

文章编号 1004-924X(2024)04-0565-13

语义流引导采样结合注意力机制的 脑肿瘤图像分割

宋建丽¹, 吕晓琪^{1,2*}, 谷 宇¹

(1. 内蒙古科技大学 信息工程学院, 内蒙古 包头市 014010;

2. 内蒙古工业大学 信息工程学院, 内蒙古 呼和浩特 010051)

摘要: U型网络结构的脑肿瘤自动分割方法由于多次卷积和采样操作会造成信息损失, 导致分割效果不佳。为解决这一问题, 提出了能够利用语义信息流引导上采样特征恢复的特征对齐单元, 并在此基础上设计轻量级的双重注意力特征对齐网络(DAFANet)。首先, 将特征对齐单元分别引入 3D UNet、DMFNet 和 HDCNet 三个经典网络, 以验证其有效性和泛化性。其次, 在 DMFNet 基础上构造轻量级的双重注意力特征对齐网络 DAFANet, 利用特征对齐单元强化上采样过程中的特征恢复, 3D 期望最大化注意力机制同时作用于特征对齐路径和级联路径, 用于重点获取上下文的全程依赖关系。同时使用广义 Dice 损失函数提升数据不平衡时的分割精度并加快模型收敛。最后, 在 BraTS2018 和 BraTS2019 公开数据集进行验证, 文中所提算法在 ET, WT 和 TC 区域的分割精度分别达到 80.44%, 90.07%, 84.57% 和 78.11%, 90.10%, 82.21%。相较于当前流行的分割网络, 具有对增强肿瘤区域更好的分割效果, 更擅长处理细节和边缘信息。

关键词: 脑肿瘤; 图像分割; 特征对齐; 注意力机制; 轻量化

中图分类号: TP394.1; TH691.9 **文献标识码:** A **doi:** 10.37188/OPE.20243204.0565

Brain tumor image segmentation based on Semantic Flow Guided Sampling and Attention Mechanism

SONG Jianli¹, LÜ Xiaoqi^{1,2*}, GU Yu¹

(1. School of Information Engineering, Inner Mongolia University of Science and Technology,
Baotou 014010, China;

2. School of Information Engineering, Inner Mongolia University of Technology,
Hohhot 010051, China)

* Corresponding author, E-mail: lxiaoqi@imut.edu.cn

Abstract: The automatic segmentation method for brain tumors based on a U-shaped network structure often suffers from information loss due to multiple convolution and sampling operations, resulting in suboptimal segmentation results. To address this issue, this study proposed a feature alignment unit that utilizes

收稿日期: 2023-04-05; **修订日期:** 2023-05-09.

基金项目: 国家自然科学基金项目 (No. 62001255, No. 61841204, No. 61771266); 中央引导地方科技发展资金项目 (No. 2021ZY0004); 内蒙古自治区高等学校青年科技英才支持项目 (No. NJYT23057); 内蒙古自治区自然科学基金 (No. 2019MS06003, No. 2015MS0604); 内蒙古科技大学基本科研业务费专项资金优秀青年基金 (No. 042); 内蒙古自治区高等学科科学技术研究项目 (No. NJZY145); 教育部“春晖计划”合作科研项目 (No. 教外司留[2019]1383 号)

semantic information flow to guide the up-sampling feature recovery and design designed a lightweight Dual Attention Feature Alignment Network (DAFANet) based on this unit. Firstly, to validate its effectiveness and generalization, the feature alignment unit was introduced separately into three classic networks, namely 3D UNet, DMFNet, and HDCNet. Secondly, a lightweight dual-attention feature alignment network named DAFANet was proposed based on DMFNet. The feature alignment unit enhanced feature restoration in the up-sampling process, and a 3D Expectation-Maximization attention mechanism was applied to both the feature alignment path and cascade path to capture the full contextual dependency. The generalized Dice loss function was also used to improve segmentation accuracy in the case of data imbalance and accelerate model convergence. Finally, the proposed algorithm is validated on the BraTS2018 and BraTS2019 public datasets, achieving segmentation accuracies of 80.44%, 90.07%, 84.57% and 78.11%, 90.10%, 82.21% in the ET, WT, and TC regions, respectively. Compared to current popular segmentation networks, the proposed algorithm demonstrates better segmentation performance in enhancing tumor regions and is more adept at handling details and edge information.

Key words: brain tumors; image segmentation; feature alignment; attention mechanism; lightweight

1 引 言

脑肿瘤是颅内肿瘤的简称,由脑部非正常增长的神经细胞组成,相较于其他肿瘤具有更高的致命性。脑胶质瘤是所有脑肿瘤中最常见、最具有侵袭性的恶性肿瘤^[1]。据世界卫生组织报道,根据其显微图像和肿瘤行为可分为四个等级^[2],I,II级为低级别胶质瘤(Low Gliomas grade, LGG),III级和IV级为高级别胶质瘤(High Gliomas grade, HGG),级别越高,致命性越强。磁共振成像(Magnetic Resonance Imaging, MRI)作为一种高性能的软组织成像手段,能够显示不同模式的信息,现已成为脑肿瘤检测和诊断的重要工具^[3]。

在临床实验中,脑肿瘤的精细分割主要依靠经验丰富的医生,注意力不集中、疲劳以及缺乏相关工作经验等因素会影响分割结果^[4]。尽管研究人员已经进行了深入研究,但是由于脑肿瘤的位置、大小和形状具有不确定性,病变组织和正常组织之间的低质量成像和边界模糊使得在实际操作中难以获得足够的分割精度,脑肿瘤分割仍然面临很大挑战^[5]。近年来,深度学习技术广泛应用于图像分析与处理、语音识别等领域,并取得突破性进展,现已成为脑肿瘤图像分割的主要技术手段^[6]。

卷积神经网络(Convolutional Neural Networks, CNN)因其强大的特征提取能力,广泛应

用于图像处理领域^[7]。Ronneberger等^[8]在全卷积神经网络(Fully Convolutional Networks, FCN)^[9]的基础上首次提出UNet模型,实现了端到端的自动分割,已成为医学图像分割领域的主流网络。Çiçek等^[10]将UNet扩展到3D UNet,改进后的网络能够更好地学习MRI图像不同维度的信息,并且各项指标的参数均优于二维的UNet网络。MultiResUNet^[11]作为UNet的增强版本,利用具有学习能力的残差块代替简单的跳跃连接,能同时对3D和2D图像产生更好的分割效果,尤其是对多模态数据集的处理。随着注意力机制有效性被证实,研究者们尝试将注意力机制引入卷积神经网络,Oktay等^[12]通过将注意力门控机制嵌入跳跃连接,实现了对特定目标结构的自动学习。Wang^[13]和Jiang^[14]将基于自注意力机制的Transformer和SwinTransformer与UNet融合,进一步实现了对脑肿瘤的精确分割。

