighboring normal brain tissue. Since 2012 , with the development of deep learning and the improvement of related hardware performance , segmentation methods based on neural networks have gradually become the mainstream. In particular , 3D convolutional neural networks are widely used in the field of brain tumor segmentation because of their advantages of sufficient spatial feature extraction and high segmentation effect. Nonetheless , their large memory consumption and high requirements on hardware resources usually require making a compromise in the network structure that adapts to the given memory budget at the expense of accuracy or training speed. To address such problems , we propose a lightweight segmentation algorithm in this paper. **Method** First , group convolution was used to replace conventional convolution for significantly reducing the parameters and improving segmentation accuracy because memory consumption is negatively correlated with batch size. A large batch size usually means enhanced convergence stability and training effect in 3D convolutional neural networks. Then , multifiber and channel shuffle units were used to enhance the information fusion among the groups and compensate for the poor communication caused by group convolution. Synchronized cross-GPU batch normalization was used to alleviate the poor training performance of 3D convolutional neural networks due to the small batch size and utilize the advantages of multigraphics collaborative computing. Aiming at the case in which the subregions have different difficulties in segmentation , a weighted mixed-loss function consisting of Dice and Jaccard losses was proposed to improve the segmentation accuracy of the subregions that are difficult to segment under the premise of maintaining the high precision of the easily segmented subregions and accelerate the model convergence speed. One of the most challenging parts of the task is to distinguish between small blood vessels in the tumor core and enhanced-tumor areas. This process is particularly difficult for the labels that may not have enhanced tumor at all. If neither the ground truth nor the prediction has an enhanced area , the Dice score of the enhancement area is 1. Conversely , in patients who did not have enhanced tumors in the ground truth , only a single false-positive voxel would result in a Dice score of 0. Hence , we post-processed the prediction results , that is , we set a threshold for the number of voxels in the tumor-enhanced area. When the number of voxels in the tumor-enhanced area is less than the threshold , these voxels would be merged into the tumor core area , thereby improving the Dice score of the tumor-enhanced and tumor core areas. **Result** To verify the overall performance of the algorithm , we first conducted a fivefold cross-validation evaluation on the training set of the public brain tumor dataset BraTS2018. The average Dice scores of the proposed algorithm in the entire tumor , tumor core , and enhanced tumor areas can reach 89.52% , 82.74% , and 77.19% , respectively. For fairness , an experiment was also conducted on the BraTS2018 validation set. We used the trained network to segment the unlabeled samples for prediction , converted them into the corresponding format , and uploaded them to the BraTS online server. The segmentation results were provided by the server after calculation and analysis. The proposed algorithm achieves average Dice scores of 90.67% , 85.06% , and 80.41% . The parameters and floating point operations are 3.2 M and 20.51 G , respectively. Compared with the classic 3D U-Net , our algorithm shows higher average Dice scores by 2.14% , 13.29% , and 4.45% . Moreover , the parameters and floating point operations are reduced by 5 and 81 times , respectively. Compared with the state-of-the-art approach that won the first place in the 2018 Multimodal Brain Tumor Segmentation Challenge , the average Dice scores are reduced by only 0.01% , 0.96% , and 1.32% . Nevertheless , the parameters and floating point operations are reduced by 12 and 73 times , respectively , indicating a more practical value. **Conclusion** Aiming at the problems of large memory consumption and slow segmentation speed in the field of computer-aided brain tumor segmentation , an algorithm combining group convolution and channel shuffle unit is proposed. The punishment intensity of sparse class classification error to model is improved using the weighted mixed loss function to balance the training intensity of different segmentation difficulty categories effectively. The experimental results show that the algorithm can significantly reduce the computational cost while maintaining high accuracy and provide a powerful reference for clinicians in brain tumor segmentation.

Key words: magnetic resonance imaging (MRI) ; brain tumor segmentation; deep learning; group convolution; weighted mixed loss function

0 引言

脑肿瘤为生长于颅内的不正常细胞群(Deangelis 2001),是一种严重危害患者生命的肿瘤。根据《临床医师癌症杂志》(A Cancer Journal for Clinicians, CA)全球癌症统计报告(Bray等2018),截至2018年,脑肿瘤新发病例约为29.7万,约占所有新发病例的1.6%,死亡病例约为24.1万,约占所有癌症死亡病例的2.5%。脑肿瘤主要分为在脑中或在源自大脑神经中形成的原发性脑肿瘤和由身体其他部位转移至颅内的继发性脑肿瘤。成人最常见的原发性脑肿瘤是原发性中枢神经系统淋巴瘤和胶质瘤,其中胶质瘤起源于神经胶质细胞周围组织,占恶性脑肿瘤的80%以上,随病变区域不同而出现不同症状如头痛、呕吐、视力衰退、癫痫和意识模糊等。

根据侵入程度及患者预后,可将胶质瘤分为高级别胶质瘤(high-grade gliomas, HGG)和低级别胶质瘤(low-grade gliomas, LGG)。高级别胶质瘤患者死亡率较高,中位生存时间仅为15个月,存活两年的患者数量较少(Menze等2015);低级别胶质瘤发展较为缓慢,治疗效果较好,患者预期寿命通常在十年以上,但部分患者会在肿瘤切除后的数年内发生肿瘤恶变,若恶变没有及时发现,短期内肿瘤可迅速增大,对患者健康造成巨大威胁(李锦等2020)。由于胶质瘤呈浸润性生长,当其在颅内占据一定空间时,无论何种类别的胶质瘤,都会导致颅内压升

高,压迫脑组织等情况,使得中枢神经受损,严重时危及患者生命,因此有效的脑肿瘤诊断和及时的治疗对于提升患者生存质量、延长患者寿命有着重大意义。

随着医疗水平的不断提高,医学成像技术(medical imaging techniques)、肿瘤监测(tumor monitoring)和患者预后预测(patient outcome prediction)等在脑肿瘤治疗方面发挥了越来越重要的作用。由于磁共振成像(magnetic resonance imaging, MRI)是一种非侵入性体内成像技术,使用射频信号在非常强大的磁场影响下激发目标组织以产生其内部图像,具有良好的软组织对比度、无放射线损害、通过改变造影剂等辅助条件来实现多模态多参数成像等独特的优点(Liang和Lauterbur 2000),因此磁共振成像较适于脑部病变检测。对于脑部MRI来说,常见的模态有如下4种(见图1):FLAIR(fluid attenuated inversion recovery)模态图像、T1(T1-weighted image)模态图像、T1C(contrast-enhanced T1-weighted image)模态图像以及T2(T2-weighted image)模态图像,不同模态对于肿瘤的各个子区域有不同的成像效果,例如FLAIR模态凸显脑肿瘤周边水肿区域(peritumoral edema),T1C模态图像凸显增强肿瘤区域(enhancing tumor, ET),T1模态能较准确识别由肿瘤增强区及坏死(necrotic)和非增强区(non-enhancing tumor)组成的肿瘤核心区轮廓(tumor core, TC),T2模态图像可较清晰地将由肿瘤核心区及其周边水肿所构成的整体区(whole tumor, WT)和健康组织区分开。

