

文章编号:1672-3961(2024)02-0080-10 DOI:10.6040/j.issn.1672-3961.0.2023.021

基于混合偏移轴向自注意力机制的脑胶质瘤分割算法

高泽文^{1,2}, 王建³, 魏本征^{1,2*}

(1. 山东中医药大学医学人工智能研究中心, 山东 青岛 266112; 2. 山东中医药大学青岛中医药科学院, 山东 青岛 266112; 3. 山东交通学院理学院, 山东 济南 250357)

摘要:为提高脑胶质瘤核磁共振成像(magnetic resonance imaging, MRI)图像分割精度及质量,设计一种混合偏移轴向自注意力机制的脑胶质瘤分割多层次轴向注意力网络(multi-level axial-attention net, MLA-Net)算法。MLA-Net算法框架中设计的混合偏移轴向自注意力机制和混合损失函数,分别用于提取更精确的全局相对位置关系、提升网络对细节结构特征的敏感程度和实现精确地分割胶质瘤模糊边界。试验结果表明,在BraTS 2018和2019的混合数据上,MLA-Net算法的dice系数可达到0.843 3, Hausdorff距离为2.587。MLA-Net算法的MRI图像脑胶质瘤分割性能优良,可以融合全局相对位置特征和局部细节特征,更好地分割出脑胶质瘤感兴趣区域。

关键词:脑胶质瘤; 图像分割; MRI; 深度学习; 轴向自注意力机制

中图分类号:TP18 **文献标志码:**A

引用格式:高泽文,王建,魏本征. 基于混合偏移轴向自注意力机制的脑胶质瘤分割算法[J]. 山东大学学报(工学版), 2024, 54(2): 80-89.

GAO Zewen, WANG Jian, WEI Benzhen. Glioma segmentation algorithm based on hybrid offset axial self-attention mechanism[J]. Journal of Shandong University (Engineering Science), 2024, 54(2): 80-89.

Glioma segmentation algorithm based on hybrid offset axial self-attention mechanism

GAO Zewen^{1,2}, WANG Jian³, WEI Benzhen^{1,2*}

(1. Medical Artificial Intelligence Research Center, Shandong University of Traditional Chinese Medicine, Jinan 266112, Shandong, China; 2. Qingdao Academy of Chinese Medicine, Shandong University of Traditional Chinese Medicine, Qingdao 266112, Shandong, China; 3. School of Science, Shandong Jiaotong University, Jinan 250357, Shandong, China)

Abstract: To improve the accuracy and quality of glioma MRI image segmentation, this paper designed a hybrid offset axial self-attention mechanism for the glioma segmentation algorithm MLA-Net (Multi-level axial-attention net). The offset axial self-attention mechanism and hybrid loss function designed in this algorithmic framework could be used to extract more accurate global relative position relationships, as well as to enhance the sensitivity of the net to detailed structural features and to achieve the role of accurately segmenting the fuzzy boundaries of gliomas, respectively. The experimental results showed that the dice coefficient of MLA-Net could reach 0.843 3 and the Hausdorff distance was 2.587 on the mixed data of BraTS 2018 and 2019. The MRI image glioma segmentation performance of MLA-Net was excellent, and could fuse the global relative position features and local detail features to better segment the lesion region.

Keywords: brain glioma; image segmentation; MRI; deep learning; axial self-attention mechanism

收稿日期:2023-01-31

基金项目:国家自然科学基金资助项目(61872225);山东省自然科学基金资助项目(ZR2019ZD04, ZR2020KF013, ZR2020ZD44, ZR2020QF043);山东省高校青创引才育才计划项目(2019-173)

第一作者简介:高泽文(1997—),男,山东泰安人,硕士研究生,主要研究方向为医学人工智能。E-mail:yzew@foxmail.com

* 通信作者简介:魏本征(1976—),男,山东莒南人,教授,博士生导师,博士,主要研究方向为医学人工智能、计算医学、机器学习及医学信息工程。E-mail:wzb99@sina.com

0 引言

脑胶质瘤是一种严重威胁人类健康的中枢神经系统疾病^[1],由大脑和脊髓胶质细胞癌变导致,其发病率约占所有脑肿瘤的70%~80%^[2-3]。根据肿瘤侵入程度和患者预后情况,脑胶质瘤分为高级别胶质瘤和低级别胶质瘤。高级别胶质瘤具有较强的浸润性和侵袭性^[4],其患者死亡率较高,存活期较短。相对而言,低级别胶质瘤生长较为缓慢,患者有较长的存活期。然而,低级别胶质瘤如果没有及时治疗,往往会演变为高级别胶质瘤。因此,脑胶质瘤的精准影像学诊断具有重要临床价值。

核磁共振成像(magnetic resonance imaging, MRI)^[5]作为一种非侵入式的成像技术,具有软组织对比度强、分辨率高等特点^[6-7],是目前检查脑胶质瘤的主流方式。临床医生通常会结合核磁共振成像对肿瘤区域进行手动分割。然而,脑组织种类复杂,结构之间边界模糊、相似性高,脑胶质瘤的空间位置、尺寸和形状多变,导致手动分割感兴趣区域(region of interest, ROI)工作耗时费力且分割结果一致性较差。因此,设计一种精确度较高的分割算法辅助脑胶质瘤分割,对临床的早期诊断和干预治疗具有重要作用。

近年来,脑胶质瘤辅助诊断算法的研究取得了一些重要进展。早期的分割算法主要利用阈值、像素相似度等先验知识对胶质瘤区域进行分割^[8-10],但由于胶质瘤患者脑部各组织灰度分布杂乱并且混叠,很难选择合适的先验知识对病灶区域进行划分,因而此类方法往往分割性能不佳。卷积神经网络(convolutional neural network, CNN)^[11-13]因其出色的局部采样和特征提取能力,广泛应用于脑胶质瘤分割任务中。文献[14]提出一种双通道CNN结构用于脑胶质瘤分割,采用级联结构对输入的图像块进行特征提取,分别关注胶质瘤的细节特征和背景区域;文献[15]将注意力门机制添加至跳连接操作中以扩展基础U-Net,通过自动学习参数提高脑组织中重要目标的权重;文献[16]结合密集连接块和U-Net中跳连接层的优点,更有效地聚合胶质瘤中低层次的视觉特征和高层次的语义特征;文献[17]通过在解码器中添加变分自编码器分支修改U-net,提高了泛化性能,在Brats 2018(brain tumor segmentation 2018)挑战赛中取得第1名。