现有模型在获得高级语义信息时会损失重要的边缘和细节信息,杨坚华等^[15]构建了并行多路径网络,使用同时进行特征提取和融合策略减少语义信息丢失;梁礼明等^[16]通过在编码部分级联高分辨率网络减少特征缺失,增强病变像素传递,但是过多的卷积操作导致网络结构冗余复杂。Huang等人^[17]在编码阶段设置并行路径,在下采样的同时融合细节信息,从输入端降低采样造成的信息损失,但忽视了不同层级的语义信息差。Aboelenein等^[18]针对肿瘤形式和大小差异

性,设计层数不同、卷积核大小不同的双轨道UNet结构,降低了小尺度特征在训练中的损失,但缺少对边缘细节信息的关注。Lu等^[19]则利用多层的DenseNet实现特征重用,保留低级信息。

针对传统卷积带来的模型参数量大,计算消耗高等问题,Chen等^[20]提出DMFNet网络,使用多纤维单元构成基础卷积层,利用不同扩张率的空洞卷积获取多尺度特征表达,显著降低了参数量和计算量,但是过多空洞卷积造成了图像特征丢失。Luo等^[21]通过对标准卷积进行跨信道和空间域的分层解耦,提出了一种轻量级的伪3D模型HDCNet,在减少信息冗余的同时完成对空间上下文信息的学习,极大地降低了参数量和计算复杂度,但是对增强区域分割欠佳。Zhou等^[22]利用计算效率较高的3D ShuffleNetV2作为编码器,利用逐点卷积和通道重排,降低计算复杂度和内存消耗,不足之处是忽略了解码器对模型效率和精度的影响。

由上述研究内容可知,通过简单地叠加模型结构,增加额外路径或者多尺度融合等方法,无法很好地缓解因网络层次加深导致的信息损失问题,同时还会造成网络自身的复杂度提高。为更好地解决特征信息损失和模型复杂度过高的问题,本文提出了一种轻量级的双重注意力特征对齐网络(Dual Attention Feature Align Net, DAFANet)。首先,在网络解码阶段使用特征对齐单元减少上采样造成的信息损失,改变传统分割

模型对上采样阶段特征恢复的直接操作,显示地学习不同层级之间存在的语义流差异性,校正语义信息差,实现更准确的特征恢复。其次,针对脑肿瘤图像的三维数据形式,将2D期望最大化注意力机制扩展为3D变体,分别作用于特征对齐单元和级联路径,有助于网络更好地利用图像的体素信息,同时促进网络对全局依赖关系的获取,双路径形式能更好地减少输入噪声,降低输入特征的差异性,强化特征恢复过程。最后,使用广义Dice损失函数提升数据不平衡条件下的分割精度,加快网络收敛。

2 脑肿瘤图像分割方法

本文提出的DAFANet模型是基于实时分割网络DMFNet改进,其网络结构如图1所示。4通道对应脑肿瘤四个模态。模型采用U型编码解码结构,在编码阶段首先采用步长为2的 $3 \times 3 \times 3$ 卷积初步采样,使用如图2所示的多纤维(Multi-Fiber, MF)单元和扩张多纤维(Dilated Multi-Fiber, DMF)单元作为卷积块,并在下采样中设置组卷积数为16,其中扩张多纤维单元用于解决在分辨率较低时感受野较小的问题,可学习的权重系数 $\omega_1, \omega_2, \omega_3$ 分配给每个扩张分支,最后三个分支相加。多纤维结构通过使用3个并行的纤维单元将具有 C_m 通道的输入分为三组,将复杂的神经网络切片替换成由多个纤维组成的集合,

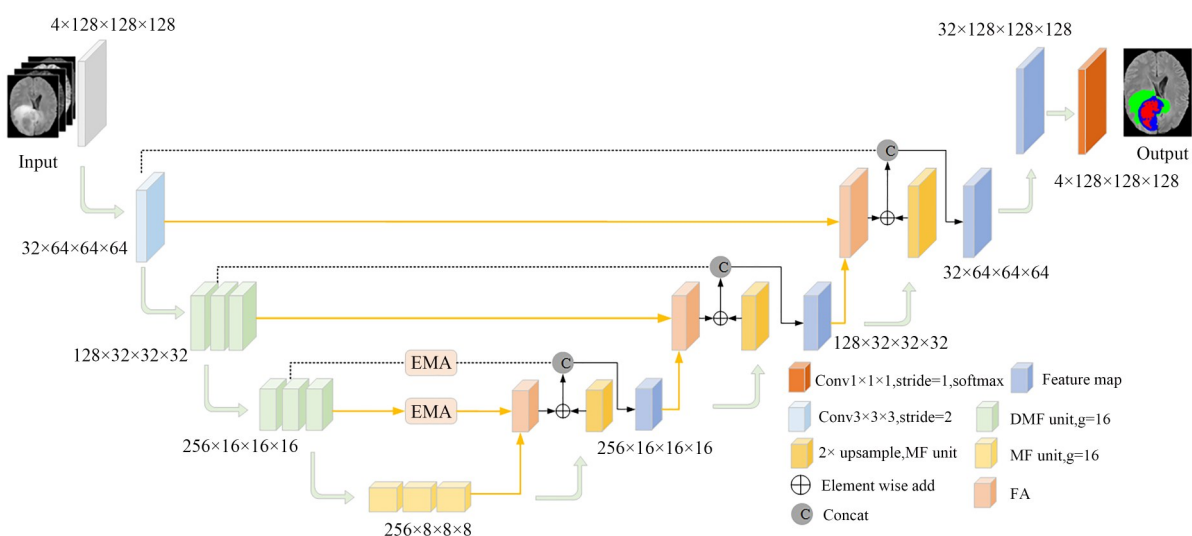


图1 DAFANet整体结构图

Fig. 1 Overall network architecture of our DAFANet model

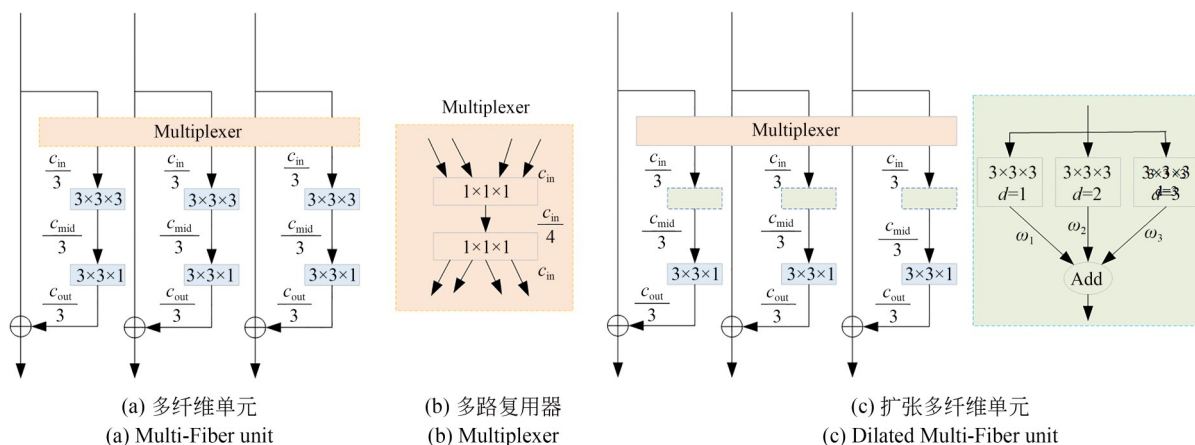


图 2 多纤维单元和扩张多纤维单元结构图

Fig. 2 Structural diagram of Multi-Fiber unit and Dilated Multi-Fiber unit

结合具有通道变换功能的多路复用器,极大地提高模型的效率,最终形成轻量级的 3D 网络结构。其中多路复用器的具体结构如图 2(b)所示,由两层 $1 \times 1 \times 1$ 卷积组成,用于不同纤维块之间的信息交互。

