

# 跨模态融合的双注意力脑肿瘤分割算法<sup>①</sup>

张鹏跃<sup>1</sup>, 马巧梅<sup>1,2</sup>

<sup>1</sup>(中北大学 软件学院, 太原 030051)

<sup>2</sup>(山西省医学影像人工智能工程技术研究中心 (中北大学), 太原 030051)

通信作者: 马巧梅, E-mail: maqiaomei@nuc.edu.cn



**摘要:** 针对脑肿瘤多模态信息融合不充分以及肿瘤区域细节信息丢失等问题, 提出了一种跨模态融合的双注意力脑肿瘤图像分割网络 (CFDA-Net). 在编码器-解码器的基础结构上, 首先在编码器分支采用密集块与大内核注意力并行的新卷积块, 可以使全局和局部信息有效融合且可以防止反向传播时梯度消失的问题; 其次在编码器的第 2、3 和 4 层的左侧加入多模态深度融合模块, 有效地利用不同模态间的互补信息; 然后在解码器分支使用 Shuffle Attention 注意力将特征图分组处理后再聚合, 其中分组的子特征一分为二地获取空间与通道的重要注意特征. 最后使用二进制交叉熵 (binary cross entropy, BCE)、Dice Loss 与  $L_2$  Loss 组成新的混合损失函数, 缓解了脑肿瘤数据的类别不平衡问题, 进一步提升分割性能. 在 BraTS2019 脑肿瘤数据集上的实验结果表明, 该模型在整体肿瘤区域、肿瘤核心区域和肿瘤增强区域的平均 Dice 系数值分别为 0.887、0.892 和 0.815. 与其他先进的分割方法 ADHDC-Net、SDS-MSA-Net 等相比, 该模型在肿瘤核心区域和增强区域具有更好的分割效果.

**关键词:** 脑肿瘤; 多模态; 深度融合; 注意力机制; 图像分割

引用格式: 张鹏跃, 马巧梅. 跨模态融合的双注意力脑肿瘤分割算法. 计算机系统应用, 2024, 33(1): 119-126. <http://www.c-s-a.org.cn/1003-3254/9361.html>

## Brain Tumor Segmentation Algorithm Based on Cross-modal Fusion Dual Attention

ZHANG Peng-Yue<sup>1</sup>, MA Qiao-Mei<sup>1,2</sup>

<sup>1</sup>(School of Software, North University of China, Taiyuan 030051, China)

<sup>2</sup>(Shanxi Medical Imaging Artificial Intelligence Engineering Technology Research Center (North University of China), Taiyuan 030051, China)

**Abstract:** This study proposes a cross-modal fusion dual attention net (CFDA-Net) for brain tumor image segmentation to solve the insufficient multi-modal information fusion of brain tumors and detail loss of the tumor regions. Based on the encoder-decoder architecture, a new convolutional block with dense blocks and large kernel attention parallel is first adopted in the encoder branch, which can effectively fuse global and local information and prevent the gradient vanishing during backpropagation. Secondly, a multi-modal deep fusion module is added to the left sides of the second, third, and fourth layers of the encoder to effectively utilize the complementary information among different modalities. Then, in the decoder branch, Shuffle Attention is adopted to group the feature maps and aggregate them, and the subfeatures of the group are divided into two parts to obtain important attention features of space and channels. Finally, binary cross entropy (BCE), Dice Loss, and  $L_2$  Loss are employed to form a new hybrid loss function, which alleviates the category imbalance of brain tumor data and further improves the segmentation performance. The experimental results on the BraTS2019 brain tumor dataset show that the average Dice coefficient values of the model in the whole tumor region, tumor core region, and tumor enhancement region are 0.887, 0.892, and 0.815 respectively. The proposed model has better segmentation performance in the core and enhanced regions of tumors than other advanced segmentation methods such as ADHDC-Net

① 基金项目: 山西省自然科学基金 (20210302123019)

收稿时间: 2023-07-13; 修改时间: 2023-08-11; 采用时间: 2023-08-18; csa 在线出版时间: 2023-11-24

CNKI 网络首发时间: 2023-11-29

and SDS-MSA-Net.

**Key words:** brain tumor; multimodal; deep fusion; attention mechanism; image segmentation

脑肿瘤是大脑中异常细胞的团块或集合体,其反过来破坏大脑的组织功能.脑肿瘤可主要分为恶性肿瘤和良性肿瘤.恶性脑肿瘤是目前最严重的癌症之一,日益威胁着人类的健康<sup>[1]</sup>.脑肿瘤被世界卫生组织划分为I-IV级,脑肿瘤的级别越高,患者的生存时间越短<sup>[2,3]</sup>,所以对脑肿瘤早期的诊断和治疗十分重要.手动分割脑肿瘤耗时长而且不同医生对同一病患的分割结果也参差不齐,因此脑肿瘤算法的研究对高效地诊断脑肿瘤具有重要的科学价值和临床意义.

随着深度学习方法的发展,卷积神经网络成为诊断脑肿瘤区域的主要方法.尤其是基于编码-解码结构的UNet<sup>[4]</sup>经典分割网络.之后UNet类型的结构得到进一步发展,如嵌套和密集连接的UNet++<sup>[5]</sup>、使用数据相关型上采样与残差块的改进UNet++<sup>[6]</sup>、DenseNet<sup>[7]</sup>网络与UNet结合的DenseUNet<sup>[8]</sup>、用于体积分割Vnet<sup>[9]</sup>结构.这些结构虽加强了对局部特征的提取,但缺乏获取远程特征依赖和多尺度信息的能力.