虽然上述方法在胶质瘤分割任务中取得了较好的效果,但卷积对图像信息的提取是区域性的,

每个卷积核只能观察到特定区域,这使得CNN网络学习全局相关性的能力较弱。如果增大卷积特征提取的视野,则需要加深网络以提取高维的特征,但这又会更加难以区分模糊的组织边界。如何构造更大的感受野进行上下文建模,既能提取全局特征又尽量不损失细节信息,一直是医学图像分割领域中研究的重要问题。

自注意力机制是一种用于提取数据内部全局相关性的特征提取方法^[18-20]。与CNN相比,自注意力机制可以在任意长度的序列中对特征之间的相关性建模,许多研究将其与CNN结合以弥补CNN全局特征提取能力不足的缺陷^[21]。文献[22]提出一种基于Unet的自注意分割网络TransUnet,将自注意力机制添加至编码器末端,以较小的计算量提取病灶区的全局特征,但这种方法只将自注意力应用于深度网络后期的低分辨率特征图中,丢失了细节信息;文献[23]提出了TransFuse,以并行的方式将自注意力机制与CNN分层融合弥补CNN全局信息提取的不足,但将单层Transformer提取到的不完整位置信息与CNN分层结合,可能会降低特征图的准确性;文献[24]提出一种基于编码器解码器结构TransBTS,将自注意力机制应用于3D脑胶质瘤分割中,在挖掘和建模局部全局信息的同时,将CNN和自注意力机制结合起来提高性能,但这种方法只在多次下采样后使用自注意机制,提取全局信息能力不足;文献[25]提出一种基于轴向自注意力机制的模型MedT,在提取全局特征的同时降低了模型的复杂度,较好地学习了病灶区域与正常组织之间的全局相关性,然而该方法将图像分割成块提取特征,会导致丢失图像块之间的联系与边缘的细节特征。

为提取脑胶质瘤的细节特征的同时,更精准地提取全局特征,本研究提出一种用于脑胶质瘤分割的多尺度融合注意网络MLA-Net(multi-level axial-attention network)。网络基于MRI图像自动分割肿瘤区域,采用全局特征融合模块提取多尺度的全局信息,该模块使用多层级的轴向自注意模块不断聚合临近的特征,并在不同尺度学习它们的相对位置关系。为了适应脑胶质瘤多变的形态特征,更精确地指导特征重构,本研究设计的混合偏移轴向自注意力模块,可根据相对位置关系的重要程度调整可学习的相对位置编码,更准确地学习特征之间的全局相关性。另外,MLA-Net并行地使用基于CNN的局部特征分支学习细节特征,并将其与全局特征融合,减少下采样造成的特征丢失。最后,为解决

脑胶质瘤图像边界模糊和数据样本不均衡的问题,还设计了一个混合损失函数。试验证明 MLA-Net 在脑胶质瘤 MRI 图像分割任务中,可获得优良的分割性能,算法具有一定的优越性。

1 方法

本研究设计的 MLA-Net 整体网络架构如图 1 所示。由图 1 可知,MLA-Net 有 2 个分支:基于自注意力机制的全局特征融合分支和基于 CNN 的局部特征提取分支。全局特征融合分支由 2 条支路构

成,分别为:(1)使用混合偏移轴向自注意机制的特征提取支路;(2)使用轴向自注意机制的特征融合支路。全局分支在输入特征图后,经过不同尺度的混合偏移轴向注意层提取相对位置关系,并行通过八层轴向注意层聚合全局的特征,最后将多尺度特征图融合得到丰富的全局特征信息。局部分支利用由多个 RSU 模块^[26]构成的 U 型结构提取细节特征,最后将全局特征和细节特征融合,经过混和损失函数训练和输出预测图,RSU 结构如图 1 中 RSU 模块所示,其中, n 为编码器层数, m 为 RSU 内部层通道数。

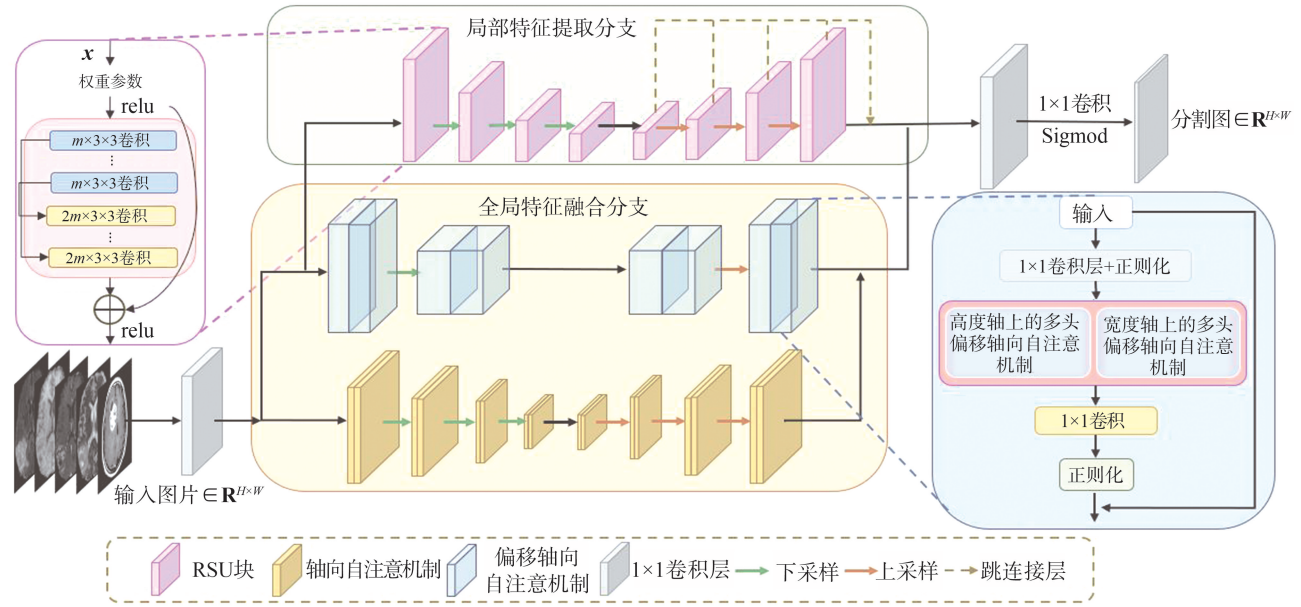


图 1 MLA-Net 整体架构
Fig.1 Overall architecture of MLA-Net

1.1 全局特征融合分支

1.1.1 轴向注意力机制

轴向注意力机制在文献[27]中提出,它应用了局部约束^[28]的思想,显著降低了视觉自注意力机制的计算成本。轴向注意力机制首先将图像宽度轴上的轴向注意层定义为简单的一维位置敏感的自我注意,并对高度轴使用相同的定义,这种计算方式保留了自注意力机制对全局的视野,同时提高了注意力机制的运算效率,将模型复杂度从原始自注意力机制的 $O(h^2w^2)$ 优化至 $O(hw)$ 。具体来说,轴向注意力机制的输入通常为一个高为 h , 宽为 w , 通道数为 C 的输入特征图 $x_0 \in \mathbf{R}^{h \times w \times C_{in}}$, 经过计算可得到经过自注意层后的输出 $y_0 \in \mathbf{R}^{h \times w \times C_{out}}$, 沿宽度轴的轴向注意层的定义为:

$$y_0 = \sum_{w=1}^w \text{softmax}(q_{iw}^T k_{iw} + q_{iw}^T r_{iw}^q + k_{iw}^T r_{iw}^k)(v_{iw} + r_{iw}^v),$$

式中: q_{iw}^T 为查询向量, $q_{iw}^T = W_Q x_{iw}$; k_{iw}^T 为键值向量, $k_{iw}^T = W_K x_{iw}$; v_{iw} 为特征值向量, 都是基于输入特征 x

在宽度轴上的线性投影, $v_{iw} = W_V x_{iw}$; W_Q, W_K, W_V 皆为可学习矩阵; $r_{iw}^q, r_{iw}^k, r_{iw}^v$ 表示基于 q_{iw}, k_{iw}, v_{iw} 生成的权重矩阵, 这个矩阵额外考虑了 q_{iw}, k_{iw}, v_{iw} 的相对位置关系, 并将其融入进位置编码中。这些位置编码是可学习的, 并已证明具有编码图像空间结构的能力^[29]。类似的公式也用于沿高度轴应用的轴向注意力, 它们共同组成了一个高计算效率的自注意模型。

1.1.2 混合偏移轴向自注意力机制

在医学领域的脑胶质瘤分割任务中, 不同患者之间的病灶形态差异很大, 很难学习固定的相对位置关系, 利用位置信息编码进行交互时也会具有不确定性, 在学习到的相对位置编码不够准确的情况下, 利用位置信息指导特征重构反而降低了分割的准确率。为了更好地学习特征之间的相对位置关系, 指导特征重构, 本研究提出了一种改进的轴向注意模块, 可以根据相对位置关系的重要程度调节

可学习偏移参数,具体结构如图 2 所示。

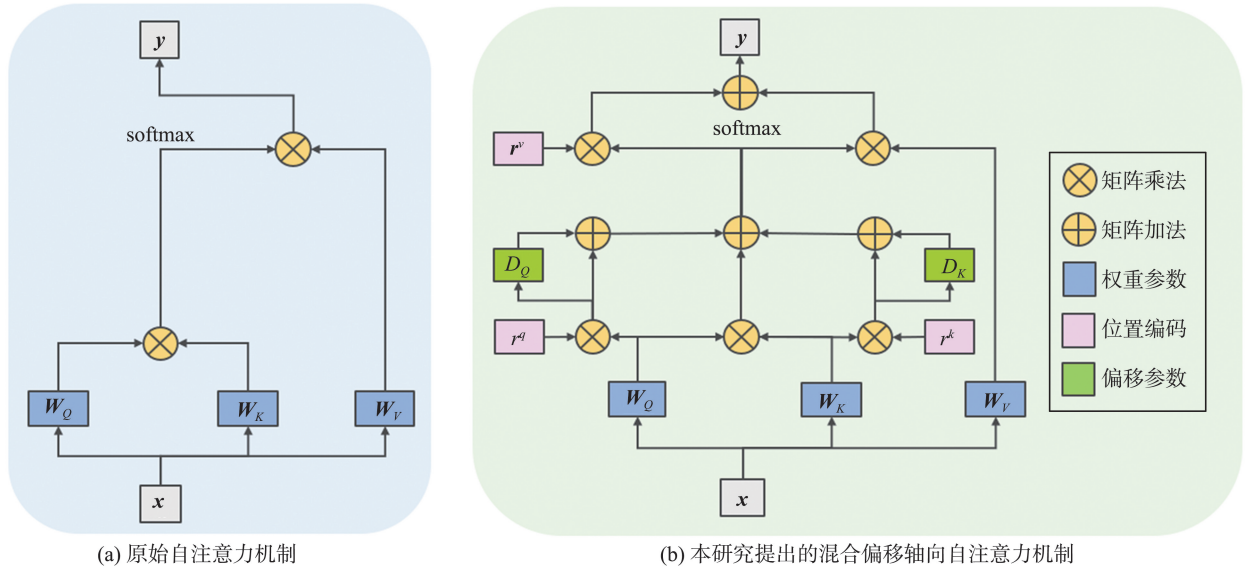


图 2 自注意力机制对比示意图

Fig.2 Self-attention mechanism comparison diagram

根据所提出的修改,应用在宽度轴上的自注意力机制可以正式写成:

$$y_0 = \sum_{w=1}^W \text{softmax}(q_{iw}^T k_{iw} + (1+D_Q)q_{iw}^T r_{iw}^q + (1+D_K)k_{iw}^T r_{iw}^k)(v_{iw} + r_{iw}^v),$$

式中, D_Q, D_K 为可学习的偏移参数, $D_Q, D_K \in \mathbf{R}$, 它们分别对 $q_{iw}^T r_{iw}^q, k_{iw}^T r_{iw}^k$ 实现偏移机制,根据位置重要程度使学习到的位置编码权重 r_{iw}^q, r_{iw}^k 发生偏移。通常情况下,如果一个特征点的位置非常重要,那么

其相对位置编码会在特征重构时赋予更高的权重,而如果某个相对位置关系出现次数较少且学习困难,其相对位置关系就会赋予低权重。

1.1.3 多级特征融合模块

在数据尺度较大的图像分割任务中,轴向注意力机制相对于原始自注意机制具有更高的计算效率和运算速度,它在理论上保留了全局视野,但是浅层轴向自注意机制无法提取在斜角方向上的特征,这会导致部分位置相关性的缺失,如图 3 所示。