解码阶段由特征对齐(Feature Align, FA)单元和多纤维单元构成,最后经过一层 $1 \times 1 \times 1$ 卷积和 softmax 层得到最后的分割图。特征图在执行上采样和级联操作之前,特征对齐单元会显示地学习编码阶段和解码阶段特征映射之间的语义信息差,利用信息差形成第一层采样层,随后和上采样特征相加,最后和编码路径的信息级联生成最后的特征信息。第三采样层在跳跃连接路径和特征对齐路径融入 3D 期望最大化注意力机制(Expectation-Maximization Attention, EMA),用于获取上下文信息,降低输入噪声。

2.1 特征对齐单元

医学图像分割更加注重图像细节和边缘,脑肿瘤 MRI 图像中的细节部分在临床中包含更多分割信息,单纯的上采样只能处理一些固定的和预定义的特征错位问题,缺少对细节和边缘的特殊处理。而多次重复的上采样和下采样会造成特征图之间信息失调更加复杂,导致高级语义信息和低级语义信息简单级联会产生信息差,这种信息差在 UNet 的 concat 操作中就已经有体现。Li 等^[23]针对这一问题提出了能够减少采样造成的信息差问题的流对齐单元,受此启发,本文通过深入研究语义流,提出能够引导信息对齐的特征对齐单元,并应用到脑肿瘤图像分割领域。

特征对齐单元将特征的逐层传递抽象为“流”的形式,语义信息流可以简单概述为具有语义差的特征在减少错位问题时信息流动所形成的差值域,其核心思想是用深层语义信息弥补浅层语义信息的缺失。通过学习不同分辨率之间的语义信息,动态地建立特征图之间的位置对应关系,降低相邻级别特征图的语义信息差,促使网络得到更加丰富的空间和语义信息。特征对齐单元以不同层级的特征信息作为输入,通过评估不同层级之间的特征差异得到合适的语义流场,该流场会自动调整两层之间的差异,给出有效的特征对齐信息。网络构成如图 3 所示。

多次递归使用下采样操作,导致上采样的特征映射与对应层的特征映射存在可预测的空间错位,通过简单地元素添加或者通道融合都会损害目标边界的预测。特征对齐单元通过动态学习特征映射之间的语义信息差,进而减少空间错位的产生,实现更准确的特征恢复。整体流程可分为如下两步:(1)获取显著映射区域;(2)特征重采样。

实现特征对齐的关键是寻找两个相邻特征层之间的对应点,由于脑肿瘤图像前景和背景之间像素极不平衡,实际分割区域在脑肿瘤图像占比不足,那么特征图中的绝大多数显著区域可以当做信息传播的关键点,也是实现特征对齐的关键信息。

首先,对于给定的两个相邻层级的特征映射

$$F_l \in \mathbb{R}^{C_l \times D_l \times H_l \times W_l} \text{ 和 } F_h \in \mathbb{R}^{C_h \times \frac{D_l}{2} \times \frac{H_l}{2} \times \frac{W_l}{2}} \quad (C \text{ 为通道})$$

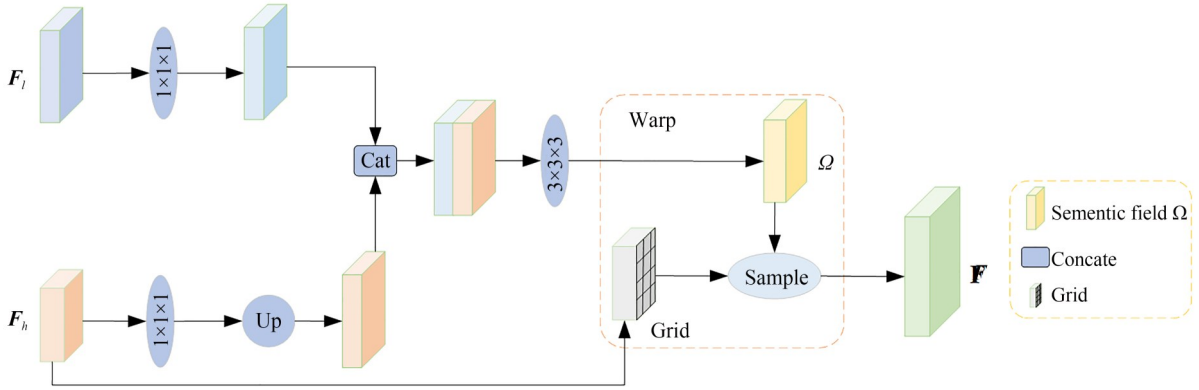


图3 特征对齐单元

Fig. 3 Feature align unit

数, D, H, W 分别代表三维图像的深度、长度和宽度值, h 代表高级别语义信息层, l 代表低级别语义信息层), 分别使用 $1 \times 1 \times 1$ 卷积调整至相同的通道数, 将 F_h 通过三线性插值上采样到与 F_l 相同的大小, 得到的输出与 F_l 连接, 连接后的特征映射输入到一个 $3 \times 3 \times 3$ 的子网络中, 通过初步特征提取获得显著区域的特征映射, 得到的输出即为预测的语义信息场 $\Omega \in \mathbb{R}^{D_l \times H_l \times W_l \times 3}$, 其中包含特征对齐所需要的关键位置信息和特征重采样需要的偏移量。数学上, 整体步骤可以通过公式 (1) 描述:

$$\Omega = \text{conv}_{3 \times 3 \times 3} [\text{cat}(F_h, F_l)], \quad (1)$$

其中: $\text{cat}(\cdot)$ 代表级联操作, $\text{conv}_{3 \times 3 \times 3}$ 代表卷积核 $3 \times 3 \times 3$ 的卷积操作。

经过 Ω 流域的预测, F_l 所处的空间灰度域中的每一个位置 p_l 都可以被映射到高级语义信息层的点 p_h 上, 最后将 F_h 和 Ω 经过 warp 操作得到最后的特征映射。对于矫正之后的深层特征, Li 等人^[23]采用与浅层特征信息融合的方式来获取最终的结果。但是对于像脑肿瘤这种实际分割占比不足而且分割区域呈现嵌套形式的图像数据, 当学习到的语义信息具备的对齐效果不佳时, 会影响各个子区域的边界分割效果。为此, 本文采用模板匹配的方式, 通过生成与 F_h 同样大小的灰度规则网格 grid, 将 Ω 中对应坐标点的像素赋值到 grid 的相应位置, 基于语义相似度将低分辨率特征图的样本提升到高分辨率的大小, 实现对高级语义信息重采样的。Warp 核心是通过

