

DOI:10.16136/j.joel.2023.08.0379

基于多模态融合的2D MR脑肿瘤图像分割算法研究

李楠, 张宏立*

(新疆大学 电气工程学院, 新疆 乌鲁木齐 830004)

摘要:针对不同模态MR脑肿瘤图像呈现的肿瘤状态差异以及卷积神经网络(convolutional neural networks, CNNs)提取特征局限性的问题,提出了一种基于多模态融合的MR脑肿瘤图像分割方法。分割模型以U-net网络为原型,创新一种多模态图像融合方式以加强特征提取能力,同时引入通道交叉注意力机制(channel cross transformer, CCT)代替U-net中的跳跃连接结构,进一步弥补深浅层次的特征差距与空间依赖性,有效融合多尺度特征,加强对肿瘤的分割能力。实验在BraTS数据集上进行了多目标分割结果验证,通过定量分析对比前沿网络分割结果,表明该方法有良好的分割性能,其分割出三种肿瘤区域的Dice系数分别达到80%、74%、71%。

关键词:脑肿瘤分割; U-net网络; 多模态融合; 通道交叉注意力机制

中图分类号: TP391.7 **文献标识码:** A **文章编号:** 1005-0086(2023)08-0890-07

Research on 2D MR brain tumor image segmentation algorithm based on multimodal fusion

LI Nan, ZHANG Hongli*

(College of Electrical Engineering, Xinjiang University, Urumqi, Xinjiang 830004, China)

Abstract: Aiming at the differences in tumor status presented by different modalities of MR brain tumor images and the limitations of feature extraction by convolutional neural networks (CNNs), a method of brain tumor image segmentation based on multimodal fusion is proposed. The segmentation model is based on the U-net network, which innovates a multimodal image fusion approach to enhance the feature extraction capability, while a channel cross transformer (CCT) module is introduced instead of the jump connection structure in the U-net to further the deep and shallow feature disparity and spatial dependency, fusing the multi-scale features effectively and enhancing the tumor segmentation capability. The results of multi-objective segmentation are verified on the BraTS dataset. Quantitative analysis and comparison of frontier network segmentation results shows that the proposed method has good segmentation performance. The Dice coefficients of three tumor regions are 80%, 74% and 71% respectively.

Key words: brain tumor segmentation; U-net; multimodal fusion; channel cross transformer (CCT)

0 引言

根据《美国医学会杂志》2020年发布的全球最新癌症负担报告^[1],脑肿瘤发病率占比约1.5%,死亡率高达3%。脑肿瘤生长侵袭速度快、组织异常且具有一定转移性。计算机辅助诊断系统(computer-aided diagnosis, CAD)可以帮助人们更好地了解病情,也是辅助医生进行诊断和

临床治疗的依据。图像分割作为CAD系统的主要支撑技术,通过算法精确分割出病变区域,完美结合了医学领域与计算机视觉领域。因此,研究脑肿瘤分割技术对于患者的早期诊断和治疗,提高患者的生存率具有重要意义。