对此,文献[10]提出了TransBTS网络,首次将Transformer集成到3D CNN的编码-解码结构中,加强对全局特征的提取.文献[11]提出了不同扩张率的分层卷积与肿瘤区域关系引导的注意力的轻量级的ADHDC-Net网络.文献[12]设计了SDS-MSA-Net,它分别从3D和2D输入中提取特征,并使用选择性深度监督来辅助输出,加快了模型的收敛速度,但同时

3D和2D资源处理,加大了计算量.以上改进的结构分别加强了全局特征和多尺度注意特征的提取能力,但目前大多数网络都局限于多模态脑肿瘤数据输入级简单的拼接融合,不能充分利用不同模态对之间互补的融合信息.因此文献[13]设计了多模态图像的像素级融合与特征级融合的两阶段网络,以实现多模态信息更精细地利用,但该网络核心肿瘤区域的分割精度还有待提升.

针对上述问题,本文提出了一种跨模态融合的双注意力脑肿瘤图像分割网络(cross-modal fusion dual attention net, CFDA-Net).为解决普通卷积感受野受限,本文设计了密集并行大核注意力的新卷积块,可以有效提取全局和局部信息;为了加强不同模态对之间的深度融合,提出了多模态深度融合模块,更好地利用模态的互补信息;为关注重要特征信息,消除对冗余特征的关注,在解码器分支使用Shuffle Attention<sup>[14]</sup>注意力以获取空间和通道重要特征.针对类别不平衡问题,本文使用了二进制交叉熵、Dice与L<sub>2</sub>组成的混合损失函数,进一步提升脑肿瘤分割精度.

## 1 本文方法

### 1.1 CFDA-Net网络总体结构

本文提出的跨模态的双注意力脑肿瘤分割网络的整体结构如图1所示.

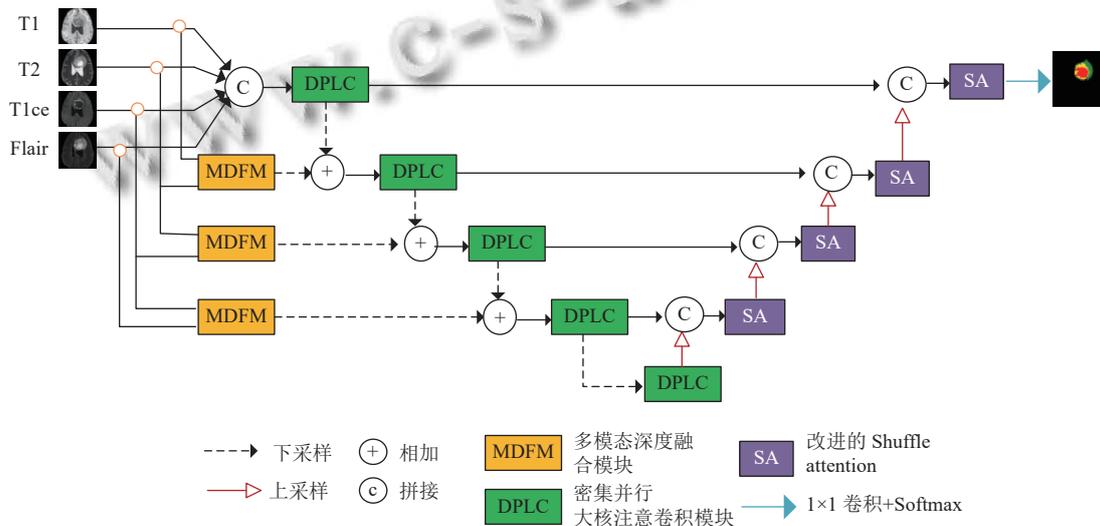


图1 CFDA-Net网络结构

图1中左侧的编码器分支有5个密集并行大核注意力卷积(dense parallel large kernel attention convolution, DPLC)模块, DPLC块一方面利用密集卷积加强特征重用,防止梯度消失;另一方面利用VAN视觉注意力网络提出的大内核注意力卷积(large kernel attention, LKA)块<sup>[15]</sup>来加强获取远距离特征依赖的能力,将这两个模块并行相加融合,可以很好解决普通卷积特征提取时的感受野不足的问题,进而使局部与全局信息融合更充分。然后在编码器的第2、3和4层左侧添加多模态深度融合(multimodal deep fusion module, MDFM)模块,有效地利用不同模态对间的互补信息,使信息融合更充分。另外除了第3、第4层MDFM模块后的下采样采用特殊的步长为4、步长为8的 $3 \times 3$ 卷积外,其余部分都使用步长为2的 $3 \times 3$ 卷积来进行下采样。

将编码器的各层级分支通过跳跃连接,与右侧对应的解码器分支经过上采样后进行通道级拼接融合,融合后使用SA(shuffle attention)注意力块进行分组后使用空间注意力和通道注意力,为了加强对重要特征的关注,最后再经过一个 $1 \times 1$ 卷积将对应通道数减半。在最后一个解码器的SA模块后使 $1 \times 1$ 卷积与Softmax激活函数得到整体肿瘤区域、核心肿瘤区域和增强肿瘤区域的3通道网络模型预测图,其中每个通道代表一个肿瘤分割区域。