使用学习到的关键位置点和偏移量调整卷积核中每个采样点的位置, 实现将上采样的特征映射对齐到一组特征映射 grid 中去, 最后输出的特征图 F 将会有特征对齐的特性。在数学上, 可以通过公式 (2) 描述:

$$F = \text{sample}[F_h, (\text{grid} + \Omega)]. \quad (2)$$

2.2 期望最大化注意力机制

在像素级别的脑肿瘤图像分割过程中, 详细的上下文信息对分割任务至关重要, 而卷积操作只能获取卷积核区域内像素的相邻关系, 无法建立与其他像素之间的联系。自注意力机制通过计算像素点之间的特征加权和获取特征表示, 以此捕获远程信息。传统自注意力机制需要在所有位置上计算所需的注意力图, 导致较高的计算复杂度。Li 等人^[24]从期望最大化迭代 (Expectation-Maximization, EM) 算法的角度对自注意力机制优化, 使用 EM 算法迭代得到一组紧凑的基, 在基上计算权重, 从而降低计算量。这一过程主要通过 E 步更新注意力图, M 步更新这组基实现, 其中 E、M 需要交替执行, 直至收敛。本文在原有的注意力机制基础上进行改进, 将其扩展为 3D 形式, 以更好地适应脑肿瘤的三维图像模式, 使其能够更好地捕获全面的空间信息。

EMA 的具体流程可以描述为, 对于给定的输入大小为 $C \times D \times H \times W$ 的特征图 X , 首先将其重塑为 $N \times C$ 大小, 其中 $N = D \times H \times W$, 选择初始基 $\mu \in \mathbb{R}^{K \times C}$ (K 为基的个数), E 步估计隐变量 $Z \in \mathbb{R}^{N \times K}$, Z 表示每个基对像素的权责, 则第 K

个基对第 n 个像素 x_n 的权责可以计算为:

$$z_{nk} = \frac{P(x_n, \mu_k)}{\sum_{j=1}^K P(x_n, \mu_j)}, \quad (3)$$

其中: $1 \leq k \leq K$, $1 \leq n \leq N$, 内核 $P(a, b)$ 选择 $\exp(a^T b)$ 的形式。则第 t 次迭代中的步骤 E 可以表示为:

$$Z^{(t)} = \text{softmax}\left(\lambda X (\mu^{(t-1)})^T\right), \quad (4)$$

其中, λ 作为超参数控制 $y_{\text{形}} = \sum_{i=0}^m y_i / m$ 的分布。

用 M 步更新基 μ , μ 被计算作为 X 的加权平均, 则第 k 个基被更新为:

$$\mu_k^{(t)} = \frac{\sum_{n=1}^N z_{nk}^{(t)} x_n}{\sum_{m=1}^K z_{mk}^{(t)}}. \quad (5)$$

EMA 算法交替执行 E 步和 M 步算法 T 次之后, 近似收敛的 μ 和 Z 对 X 进行重估计得 $\bar{X}$, 公式为:

$$\bar{X} = Z^{(T)} \mu^{(T)}. \quad (6)$$

其主要流程如图 4 所示, 对于输入的特征图 X , 交替执行 E 步、M 步直至收敛, 使用收敛的 μ 和生成的注意力图 Z 对 X 进行特征重构得到 $\bar{X}$, 两个 $1 \times 1 \times 1$ 卷积分别用于变换输入的值域和将估计值 $\bar{X}$ 映射到残差空间, 最后经由残差运算得到加强后的特征信息。

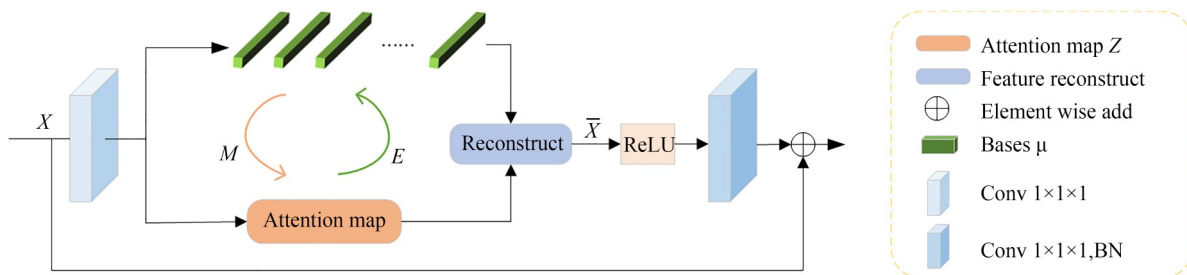


图 4 期望最大化注意力机制整体架构

Fig. 4 Structure of expectation-maximization attention

2.3 损失函数

脑肿瘤的病变区域占整个输入空间的一小部分, 存在目标区域和背景区域分布占比不平衡的问题, 而传统的 Dice 损失函数不利于小目标区域的分割, 小目标区域部分像素的错误预测会引起 Dice 系数大幅度波动, 不利于模型收敛。为此, 本文选择使用广义 Dice 损失函数 (Generalized Dice Loss, GDL) 缓解小目标区域预测效果不佳时对模型的影响。GDL 在 Dice Loss 的基础上引入权重以改善对小目标检测不利的情况, 同时将多个类别预测情况下的结果进行整合, 使用一个指标作为分割结果的预测指标。脑肿瘤图像分割作为对健康组织和病灶的二分类任务, 此时广义 Dice 损失函数如公式 (7) 所示:

$$L_{\text{GDL}} = 1 - \frac{2 \sum_{i=1}^m \omega_i \sum_{j=1}^N r_{ij} p_{ij}}{\sum_{i=1}^m \omega_i \sum_{j=1}^N (r_{ij} + p_{ij})}, \quad (7)$$

其中: m 为类别数, r_{ij} 为类别 i 在第 j 个像素处的标准值, p_{ij} 为类别数 i 在第 j 个像素处的预测值, ω

为每个类别的权重, 由类别中的像素个数决定, 具体如式 (8) 所示:

$$\omega_i = \frac{1}{\left(\sum_{j=1}^N r_{ij}\right)^2}. \quad (8)$$

3 实验结果与分析

3.1 实验环境

本文所用实验环境为: CPU: 15 vCPU AMD EPYC 7543 32-Core Processor, GPU: RTX 3090 (24 GB) * 1; ubuntu18.04 操作系统, Pytorch 深度学习框架。

3.2 数据集

实验数据集使用脑肿瘤 (BraTS) 分割挑战赛 BraTS2018 和 BraTS2019 数据集, 数据集的数据构成如表 1 所示, 其中使用 BraTS2019 做主要实验, BraTS2018 用于不同数据集的泛化验证。每个病例都有四种模态, 包括 T1 加权 (T1)、T2