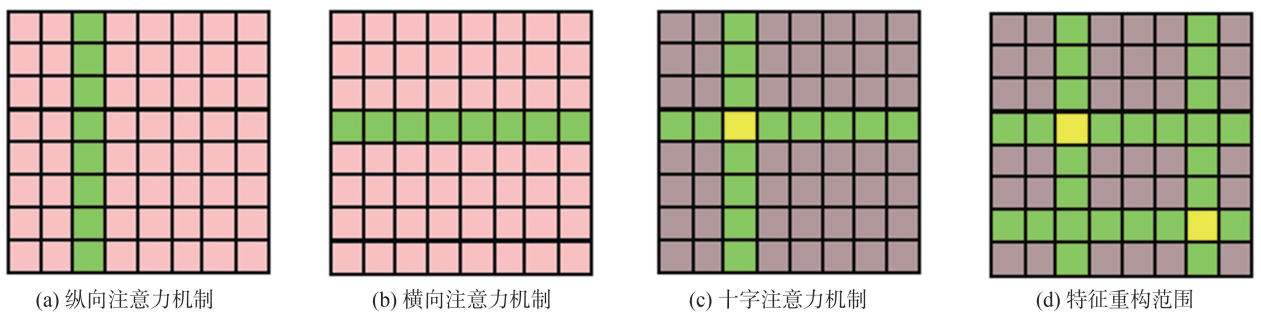


图 3 轴向注意力机制视野范围

Fig.3 Axial attention mechanism field of view

针对脑胶质瘤分割任务,斜角特征和临近的特征应该赋予同样的权重。为了提高对病灶区域整体的理解,本研究设计了基于轴向自注意力机制的多级特征融合模块。该模块分为 2 条支路,特征提取支路使用轴向自注意模块,在不同尺度下进行多次注意力计算,以提取完整的全局特征,并经过 3 次特征重构使相邻的特征汇聚起来;特征修正支路使用混合偏移轴向自注意模块在原始尺度和单次下采样后提取更精确的位置信息,并经过混合偏移模

块修正位置信息。最后,将 2 个支路的输出特征融合,传递给 1×1 的卷积层。此模块利用低计算量的轴向注意堆叠和聚合特征,利用偏移轴向注意在全局提取准确的相对位置关系,然后在多尺度下进行特征重构,在减少轴向自注意机制斜角特征丢失问题的同时,提高了性能。

1.2 局部特征提取分支

在脑肿瘤图像分割任务中,若仅使用注意力机制提取特征,可能会造成部分细节信息遗漏,最终

导致算法性能欠佳。为了更好地提取局部细节特征,网络使用 RSU(ReSidual U)特征提取模块提取多尺度的特征。RSU 主要由输入卷积层、U 型编解码器结构和残差连接组成。输入卷积层首先提取局部特征信息,将输入特征转换为中间特征映射,U 型结构用来学习提取和编码多尺度上下文信息,其中网络的深度越深,U 型结构的层数越少;网络深度

越浅,池化操作越多,接受域范围越大,局部和全局特征越丰富。此模块从多尺度获取了全方位的特征,减少了下采样造成的细微细节损失。最后通过残差连接将多尺度特征和初始局部特征融合。

1.3 混合损失函数

胶质瘤区域边界模糊且形状不规则,绝大多数图像中病灶区域较小,样本不平衡,如图 4 所示。

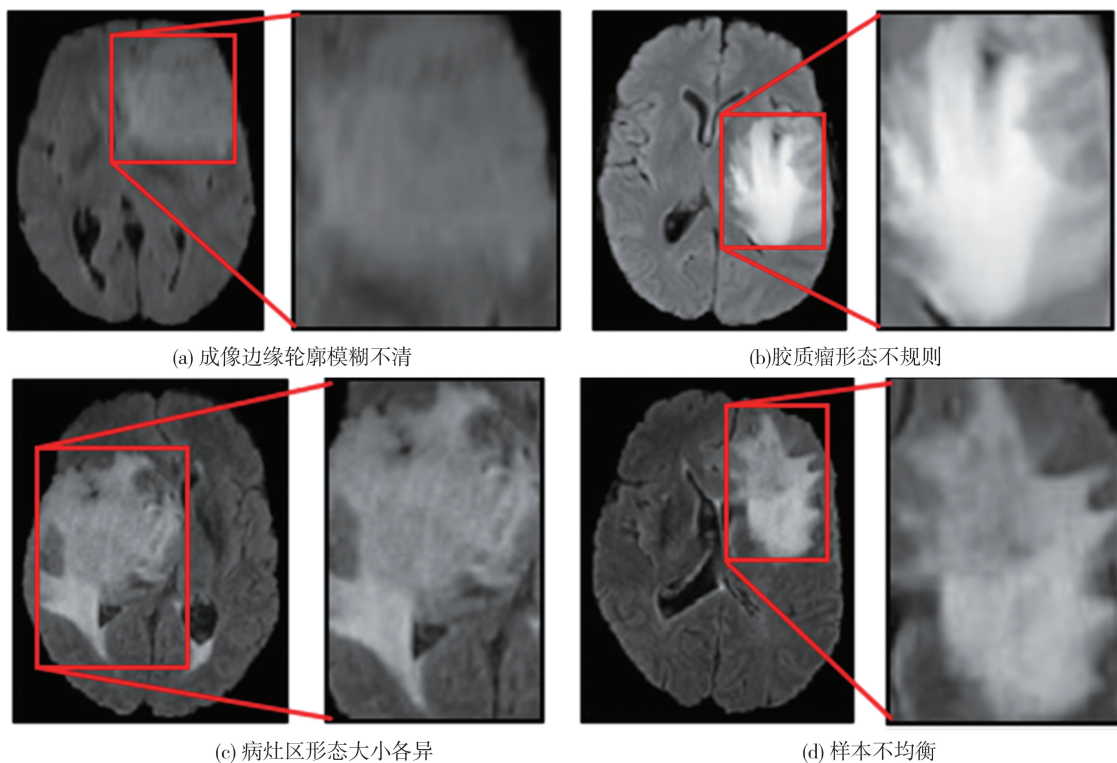


图4 脑胶质瘤 MRI 成像示意图
Fig.4 MRI imaging diagram of glioma

本研究使用交叉熵损失函数作为基准,引入 dice 和 iou Loss(intersection over union loss)解决样本不平衡的问题^[30-31]。引入 ssim(structural similarity)函数解决胶质瘤边界不清晰的问题^[32]。在网络训练阶段,将损失函数 l 定义为混合损失,即:

$$l = \alpha l_{\text{bce}} + \beta l_{\text{ssim}} + \lambda l_{\text{dice}} + \theta l_{\text{iou}}, \quad (1)$$