## 1.2 密集并行大核注意力卷积模块

为了解决普通卷积特征提取时感受野受限的问题,本文设计了密集并行大核注意力卷积(DPLC)模块,如图2所示。

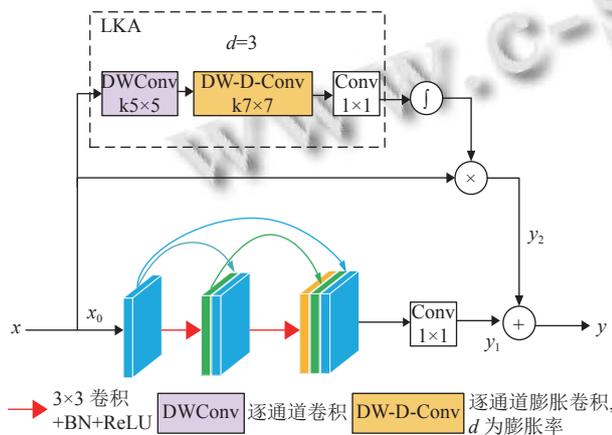


图2 密集卷积并行大核注意力模块结构图

首先第1个分支引用了经典的DenseNet网络的密集块,它可以加强特征重用且能防止梯度消失,本文

为了减少参数和计算量,将密集块中的层数设定为2层,每个层将所有先前层的输出作为其输入,并将其自身的特征传递给后续层。然后再通过 $1 \times 1$ 卷积来降低通道数,该密集块如式(1)所示:

$$X_l = H([X_0, \dots, X_{l-1}]) \quad (1)$$

其中, $[X_0, \dots, X_{l-1}]$ 代表第 $0, \dots, l-1$ 层输出特征图在通道维度进行拼合,密集块分支的 $X_0$ 即为该模块初始的输入 $X$ 。非线性变换 $H(\cdot)$ 由 $3 \times 3$ 卷积、批归一化和ReLU激活函数组成。然后将密集特征再经过 $1 \times 1$ 卷积降维通道以得到第1分支的输出特征图 $y_1$ 。

第2个分支使用视觉注意力网络VAN<sup>[15]</sup>中提出降低参数量的大核注意力卷积(LKA)块,该网络提出了一种适用于大核卷积的分解方式,一个 $K \times K$ 的卷积核可以分解为3部分:一个 $(2d-1) \times (2d-1)$ 的逐通道(DWConv)卷积;一个 $K/d \times K/d$ 的逐通道膨胀(DW-D-Conv)卷积;一个 $1 \times 1$ 卷积。其中 $K$ 为卷积核大小, $d$ 为膨胀率。LKA计算公式如式(2)所示:

$$LKA = \text{Conv}_{1 \times 1}(\text{DW-D-Conv}(\text{DWConv}(X))) \quad (2)$$

然后将LKA经过Sigmoid激活函数得到包含全局信息的注意力系数,将注意力系数与第3个恒等连接分支进行元素级相乘得到特征图 $y_2$ ,如式(3)所示。其中 $\sigma$ 为Sigmoid激活函数。

$$y_2 = X \cdot \sigma(LKA) \quad (3)$$

最后将 $y_1$ 与 $y_2$ 特征图进行元素级相加生成该模块的最终输出,如式(4)所示:

$$y = y_1 + y_2 \quad (4)$$

## 1.3 多模态深度融合模块

如今大多数网络模型针对多模态数据只是输入级的简单拼接,不能有效利用模态间的互补信息,为了更好地利用模态对之间的互补信息,本文提出了多模态深度融合(MDFM)模块。并在编码器的第2、3和4分支左侧分别添加了T1-T2、T2-T1ce与T1ce-Flair的模态对作为MDFM模块的输入,这样可以对编码器主分支不断地补充空间特征信息,使特征信息融合更加充分<sup>[16]</sup>,有利于提取脑肿瘤的增强和坏死区域的细节特征。图3为MDFM多模态深度融合模块结构图。

图3以图1中第2层编码器左侧的MDFM模块为例,该模块输入的ModelA为T1模态、ModelB为

T2 模态, 这两个模态的空间大小都为  $160 \times 160$ 、通道数为 1, 本文首先将这两个模态进行拼接融合得到 T1-T2 的双通道特征, 将 T1-T2 的双通道特征分别输入到 3 个分支, 对 3 个分支开始处依次使用逐通道卷积、全局平均池化、全局最大池化, 同时对三分支的结尾处使用 ReLU 激活函数、 $1 \times 1$  卷积进行通道升维得到特征图  $F_1$ 、 $F_2$  和  $F_3$ . 最后将 3 个分支的特征图相加, 再将相加后的特征图使用 Sigmoid 函数得到融合权重系数  $\alpha$ . 将模态 A 使用  $1 \times 1$  卷积升为 C 维后再与  $\alpha$  相乘, 同理另一个模态 B 升维后与  $(1-\alpha)$  乘积, 将两个乘积结果进行相加后, 得到该模块的输出特征  $S$ . 该模块如式 (5)–式 (9) 所示:

$$F_1 = \text{Conv}_{1 \times 1}(\delta(\text{DWConv}[A, B])) \quad (5)$$

$$F_2 = \text{Conv}_{1 \times 1}(\delta(\text{GAP}[A, B])) \quad (6)$$

$$F_3 = \text{Conv}_{1 \times 1}(\delta(\text{GMP}[A, B])) \quad (7)$$

$$\alpha = \sigma(F_1 + F_2 + F_3) \quad (8)$$

$$S = \text{Conv}_{1 \times 1}(A) \cdot \alpha + \text{Conv}_{1 \times 1}(B) \cdot (1 - \alpha) \quad (9)$$

其中,  $\text{DWConv}(\cdot)$  是逐通道卷积,  $\text{Conv}_{1 \times 1}(\cdot)$  目的是进行通道升维,  $\delta(\cdot)$  为 ReLU 激活函数,  $\sigma(\cdot)$  为 Sigmoid

激活函数,  $\text{GAP}(\cdot)$  为全局平均池化,  $\text{GMP}(\cdot)$  为全局最大池化,  $\alpha$  为从激活函数中得到的融合权重系数.