磁共振成像(magnetic resonance imaging, MRI)具有对软组织的敏感性和优越的图像对比度,被认为是处理脑诊断的首选成像类型。MRI

* E-mail: 1606829274@qq.com

收稿日期: 2022-05-21 修订日期: 2022-07-28

基金项目: 国家自然科学基金(62162058)资助项目

由3D灰度体素组成,并呈现出T1、T2、T1c、Flair多模态图像,同时,脑影像中组织结构复杂,边界模糊且存在偏移场效应、容积效应等因素,导致分割难度加大。

图像分割算法整体分为传统图像分割算法与基于深度学习的图像分割算法。前者主要包括阈值、区域生长和合并、聚类、边缘检测等方法,后者主要依靠深度卷积网络实现分割。然而传统图像分割算法已经不能满足当前临床需求,基于神经网络的图像分割手段由于其优秀的表现能力成为当今智能医疗领域的主流趋势。自LONG等^[2]提出了全卷积网络框架后,该架构成为了图像分割领域的主流架构,但其不能满足医学数据对分割精度的要求。针对此问题,RONNEBERGER等^[3]提出了U-net网络,有效的编码-解码结构与跳跃连接结构,有效地防止了信息丢失,但U-net在长期依赖性方面存在局限性。ZHOU等^[4]通过改进U-net的跳跃连接结构提出UNet++网络,使用嵌套和密集连接使编码器和解码器的语义特征更接近,但计算参数较多且容易丢失边缘信息。CHEN等^[5]提出DeeplabV3+网络更加细化了分割结果边界。PENG等^[6]提出了一种多尺度的三维U网络体系结构,使用多个U网络捕获不同分辨率下的远距离空间信息,提高分割精度的同时产生了大量计算参数。OZAN等^[7]提出了一种集成注意力机制的Attention U-net网络,通过加入注意力模块自动聚焦目标,提高模型预测精度,但模型参数量大、耗费时间长。随着深度学习的不断进步,Transformer一种自我注意机制的替代结构促使分割任务迈向新的台阶。CAO等^[8]提出一种U型Transformer网络,虽在一定程度上提

高了分割准确率,但计算量爆炸且兼容性差。CHEN等^[9]将卷积神经网络(convolutional neural network,CNN)与Transformer结合提出了TransUNet模型,恢复了局部空间信息,更加关注细节特征,分割结果证明了CNN与Transformer结合的有效性,但仍会出现低层次特征提取不足、定位能力较差的问题。WANG等^[10]提出了使用通道交叉注意力机制(channel cross transformer,CCT)模块代替U-net跳跃连接结构的UCTransNet网络,弥补了深浅层特征间的语义差距,实现了多尺度特征间的融合,但仍会造成特征冗余。

目前,国内外大多数有关MR脑肿瘤的图像分割方法均采用并行的多模态融合方式处理MR图像,并未充分了解MR多模态图像的成像原理,因此,本文依据MR的成像特点,创新多模态图像融合方法,丰富了图像的特征表达,提取到更细致的特征图像。同时,基于神经网络的医学图像分割算法大多依靠CNN提取特征,并没有考虑到特征之间的长距离依赖关系以及深浅层次特征间的差距,因此,文章在U-net跳跃连接部分引入CCT模块代替传统CNN网络进行多尺度信息融合,缩短了特征间差异。现阶段,为解决3D MR图像分割算法产生大量参数而导致存储空间不足的问题,文中实验将3D MR图像全部切片成2D MR图像进行处理,节省了设备资源与运行时间。

1 本文方法介绍

基于深度学习的脑肿瘤分割分为3个步骤,首先对数据集进行预处理与切片操作,以适应设备和网络的需求,减少计算量;随后搭建分割网络模型,提取肿瘤区域特征,实现精准分割;最后将图像数据可视化并对前沿算法验证本文算法性能。