表1 数据集构成

Tab. 1 Datasets composition

Datasets	Train	Valid
BraTS2018	285	66
BraTS2019	335	125

加权(T2)、T1对比(T1ce)和液体衰减向内恢复(FLARE)序列,每种模态的图像大小都是 $240 \times 240 \times 155$ 。所有训练集均有标签,验证集没有标签,病例的真实标签均有多位经验丰富的医师标注。论文中的所有结果均在BraTS挑战赛的在线评估服务器进行测评,验证算法的有效性。图5为同一病例不同模态的脑肿瘤图像和真实分割标签。

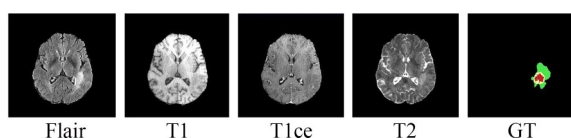


图5 不同模态的脑肿瘤图像与真实分割标签

Fig. 5 Brain tumor image and real segmentation tags with different modalities

3.3 评价指标

为对分割结果进行有效评估,采用多指标综合评判方法。选择Dice相似系数、灵敏度(Sensitivity)、特异性(Specificity)和豪斯多夫距离(Hausdorff95 distance)对脑肿瘤的浮肿(Whole Tumor, WT)区域、坏疽(Tumor Core, TC)区域和增强肿瘤(Enhance Tumor, ET)区域进行评价。其中Dice相似系数用来计算分割结果和图像真实标签的相似程度,取值范围为 $[0, 1]$, Dice值越大说明分割精度越高。计算公式为:

$$Dice = \frac{2TP}{FN + FP + 2TP}. \quad (9)$$

灵敏度和特异性用于衡量结果和真实值之间的体素重叠程度,具体计算方式为:

$$Sensitivity = \frac{TP}{TP + FN}, \quad (10)$$

$$Specificity = \frac{TP}{TP + FP}, \quad (11)$$

其中: TP 表示真阳性, FP 为假阳性, FN 为假阴性。

Hausdorff95距离用来评估模型的分割结果

和真实标签边界之间的距离,是分割结果边界中的某个点到真实标签边界中的最近点的最大值,距离越小说明分割精度越高。计算公式为:

$$HD(X, Y) =$$

$$\max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}, \quad (12)$$

其中: $\sup(\cdot)$ 用于寻找集合下确界, $\inf(\cdot)$ 寻找集合上确界, $d(x, y)$ 表示 x, y 两点之间的欧式距离, X 和 Y 分别为网络预测的脑肿瘤区域和真实标签区域的体素集, x, y 分别为两个体素点的距离。

3.4 参数设置

实验过程中使用自适应矩阵估计(Adam)进行训练,设置初始学习率为 10^{-3} ,使用衰减率为 10^{-5} 的L2范数作为正则化系数;使用数据增强技术扩充数据集,包括将图像从 $240 \times 240 \times 155$ 随机剪裁为 $128 \times 128 \times 128$,以0.5的概率沿任意轴向对图片进行随机镜像翻转,对剪裁后的图像在 $[-10^\circ, 10^\circ]$ 范围内随机旋转等操作。在EMA模块中设置基集数 k 的值为64,迭代次数 T 为3,超参数 λ 为1。根据实验设备资源情况,设置批处理大小为4,训练500轮基本收敛。

3.5 特征对齐单元的有效性验证

为更好地证明所提模块的有效性和泛化性,分别在DMFNet、HDCNet和3D UNet网络进行实验,结果如表2所示。实验结果表明,结合FA单元后的三个基础网络综合性能均得到提升,其中Dice相似系数的效果最为明显,表明经过处理之后的特征映射具备信息对齐的效果,证明了特征对齐单元在减少特征错位问题上的有效性。尤其是改进后的3D UNet在Dice系数和灵敏度指标上较原网络都有大幅提升,嵌入FA的3D UNet在ET、WT和TC三个区域的Dice相似系数分别提升了1.97%,1.03%,1.74%,灵敏度分别提升了4.13%,1.06%,3.23%。DMFNet在TC区域的Dice相似系数增加了0.79%,HDCNet在WT区域的Dice精度增加了0.41%,在其他区域的精度也都有小幅提升。同时改进后的三个网络在特异性和Hausdorff95距离方面也有不同程度的改善,充分证明了特征对齐单元的有效性和泛化性。

表 2 不同模型使用 FA 后的效果

Tab. 2 Comparison of effects of different models using FA

Methods	Dice/%			Sensitivity/%			Specificity/%			Hausdorff95/mm		
	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC
DMFNet	76.68	88.47	80.19	77.79	88.31	79.14	99.81	99.42	99.73	3.341	5.407	6.138
FA+DMFNet	76.78	88.54	80.98	77.53	88.70	79.49	99.82	99.52	99.75	3.134	5.385	6.148
3D UNet	72.68	87.18	70.88	70.15	86.72	65.81	99.84	99.49	99.89	5.341	9.172	9.461
FA+UNet	74.65	88.21	72.62	74.28	87.78	69.04	99.88	99.47	99.83	5.422	5.430	9.652
HDCNet	76.98	89.15	81.51	77.63	92.26	80.61	99.83	99.26	99.75	4.159	6.433	8.009
FA+HDCNet	77.17	89.56	81.55	78.94	92.22	80.78	99.81	99.28	99.71	3.079	5.769	7.133

3.6 DAFANet 的消融实验

为验证本文所提模块的有效性,在 DMFNet 网络的基础上分别使用特征对齐单元(FA),期

望最大化注意力机制(EMA),对比不同网络结构下 Dice 相似系数和 Hausdorff95 距离,验证不同模块对分割结果的有效性,结果如表 3 所示。

表 3 DAFANet 模型的消融实验

Tab. 3 Ablation experiment of DAFANet model

Methods	Dice/%			Hausdorff95/mm		
	ET	WT	TC	ET	WT	TC
DMFNet	76.68	88.47	80.19	3.341	5.407	6.138
DMFNet+EMA	77.14	89.69	81.25	3.960	5.181	7.091
FA+DMFNet	76.78	88.54	80.98	3.134	5.385	6.148
FA+DMFNet+EMA(up)	76.81	89.78	81.61	3.742	4.863	6.954
FA+DMFNet+EMA(down)	77.09	89.59	81.54	2.677	4.959	6.125
FA+DMFNet+2EMA	78.11	90.10	82.21	3.072	4.360	6.229

表中的“up”和“down”分别表示在特征对齐路径和跳跃连接路径使用 EMA,“2EMA”表示在两条路径同时使用 EMA。结果表明,不同结构下的 Dice 相似系数较基础网络都有不同程度的提升,本文提出的 DAFANet 模型 Dice 相似系数达到了 78.11%,90.10%,82.21%,对比原网络分别提升了 1.43%,1.63%,2.02%。Hausdorff95 距离在 ET 和 WT 区域均有减少,但在 TC 区域略有增加;引入 EMA 之后的效果更佳,