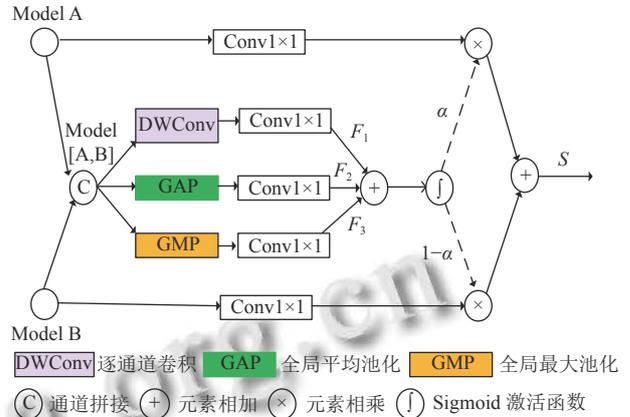


图3 多模态深度融合模块结构图

### 1.4 改进的 Shuffle Attention 注意力模块

之前的大多数网络都是在跳跃连接处添加相应的注意力机制, 但是这样只加强了对应编码器分支重要的特征关注, 而忽视了对应解码器分支对后续特征的影响, 因此本文在解码器分支引入了轻量级的 SA<sup>[14]</sup> 注意力模块, 本文在 SA 模块的最后增加了  $1 \times 1$  卷积、BN 归一化和 ReLU 激活函数来对解码器分支进行降维. 改进后的 SA 模块结构如图 4.

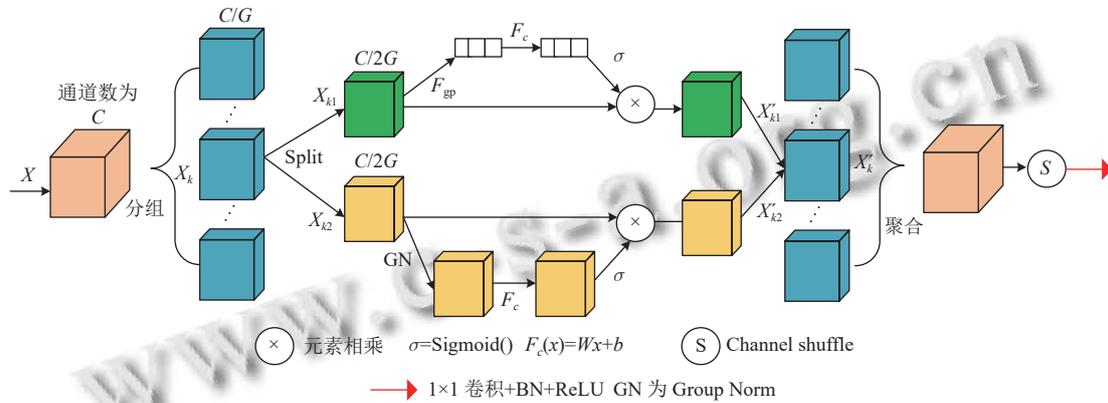


图4 Shuffle Attention 注意力结构图

对于输入特征映射  $X \in R^{C \times H \times W}$ , 其中  $C$  是通道的数量,  $H$ 、 $W$  分别为空间的高度和宽度, SA 最初在通道维度上将  $X$  划分为  $G$  个组, 即  $X = [X_1, \dots, X_k, \dots, X_G]$ ,  $X_k \in R^{C/G \times H \times W}$ . 然后, 分组的子特征  $X_k$  在信道维度中被分成两个分支:  $X_{k1}, X_{k2} \in R^{C/2G \times H \times W}$ . 一个分支通过获取重要关注的信道序列, 另一个分支获取空间重点目标区域. 对于通道注意分支, 先使用全局平均池化来嵌入全局信息  $s \in R^{C/2G \times 1 \times 1}$  来产生信道统计量, 如式 (10) 所示.

$$s = F_{gp}(X_{k1}) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{k1}(i, j) \quad (10)$$

这里通道注意分支结果为  $X'_{k1}$ , 其中  $W_1 \in R^{C/2G \times 1 \times 1}$ ,  $b_1 \in R^{C/2G \times 1 \times 1}$  用于缩放和移位来增强  $X_{k1}$  的特征表示, 如式 (11) 所示. 这里的  $\sigma(\cdot)$  是 Sigmoid 激活函数.

$$X'_{k1} = \sigma(F_c(s)) \cdot X_{k1} = \sigma(W_1 s + b_1) \cdot X_{k1} \quad (11)$$

对于空间注意分支, 首先对  $X_{k2}$  进行分组归一化

(GN) 得到空间统计量, 其中  $W_2 \in R^{C/2G \times 1 \times 1}$ ,  $b_2 \in R^{C/2G \times 1 \times 1}$  用于缩放和移位来增强  $X_{k2}$  的特征表示如式 (12) 所示:

$$X'_{k2} = \sigma(W_2 \cdot GN(X_{k2}) + b_2) \cdot X_{k2} \quad (12)$$

然后将通道和空间分支拼接得到子特征  $X'_k = [X'_{k1}, X'_{k2}] \in R^{C/G \times H \times W}$  得, 将子特征聚合后使用 Channel shuffle 使跨组信息能够沿着通道维度流动, 最后使用  $1 \times 1$  卷积降低通道维数, 并对特征图进行批归一化和用 ReLU 激活函数提高非线性特征.