1.1 总体框架

本文总体框架如图1所示,以U-net网络作为本

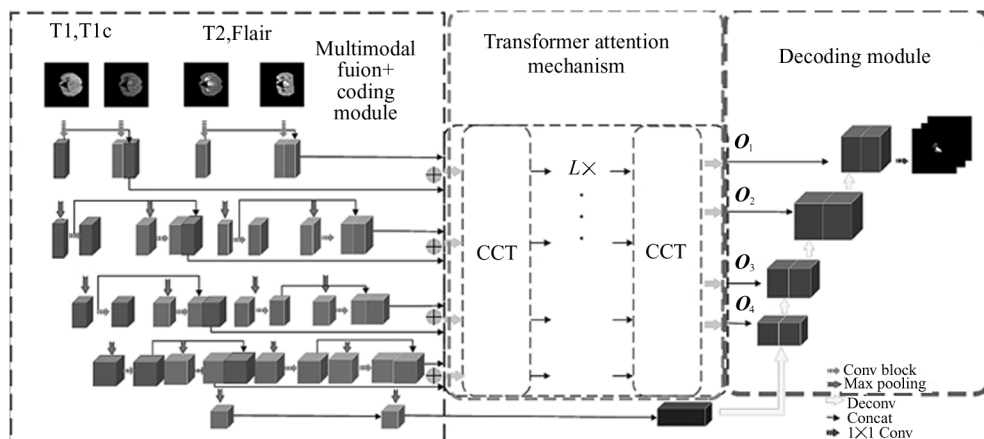


图1 本文网络框架

Fig. 1 Network framework

文基础网络框架,图1左侧部分为多模态融合与下采样模块,根据MR成像原理,在编码阶段不断进行下采样与特征融合操作,通过残差分组融合增加网络特征提取能力;图1中间部分为CCT注意力机制模块,使用Transformer结构代替U-net跳跃连接结构以加强各切片间的空间联系,实现浅层次特征融合、上采样恢复图像分辨率,完成对脑肿瘤区域的分割。

本文提出了一种具有Transformer的多模态融合U网络(multimodal fusion with transformer based on U-net network,MTU-net),有效结合了多模态MR图像信息,残差分组融合多模态病灶特征,提取多尺度特征图像,并通过CCT注意力机制增强切片间的空间与全局联系,更好地分割目标肿瘤区域。

1.2 多模态图像融合

单模态医学图像所提供的信息有限,不能满足临床诊断需求。为获得更好的图像对比度和高质量图像,需进行图像的多模态融合来提高图像的临床适用性。一般来说,按融合时间的不同可分为:早期融合和晚期融合。早期融合实现在早期输入的数据层面,最常用的方法是对各模态信息进行叠加。晚期融合发生在中后期环节,即融合不同结合策略的输出结果。晚期融合图像特征复杂多样,可以产生相比于早期融合好的效果。区别于其他多模态融合方法:融合单模态分割结果^[11]、多模图像早期融合、多路径提取多模态特征^[12]等方法,本文通过分析MR成像原理,组合早晚期融合方式,实现多模态的分组融合,丰富图像特征,提高网络识别能力。

如图1所示,多模态融合发生于整个网络的编码阶段,通过残差连接不断叠加特征图像。根据MR成像原理^[13]可知,T1、T2模态作为分析脑部状况的主要图像,成像相对恒定,T1组织纵向弛豫,可以清晰分辨脑部白质、灰质和脑脊液的分布,水肿区域在T1模态上呈深色,T1模态更注重组织的解剖细节。T2组织横向弛豫,T2模态更能突显病灶区域。Flair模态利用水抑制技术成像,水肿区域呈低信号。T1c作为T1的特殊成像方式,可以显示血液供氧情况,判定肿瘤边缘,因此,通过T1和T1c可以清晰观察脑的解剖结构和相关肿瘤情况,对鉴别肿瘤和非肿瘤区域具有很强的参考价值;Flair和T2模态成像均与液体信号有关,通过像素值亮度差异反映出水肿区域与水肿区域的轮廓、位置及大小。

本文提出的多模态融合方法依据MR成像原理

将T1与T1c、T2与Flair模态分成两组后并行输入网络,通过4次卷积和池化操作,残差叠加相应模态特征图,最终融合两组图像所有特征图,实现特征互补。多模态图像融合参数如表1所示。

表1 多模态图像融合参数表

Tab.1 Multimodal image fusion parameter table

Coding	Input size	Output size
Conv1	224×224×1	224×224×32
Max pooling layer 1 (after merging)	224×224×32 224×224×64	112×112×32 112×112×64
Conv2	112×112×32 112×112×64	112×112×64 112×112×128
Max pooling layer 2	112×112×64 112×112×192	56×56×64 56×56×192
Conv3	56×56×64 56×56×192	56×56×128 56×56×384
Max pooling layer 3	56×56×128 56×56×512	28×28×128 28×28×512
Conv4	28×28×128 28×28×512	28×28×256 28×28×1024
Max pooling layer 4	28×28×1280	14×14×1280
Middle bridge	14×14×2560	