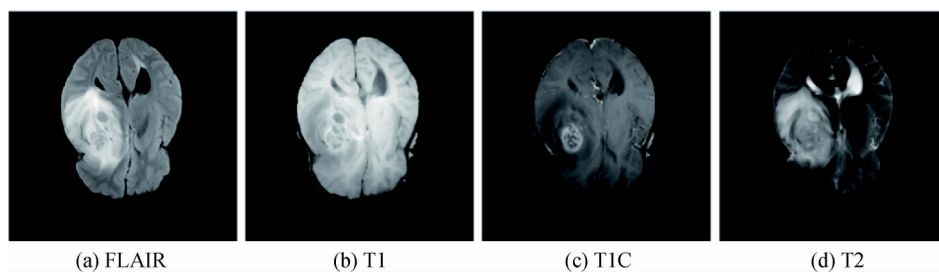


图1 不同模态下的脑肿瘤磁共振图像

Fig. 1 MRI images in different modes((a) FLAIR; (b) T1; (c) T1C; (d) T2)

自2012年以来,随着深度学习的发展和硬件性能的提升,基于神经网络的分割方法逐渐成为主流(江宗康等2020)。由于脑肿瘤为渗透性肿瘤,其外观高度异质、位置随机和各子区域体素数量

相差悬殊,且肿瘤组织和近邻的正常脑组织具有高度的灰度近似性,尽管脑肿瘤研究已取得很大进展,在多模态MRI中自动识别肿瘤的轮廓信息及其对各子区域进行精细划分仍是一项艰巨的任务(童云

飞等 2018; 雷晓亮等 2019)。目前利用深度学习进行 3D 脑肿瘤图像分割主要有两种方法: 一种是使用 2D 卷积网络, Dong 等人(2017) 将 3D 影像进行切片扫描后转换为多幅 2D 图像, 通过翻转、旋转、平移、弹性变形等方法进行图像增强, 并用 2D 卷积核提取图像特征, 实现了全自动脑肿瘤分割。Havaei 等人(2017) 实现了一种双通道结构以获取更多的上下文信息, 并针对肿瘤类别分布不平衡的问题提出了两阶段训练方式, 取得了较好的分割效果。但 2D 网络通常是将多幅输出图像机械性拼接成 3D 图像作为分割结果, 未能有效利用来自相邻切片的上下文信息, 经常出现锯齿、断层等问题而影响分割精度(褚晶辉等 2019)。另一种是使用 3D 卷积网络, Kamnitsas 等人(2018) 采用 DeepMedic (Kamnitsas 等 2017)、FCNN(fully convolutional neural networks) (Long 等, 2015)、U-Net (Ronneberger 等 2015) 多模型集成方法, 提升了网络的泛化能力, 防止对一些特定数据集的过拟合。Isensee 等人(2019) 通过减少上采样前的卷积核数量以提升下采样分辨率, 通过对肿瘤增强区设置体素数量阈值以增强其分割精度。Myronenko (2019) 在经典 U-Net 的基础上通过添加变分自动编码分支来重建输入图像, 使共享编码器正则化, 取得了 BraTS2018 挑战的第 1 名。对于 3D 卷积网络来说, 其数据量远小于 2D 卷积网络, 加之医学图像数据普遍不足等问题, 3D 卷积网络将更容易产生过拟合风险, 同时其参数量又大于 2D 卷积网络, 因此需要更大的 GPU 显存空间、更长的训练时间, 其在硬件方面巨大的开销在很大程度上制约着相关研究的发展。

为了降低显存占用等计算开销, Chen 等人(2019) 采用组卷积搭配多尺寸的膨胀卷积来获取不同大小的感受野, 有效缓解了脑肿瘤各子区域因体积相差悬殊而影响分割精度的问题, 但未能进一步克服由组卷积带来的跨通道信息交流不畅等问题; Chen 等人(2018a) 针对不同肿瘤类别的组织相似性所导致的类间干扰的问题, 将多分割任务拆分为由聚焦、分割、擦除组成的多次二分割任务来解决显存占用和类间干扰等问题, 但在核心区 and 增强区的分割精度较低。Brügger 等人(2019) 通过引入可逆卷积块, 降低了网络前向传播时需要储存的参数量, 但在计算量方面牺牲较大。

基于以上问题, 本文提出一种基于组卷积的轻量级 3D 脑肿瘤分割方法, 该方法在维持较高分割精度的同时, 显著降低显存占用和计算量, 为医学影像实现实时分割提供了一种可行性方案。相关代码已经开源(<https://github.com/easthorse/brain-tumor-segmentation-based-on-group-convolution>)。

1 网络框架及算法原理

1.1 完整网络框架

以多纤混合单元(multi-fiber shuffle unit, MFS)作为基本结构, 完整分割网络框架如图 2 所示。网络分为上采样与下采样两部分, 其中下采样部分由编码卷积块组成, 每个卷积块由一个步长为 2 的 MFS 单元、两个步长为 1 的 MFS 单元组成, 以降低特征层尺寸并提取深层信息。上采样部分由解码卷积块组成, 每个解码卷积块由一次三线性插值和一个步长为 1 的 MFS 单元组成, 以增大特征层尺寸并压缩其深度。类似于 U-Net 结构, 在每个解码卷积块后拼接来自编码卷积块对应尺寸的特征层, 以融合深浅层特征信息。在网络的最后使用卷积核尺寸为 $1 \times 1 \times 1$ 卷积层和 Softmax 激活函数, 生成一个与网络输入同尺寸的 4 维矩阵作为分割结果。其中 MFS 模块由多纤单元(multi-fiber unit) (Chen 等, 2018b)、通道混合单元(channel shuffle unit) (Zhang 等 2018b)、残差单元(He 等 2016) 以及多次组卷积操作等构成。如图 3 所示, 为了平衡网络效率与分割精度, 本文仅在 MFS 单元中使用一次多纤单元, 而在各组卷积之间使用参数量和计算量更低的通道混合单元, 用残差单元来缓解因网络深度过深而产生的网络退化问题, 通过额外增加一次组卷积操作来解决残差单元中输入层与输出层在通道数或尺寸方面不匹配的问题, 并在每次卷积操作前使用跨卡同步批量归一化(synchronized cross-GPU batch normalization) (Zhang 等 2018a) 以及 ReLU 激活函数去线性化。