尤其是在 ET 和 TC 区域的分割,对跳跃连接和特征对齐引导路径同时使用 EMA 的综合效果优于单独使用 EMA。为进一步验证 EMA 位置对模型整体的影响,分别在网络第一层、第二层和第三层使用 EMA,验证其位置对分割精度和模型参数的影响,结果如表 4 所示。由表 4 可以看出,在第三层使用 EMA 的分割效果优于在其他位置的分割值,同时计算量低于其他位置,此时模型的整体性能达到最优。

表 4 EMA 位置的消融实验

Tab. 4 Ablation experiment of EMA position

Methods	Dice/%			Hausdorff95/mm			Params/M	FLOPs/G
	ET	WT	TC	ET	WT	TC		
Layer 1	76.72	88.35	78.67	4.316	5.972	6.871	4.11	30.51
Layer 2	74.85	87.66	76.25	3.241	8.451	7.628	4.13	31.04
Layer 3	78.11	90.10	82.21	3.072	4.360	6.229	4.23	30.50

3.7 基数 k 值选择

基集的初始值会影响EMA收敛到全局最大值的进程,在迭代过程中尤为重要。为更好地发挥算法性能,本文针对 k 的取值做如表5所示的对比实验,验证 k 取不同数值的分割效果,Dice系数的显式对比如图6所示。结果表明,分割效果随 k 的取值表现出明显的差异性, $k=16$ 和 $k=64$ 优于 $k=32$ 的分割效果,特别地,当 $k=64$ 时,在TC区域的分割精度值较 $k=32$ 增加了2%,能够实现整体效果最佳。

表 5 超参数 k 的对比实验

Tab. 5 Comparative experiment of superparameter k

k	Dice/%			Hausdorff95/mm		
	ET	WT	TC	ET	WT	TC
64	78.11	90.10	82.21	3.072	4.306	6.229
32	76.01	89.38	80.21	3.246	5.027	6.071
16	78.16	90.03	80.73	3.022	4.682	6.326

3.8 与经典模型的结果对比

本文的DAFANet模型与3D UNet、注意力UNet、DMFNet等经典脑肿瘤分割模型的性能指标对比结果如表6所示,从表中可知DAFANet网络的整体分割效果优于其他六个模型,Dice分割精度有不同程度提高,Hausdorff95距离也有不同程度减少。在参数量和计算量(每秒浮点运算次数)上,DAFANet的整体参数量为4.23 M,比原网络和3D ESPNet略有增加,远小于注意力UNet和TransBTS,约为注意力UNet和TransBTS参数的1/8,约为3DUNet参数的1/3,较

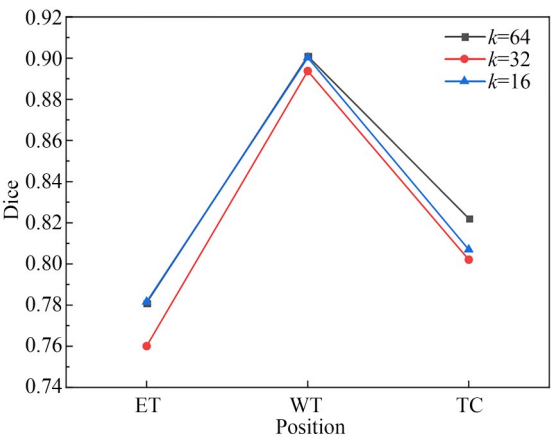


图 6 k 取不同值的对比图

Fig. 6 Comparison chart of different values of k

HDCNet增加了3.94 M;在计算量上,约为TransBTS的1/11,约为3D UNet的1/6,约为注意力UNet和3D ESPNet的1/2。相较于HDCNet,DMFNet和3D ESPNet,DAFANet能够在较少增加计算量或者参数量的基础上,分割精度得到有效提升。

为更好地显示数据分散性,对DAFANet和DMFNet的Dice相似系数做箱线图比较。图7中最上方和最下方的线段分别代表最大值和最小值,箱体自上而下分别代表第三四分位数、中位数和第一四分位数。由图7可知DAFANet整体数据的平均水平较高,在ET和TC区域的Dice相似系数稳定性高于DMFNet,离散数据值更接近第一四分位数值。综合而言,DAFANet相对于DMFNet能够在保持分割稳定性的同时拥有更好的分割效果。

表 6 与经典模型的分割结果对比

Tab. 6 Comparison of segmentation results with classical models

Methods	Dice/%			Hausdorff95/mm			Params/M	FLOPs/G
	ET	WT	TC	ET	WT	TC		
TransBTS ^[13]	78.93	90.00	81.94	3.736	5.644	6.049	32.99	333.00
3D UNet ^[10]	72.68	87.18	70.88	5.341	9.172	9.461	13.12	176.61
3D ESPNet ^[25]	77.35	89.30	81.83	4.190	6.740	7.980	3.63	76.51
Attention UNet ^[12]	75.96	88.81	77.20	5.020	7.756	8.258	34.90	51.30
DMFNet ^[20]	76.68	88.47	80.19	3.341	5.407	6.138	3.88	27.04
HDCNet ^[21]	76.98	89.15	81.51	4.159	6.433	8.009	0.29	24.00
DAFANet	78.11	90.10	82.21	3.072	4.360	6.229	4.23	30.50

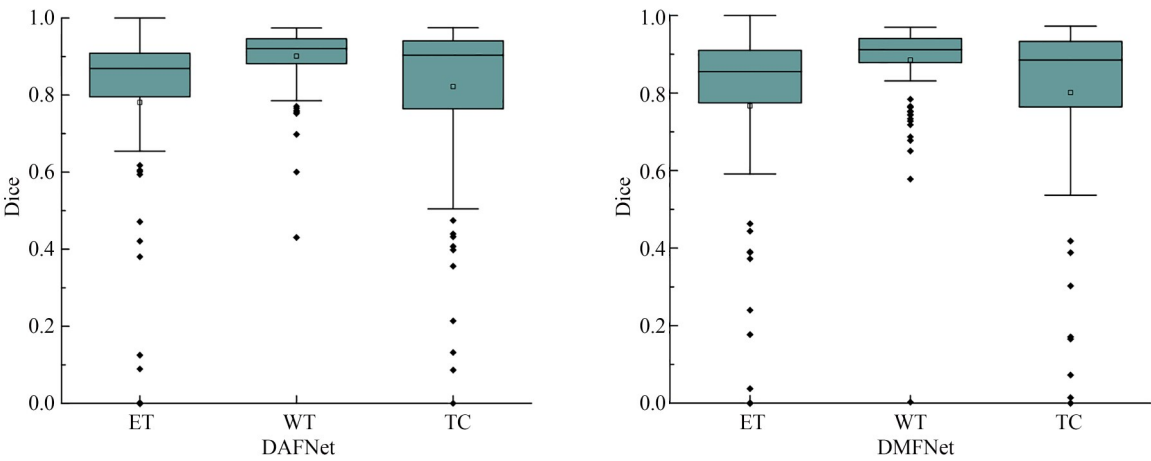


图 7 DAFANet和DMFNet的箱线图对比

Fig. 7 Comparison of box diagram between DAFANet and DMFNet

表 7 不同数据集下的 Dice 系数和 Hausdorff95 距离同其他模型的结果对比

Tab. 7 Comparison of Dice coefficient and Hausdorff95 distance with other models under different datasets