1.3 CCT

为提取多尺度图像特征,有效融合深浅层信息,减小特征差异,引入CCT注意力机制代替U-net跳跃连接结构,CCT具体结构如图2所示。输入图片经卷积块(两个3×3卷积、批量归一化层和Relu函数)和池化(2×2,步长为2)后,分别生成112×112、56×56、28×28、14×14大小的特征图像。CCT将生成的4种尺寸特征图像融合后与相应解码层的解码图像连接起来,上采样还原图像分辨率,最后通过1×1卷积和Softmax函数确定肿瘤区域。

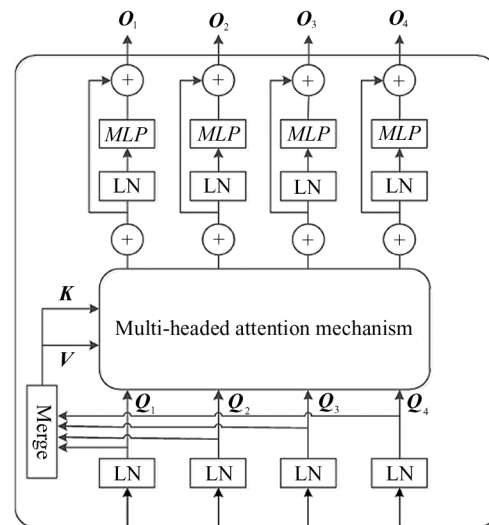


图2 CCT结构图

Fig.2 Structure diagram of CCT

为保留二维切片的空间联系,引入 CCT 模块融合 4 种尺度图像。Transformer 最早应用在语义分割领域,由于其良好的性能后被广泛应用在图像分割领域,但仍需要大量的数据训练作为支撑,将 MR 图像转换成二维切片图像恰恰解决了医学数据稀缺以及 Transformer 所需训练数据量的问题。本文的 Transformer 注意力机制模块内部主要由多尺度特征嵌入、多头通道交叉注意模块以及多层感知器 (multi-layer perceptron, MLP) 构成。其特征图的嵌入为二维矩阵,因此首先将图像变为 $N \times (P^2 \times C)$ 序列,即将图像切分成 N 小块后展平。其中, P 是图像块的大小,本文将特征图重塑为大小 $P, P/2, P/4, P/8$ 的二维平面序列作为注意力机制的输入,保留了原始通道尺寸。随后,将 4 层 $T_i (i=1, 2, 3, 4)$ 标记相连记为 T_Σ , 4 个 T_i 作为多头注意力查询 Q (query), T_Σ 作为注意力地址 K (key) 和值 V (value)。通过式 (1)、(2)、(3), 可计算出 Q_i, K, V 为:

$$Q_i = T_i W_Q, \quad (1)$$

$$K = T_\Sigma W_K, \quad (2)$$

$$V = T_\Sigma W_V. \quad (3)$$

本文 $C_1=128, C_2=384, C_3=1024, C_4=2560$ 通过 CCT 得到相似矩阵 M_i , 通过 M_i 对 V 特征值进行加权, 通过式 (4) 得到:

$$CA_i = M_i V^T = \sigma \left[\varphi \left(\frac{Q_i^T K}{\sqrt{C_\Sigma}} \right) \right] V^T = \sigma \left[\varphi \left(\frac{W_{Q_i}^T T_i^T T_\Sigma W_K}{\sqrt{C_\Sigma}} \right) \right] W_V^T T_\Sigma, \quad (4)$$

式中, $\varphi(\cdot)$ 和 $\sigma(\cdot)$ 分别代表实例归一化和 Softmax 函数, 多头交叉注意后的输出 MCA_i 计算如式 (5) 所示:

$$MCA_i = \frac{CA_i^1 + CA_i^2 + \dots + CA_i^N}{n}, \quad (5)$$

式中, n 代表头的数量, 设置为 4。随后应用简单的 MLP 模块和残差算子, 由式 (6) 得到输出 O_i :

$$O_i = MCA_i + MLP(Q_i + MCA_i). \quad (6)$$

将输出 O_1, O_2, O_3, O_4 输入到解码层中, 通过与上采样特征相连接, 更好地融合多尺度特征图像, 消除编码与解码特征间的歧义。

2 实验过程与分析

2.1 数据集及预处理

本实验使用的数据集为 BraTS 2018 数据集, 是由 MICCAI 协会举办的肿瘤分割比赛数据集, 该数据集包括 285 个病例。数据格式为 .nii.gz, 分别包含