1.2 算法原理

1.2.1 组卷积

在常规卷积中, 输出层的每一通道均由输入层的所有通道经过卷积运算得到, 而组卷积将通道分为多组后进行卷积操作, 减少了特征层和卷积核之

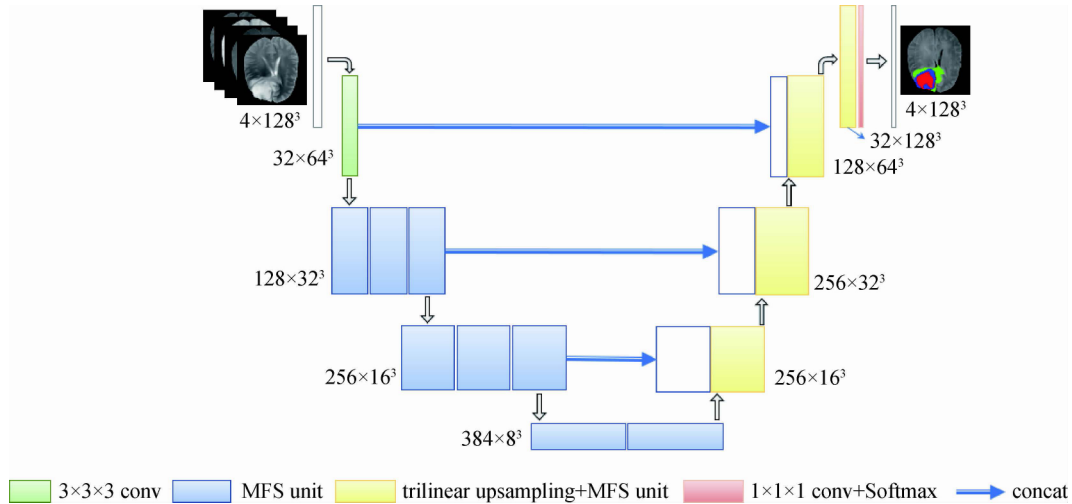


图2 完整网络结构示意图

Fig. 2 Schematic diagram of the complete network structure

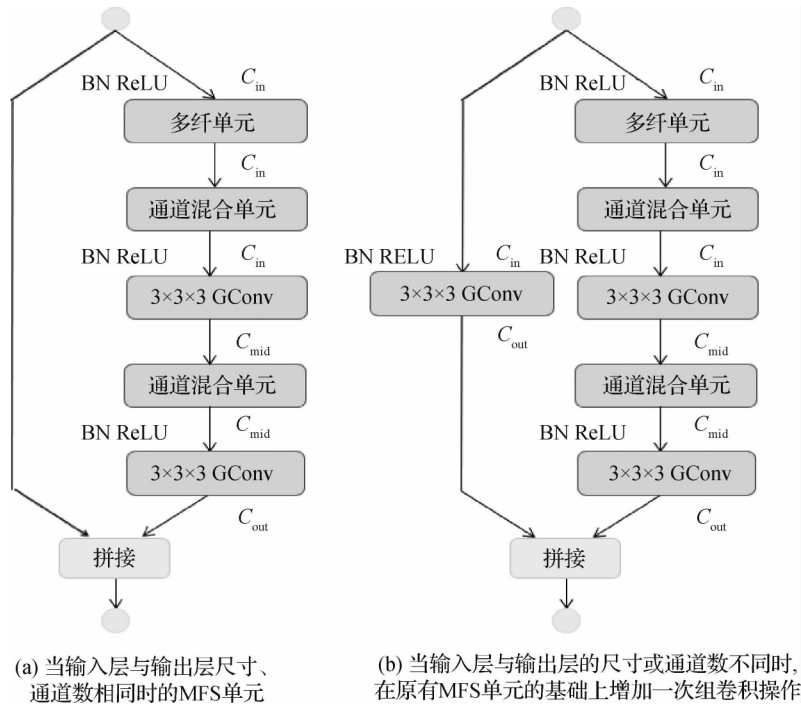


图3 基于组卷积和残差结构的MFS单元示意图

Fig. 3 Schematic diagram of MFS unit based on group convolution and residual structure

((a) MFS unit when channel numbers and size of the input layer are the same as the output layer's;

(b) MFS unit when channel numbers or size of the input layer are the different as the output layer's)

间的联系,大大降低参数量。如图4(b)所示,以组数为3、卷积核尺寸为 $3 \times 3 \times 3$ 的组卷积为例,首先将具有 C_{in} 通道的输入层分为3组,每组通道数为 $C_{in}/3$,之后每组进行两次卷积操作,第1次输出通道数为 $C_{mid}/3$,第2次输出通道数为 $C_{out}/3$,因此组卷积的参数量为

$$p_a = k \times (4 \times C_{in}/4 \times C_{mid}/4 + 4 \times C_{mid}/4 \times C_{out}/4) \quad (1)$$