式中, l_{bce} 、 l_{ssim} 、 l_{dice} 、 l_{iou} 分别表示 bce loss、ssim Loss、dice Loss 和 iou Loss 的损失函数, α 、 β 、 λ 、 θ 为权重。其中 bce Loss^[33] 是二值分类和分割中应用最广泛的损失函数,定义为:

$$l_{\text{bce}} = - \sum_{(x,y)} [G(x,y) \log(S(x,y)) + (1-G(x,y)) \log(1-S(x,y))], \quad (2)$$

式中: $G(x,y)$ 为像素 (x,y) 的真实值标签, $G(x,y) \in \{0,1\}$; $S(x,y)$ 为显著目标的预测概率。

iou 函数最初用作度量 2 个集合的相似度,后来作为一个标准的评价指标用于目标检测和分割。最近,它用作训练损失,定义为:

$$l_{\text{iou}} = 1 - \frac{\sum_{x=1}^H \sum_{y=1}^W S(x,y)G(x,y)}{\sum_{x=1}^H \sum_{y=1}^W [S(x,y) + G(x,y) - S(x,y)G(x,y)]}. \quad (3)$$

dice 函数的设计思想是加大前景区域的权重以避免学习过程陷入损失函数的局部最小值,定义为:

$$l_{\text{dice}} = 1 - \frac{\sum_{x=1}^H \sum_{y=1}^W S(x,y)G(x,y)}{\sum_{x=1}^H \sum_{y=1}^W [S(x,y) + G(x,y)]}. \quad (4)$$

MRI 成像的肿瘤区边界较为模糊,为了得到肿瘤区清晰的边界、提升区域分割的准确性,使用 ssim 函数捕捉胶质瘤图像中的结构信息。

令 x 和 y 分别是由预测概率图 S 和二值标签 G 裁剪的 2 个对应 patch 的像素值,定义 x 和 y 的 ssim 函数值为:

$$l_{\text{ssim}} = 1 - \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\mu_x^2 + \mu_y^2 + C_2)}, \quad (5)$$

式中, μ_x 、 μ_y 和 σ_x 、 σ_y 分别为 x 、 y 的均值和标准差, σ_{xy} 为它们的协方差, 设置 C_1 、 C_2 为极小值以避免分母为零。

2 试验

2.1 试验环境与数据集

MLA-Net 算法采用 Pytorch1.7.1 深度学习框架编写。算法在 Intel 8268 CPU、NVIDIA Tesla V100 32 GB GPU 上进行训练和测试。算法初始学习率设置为 0.001, 使用 Adam 作为优化器, 动量参数为 0.9。Batchsize 设置为 64。网络训练约在 200 轮后达到收敛, 每轮训练时间约 500 s。采用了以下数据增强技术: (1) 将 3D 数据从 240 像素×240 像素×155 像素×4 像素随机裁剪为 2D 数据, 大小为 128 像素×128 像素×4 像素; (2) 随机镜像以概率 0.5 翻转冠状面。

本研究使用的数据集来源于脑肿瘤分割 (BraTS) 2017、2018 和 2019 挑战赛^[34-35], 其中 2017 年数据集包括高级别胶质瘤病例 210 例和低级别胶质瘤病例 75 例, 2018 年数据集包括胶质瘤病例 285 例, 2019 年数据集在 2018 年的基础上增加了 50 例。每个样本由 4 种方式的 MRI 扫描组成, 即 T1 加权、T1 序列、T2 加权和液体衰减反转序列, 每个模态都有一个 240 像素×240 像素×155 像素的 3D 序列以及专家手工标注的标签。BraTS 挑战赛数据集包括训练集、验证集和测试集, 其中测试集和验证集只在比赛中使用, 不对外公开。本研究将 2017 年数据集随机分为 245 例的训练集和 40 例数据的测试集, 将 2018 年数据集中训练集的 70% 用于训练, 剩余 30% 用于验证, 2019 年新增数据的训练集用于测试。

2.2 试验设计与评价标准

为了验证 MLA-Net 各组件的有效性, 本研究分别针对多级特征融合模块、局部特征提取分支、混合偏移轴向自注意力机制和混和损失函数设计了消融试验。在针对局部特征提取分支的试验设计中, 本研究分别测试了去除局部特征提取分支的 MLA-Net-n 和使用 Unet 替换细节分支后的 MLA-net。在针对多级特征融合模块的试验设计中, 本研究测试了去除多级特征融合模块的 AL-net, 并分别在 Unet 和 Unet++ 网络中添加多级特征融合模块对比前后性能。为了测试混合偏移轴向自注意力

机制的有效性, 本研究测试了将混合偏移轴向自注意力机制替换为原始的轴向自注意机制的网络。为了验证混和损失函数的有效性, 依次对比使用混合损失函数和使用添加 dice、bce、iou 和 ssim 后 MLA-Net 的效果。为了确定损失函数各部分权重参数的最优值, 本研究分别设计了针对损失函数各部分权重的对比试验。

为了对比 MLA-Net 在脑胶质瘤分割任务中与卷积网络和其他自注意力网络的效果, 本研究设计了对比试验。对于卷积基线, 试验比较了 U-Net、U-Net++、Attention Unet、DARTS、SegResNetVAE 和 U²net, 对于基于自注意机制的基线, 使用了医学图像分割网络 MedT 和基于 R-Transformer 的脑胶质瘤分割网络 ERTN。

对于分割结果, 本研究使用 dice、miou 和 Hausdorff (HD) 距离验证分割方法的精度, 使用 D 表示 dice 系数, I 表示 miou 系数, H 表示 Hausdorff 距离。其定义分别为:

$$D(A, B) = \frac{2S(A \cap B)}{S(A) + S(B)},$$

$$I(A, B) = \frac{S(A \cap B)}{S(A \cup B)},$$

$$H(A, B) = \max(\max_{a \in A}(\min_{b \in B} d(a, b)), \max_{b \in B}(\min_{a \in A} d(b, a))),$$

式中, $S(A)$ 、 $S(B)$ 分别表示自动分割结果和专家手工标注结果的面积, $S(A \cap B)$ 表示自动分割结果和专家手工标注结果重叠部分的面积, $S(A \cup B)$ 表示自动分割结果和专家手工标注结果并集的面积, $d(a, b)$ 表示 a 、 b 两个点的欧氏距离。