Dataset	Methods	Dice/%			Hausdorff95/mm		
		ET	WT	TC	ET	WT	TC
BraTS2019	Nuechterlein ^[25]	77.35	89.30	81.83	4.19	6.74	7.98
	Chen ^[20]	76.68	88.47	80.19	3.34	5.41	6.14
	Zhang ^[26]	70.90	87.00	70.90	N/A	N/A	N/A
	Sheng ^[27]	72.40	87.50	78.80	5.79	9.35	11.47
	Akbar ^[28]	74.20	88.48	80.98	6.67	10.25	10.83
	Liu ^[29]	77.91	89.94	83.89	4.03	5.45	6.56
	Chang ^[30]	78.20	89.00	81.20	3.82	8.53	7.43
	Ours	78.11	90.10	82.21	3.07	4.36	6.23
BraTS2018	Nuechterlein ^[25]	73.70	88.30	81.40	5.30	5.46	7.85
	Chen ^[20]	80.11	90.61	84.54	3.06	4.66	6.44
	Zhang ^[26]	77.20	87.20	80.80	5.69	12.60	9.62
	Sheng ^[27]	77.10	87.60	81.10	3.22	8.42	10.56
	Akbar ^[28]	77.71	89.59	79.77	3.90	9.13	8.67
	Liu ^[29]	80.41	89.85	85.44	2.24	4.05	5.76
	Chang ^[30]	79.50	90.00	83.90	2.92	6.51	5.71
	Ours	80.44	90.07	84.57	2.75	4.70	5.49

3.9 不同数据集的泛化实验

为验证该模型在其他数据集上的效果,基于 BraTS2018 验证集开展进一步实验,在三个区域最终的 Dice 分割精度分别达到 80.44%, 90.07%, 84.57%, 对应的 Hausdorff95 距离分别达到 2.75 mm, 4.70 mm, 5.49 mm。将两个数据集的分割结果和当前流行方法进行比较,结果如表 7 所示。本文所提算法在两个数据集

ET 区域的分割几乎都可以获得可观的 Dice 精度,侧面验证了算法对细小区域分割的有效性。相对最新的 Liu 等^[29]和 Chang 等^[30]的模型,DAFANet 的结果虽然在 TC 或者 WT 的分割值略低,但在其他区域都有不同程度的提升,进一步说明在获取细节和边缘轮廓信息上有一定的可行性。综合对比,本文的综合效果更佳。

3.10 结果的可视化比较

为更直观地显示分割效果,对部分分割结果进行可视化,如图8所示(彩图见期刊电子版)。本文DAFANet网络相比其他网络更接近真实值,能显著减少分类错误。DMFNet、注意力UNet和UNet的分割结果图比较光滑,但是丢失了许多细节信息,导致轮廓分割模糊。3D ESPNet虽

然对小目标区域有一定优势,但在分割子区域的重叠部分会产生较多错误分类,HDCNet和TransBTS的分割效果虽然优于其他网络,但是在处理小目标区域以及核心区域细节部分上仍有不足。DAFANet相对于其他网络,对肿瘤的小目标区域识别更加细腻,轮廓分割也更加准确,尤其是在红色ET区域的分割效果,优于其他网络。

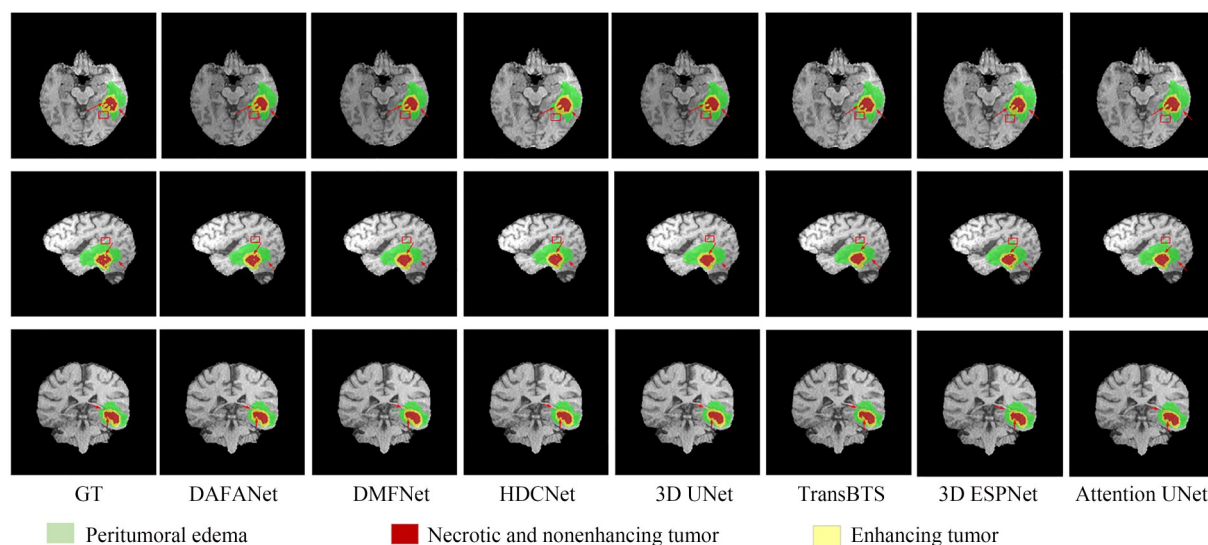


图8 分割结果的可视化对比

Fig. 8 Visual comparison of segmentation result

4 结 论

为了获取更加丰富的细节信息,降低临床操作中因边缘模糊导致的分割失误,本文通过对语义信息流概念的深入分析,设计了能够减少因多次采样造成信息损失的特征对齐引导单元,在探索特征对齐单元可行性的基础上,提出了一种轻量级融合注意力的特征对齐引导网络DAFANet。首先使用特征对齐单元缓解因网络层次加深造成的信息退化,减少了网络上采样过程中的信息损失,使得模型本身能够兼顾特征的深度获取和细节保留;其次,在跳跃连接和特征对

齐路径中融入3D期望最大化注意力机制,既能获取全局依赖关系,又能降低输入噪声。最后,利用广义Dice损失函数降低类别不均衡对分割精度的影响。实验结果显示,本文所提方法在BraTS2018和BraTS2019数据集的平均分割精度达到了85.03%,83.47%,表明DAFANet模型具有一定的优越性和较高的临床价值,能够更好地帮助医生诊断病情。而特征对齐单元作为一种即用即改的模块,目前主要在U型网络进行测试,未来会尝试将其应用到其他网络架构和其他医学图像数据集。

参考文献:

- [1] BALWANT M K. A review on convolutional neural networks for brain tumor segmentation: meth-

ods, datasets, libraries, and future directions [J]. *IRBM*, 2022, 43(6): 521-537.