式中, k 为卷积核尺寸。整个网络呈稀疏连接,普通卷积的参数量

$$p_b = k \times (C_{in} \times C_{mid} + C_{mid} \times C_{out}) \quad (2)$$

从式(1)(2)可知,组卷积的参数量小至1/4,大幅

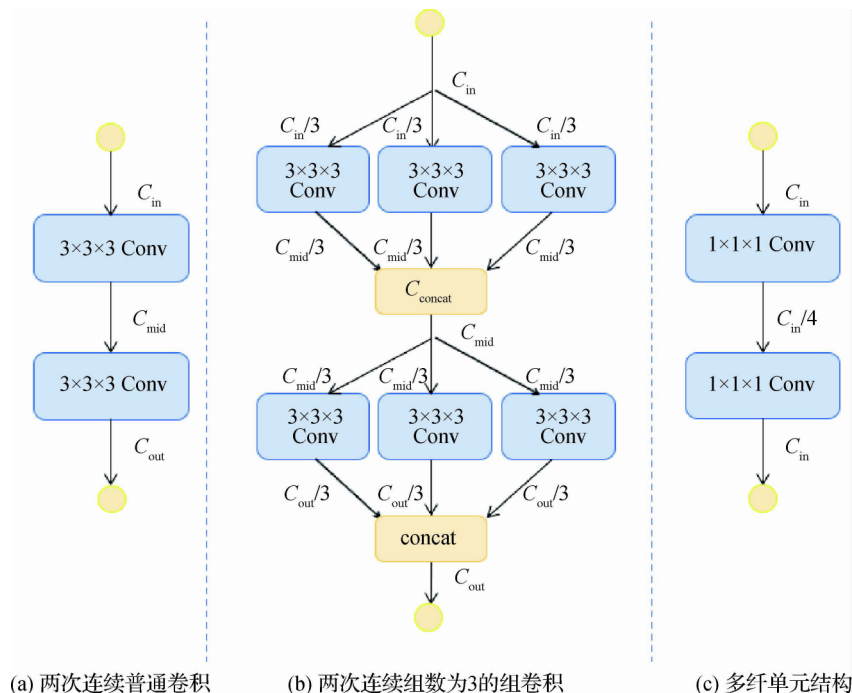


图4 组卷积与多纤单元的结构示意图

Fig. 4 A structural diagram of group convolution and multi-fiber units ((a) schematic diagram of two consecutive convolutions; (b) schematic diagram of two group convolution layers with a number of groups of three; (c) architecture details of multi-fiber unit)

提高3D卷积神经网络的效率,为3D图像显存占用大、训练时间长等问题提供了一种有效的解决方案。但因每个通道的输出仅来自部分输入通道,无法有效学习到所有输入层的信息,简单地利用组卷积替换常规卷积可能会因影响通道之间的信息交换能力而损害神经网络的学习能力,且若多次连续进行组卷积操作,不同组的通道信息交换不畅问题会愈发严重。

1.2.2 多纤单元

为了改善跨组间的信息流,采用了一种轻量级的多路复用器模块,在保持特征层尺寸不变的前提下,大幅增加非线性特征,实现跨通道交互和信息整合,有助于网络的进一步加深。如图4(c)所示,通过两次卷积核尺寸为 $1 \times 1 \times 1$ 的卷积操作,第1次卷积先将输入层通道数 C_{in} 缩至 $C_{in}/4$ 以从所有组中压缩并整合特征,再将通道数还原至 C_{in} ,从而将整合好的特征信息重新反馈到各组中,使各个通道充分融合其他通道信息,促进各组间的信息交换。该方法的参数量为

$$p_a = C_{in} \times C_{in}/4 + C_{in}/4 \times C_{in} = C_{in}^2/2 \quad (3)$$

相比于一次卷积核尺寸为 $1 \times 1 \times 1$ 的卷积操作的参数量为

$$p_b = C_{in} \times C_{in} = C_{in}^2 \quad (4)$$

参数量降低至 $1/2$ 。

1.2.3 通道混合单元

组卷积层如从不同组获得输入数据,则每组的输出层将和其他的输入层更好地关联。以两次的组数为4的组卷积操作为例,如图5(a)所示,普通组卷积每次的输出层仅与同组的输入层有关,加入了通道混合单元后如图5(b)所示,对于第1次的输出层,先将每个组中的通道分为4个子组,将所有子组逐一提供给下一层的每个组作为第2次组卷积的输入,有效增强了各组间信息交流。通道混合单元实质上是一种矩阵的转置操作,仅轻微增加神经网络的计算量且不额外占用显存(Zhang等,2018b),搭配多纤单元后为基于组卷积的高精度、轻量化的分割网络结构奠定了基础。

1.2.4 跨卡同步批量归一化

经典的批量归一化(batch normalization)因为使用数据并行(data parallel),只在单张显卡上对样本进行归一化,相当于减小了批量大小(batch size)。在某些视觉任务如ImageNet分类中,单样本占用显存较小,批量归一化处理的样本数量足够大,因此可以获得较好的统计信息。但在单样本显存占用较大

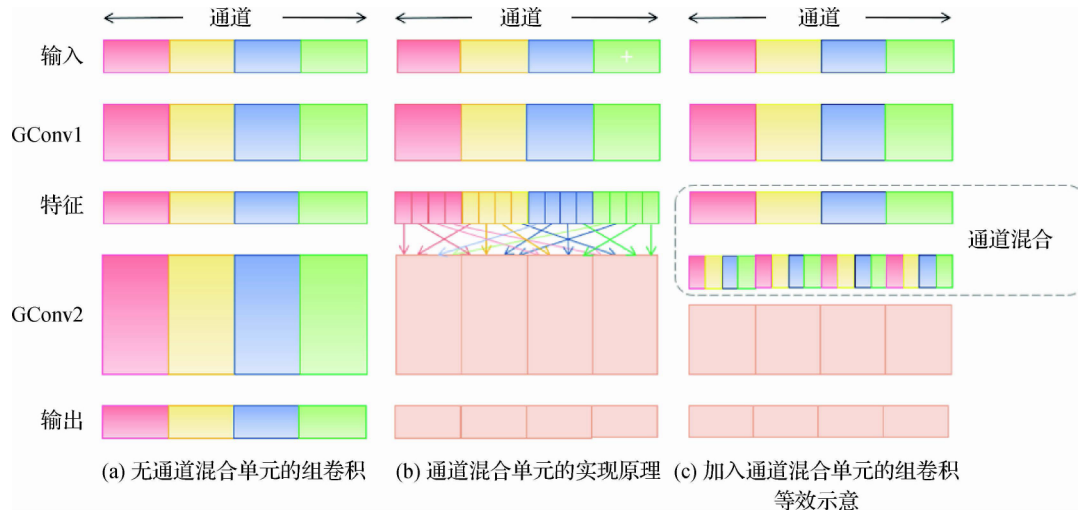


图 5 通道混合单元的原理细节图

Fig. 5 Schematic diagram of the channel shuffle unit ((a) schematic diagram of group convolution without channel shuffle unit; (b) implementation principle diagram of channel shuffle unit; (c) equivalent schematic diagram of group convolution added to channel shuffle unit)