## 2 实验设置

### 2.1 数据集

本文数据来源为 (brain tumor segmentation, BraTS) 竞赛 2019 年公开的数据集, 其中 BraTS2019 数据集有 76 例低级别脑胶质瘤病例与 259 例高级别脑胶质瘤病例, 本文将 BraTS2019 数据集按照 8:2 的比例对训练集和测试集的数据进行划分.

### 2.2 数据预处理

由于使用 3D 格式图像进行训练时间周期长, 并且需要更好的 GPU 和更多内存, 因此选择使用 2D 切片训练本文提出的网络. 每个模态的 3D 数据大小为  $240 \times 240 \times 155$ , 由于脑肿瘤数据的外边缘区域有很多无用的背景信息, 会出现数据类别不平衡的问题. 所以本文首先将 3D 数据的空间大小裁减为  $160 \times 160 \times 155$ , 以消除无用空间背景信息, 然后再沿着通道方向对数据切片, 将 3D 数据变为 155 张  $160 \times 160$  的 2D 切片以满足本文模型训练的需要.

因其成像机制不同, 不同模态的影像对比度也不同, 所以使用归一化操作让不同模态的数据强度变得均衡, 有利于模型利用不同模态间的互补信息. 归一化操作如式 (13) 所示:

$$z = \frac{x - \bar{x}}{\lambda} \quad (13)$$

其中,  $x$  是裁减后的 2D 切片图像,  $\bar{x}$  是输入图像的平均值,  $\lambda$  是输入图像的标准差,  $z$  是经过归一化处理后的图像.

如图 5 所示, 经过本文预处理后的某一病例切片的 4 种模态和真实标签的图像, 从左到右依次为 T1、T1ce、T2、Flair 和 GT (ground truth) 标签. 其中网络模型预测图和真实标签中的绿色为整体肿瘤区、黄色为增强肿瘤区与红色为坏死和非增强肿瘤区域.

### 2.3 实验环境配置

本文模型使用的深度学习框架是 PyTorch 1.11.0,

Cuda 11.3 的软件版本, 硬件环境为 CPU 是 15 vCPU AMD EPYC 7543 32-Core Processor、30 GB 内存与使用的 GPU 是 NVIDIA RTX A5000, 显存大小是 24 GB. 网络模型使用 Adam (adaptive moment estimation) 算法<sup>[17]</sup> 作为优化器来更新模型的权值, 初始学习率为 0.003, 动量值设定为 0.09, 权重衰减系数为  $1E-5$ . 设定的批量大小为 20, 训练周期数为 100.

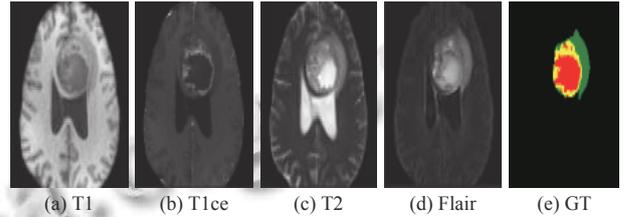


图 5 经过预处理后的 4 种模态图像与 GT 真实标签

### 2.4 评价指标

为评测模型的性能, 本文对整体肿瘤区域、核心肿瘤区域和增强肿瘤区域均使用 Dice 相似系数 (Dice similarity coefficient) 来评测模型性能, 如式 (14) 所示. Dice 数值越趋近于 1, 说明脑肿瘤分割结果与人工标签结果越相似, 分割效果也就越好. 而 Dice 数值越趋近于 0, 分割效果则越差.

$$Dice = \frac{2|Y \cap P|}{|Y| + |P|} \quad (14)$$

其中,  $Y$  为人工分割标签的集合,  $P$  为模型预测值的集合,  $Y \cap P$  表示为人工标签和模型预测分割集合的交集.

### 2.5 损失函数

在不同医学数据集的分割任务上常使用 (binary cross entropy, BCE) 二进制交叉熵作为损失函数, 如式 (15) 所示:

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (15)$$

其中,  $N$  是输出像素的总数,  $y_i$  是对应于第  $i$  个像素的标签值, 并且  $p_i$  是第  $i$  个像素的模型预测值. 由于 BCE 损失函数对前景像素和背景像素赋予相同权重, 但是多模态脑肿瘤的前景像素和背景像素比例差别大, 前景像素仅占其中的少数, 数据存在类别不平衡的问题, 为解决数据类别不平衡的问题, 本文在 BCE 损失函数的基础上增加了 Dice 损失函数, Dice 损失函数如式 (16) 所示:

$$L_{Dice} = 1 - \frac{2 \sum_i p_i y_i + \varepsilon}{\sum_i p_i + \sum_i y_i + \varepsilon} \quad (16)$$

为了保证数据稳定性,将参数  $\varepsilon$  的值设为  $10^{-6}$ . 另外在模型训练过程中容易出现过拟合,所以在此基础上又引入了  $L_2$  损失来缓解训练模型过程中的过拟合问题,有利于网络的收敛.  $L_2$  损失函数如式 (17) 所示:

$$L_2 = \frac{1}{N} \sum_i^N (y_i - p_i)^2 \quad (17)$$

综上,本文的混合损失函数如式 (18) 所示:

$$L = L_{Dice} + \alpha L_{BCE} + \beta L_2 \quad (18)$$

本文使用控制超参变量的方法来试验最佳参数,首先将  $\beta$  设置为 0,以试验最佳的超参  $\alpha$ ,如图 6 所示,  $\alpha$  的波动范围在 0-1 之间,当  $\alpha$  为 0.5 时,本文模型在整体肿瘤 (whole tumor, WT)、核心肿瘤 (tumor core, TC) 与增强肿瘤 (enhancing tumor, ET) 区域上达到最高峰值点,说明此时模型性能达到最优.