3 试验结果与分析

3.1 消融试验

MLA-Net 各组件消融试验如表 1 所示。在表 1 中, 由针对多级特征融合模块的试验可知, 在对 Unet、Unet++ 添加多级特征融合模块后^[36], dice 系数分别由 82.04、81.73 提高至 83.27、82.91, Hausdorff 距离分别由 2.617、2.642 降至 2.596、2.631。对 LA-net 添加混合偏移轴向自注意机制形成多级特征融合模块后, dice 系数由 82.97 提高至 84.33。由针对局部特征提取分支的试验可知, 在 MLA-Net 中删除细节分支后, dice 系数由 84.33 降至 82.51, 使用 Unet 替换细节分支后, dice 系数由 84.33 降至 83.27, 证明了局部特征提取分支的有效性。在针对混合偏移轴向自注意力机制的试验中, 使用原始轴向自注意机制替换混合偏移轴向自注意机制后, dice 系数由 84.33 降至 83.67, Hausdorff 距离由 2.587 增加至 2.610。

表1 算法各组件消融试验结果
Table 1 The results of the ablation test of each component of the algorithm

试验模块	方法	D	I	H
多级特征融合模块	Unet ^[37]	82.04	74.16	2.617
	MLA-unet	83.27	75.08	2.596
	Unet++ ^[36]	81.73	73.71	2.642
	MLA-unet++	82.91	74.63	2.631
	LA-net	82.97	74.72	2.628
局部特征提取分支	MLA-Net-n	82.51	74.48	2.660
	MLA-unet	83.27	75.08	2.596
混合偏移轴向自注意力机制	MLA-Net-origin	83.67	75.42	2.610
本研究方法	MLA-Net	84.33	76.01	2.587

MLA-Net 的 2 个分支都对网络分割结果做出贡献。对于多级特征融合模块,试验将其作为组件添加到 Unet 和 Unet++ 网络中,分割性能提升。Unet 和 Unet++ 网络都通过下采样扩大卷积视野,这导致了全局信息提取不足,试验将多级特征融合模块添加至网络中后,丰富了网络提取的全局特征,证明了全局特征的重要性。此外,试验将多级特征融合模块拆分并删除特征融合层分支得到 LA-net,分割效果变差,证明了轴向自注意机制需要多次特征重构以获取全局的视野,也证明了多级特征融合模块的有效性。在局部特征提取分支的试验中,结果表明,该分支对网络分割效果提高做出了贡献,且分割效果优于 Unet,证明了局部特征提取分支的有效性。在混合偏移轴向自注意力机制的

试验中,将混合偏移轴向自注意力机制替换为原始轴向自注意机制后,分割效果降低,证明原始轴向自注意机制不能很好地提取脑胶质瘤的复杂结构特征,本研究提出的混合偏移轴向自注意机制对其作出了改进,使其更适合分割物形态不固定的任务。

损失函数消融试验结果如表 2 所示。在组件有效性试验中依次添加 dice、bce、iou 和 ssim 函数并测试效果,网络性能依次提升,证明了 dice、bce、iou 和 ssim 函数在混和损失函数中的有效性。添加 bce loss 函数后 dice 系数提高了 0.86;添加 iou 函数后, dice 系数提高了 0.19;添加 ssim 函数后, dice 系数提高了 0.34, Hausdorff 距离降低了 0.21,证明了 ssim 函数可提高分割网络对于边界的敏感程度。

表2 损失函数消融试验结果
Table 2 Loss function ablation test results

试验类型	参数选择	D	I	H
组件有效性试验	dice loss ($\alpha=0, \beta=0, \lambda=1, \theta=0$)	82.84	74.61	2.640
	dice+bce loss ($\alpha=1, \beta=0, \lambda=1, \theta=0$)	83.70	75.51	2.628
	dice+bce+iou loss ($\alpha=1, \beta=0, \lambda=1, \theta=1$)	83.89	75.74	2.615
	dice+ssim+bce+iou loss ($\alpha=1, \beta=1, \lambda=1, \theta=1$)	84.23	75.93	2.594
权重参数试验	(1) Mix loss ($\alpha=1, \beta=1, \lambda=0.2, \theta=0.2$)	83.90	75.69	2.602
	(2) Mix loss ($\alpha=3.2, \beta=4, \lambda=1, \theta=1$)	83.97	75.77	2.598
	(3) Mix loss ($\alpha=1, \beta=2, \lambda=1, \theta=1$)	83.21	75.02	2.652
	(4) Mix loss ($\alpha=1, \beta=1.5, \lambda=1, \theta=1$)	83.67	75.43	2.634
	(5) Mix loss ($\alpha=1, \beta=0.9, \lambda=1, \theta=1$)	83.80	75.58	2.616
	(6) Mix loss ($\alpha=1, \beta=0.5, \lambda=1, \theta=1$)	83.67	75.45	2.637
	(最终参数) Mix loss ($\alpha=1, \beta=1.2, \lambda=1, \theta=1$)	84.33	76.01	2.587

混合损失函数使用 dice 和 iou Loss 解决胶质瘤图像前后景不均衡情况,混合损失函数使用 dice 和 iou Loss 解决胶质瘤图像前后景不均衡情况,但在极端情况下即标签和预测值极端小时,梯度会非常大,这导致训练不稳定,故在权重参数试验(1、2)中,测试能否在降低 dice 和 iou Loss 权重的同时,解

决前后景不均衡的问题。试验证明,降低 dice 和 iou Loss 权重分割性能下降。脑胶质瘤形态多变,边界模糊,在混合损失函数中添加 ssim 函数提高网络对边界信息的关注程度,试验(3~6)调整 ssim 函数权重,测试分割效果最优时 β 取值。试验证明, β 取值从两侧趋近于 1.2 时,网络性能提升,

在 $\beta=1.2$ 时网络有最优分割效果。

3.2 算法分割结果对比

外部对比试验结果如表 3、4 所示,在脑胶质瘤分割任务中,U²net 算法在卷积基线取得了最好性能,在 Brats2018&2019 中 dice 系数为 83.4,在 Brats2017 中 dice 系数为 82.20,该算法在多级优化的过程中关注到了不同尺度的特征。在注意力基线中,MedT 利用轴向注意网络实现多分支分割,提取了全局的位置信息,但其使用 Transformer 切分图片作为输入数据,带来切分后各图像块间的位置关系被忽略的问题。ERTN 结合 R-Transformer 与 U-Net 构建一个双分支网络,以此捕获复杂的语义特征和全局上下文信息,分别在两数据集中实现了 dice 系数为 83.27 和 83.10,在鲁棒性上具有优越性。本研究的架构结合了卷积网络和注意力网络的优势,融合了细节特征与全局特征,同时没有将图片分片输入,保证了特征的完整性,在两数据集中,与最优卷积基线网络相比,dice 系数分别提高了 1.20 和 1.01,与全注意网络相比分别提高了 0.49 和 0.11。