的任务中如 3D 语义分割,不同显卡内模型运算相互独立且单卡内批量过小,严重影响了收敛效果。如图 6 所示,跨卡同步归一化充分利用了多张显卡协同计算的优势,使用全局样本进行归一化,间接增大了批量大小,缓解了显存过大的 3D 医学影像因批量值过小所导致的精度下降问题。图中 x_i 与 y_i 分别为第 i 张显卡上在归一化前后的样本点。此外,将先同步求均值 μ 再发回各卡同步利用 μ 求方差 σ^2 ,替换为将方差 σ^2 通过 $\sum x_i$ 和 $\sum x_i^2$ 定量表

示,可将跨卡同步次数从两次降低至一次,提高归一化同步计算速度,即

$$\mu = \frac{1}{m} \sum_{i=1}^m x_i \quad (5)$$

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)^2 = \frac{1}{m} \sum_{i=1}^m x_i^2 - \mu^2 = \frac{1}{m} \sum_{i=1}^m x_i^2 - \left(\frac{1}{m} \sum_{i=1}^m x_i \right)^2 \quad (6)$$

式中, x_i 为样本点, m 为多张显卡的样本点总数。

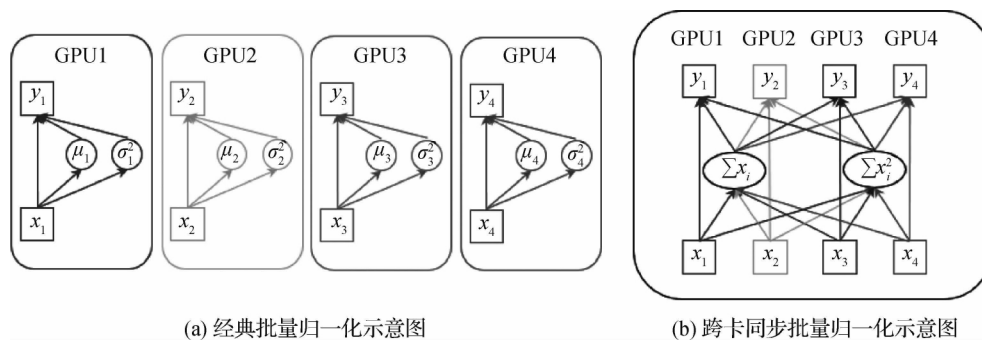


图 6 经典批量归一化与跨卡同步归一化的结构对比图

Fig. 6 A structural contrast diagram of classic batch normalization and synchronized cross-GPU batch normalization ((a) classic batch normalization diagram; (b) synchronized cross-GPU batch normalization diagram)

1.3 加权混合损失函数

由于脑肿瘤体积在整个脑组织中占比相对较小,这通常会使得损失函数达到局部最小值,导致稀疏类预测错误对于整体损失来说影响甚小,从而

得到一个预测强烈偏向背景的网络, Dice Loss 函数 (Milletari 等, 2016) 能较好地平衡前景背景信息,广泛应用于脑肿瘤分割领域,其公式及偏导式为

$$D = 1 - \frac{1}{K} \sum_{k=1}^K \frac{2 \sum_{i=1}^N p_{ik} g_{ik}}{\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik}} \quad (7)$$

$$\frac{\partial D}{\partial p_j} = -\frac{2}{K} \sum_{k=1}^K \frac{g_{jk} \left(\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik} \right) - \sum_{i=1}^N p_{ik} g_{ik}}{\left(\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik} \right)^2} \quad (8)$$

式中, K 为标签类别总数, 本文中为 4, N 为一个训练样本中像素点总数, p_j 为第 j 个像素点的概率分布, p_{ik} 为第 i 个体素预测为类别 k 的概率, g_{ik} 为其对应的标签值, 当该体素的真值为第 k 类时 $g_{ik} = 1$, 其余情况 $g_{ik} = 0$ 。

Jaccard Loss (Cai 等 2017) 在 Dice Loss 的基础上进行优化, 在分子分母中同时减去 $\sum_{i=1}^N p_{ik} g_{ik}$, 其公式及其偏导式为

$$J = 1 - \frac{1}{K} \sum_{k=1}^K \frac{\sum_{i=1}^N p_{ik} g_{ik}}{\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik} - \sum_{i=1}^N p_{ik} g_{ik}} \quad (9)$$

$$\frac{\partial J}{\partial p_j} = -\frac{1}{K} \sum_{k=1}^K \frac{g_{jk} \left(\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik} \right) - \sum_{i=1}^N p_{ik} g_{ik}}{\left(\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik} - \sum_{i=1}^N p_{ik} g_{ik} \right)^2} \quad (10)$$

Dice Loss 具有收敛快、精度高等优点, 但在训练中后期随着预测准确率的提升, 分母数值的增加, 收敛速度迅速降低, 不利于精度的进一步提升。Jaccard Loss 有效避免了该问题, 但该损失函数在前期收敛较慢, 且容易造成分割结果较大、边界不准等问题, 为有效兼顾两种损失函数的优势, 现提出一种混合损失函数, 其表达式为

$$L = (1 - \alpha) \left(1 - \frac{1}{K} \sum_{k=1}^K \frac{2 \sum_{i=1}^N p_{ik} g_{ik}}{\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik}} \right) + \beta \alpha \left(1 - \frac{1}{K} \sum_{k=1}^K \frac{\sum_{i=1}^N p_{ik} g_{ik}}{\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik} - \sum_{i=1}^N p_{ik} g_{ik}} \right) \quad (11)$$

式中, α 和 β 用来平衡 Dice Loss 和 Jaccard Loss 的

权值, 经实验 $\alpha = t_{ep}/N_{ep}$, $\beta = 2$ 时效果较好, N_{ep} 为总训练轮数, t_{ep} 为当前训练轮数。

此外, 由于脑肿瘤图像不仅在患病区和正常区的体素数量相差悬殊, 不同种患病区体素数量相差亦较大, 且某些患病区与分割边界灰度高度近似, 导致核心区、增强区的分割精度较低且难以提升, 基于以上问题, 进一步改进得

$$L = (1 - \alpha) \left(1 - \frac{2 \sum_{k=1}^K w_k \sum_{i=1}^N p_{ik} g_{ik}}{\sum_{k=1}^K w_k \left(\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik} \right)} \right) + \beta \alpha \left(1 - \frac{\sum_{k=1}^K w_k \sum_{i=1}^N p_{ik} g_{ik}}{\sum_{k=1}^K w_k \left(\sum_{i=1}^N p_{ik} + \sum_{i=1}^N g_{ik} - \sum_{i=1}^N p_{ik} g_{ik} \right)} \right) \quad (12)$$