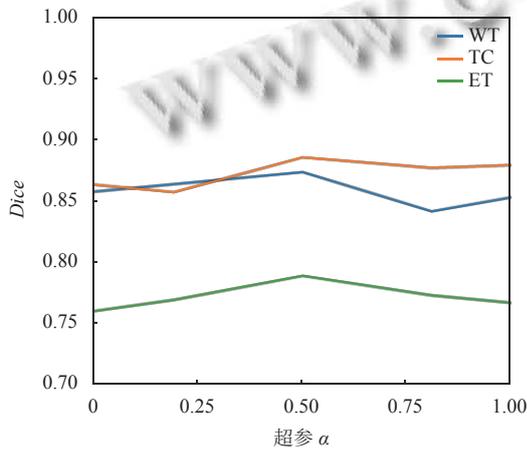


图 6 超参  $\alpha$  对模型性能的影响

在此基础上,继续试验最佳的超参  $\beta$ ,如图 7 所示,其中  $\beta$  的波动范围在 0-1 之间. 最终确定参数  $\alpha$  为 0.5,  $\beta$  为 0.05 时本文模型在 WT、TC 和 ET 区域上的 Dice 系数达到 0.887、0.892 与 0.815,模型预测性能达到了最好.

将最佳超参  $\alpha=0.5$  与  $\beta=0.05$  带入混合损失函数中,当模型训练迭代次数为 100 轮时,模型的训练损失值和验证损失值达到最佳,如图 8 所示.

### 2.6 消融实验

为了验证本文设计和添加模块的有效性,在 UNet 的架构上添加不同的模块,使用相同的损失函数和参数实验. 结果如表 1 所示.

表 1 中 DPLC 代表密集卷积并行大核注意模块、MDFM 代表多模态深度融合模块、SA 代表 Shuffle Attention 注意力模块. 在将 DLCCB 模块添加到 UNet

后,在模型的整体肿瘤、核心肿瘤与增强肿瘤区域中分别增加了 1.7%、3.6% 和 1.4%. 随着 MDFM 模块的加入,模型在核心肿瘤区和增强区的表现也得到了进一步地提高. 最后,增加了改进的 SA 注意力. 该方法在 3 种不同类型的肿瘤中均取得了良好的分割效果,证明了该方法的有效性.

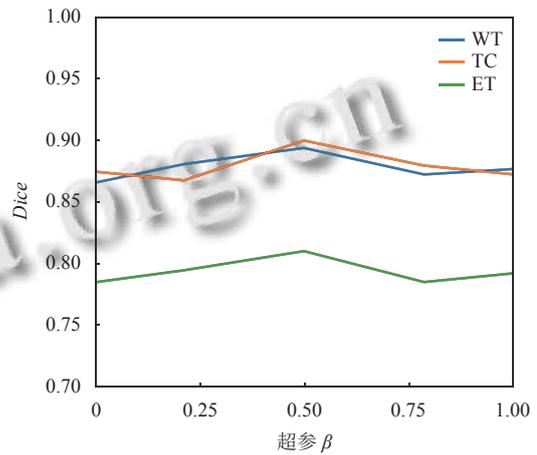


图 7 超参  $\beta$  对模型性能的影响

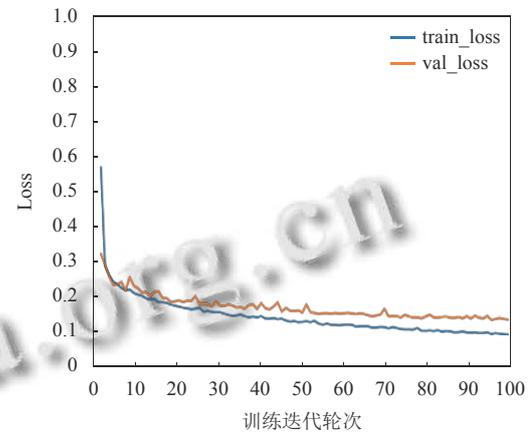


图 8 模型训练 Loss 变化

表 1 模块消融对比实验

网络模型	Dice精度		
	WT	TC	ET
UNet	0.834	0.823	0.769
UNet+DPLC	0.851	0.859	0.783
UNet+DPLC+MDFM	0.879	0.873	0.791
UNet+DPLC+MDFM+SA (本文)	<b>0.887</b>	<b>0.892</b>	<b>0.815</b>

为了测试在解码器分支添加不同注意力的有效性,本文将 SA 注意力机制分别与 SE 通道注意力<sup>[18]</sup>、CBAM 通道空间注意力<sup>[19]</sup>、ECA 高效通道注意力<sup>[20]</sup>进行对比,如表 2 所示,在解码器分支使用改进的 SA 注意力机制相比使用 SE 注意力,在模型的整体肿瘤、