- [2] LOUIS D N, PERRY A, WESSELING P, *et al.* The 2021 WHO Classification of Tumors of the

- Central Nervous System: a summary[J]. *Neuro-oncology*, 2021, 23(8): 1231-1251.
- [3] MOHAMMED Y M A, GAROUANI SEL, 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.
- [4] GU Y, CHI J Q, LIU J Q, *et al.* A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning[J]. *Computers in Biology and Medicine*, 2021, 137: 104806.
- [5] LIU Z H, TONG L, CHEN L, *et al.* Deep learning based brain tumor segmentation: a survey[J]. *Complex & Intelligent Systems*, 2023, 9(1): 1001-1026.
- [6] 谷宇. 基于深度卷积神经网络的CT影像肺结节检测技术研究[D]. 上海: 上海大学, 2019.
- GU Y. *Research on Lung Nodule Detection Based on Deep Convolutional Neural Network in Computed Tomography*[D]. Shanghai: Shanghai University, 2019. (in Chinese)
- [7] 徐胜军, 张若暄, 孟月波, 等. 融合分形几何特征 Resnet 遥感图像建筑物分割[J]. *光学精密工程*, 2022, 30(16): 2006-2020.
- XU S J, ZHANG R X, MENG Y B, *et al.* Fusion of fractal geometric features Resnet remote sensing image building segmentation [J]. *Opt. Precision Eng.*, 2022, 30(16): 2006-2020. (in Chinese)
- [8] RONNEBERGER O, FISCHER P, BROX T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*[M]. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241.
- [9] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [C]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. IEEE, 2017: 640-651.
- [10] ÇIÇEK Ö, ABDULKADIR A, LIENKAMP S S, *et al.* 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation[M]. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016. Cham: Springer International Publishing, 2016: 424-432.
- [11] IBTEHAZ N, RAHMAN M S. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. *Neural Networks*, 2020, 121: 74-87.
- [12] OKTAY O, SCHLEMPER J, LE FOLGOC L, *et al.* Attention U-Net: Learning Where to Look for the Pancreas [EB/OL]. 2018: *arXiv*: 1804.03999. <http://arxiv.org/abs/1804.03999.pdf>
- [13] WANG W, CHEN C, DING M, *et al.* Transbts: multimodal brain tumor segmentation using transformer[C]. *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021: 24th International Conference, Strasbourg, France, September 27-October 1, 2021, Proceedings, Part I* 24. Springer International Publishing, 2021: 109-119.
- [14] JIANG Y, ZHANG Y, LIN X, *et al.* SwinBTS: a method for 3D multimodal brain tumor segmentation using swin transformer [J]. *Brain Sciences*, 2022, 12(6): 797.
- [15] 杨坚华, 张浩, 花海洋. 并行路径与强注意力机制遥感图像建筑物分割[J]. *光学精密工程*, 2023, 31(2): 234-245.
- YANG J H, ZHANG H, HUA H Y. Parallel path and strong attention mechanism for building segmentation in remote sensing images [J]. *Opt. Precision Eng.*, 2023, 31(2): 234-245. (in Chinese)
- [16] 梁礼明, 周琬颂, 冯骏, 等. 基于高分辨率复合网络的皮肤病变分割[J]. *光学精密工程*, 2022, 30(16): 2021-2038.
- LIANG L M, ZHOU L S, FENG J, *et al.* Skin lesion segmentation based on high-resolution composite network[J]. *Opt. Precision Eng.*, 2022, 30(16): 2021-2038. (in Chinese)
- [17] HUANG Z, ZHAO Y W, LIU Y H, *et al.* GCAUNet: a group cross-channel attention residual UNet for slice based brain tumor segmentation [J]. *Biomedical Signal Processing and Control*, 2021, 70: 102958.
- [18] ABOELENEIN N M, PIAO S H, KOU BAA A, *et al.* HTTU-net: hybrid two track U-net for automatic brain tumor segmentation[J]. *IEEE Access*, 2020, 8: 101406-101415.
- [19] LU X Q, GU Y, YANG L D, *et al.* Multi-level 3D densenets for false-positive reduction in lung nodule detection based on chest computed tomography[J]. *Current Medical Imaging*, 2020, 16(8): 1004-1021.
- [20] CHEN C, LIU X P, DING M, *et al.* 3D Dilated

- Multi-Fiber Network for Real-Time Brain Tumor Segmentation in MRI*[M]. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019: 184-192.
- [21] LUO Z R, JIA Z D, YUAN Z M, *et al.* HDC-net: hierarchical decoupled convolution network for brain tumor segmentation [J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(3): 737-745.
- [22] ZHOU X Y, LI X Y, HU K, *et al.* ERV-Net: an efficient 3D residual neural network for brain tumor segmentation [J]. *Expert Systems with Applications*, 2021, 170: 114566.
- [23] LI X T, YOU A S, ZHU Z, *et al.* *Semantic Flow for Fast and Accurate Scene Parsing*[M]. Computer Vision - ECCV 2020. Cham: Springer International Publishing, 2020: 775-793.
- [24] LI X, ZHONG Z S, WU J L, *et al.* Expectation-maximization attention networks for semantic segmentation [C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea (South). IEEE, 2019: 9166-9175.
- [25] NUECHTERLEIN N, MEHTA S. 3D-ESPNet with pyramidal refinement for volumetric brain tumor image segmentation [C]. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, Springer International Publishing*, 2019: 245-253.
- [26] ZHANG J X, JIANG Z K, DONG J, *et al.* Attention gate ResU-net for automatic MRI brain tumor segmentation[J]. *IEEE Access*, 2020, 8: 58533-58545.
- [27] SHENG N, LIU D W, ZHANG J X, *et al.* Second-order ResU-Net for automatic MRI brain tumor segmentation [J]. *Mathematical Biosciences and Engineering: MBE*, 2021, 18(5): 4943-4960.
- [28] AKBAR A S, FATICHAH C, SUCIATIN. Single level UNet3D with multipath residual attention block for brain tumor segmentation[J]. *Journal of King Saud University-Computer and Information Sciences*, 2022, 34(6): 3247-3258.
- [29] LIU H X, HUO G Q, LI Q, *et al.* Multiscale lightweight 3D segmentation algorithm with attention mechanism: brain tumor image segmentation [J]. *Expert Systems with Applications*, 2023, 214: 119166.
- [30] CHANG Y K, ZHENG Z Z, SUN Y W, *et al.* DPAFNet: a residual dual-path attention-fusion convolutional neural network for multimodal brain tumor segmentation [J]. *Biomedical Signal Processing and Control*, 2023, 79: 104037.

作者简介:



宋建丽(1995—),女,山东枣庄人,硕士研究生,2021年于菏泽学院获得学士学位,主要从事智能图像处理、深度学习和图像分割方面研究。E-mail: 1960482133@qq.com

通讯作者:



吕晓琪(1963—),男,内蒙古包头人,博士,教授,博士生导师,1984年于内蒙古大学获得学士学位,1989年于西安交通大学获得硕士学位,2003年于北京科技大学获得博士学位,主要从事智能信息处理、医学图像处理、数字化医疗相关技术等方面的研究。E-mail: lxiaoqi@imut.edu.cn