式中, 加权系数 $w_k = 1/N_k^2$, N_k 为标签图像中某类别体素的数量, 若 N_k 越小, w_k 相对越大, 该类别的损失值占整体的比重就越大, 提高了稀疏类分类错误对模型的惩罚, 有效平衡了不同分割难度类别的训练强度。

2 实验和结果

实验环境为 CPU Intel® Core i9-9900X 3.5 GHz、GPU Nvidia GTX2080Ti (11GB) × 4、Ubuntu 16.04 位操作系统, 使用 PyTorch 开源深度学习框架。

2.1 实验评价指标

以分割准确度和计算复杂度来定量评价模型性能, 计算复杂度由模型参数量 (parameters) 以及浮点运算次数 (floating point of operations, FLOPs) 决定, 分割准确度由 Dice 相似系数 (Dice similarity coefficient) 以及豪斯多夫距离 (Hausdorff95 distance, HD95) 定量表示。其中相似系数为预测与真实肿瘤区域的相似程度为

$$Dice(\mathbf{P}, \mathbf{T}) = \frac{|\mathbf{P} \cap \mathbf{T}|}{(|\mathbf{P}| + |\mathbf{T}|)/2} \quad (13)$$

豪斯多夫距离为两个点集的最大不匹配程度为

$$Haus(\mathbf{P}, \mathbf{T}) = \max \left\{ \sup_{p \in \mathbf{P}} \inf_{t \in \mathbf{T}} d(p, t), \sup_{t \in \mathbf{T}} \inf_{p \in \mathbf{P}} d(t, p) \right\} \quad (14)$$

式中, \mathbf{P} 、 \mathbf{T} 分别为模型预测的肿瘤区和真实标注

区域的体素集, p 、 t 分别为两个体素集中的体素点, $d(p, t)$ 为两个体素点距离, \cap 为两个体素集对应点的与运算, $|\cdot|$ 为体素集中体素数量的代数求和运算, \inf 表示取最小值。

2.2 数据处理及实验过程

实验数据选取自 BraTS2018 数据集 (Menze 等, 2015), 其中训练集包含 210 个高级胶质瘤患者样本, 75 个低级胶质瘤患者样本, 验证集包含 66 个无标签患者样本。在训练集中每个样本包含 4 种 MRI 模态和由多位专业医师手动标注的真值标签图, 通过一系列数据增强方法, 如对尺寸为 $240 \text{ mm} \times 240 \text{ mm} \times 155 \text{ mm}$ 的脑肿瘤图像随机截取至 $128 \text{ mm} \times 128 \text{ mm} \times 128 \text{ mm}$, 随机在轴向、冠状、矢状等方向翻转, 在 $[-10^\circ, 10^\circ]$ 范围内随机旋转等, 避免因训练集数据不足导致的过拟合问题。脑肿瘤 MRI 的 4 模态以 4 通道的形式作为输入, 通过编解码网络进行训练, 并使用加权混合损失函数计算损失值, 优化器采用 Adam 优化器, 初始学习率 $l_0 = 0.001$, 训练 400 轮后模型基本收敛。

在脑肿瘤分割任务中, 最具挑战性的任务之一是将增强肿瘤区与肿瘤核心区的小血管部分进行区分, 这对于没有增强肿瘤区的 LGG 患者尤其困难。由式 (13) 可知, 如果真值图和预测图均无增强区, 则其增强区的相似系数为 1; 相反, 若真值图中患者不存在增强肿瘤区, 但在预测图中有一个假阳性

体素, 会导致其相似系数为 0。因此本文对预测结果进行后处理: 为肿瘤增强区设置体素数量阈值, 当某样本的肿瘤增强区预测体素数量低于 400 时将该部分体素并入肿瘤核心区, 在一定程度上提高了肿瘤增强区和核心区的分割精度。

2.3 结果和分析

首先对训练集进行五重交叉验证评估, 本文算法的 Dice 值在整体区、核心区和增强区的平均 Dice 值分别可达 89.52%、82.74% 和 77.19%, 如图 7 所示, 与经典 3D U-Net 的分割图像进行对比, 本文算法通过使用组卷积、跨卡同步批量归一化以及加权混合损失函数等方法显著降低了假阳性体素数量, 在增强区、核心区等分割难度较大的区域具有明显的优势。此外还以 BraTS2018 提供的 66 个无标签患者样本作为验证集, 用训练后的网络对样本进行分割预测, 后上传至 BraTS 线上服务器, 其分割结果亦由该服务器计算分析后提供。如表 1 所示, 本文算法的 Dice 值在整体区、核心区和增强区分别为 90.67%、85.06% 和 80.41%, 模型的参数量和计算量分别为 3.2 M 和 20.51 G, 与 Chen 等人 (2019) 方法相比, 其参数量和计算量分别降低了 17% 和 24%, 同时在整体区、核心区、增强区的 Dice 值分别提高了 0.05%、0.53% 和 0.29%。与 BraTS2018 比赛第 1 名 Myronenko 等人 (2019) 方法相比, 本文算法的 Dice 值在整体区、核心区、增强区分别比其低