核心肿瘤与增强肿瘤区域中分别增加了 4.2%、3.5% 和 4.4%，加强了模型捕获重要特征信息的能力。

表 2 添加不同注意力模块的实验

网络模型	Dice精度		
	WT	TC	ET
(UNet+DPLC+MDFM)+SE <sup>[18]</sup>	0.845	0.857	0.771
(UNet+DPLC+MDFM)+ECA <sup>[20]</sup>	0.858	0.862	0.797
(UNet+DPLC+MDFM)+CBAM <sup>[19]</sup>	0.867	0.882	0.785
(UNet+DPLC+MDFM)+SA (本文)	<b>0.887</b>	<b>0.892</b>	<b>0.815</b>

## 2.7 实验结果

为了进一步验证本文方法对多模态脑肿瘤 MR 图像分割的有效性, BraTS2019 部分数据集作为测试集来与其他先进方法进行比较, 结果如表 3 所示。

相比与 UNet、UNet++ 与 DenseUNet 的经典 2D 网络, 本文模型在整体、核心与增强肿瘤区域都有较大的性能提升。而与先进的 TransBTS、ADHDC-Net 与 SDS-MSA-Net 等 3D 模型对比时, 3D 模型因其空间连续在整体肿瘤分割中占有优势, 但本文 2D 模型在核心与增强肿瘤区具有更好的分割性能, 且相较于 3D 模型, 降低了对高显存硬件资源的依赖。

最后本文对 BraTS2019 测试数据集的样例进行

分割对比, 模型的输入图像与分割结果对比如图 9 所示, 每一行代表一个病例患者, 第 1-4 列依次为 T1、T1CE、Flair、T2 的不同脑肿瘤模态, 第 5 列为高级医师的黄金分割标准。从图 9 可以看出, UNet 对第 1 个病例的水肿区域存在过分割现象; 当前先进的 TransBTS 与 SDS-MSA-Net 的 3D 网络因其空间连续, 对绿色的整体肿瘤区域分割效果良好, 但其在第 2 与第 3 病例上将部分水肿区域误分割为红色的核心肿瘤区域, 本文提出的分割模型在 TC 和 ET 肿瘤区域上, 比先进 3D 网络的 SDS-MSA-Net 分别提升了 3.5%、1.9%, 相比之下本文方法对肿瘤核心和增强区域有更好的分割效果。

表 3 在 BraTS2019 测试集上与先进算法的比较

维度	网络模型	Dice精度		
		WT	TC	ET
2D	UNet	0.834	0.823	0.769
2D	UNet++	0.848	0.860	0.784
2D	DenseUNet	0.847	0.862	0.781
3D	TransBTS <sup>[10]</sup>	0.900	0.819	0.789
3D	Liu 等人 <sup>[13]</sup>	0.893	0.823	0.771
3D	ADHDC-Net <sup>[11]</sup>	0.899	0.838	0.779
3D	SDS-MSA-Net <sup>[12]</sup>	<b>0.902</b>	0.857	0.796
2D	CFDA-Net (本文)	0.887	<b>0.892</b>	<b>0.815</b>

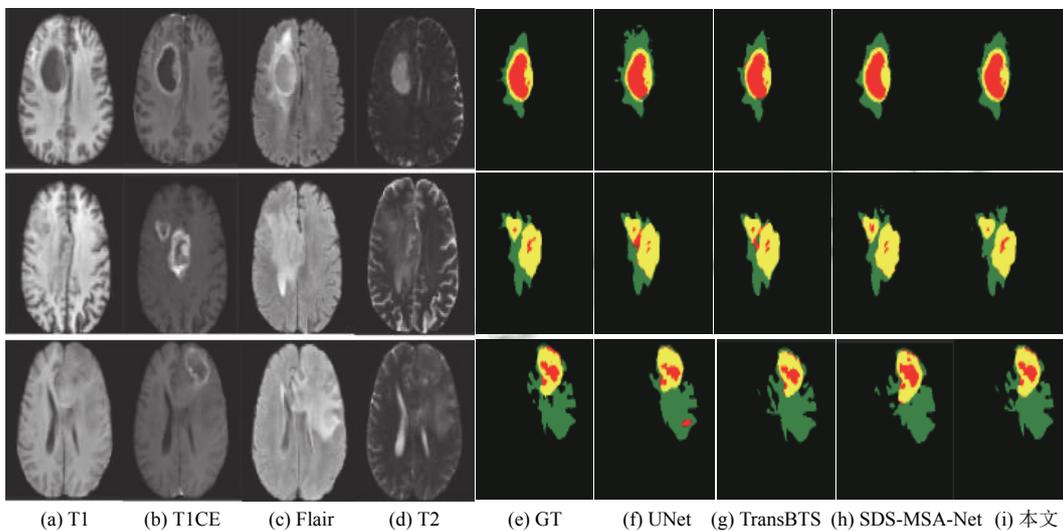


图 9 分割结果直观对比

## 2.8 网络模型复杂度对比分析

为了更好地检测模型的性能, 在相同输入大小的条件下, 对网络模型的参数量 (Params) 和计算量 (FLOPs) 进行对比, 如表 4 所示。本文模型的参数量相比基础网络 UNet 减少了约 55.17%, 计算量降低了约 19.85%。且与先进的 3D 模型 SDS-MSA-Net 相比, 本文的参数量减少了约 45.42%, 计算量降低了约 69.57%。综上所述,