4 种模态图像和一个标签图像 T1、T2、T1c、Flair, seg, 其中 seg 是由诸多专家手动标注的真实标签, 具有很强的参考性。每张图像大小为 $240 \times 240 \times 155$ 。标签按像素灰度值标记, 实验中需要分割的 3 块肿瘤区域分别记为核心肿瘤区域 (tumor core, TC)、整个肿瘤区域 (whole tumor, WT) 和增强肿瘤区域 (enhance tumor, ET)。标签标注区域如图 3 所示。

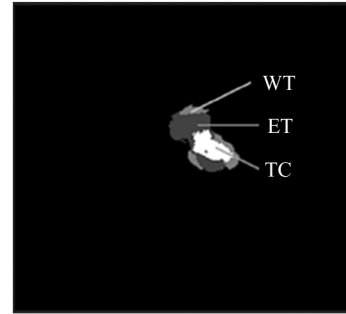


图 3 标签标注图

Fig. 3 Ground truth

为减少计算成本, 本文将 3D MR 图像及标签切成 155 张 240×240 大小的二维图片, 图 4 为 4 种模态和标签的切片图像。

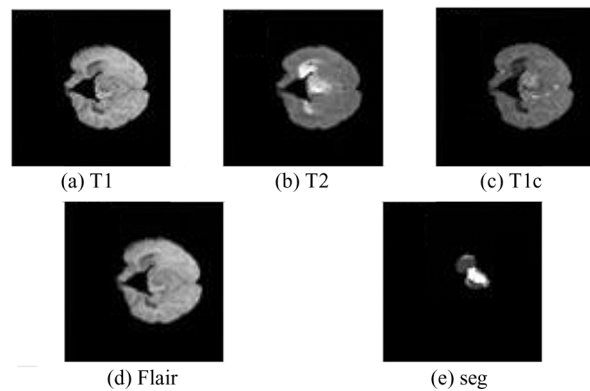


图 4 4 种模态和标签的切片图像

Fig. 4 Slices of four modalities and label

其次, 由于不同病人的肿瘤区域位置、形状以及对比度差异均会影响肿瘤在 MR 图像上的强度变化, 导致其与健康脑组织强度值重叠, 增加分割难度。因此, 采用 z-score 归一化方法将图像强度转换到固定范围, 本文首先去除图像顶部和底部 1% 的强度, 随后对非零区域归一化, z-score 归一化公式如 (7) 所示:

$$F(x, y) = \frac{f(x, y) - \text{mean}[f(x, y)]}{\text{std}[f(x, y)]}, \quad (7)$$

式中, $F(\cdot)$ 与 $f(\cdot)$ 代表像素值, x, y 代表像素坐标。

同时, 在 MRI 机器扫描过程中, 不均匀磁场会产

生伪影,造成信号强度低频变化。由此会导致医学图像处理性能显著下降,被称为偏置场^[14]效应。因此,在进行分割之前,需要校正 MR 图像的偏置场。N4ITK 算法^[15]是目前校正医学图像偏置场最流行、最有效的算法。本文采用 N4ITK 算法对 MR 图像进行校正。除此之外,为减少无关背景影响,本文将 2D MR 图像统一裁剪为 224×224×155 大小,并进行翻转、平移等数据增强操作以防止模型过拟合。

2.2 实验内容

2.2.1 评价指标

为验证脑肿瘤分割结果的准确性以及本文网络的可靠性,本文采用 4 种经典指标对脑肿瘤分割结果进行测评。

1) Dice 相似性系数 (dice similarity coefficient, DSC),用来判断分割结果与真实标签间元素的相似度,是一种比较综合的评价指标,适用于医学图像分割,取值范围为[0,1],DSC 越高代表分割结果越好,如式(8)所示:

$$DSC = \frac{2TP}{FP + 2TP + FN}, \quad (8)$$

式中,TP 代表预测结果完全正确,即预测结果是正类,真实也是正类,FP 代表预测结果错误,预测结果是正类,真实为负类,FN 代表预测结果错误,预测结果是负类,真实为正类。