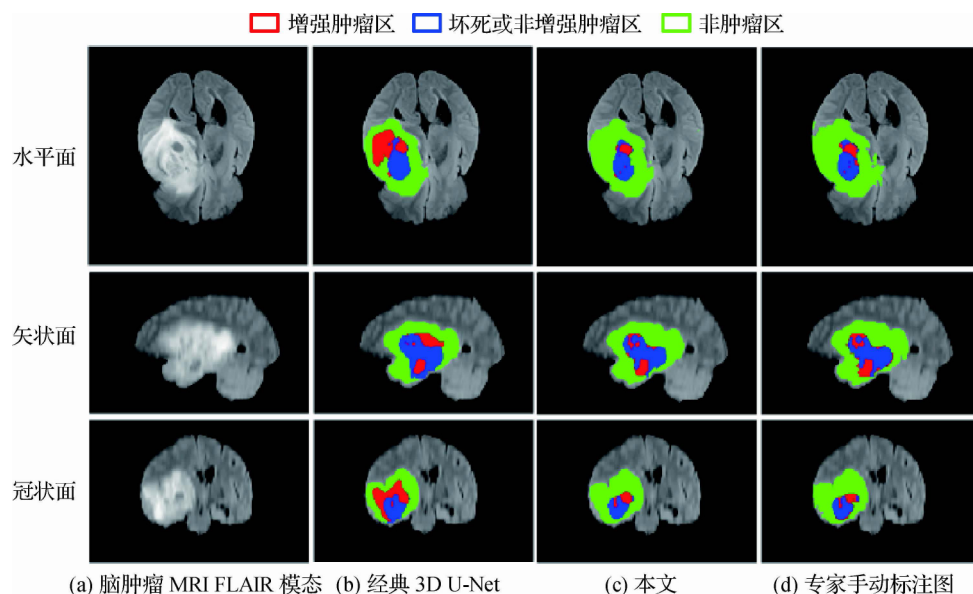


图7 脑肿瘤 MRI 分别在水平面、矢状面、冠状面的分割效果对比图

Fig. 7 The visual comparison of MRI brain tumor segmentation results in horizontal plane, sagittal plane, and coronal plane ((a) FLAIR modality of the brain tumor MRI; (b) 3D U-Net; (c) ours; (d) ground truth)

表 1 各类算法在 BraTS2018 验证集上的分割效果对比
Table 1 Comparison of various algorithms on the BraTS2018 validation set

脑肿瘤分割算法	参数量(parameters) /M	计算量(FLOPs) /G	相似系数(Dice) /%			豪斯多夫距离(HD95)		
			增强区	整体区	核心区	增强区	整体区	核心区
本文	3. 2	20. 51	80. 41	90. 67	85. 06	2. 51	4. 13	5. 79
Çiçek 等人(2016)	16. 21	1 669. 53	75. 96	88. 53	71. 77	6. 04	17. 1	11. 62
Kao 等人(2019)	9. 45	203. 04	78. 75	90. 47	81. 35	3. 81	4. 32	7. 56
Chen 等人(2019)	3. 86	26. 93	80. 12	90. 62	84. 54	3. 06	4. 66	6. 31
Brügger 等人(2019)	3. 01	956. 2	80. 56	90. 61	85. 71	3. 35	5. 61	7. 83
Isensee 等人(2019)	12. 43	296. 83	80. 66	90. 92	85. 22	2. 74	5. 83	7. 2
Myronenko 等人(2019)	40. 06	1 495. 53	81. 73	90. 68	86. 02	3. 82	4. 41	6. 84

0. 01%、0. 96%、1. 32% ,但模型的参数量和计算量分别降低了 12. 5 倍和 73 倍 ,在 4 张 Nvidia GTX 2080Ti 下仅需 4 h 即可完成训练 ,单样本 1. 6 s 可完成分割预测。因此 ,本文算法在保持与 NVDLMED 相当的分割精度的基础上 ,显著降低了显存占用 ,提升了训练速度 ,在分割效率方面有更大的优势。

3 结 论

利用 3D 卷积神经网络进行脑肿瘤 MRI 分割通常存在着显存占用大、硬件资源消耗高等问题 ,在很大程度上制约着计算机辅助诊断在临床医学上的应用。基于以上问题 ,提出一种基于组卷积的轻量级脑肿瘤分割算法 ,通过使用组卷积代替常规卷积 ,有效降低了 3D 卷积网络的显存占用 ,同时也使网络深度得以更深。利用通道混合单元与多纤单元 ,以削弱由组卷积带来的各组间信息交流不畅等问题 ,并使用跨显卡同步批量归一化发挥多显卡协同计算的优势。最后提出一种加权混合损失函数 ,在提高收敛速度的同时进一步提升分割效果。与经典 3D U-Net 相比 ,本文算法显著降低了假阳性样本数量 ,在增强肿瘤区以及核心区等分割难度较大的区域具有巨大优势。与当前最佳分割算法相比 ,本文算法亦具有分割精度高、算力资源占用低等特点 ,可为临床医师进行脑肿瘤分割提供有力参考。但该算法的适用范围有一定的局限性 ,在未来的研究中 ,将验证是否适用于其他医学图像处理任务如肺结节检测、肝肿瘤分割等。另外 ,尽管使用了数据增强方法 ,但是网络模型训练的数据量仍然较小 ,是否可以通过

自监督学习等方法增大数据量 ,进一步提高分割准确度 ,这也是今后工作的重点。

参考文献(References)

Bray F , Ferlay J and Soerjomataram I. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a Cancer Journal for Clinicians , 68(6) : 394-424 [DOI: 10. 3322/caac. 21492]

Brügger R , Baumgartner C F and Konukoglu E. 2019. A partially reversible U-Net for memory-efficient volumetric image segmentation // Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen: Springer: 429-437 [DOI: 10. 1007/978-3-030-32248-9_48]

Cai J Z , Lu L , Xie Y P , Xing F Y and Yang L. 2017. Improving deep pancreas segmentation in CT and MRI images via recurrent neural contextual learning and direct loss function [EB/OL]. [2020-05-15]. <https://arxiv.org/pdf/1707.04912.pdf>

Chen C , Liu X P , Ding M , Zheng J F and Li J Y. 2019. 3D dilated multi-fiber network for real-time brain tumor segmentation in MRI // Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen: Springer: 184-192 [DOI: 10. 1007/978-3-030-32248-9_21]

Chen X , Liew J H , Xiong W , Chui C K and Ong S H. 2018a. Focus , segment and erase: an efficient network for multi-label brain tumor segmentation // Proceedings of the 15th European Conference on Computer Vision. Munich: Springer: 674-689 [DOI: 10. 1007/978-3-030-01261-8_40]

Chen Y P , Kalantidis Y , Li J S , Yan S C and Feng J S. 2018b. Multi-fiber networks for video recognition // Proceedings of the 15th European Conference on Computer Vision. Munich: Springer: 364-380 [DOI: 10. 1007/978-3-030-01246-5_22]

Chu J H , Li X C , Zhang J Q and Lyu W. 2019. Fine-grained segmentation method for three-dimensional brain tumors using cascaded conv-