本文模型在具有较低参数量和计算量的条件下, 对脑肿瘤的核心和增强区域的分割性能达到了最优状态。

## 3 结束语

为了解决脑肿瘤多模态信息融合不充分以及肿瘤区域细节信息丢失的问题, 本文在编码器-解码器结构的基础上进行改进, 将本文设计的多模态深度融合

MDFM 模块、密集并行大核注意卷积 DPLC 模块与改进的 Shuffle Attention 引入到架构中, 提出了一种跨模态融合的双注意力脑肿瘤分割方法。其中 MDFM 模块可以增强不同模态对间的互补信息, 使肿瘤区域特征融合更充分; 编码器分支使用 DPLC 模块有效提取了全局和局部信息; 解码器分支使用改进的 SA 模块, 通道分组后一分为二地捕获重要的空间和通道特征; 设计了新混合函数, 缓解了脑肿瘤数据类别不平衡问题与网络训练时存在的过拟合问题。实验结果证明, 本文模型在核心肿瘤和增强肿瘤区域优于其他先进模型; 在整体肿瘤区域, 虽然提高了检测效率, 但切片间的连续信息会有丢失。因此在后续工作中, 将使用 2.5D 网络或者低参数的 3D 网络来进一步加强对整体肿瘤区域的分割, 提升对脑肿瘤的分割精度。

表4 不同网络模型的复杂度对比

网络模型	Params (M)	FLOPs (G)
UNet	69.71	28.46
UNet++	34.96	23.53
TransBTS <sup>[10]</sup>	32.99	333.00
Liu等人 <sup>[13]</sup>	66.97	117.65
SDS-MSA-Net <sup>[12]</sup>	57.26	74.95
CFDA-Net (本文)	<b>31.25</b>	<b>22.81</b>

## 参考文献

- Ostrom QT, Cioffi G, Waite K, *et al.* CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2014–2018. *Neuro-Oncology*, 2021, 23(S3): iii1–iii105.
- Zhou TX. Modality-level cross-connection and attentional feature fusion based deep neural network for multi-modal brain tumor segmentation. *Biomedical Signal Processing and Control*, 2023, 81: 104524. [doi: 10.1016/j.bspc.2022.104524]
- Wang JJ, Yu ZS, Luan ZY, *et al.* RDAU-net: Based on a residual convolutional neural network with DFP and CBAM for brain tumor segmentation. *Frontiers in Oncology*, 2022, 12: 805263. [doi: 10.3389/fonc.2022.805263]
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention*. Munich: Springer, 2015. 234–241.
- Zhou ZW, Rahman Siddiquee MM, Tajbakhsh N, *et al.* UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 2020, 39(6): 1856–1867. [doi: 10.1109/TMI.2019.2959609]
- 侯奕辰, 彭辉, 谢俊章, 等. 改进 UNet++ 在脑肿瘤图像分割的研究. *计算机工程与设计*, 2022, 43(6): 1725–1731.
- Huang G, Liu Z, van der Maaten L, *et al.* Densely connected convolutional networks. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*.

- Honolulu: IEEE, 2017. 2261–2269.
- Kaku A, Hegde CV, Huang J, *et al.* DARTS: DenseUNet-based automatic rapid tool for brain segmentation. *arXiv:1911.05567*, 2019.
- Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. *Proceedings of the 4th International Conference on 3D Vision (3DV)*. Stanford: IEEE, 2016. 565–571.
- Wang WX, Chen C, Ding M, *et al.* TransBTS: Multimodal brain tumor segmentation using Transformer. *Proceedings of the 24th International Conference on Medical Image Computing and Computer Assisted Intervention*. Strasbourg: Springer, 2021. 109–119.
- Liu HX, Huo GQ, Li Q, *et al.* Multiscale lightweight 3D segmentation algorithm with attention mechanism: Brain tumor image segmentation. *Expert Systems with Applications*, 2023, 214: 119166. [doi: 10.1016/j.eswa.2022.119166]
- Rehman A, Usman M, Shahid A, *et al.* Selective deeply supervised multi-scale attention network for brain tumor segmentation. *Sensors*, 2023, 23(4): 2346. [doi: 10.3390/s23042346]
- Liu Y, Mu FH, Shi Y, *et al.* Brain tumor segmentation in multimodal MRI via pixel-level and feature-level image fusion. *Frontiers in Neuroscience*, 2022, 16: 1000587. [doi: 10.3389/fnins.2022.1000587]
- Zhang QL, Yang YB. SA-NET: Shuffle attention for deep convolutional neural networks. *Proceedings of the 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Toronto: IEEE, 2021. 2235–2239.
- Guo MH, Lu CZ, Liu ZN, *et al.* Visual attention network. *Computational Visual Media*, 2023, 9(4): 733–752. [doi: 10.1007/s41095-023-0364-2]
- Zhou Y, Liang XF, Gu Y, *et al.* Multi-classifier interactive learning for ambiguous speech emotion recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2022, 30: 695–705. [doi: 10.1109/TASLP.2022.3145287]
- Wu XD, Huang FH, Huang H. Fast stochastic recursive momentum methods for imbalanced data mining. *Proceedings of the 2022 IEEE International Conference on Data Mining (ICDM)*. Orlando: IEEE, 2022. 578–587.
- Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 7132–7141.
- Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 3–19.
- Wang QL, Wu BG, Zhu PF, *et al.* ECA-net: Efficient channel attention for deep convolutional neural network. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2020. 11531–11539.

(校对责编: 牛欣悦)