2) 灵敏度 (sensitivity, S),取值范围为[0,1],统计像素点正确分类的像素比例,衡量分割正确区域的能力,S 值越高代表模型越灵敏,如式(9)所示:

$$S = \frac{TP}{TP + FN}. \quad (9)$$

3) 阳性预测值 (positive predictive value, PPV),取值范围为[0,1],在医学领域应用广泛,通

常用于判断预测所有阳性病例中真阳性个数,应用在图像分割领域里表示预测为阳性的像素点占真阳性像素点的比例,PPV 越接近 1 分类正确率越高,如式(10)所示:

$$PPV = \frac{TP}{TP + FP}. \quad (10)$$

4) Jaccard 相似性指数,用于描述两个集合之间的距离,Jaccard 值越大相似度越高,如式(11)所示:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}, \quad (11)$$

式中,A 代表真实值,B 为预测结果,通过交并比反映出真实值和预测结果之间的相似性与差异性。

2.2.2 实验过程

本文实验在 NVIDIA GTX-3060Ti 的工作站进行训练,编程环境为 python 3.7。采用五折交叉验证方法,batch 设置为 20,epoch 设置为 100,使用 Adam 优化器进行权重更新,初始学习率为 0.0001,使用交叉熵损失函数作为本文损失函数 L,其如式(12)所示:

$$L = \frac{1}{N} \sum L_i = -\frac{1}{N} \sum_{c=1}^M y_k \log \sum (p_k), \quad (12)$$

式中,M 为类别数量,y_k 为符号函数,样本 i 的真实类别为 c,p_k 为 i 属于 c 的预测概率。

为验证本文多模态融合方法的有效性,本文通过消融实验对比常用并行多模态融合方法,实验对照组设置为 U-net 网络。数据结果如表 2 所示,可以看出采用本文多模态融合方法的结果数据在 4 种指标上大多高于使用并行多模态融合方法,除 S 指标表现效果略差,其他指标均高于未使用本文多模态融合方法的 1%—4%左右,尤其是在划分 TC 与 WT

表 2 消融实验

Tab. 2 Ablation experiments

Method	DSC			S			PPV			Jaccard		
	TC	WT	ET	TC	WT	ET	TC	WT	ET	TC	WT	ET
Label/%												
U-net+Parallel fusion method	68.96	75.35	60.3	77.29	74	63.1	68.51	81.25	71.72	58.98	63.12	48.97
U-net+New fusion method	69.93	76.97	60.72	78.24	73.12	64.02	69.74	83.96	73.62	61.54	66.89	52.41
MTU-net+Parallel fusion method	72.90	78.16	70.24	78.11	77.62	71.28	76.83	82.39	84.44	66.37	65.78	59.31
MTU-net+New fusion method	74.43	80.02	71.25	80.23	76.18	71.81	78.66	86.89	85.74	69.54	67.13	61.42

上表现更优,充分验证了本文多模态融合方法的有效性。

2.2.3 实验结果

本文提出的方法与主流医学图像分割算法进行对比实验,实验数据如表3所示。其中,U-net网络为本文改进的基础网络框架;DeeplabV3+网络可以有效提取多尺度特征图像,在分割领域中表现优秀;Swin-Unet为具有U型结构的Transformer新型分割网络。本实验对比以上3种前沿网络,实验数据更具说服力。

表3 实验数据

Tab.3 Experimental data

Parameters/%	Network	TC	WT	ET
DSC	DeeplabV3+	73.65	73.71	65.94
	U-net	68.96	75.35	60.3
	Swin-Unet	72.94	71.06	63.28
	MTU-net	74.43	80.02	71.25
S	DeeplabV3+	74.4	71.08	64.43
	U-net	77.29	74	63.1
	Swin-Unet	76.14	70.69	67.34
	MTU-net	80.23	76.18	71.81
PPV	DeeplabV3+	76.97	82.57	80.13
	U-net	68.51	81.25	71.72
	Swin-Unet	70.4	84.53	81.64
	MTU-net	78.66	86.89	85.74
Jaccard	DeeplabV3+	66.57	62.55	56.98
	U-net	58.98	63.12	48.97
	Swin-Unet	66.17	65.94	59.39
	MTU-net	69.54	67.13	61.42