- olutional network. *Laser and Optoelectronics Progress*, 56(10): 101001 (褚晶辉, 李晓川, 张佳祺, 吕卫. 2019. 一种基于级联卷积网络的三维脑肿瘤精细分割. *激光与光电子学进展*, 56(10): 101001) [DOI: 10.3788/LOP56.101001]
- Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T and Ronneberger O. 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation//*Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Athens: Springer: 424-432 [DOI: 10.1007/978-3-319-46723-8_49]
- Deangelis L M. 2001. Brain tumors. *New England Journal of Medicine*, 344(2): 114-123 [DOI: 10.1056/NEJM20010113440207]
- Dong H, Yang G, Liu F D, Mo Y H and Guo Y K. 2017. Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks//*Proceedings of the 21st Annual Conference on Medical Image Understanding and Analysis*. Edinburgh: Springer: 506-517 [DOI: 10.1007/978-3-319-60964-5_44]
- Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin P M and Larochelle H. 2017. Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, 35: 18-31 [DOI: 10.1016/j.media.2016.05.004]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Isensee F, Kickingereder P, Wick W, Bendszus M and Maier-Hein K H. 2019. No new-net//*Proceedings of the 4th International MICCAI Brainlesion Workshop*. Granada: Springer: 234-244 [DOI: 10.1007/978-3-030-11726-9_21]
- Jiang Z K, Lyu X G, Zhang J X, Zhang Q and Wei X P. 2020. Review of deep learning methods for MRI brain tumor image segmentation. *Journal of Image and Graphics*, 25(2): 215-228 (江宗康, 吕晓钢, 张建新, 张强, 魏小鹏. 2020. MRI脑肿瘤图像分割的深度学习方法综述. *中国图象图形学报*, 25(2): 215-228) [DOI: 10.11834/jig.190173]
- Kamnitsas K, Bai W, Ferrante E, McDonagh S, Sinclair M, Pawlowski N, Rajchl M, Lee M, Kainz B, Rueckert D and Glocker B. 2018. Ensembles of multiple models and architectures for robust brain tumour segmentation//*Proceedings of the 3rd International MICCAI Brainlesion Workshop*. Quebec City: Springer: 450-462 [DOI: 10.1007/978-3-319-75238-9_38]
- Kamnitsas K, Ledig C, Newcombe V F J, Simpson J P, Kane A D, Menon D K, Rueckert D and Glocker B. 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 36: 61-78 [DOI: 10.1016/j.media.2016.10.004]
- Kao P Y, Ngo T, Zhang A, Chen J W and Manjunath B S. 2019. Brain tumor segmentation and tractographic feature extraction from structural MR images for overall survival prediction//*Proceedings of the 4th International MICCAI Brainlesion Workshop*. Granada: Springer: 128-141 [DOI: 10.1007/978-3-030-11726-9_12]
- Lei X L, Yu X S, Chi J N, Wang Y and Wu C D. 2019. Brain tumor segmentation based on prior sparse shapes. *Journal of Image and Graphics*, 24(12): 2222-2232 (雷晓亮, 于晓升, 迟剑宁, 王莹, 吴成东. 2019. 稀疏形状先验的脑肿瘤图像分割. *中国图象图形学报*, 24(12): 2222-2232) [DOI: 10.11834/jig.190070]
- Li Q, Bai K X, Zhao L and Guan X. 2020. Progresss and challenges of MRI brain tumor image segmentation. *Journal of Image and Graphics*, 25(3): 419-431 (李锦, 白柯鑫, 赵柳, 关欣. 2020. MRI脑肿瘤图像分割研究进展及挑战. *中国图象图形学报*, 25(3): 419-431) [DOI: 10.11834/jig.190524]
- Liang Z P and Lauterbur P C. 2000. Principles of Magnetic Resonance Imaging: A Signal Processing Perspective. New York: The Institute of Electrical and Electronics Engineers Press
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE: 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Menze B H, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R, Lanczi L, Gerstner E, Weber M A, Arbel T, Avants B B, Ayache N, Buendia P, Collins D L, Cordier N, Corso J J, Criminisi A, Das T, Delingette H, Demiralp Ç, Durst C R, Dojat M, Doyle S, Festa J, Forbes F, Geremia E, Glocker B, Golland P, Guo X T, Hamamci A, Iftekharuddin K M, Jena R, John N M, Konukoglu E, Lashkari D, Mariz J A, Meier R, Pereira S, Precup D, Price S J, Raviv T R, Reza S M S, Ryan M, Sarikaya D, Schwartz L, Shin H C, Shotton J, Silva C A, Sousa N, Subbanna N K, Szekely G, Taylor T J, Thomas O M, Tustison N J, Unal G, Vasseur F, Wintermark M, Ye D H, Zhao L, Zhao B S, Zikic D, Prastawa M, Reyes M and Van Leemput K. 2015. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10): 1993-2024 [DOI: 10.1109/TMI.2014.2377694]
- Milletari F, Navab N and Ahmadi S A. 2016. V-Net: fully convolutional neural networks for volumetric medical image segmentation//*Proceedings of the 4th International Conference on 3D Vision*. Stanford: IEEE: 565-571 [DOI: 10.1109/3DV.2016.79]
- Myronenko A. 2019. 3D MRI brain tumor segmentation using autoencoder regularization//*Proceedings of the 4th International MICCAI Brainlesion Workshop*. Granada: Springer: 311-320 [DOI: 10.1007/978-3-030-11726-9_28]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Tong Y F, Li Q and Guan X. 2018. An improved multi-modal brain tumor segmentation hybrid algorithm. *Journal of Signal Processing*,

34(3): 340-346 (童云飞, 李锵, 关欣. 2018. 改进的多模式脑肿瘤图像混合分割算法. 信号处理, 34(3): 340-346) [DOI: 10.16798/j.issn.1003-0530.2018.03.011]

Zhang H, Dana K, Shi J P, Zhang Z Y, Wang X G, Tyagi A and Agrawal A. 2018a. Context encoding for semantic segmentation//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 7151-7160 [DOI: 10.1109/CVPR.2018.00747]

Zhang X Y, Zhou X Y, Lin M X and Sun J. 2018b. Shuffle-Net: an extremely efficient convolutional neural network for mobile devices//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 6848-6856 [DOI: 10.1109/cvpr.2018.00716]



李锵, 通信作者, 男, 教授, 主要研究方向为音乐信号处理、医学图像处理、滤波器设计、数字系统与微系统设计。

E-mail: liqiang@tju.edu.cn

关欣, 女, 副教授, 主要研究方向为音乐信息检索、统计学习和信息处理。E-mail: guanxin@tju.edu.cn

作者简介



赵奕名, 1995 年生, 男, 硕士研究生, 主要研究方向为 3 维医学影像处理。

E-mail: zym950408@sina.cn