3.3 试验结果的可视化分析

为直观显示本研究 MLA-Net 的分割效果,对部分分割结果进行可视化,如图 5 所示。对于边缘不规则的病灶区,Unet、U²net、MedT 等方法的分割结果会分割出错误的尖锐边界,Att-Unet、Unet++等方法会将病灶区域的延伸误识别为病灶区,MLA-Net 分割边界更加平滑,对于尖锐区域分割效果较好。

对于离散、成像模糊的病灶区,大部分卷积和注意力方法将正常组织识别为肿瘤区。本研究的方法成功消除了被错误分割的边界部分,降低了误识别率,提高了分割精度。

表 3 Brats2018&2019 中算法分割结果对比
Table 3 Comparison of segmentation results of different algorithms in Brats2018&2019

类型	方法	<i>D</i>	<i>I</i>	<i>H</i>
卷积基线	AttentionUnet ^[15]	81.36	73.28	2.664
	Unet++ ^[36]	81.73	73.71	2.642
	Unet ^[37]	82.04	74.16	2.617
	DARTS ^[16]	81.55	73.23	2.678
	SegResNetVAE ^[17]	83.01	75.07	2.605
注意力基线	U ² net ^[26]	83.40	75.56	2.584
	MedT ^[25]	81.13	72.41	2.698
本研究方法	ERTN ^[38]	83.84	75.89	2.590
	MLA-unet	83.27	75.08	2.596
	MLA-Net	84.33	76.01	2.587

表 4 Brats2017 中算法分割结果对比
Table 4 Comparison of segmentation results of different algorithms in Brats2017

类型	方法	<i>D</i>	<i>I</i>	<i>H</i>
卷积基线	Unet ^[37]	81.18	72.43	2.698
	SegResNetVAE ^[17]	81.79	73.87	2.655
	U ² net ^[26]	82.20	74.20	2.646
注意力基线	MedT ^[25]	79.65	70.13	2.728
	ERTN ^[36]	83.10	75.54	2.616
本研究方法	MLA-Net	83.21	75.41	2.601

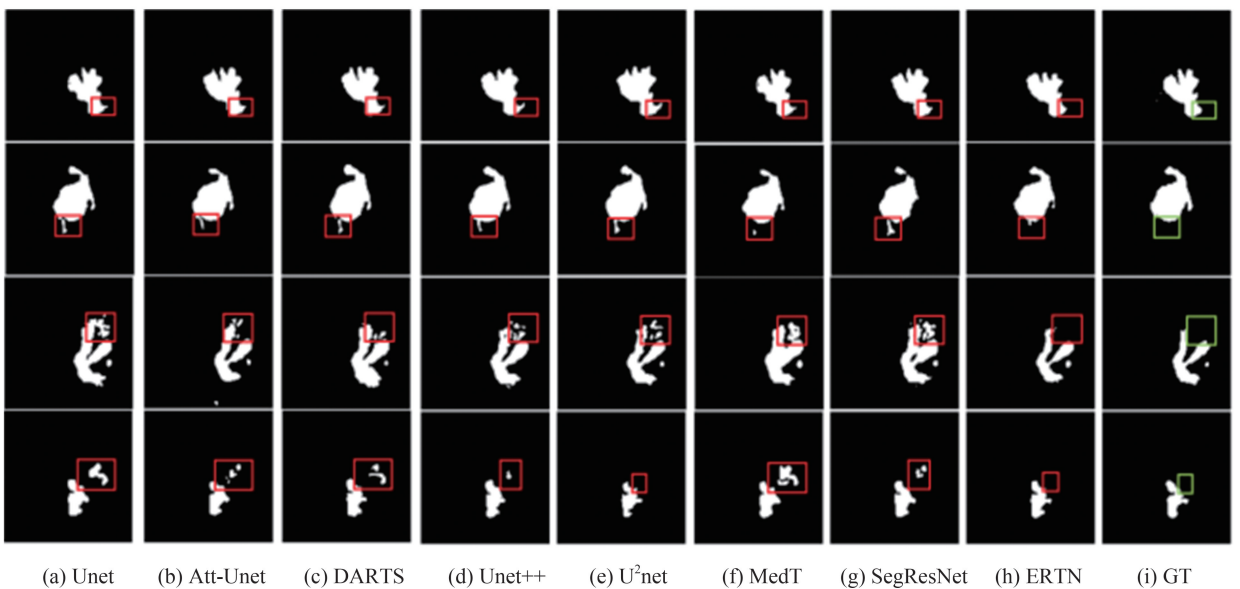


图 5 分割结果可视化对比
Fig.5 Visual comparison of segmentation results

4 结论

为提高脑胶质瘤 MRI 图像分割精度及质量,本研究提出了一种基于混合偏移轴向自注意机制的脑胶质瘤分割算法 MLA-Net。该算法在轴向注意力机制基础上进行改进,使用混合偏移轴向自注意机制提取全局特征,在降低算法复杂度的同时提取更精确的全局相对位置关系,聚合多尺度的全局特征和局部细节特征,减少下采样造成的特征丢失,促进有效特征的表达。针对脑胶质瘤边界模糊,结构不清晰,设计了一种混和损失函数,以提升网络对细节结构特征的敏感程度,精确地分割胶质瘤模糊边界。研究结果表明,本研究设计的 MLA-Net 算法有较好的分割性能,为临床医生诊断脑胶质瘤并确定患者的治疗方案提供了有价值的参考。但在本研究试验中,没有区分低级别胶质瘤与高级别胶质瘤,在后续的工作中可以根据脑胶质瘤种类的不同,设计更优的深度学习方法,进一步提升脑胶质瘤分割的精度和质量。

参考文献:

- [1] BEN K, HALL L O, GOLDFOG D B, et al. Ensembles of convolutional neural networks for survival time estimation of high-grade glioma patients from multimodal MRI [J]. *Diagnostics*, 2022, 12(2): 345-359.
- [2] AHIR B K, ENGELHARD H H, LAKKA S S. Tumor development and angiogenesis in adult brain tumor: glioblastoma [J]. *Molecular Neurobiology*, 2020, 57: 2461-2478.
- [3] LI S, WANG C, CHEN J, et al. Signaling pathways in brain tumors and therapeutic interventions [J]. *Signal Transduction and Targeted Therapy*, 2023, 8(1): 8-31.
- [4] BAI J, VARGHESE J, JAIN R. Adult glioma WHO classification update, genomics, and imaging: what the radiologists need to know [J]. *Top Magn Reson Imaging*, 2020, 29(2): 71-82.
- [5] ANTONIO Di Ieva. Application of deep learning for automatic segmentation of brain tumors on magnetic resonance imaging: a heuristic approach in the clinical scenario [J]. *Neuroradiology*, 2021, 63(8): 1-10.
- [6] ANCA L, GROS U, WOLFGANG A, et al. PET for radiation treatment planning of brain tumours [J]. *Radiotherapy and Oncology*, 2010, 96(3): 325-327.
- [7] CHU R, KIM G, TAUHID S, et al. Whole brain and deep gray matter atrophy detection over 5 years with 3T MRI in multiple sclerosis using a variety of automated segmentation pipelines [J]. *PLoS ONE*, 2018, 13(11): 1-14.
- [8] TUSTISON N J, SHRINIDHI K L, WINTERMARK M, et al. Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR [J]. *Neuroinformatics*, 2015, 13: 209-225.
- [9] KAUS M, WARFIELD S, NABAVI A, et al. Automated segmentation of MR images of brain tumors [J]. *Radiology*, 2001, 218(2): 586-591.
- [10] WU W, CHEN A, ZHAO L, et al. Brain tumor detection and segmentation in a CRF(conditional random fields) framework with pixel-pairwise affinity and super-pixel-level features [J]. *International Journal of Computer Assisted Radiology and Surgery*, 2014, 9(2): 241-253.
- [11] KHAN A, SOHAIL A, ZAHOORA U, et al. A survey of the recent architectures of deep convolutional neural networks [J]. *Artificial Intelligence Review*, 2020, 53(8): 5455-5516.
- [12] MOHAMMED Y M A, EL S, JELLOULI I. A survey of methods for brain tumor segmentation-based MRI images [J]. *Journal of Computational Design and Engineering*, 2023, 10(1): 266-293.
- [13] ZHU Z, HE X, QI G, et al. Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI [J]. *Information Fusion*, 2023, 91: 376-387.
- [14] HAVAEI M, DAVY A, WARDE-FARLEY D, et al. Brain tumor segmentation with deep neural networks [J]. *Medical Image Analysis*, 2017, 35: 18-31.
- [15] ISLAM M, VIBASHAN V S, JOSE V J M, et al. Brain tumor segmentation and survival prediction using 3D attention UNet [C]//Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019. Shenzhen, China. Springer, 2020: 262-272.
- [16] LI X, CHEN H, QI X, et al. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes [J]. *IEEE Transactions on Medical Imaging*, 2018, 37(12): 2663-2674.
- [17] MYRONENKO A. 3D MRI brain tumor segmentation using autoencoder regularization [C]//Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018. Granada, Spain: Springer, 2019: 311-320.
- [18] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. *Advances in Neural Information Processing Systems*, 2017: 30-40.

- [19] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: transformers for image recognition at scale[C]//International Conference on Learning Representations. Vienna, Austria: ICLR, 2021: 21-31.
- [20] AZIZ M J, ZADEH M, FARNIA P, et al. Accurate automatic glioma segmentation in brain MRI images based on CapsNet[C]//International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Guadalajara, Mexico: IEEE, 2021: 3882-3885.
- [21] JIA Q, SHU H. Bitr-unet: a CNN-transformer combined network for MRI brain tumor segmentation[C]//International MICCAI Brainlesion Workshop. Singapore: Springer, 2022: 3-14.
- [22] CHEN J, LU Y, YU Q, et al. Transunet: transformers make strong encoders for medical image segmentation [C]//International MICCAI Brainlesion Workshop. Strasbourg, France: Springer, 2021: 3-14.
- [23] ZHANG Y, LIU H, HU Q. Transfuse: fusing transformers and cnns for medical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Strasbourg, France: Springer, 2021: 14-24.
- [24] WANG W, CHEN C, DING M, et al. Transbts: multi-modal brain tumor segmentation using transformer[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Strasbourg, France: Springer, 2021: 109-119.
- [25] VALANARASU J M J, OZA P, HACIHALILOGLU I, et al. Medical transformer: gated axial-attention for medical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Strasbourg, France: Springer, 2021: 36-46.
- [26] QIN X, ZHANG Z, HUANG C, et al. U2-Net: going deeper with nested U-structure for salient object detection [J]. Pattern Recognition, 2020, 106: 107404.
- [27] WANG H, ZHU Y, GREEN B, et al. Axial-deeplab: stand-alone axial-attention for panoptic segmentation [C]//European Conference on Computer Vision. Scotland, UK: Springer, 2020: 108-126.
- [28] RAMACHANDRAN P, PARMAR N, VASWANI A, et al. Stand-alone self-attention in vision models[J]. Advances in Neural Information Processing Systems, 2019: 32-45.
- [29] SHAW P, USZKOREIT J, VASWANI A. Self-attention with relative position representations [C]//Proceedings of NAACL-HLT. New Orleans, USA: 2018: 464-468.
- [30] RAHMAN M A, WANG Y. Optimizing intersection-over-union in deep neural networks for image segmentation[C]//International Symposium on Visual Computing. Las Vegas, USA: Springer, 2016: 234-244.
- [31] ABDOLLAHI A, PRADHAN B, ALAMRI A. VNet: an end-to-end fully convolutional neural network for road extraction from high-resolution remote sensing data[J]. IEEE Access, 2020, 8: 179424-179436.
- [32] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment [C]//The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers. Pacific Grove, USA: IEEE, 2003: 1398-1402.
- [33] DE BOER P T, KROESE D P, MANNOR S, et al. A tutorial on the cross-entropy method[J]. Annals of Operations Research, 2005, 134(1): 19-67.
- [34] 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(1): 1-13.
- [35] MENZE B H, JAKAB A, BAUER S, et al. The multi-modal brain tumor image segmentation benchmark (BRATS)[J]. IEEE Transactions on Medical Imaging, 2014, 34(10): 1993-2024.
- [36] ZHOU Z, RAHMAN S M M, TAJBAKHSN N, et al. Unet++: a nested U-net architecture for medical image segmentation [C]//International MICCAI Brainlesion Workshop. Granada, Spain: Springer, 2018: 3-11.
- [37] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [C]//International MICCAI Brainlesion Workshop. Munich, Germany: Springer, 2015: 234-241.
- [38] HU Z, LI L, SUI A, et al. An efficient R-transformer network with dual encoders for brain glioma segmentation in MR images[J]. Biomedical Signal Processing and Control, 2023, 79: 104034.

(编辑:郭少华)