MTU-net网络在DSC、S、PPV以及Jaccard系数的TC、WT、ET表现分别为74/80/71、80/76/72、79/87/86、69/67/61百分点。3块分割区域中本文MTU-net的DSC较U-net网络提升5—11个百分点,S提升2—9百分点,PPV提升6—14百分点,Jaccard系数提升4—12百分点;MTU-net较DeeplabV3+的DSC高出1—6百分点,S高出5—8百分点,PPV高出2—6百分点,Jaccard系数高出3—4百分点;MTU-net与Swin-Unet相比,DSC提升1.5—9百分点,S提升4—5.5百分点,PPV提升2—8百分点,Jaccard系数提升1—3百分点。对比数据可知MTU-net的分割指标均优于以上3种前沿分割网络,足以证明本文网络性能的优越性。

本实验的分割结果如图5所示。其中,(a)为U-net的分割结果,(b)为DeeplabV3+的分割结果,(c)为Swin-Unet的分割结果,(d)为MTU-net的分割结果,(e)为真实标签。对比真实标签,可以看出U-net和Swin-Unet存在明显漏分和误分现象。通过量化指标分析与分割结果可视化图均可证明本文方法分割准确率高,更具临床意义。

3 结论

本文提出了一种2D MR多模态脑肿瘤分割方法。根据MR成像原理设计多模态图像融合方法,在编码阶段利用残差分组融合T1与T1c模态、T2与Flair模态,精准提取脑肿瘤目标位置及形状。引

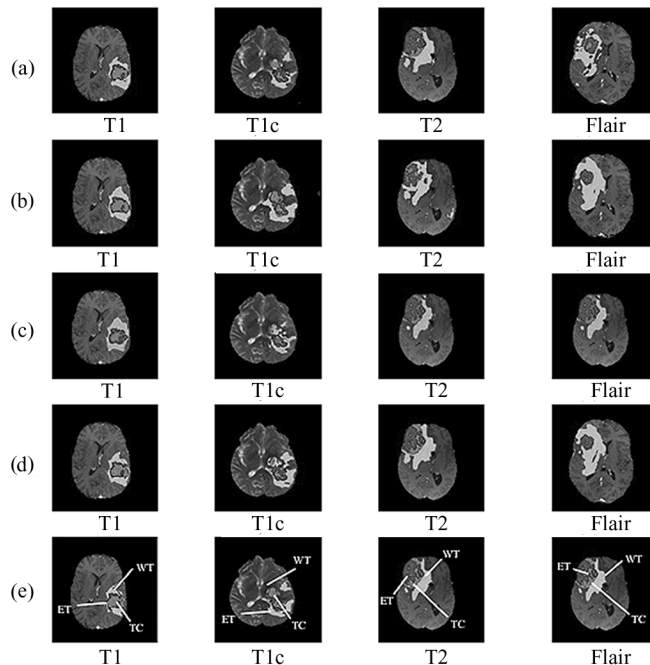


图5 4种模态分割图像:(a) U-net; (b) DeeplabV3+; (c) Swin-Unet; (d) MTU-net; (e) Ground truth

Fig.5 Four modal segmentation images:(a) U-net; (b) DeeplabV3+; (c) Swin-Unet; (d) MTU-net; (e) Ground truth

入 CCT 模块替代 U-net 的跳跃连接结构,消除深浅层次特征间差异,关注切片间的空间联系与全局依赖性。通过实验证明,本文提出的方法可以更有效地提取脑部不同区域的肿瘤信息。与其他前沿分割方法相比,本文提出的方法在性能上更优,为 MR 多模态图像分割提出了新的解决思路。未来的工作更多是探索多模态三维体素图像分割方法,寻求计算参数所需设备要求与精度之间的最优值,进一步提升网络在医学图像中的适用性。

参考文献:

- [1] FITZMAURICE C, ALLEN C, BARBER R M, et al. Global regional and national cancer incidence mortality years of life lost years lived with disability and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study[J]. *JAMA Oncology*, 2017, 3(4): 524-548.
- [2] LONG J, SHELHAMER E, DARRELL T. An adaptive sampling scheme to efficiently train fully convolutional networks for semantic segmentation[EB/OL]. (2017-12-22) [2022-05-10]. <https://arxiv.org/abs/1709.02764v1>.
- [3] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[EB/OL]. (2015-05-18) [2022-05-10]. <https://arxiv.org/abs/1505.04597>.
- [4] ZHOU Z W, SIDDIQUEE M R, TAJBAKSH N, et al. Unet++: a nested u-net architecture for medical image segmentation[C]//International Workshop on Deep Learning in Medical Image Analysis, International Workshop on Multimodal Learning for Clinical Decision Support, DL MIA, ML-CDS, 2018, September 20, 2018, Granada, Spain. Cham: Springer, 2018, 11045: 3-11.
- [5] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[EB/OL]. (2018-08-22) [2022-05-10]. <https://arxiv.org/abs/1802.02611v2>.
- [6] PENG S T, CHEN W, SUN J W, et al. Multi-scale 3D U-nets: an approach to automatic segmentation of brain tumor[J]. *International Journal of Imaging Systems and Technology*, 2020, 30(1): 5-17.
- [7] OKTAY O, SCHLEMPER J, FOLGOC L, et al. Attention U-net: learning where to look for the pancreas [EB/OL]. (2018-05-20) [2022-05-10]. <https://arxiv.org/abs/1804.03999>.
- [8] CAO H, WANG Y Y, CHEN J, et al. Swin-Unet: Unet-like pure transformer for medical image segmentation [EB/OL]. (2021-05-12) [2022-05-10]. <https://arxiv.org/abs/2105.05537>.
- [9] CHEN J N, LU Y Y, YU Q H. TransUNet: transformers make strong encoders for medical image segmentation [EB/OL]. (2021-02-8) [2022-05-10]. <https://arxiv.org/abs/2102.04306>.
- [10] WANG H N, CAO P, WANG J Q. UCTransNet: rethinking the skip connections in U-net from a channel-wise perspective with transformer[EB/OL]. (2022-09-09) [2022-05-10]. <https://arxiv.org/abs/2109.04335v3>.
- [11] SUN J, PENG Y, GUO Y, et al. Segmentation of the multi-modal brain tumor image used the multi-pathway architecture method based on 3D FCN[J]. *Neurocomputing*, 2021, 423(1): 34-45.
- [12] XIONG W, ZHOU L, YUE L, et al. Research on MRI brain tumor image segmentation based on multiscale feature extraction[J]. *Journal of Optoelectronics · Laser*, 2021, 32(11): 7-8.
熊炜,周蕾,乐玲,等.基于多尺度特征提取的MRI脑肿瘤图像分割研究[J].*光电子·激光*, 2021, 32(11): 7-8.
- [13] JEFFREY R, PETRELLA JAMES M. MR perfusion imaging of the brain techniques and applications[J]. *Provenzale American Journal of Roentgenology*, 2000, 175(1): 207-219.
- [14] UKA A, POLISI X, BARTHES J, et al. Effect of preprocessing on performance of neural networks for microscopy image classification[C]//2020 International Conference on Computing, Electronics & Communications Engineering, August 17-18, 2020, Online Via Zoom. New York: IEEE, 2020: 162-165.
- [15] TUSTISON N J, AVANTS B B, COOK P A, et al. N4ITK: improved N3 bias correction[J]. *IEEE Transactions on Medical Imaging*, 2010, 29(6): 1310-1320.

作者简介:

张宏立 (1972—),男,博士,教授,博士生导师,主要从事机器视觉方面的研究。