

## 基于级联动态注意力U-Net的脑肿瘤分割方法

陈柏年, 韩雨童, 何涛, 刘斌, 张建新

### 引用本文

陈柏年, 韩雨童, 何涛, 刘斌, 张建新. [基于级联动态注意力U-Net的脑肿瘤分割方法](#)[J]. 计算机科学, 2023, 50(11A): 221100180-7.

CHEN Bonian, HAN Yutong, HE Tao, LIU Bin, ZHANG Jianxin. [Cascade Dynamic Attention U-Net Based Brain Tumor Segmentation](#) [J]. Computer Science, 2023, 50(11A): 221100180-7.

---

### 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

#### Similar articles recommended (Please use Firefox or IE to view the article)

#### [一种融合CNN和Swin Transformer的医学显微图像分割模型](#)

Medical Microscopic Image Segmentation Model Based on CNN Structure and Swin Transformer  
计算机科学, 2023, 50(11A): 230200119-8. <https://doi.org/10.11896/jsjcx.230200119>

#### [基于多编码器的多模态MRI脑肿瘤分割](#)

Multimodal MRI Brain Tumor Segmentation Based on Multi-encoder Architecture  
计算机科学, 2023, 50(6A): 220200108-6. <https://doi.org/10.11896/jsjcx.220200108>

#### [综合应用Faster R-CNN和U-net的心脏MRI图像分割](#)

Cardiac MRI Image Segmentation Based on Faster R-CNN and U-net  
计算机科学, 2023, 50(6A): 220600047-9. <https://doi.org/10.11896/jsjcx.220600047>

#### [基于深度跨模态信息融合网络的股票走势预测](#)

Deep Cross-modal Information Fusion Network for Stock Trend Prediction  
计算机科学, 2023, 50(5): 128-136. <https://doi.org/10.11896/jsjcx.220400089>

#### [融合边缘增强与多尺度注意力的皮肤病变分割](#)

Skin Lesion Segmentation Combining Boundary Enhancement and Multi-scale Attention  
计算机科学, 2023, 50(4): 96-102. <https://doi.org/10.11896/jsjcx.220300054>

# 基于级联动态注意力 U-Net 的脑肿瘤分割方法

陈柏年<sup>1,2</sup> 韩雨童<sup>1</sup> 何涛<sup>1,2</sup> 刘斌<sup>3</sup> 张建新<sup>1,2</sup>

1 大连民族大学计算机科学与工程学院 大连 116600

2 大连民族大学机器智能与生物计算研究所 大连 116600

3 大连理工大学中日国际信息与软件学院 大连 116620

(chenbn266@163.com)

**摘要** 脑肿瘤是一种严重威胁人类健康的疾病,脑肿瘤精确分割在临床诊疗中非常重要。由于脑肿瘤形状大小各异、位置不固定和边界模糊等,实现高精度脑肿瘤自动分割仍是一项具有挑战性的任务。近年来,U-Net 凭借其简洁的架构和优秀的性能成为解决医学图像分割任务的主流模型,但其也存在局部感受野有限、空间信息丢失和未充分利用上下文信息等问题。为此,提出一种基于动态卷积和非局部注意力机制的级联 U-Net 新模型(CDAU-Net)用于脑肿瘤分割任务。首先,将两阶段级联三维 U-Net 作为主体架构,来重建更精细的高分辨率脑肿瘤空间信息;进而,在级联网络横向连接上添加期望最大化注意力,通过提高网络捕获长距离依赖能力来更好利用肿瘤上下文信息;最后,在级联网络中将普通卷积替换为具有局部自适应能力的动态卷积,可进一步增强网络局部特征捕获能力。在公开的 BraTS2019-2020 数据集上进行了大量实验并与其他代表性方法进行对比,实验结果表明了所提方法在脑肿瘤分割任务上的有效性。其中,在 BraTS2019/2020 验证集上获得的全部肿瘤、肿瘤核心和增强肿瘤分割 Dice 值分别为 0.897/0.903,0.826/0.828 和 0.781/0.786,表现出了良好的脑肿瘤分割性能。

**关键词** 脑肿瘤分割;U-Net;级联网络;动态卷积;期望最大化注意力

中图法分类号 TP391

## Cascade Dynamic Attention U-Net Based Brain Tumor Segmentation

CHEN Bonian<sup>1,2</sup>, HAN Yutong<sup>1</sup>, HE Tao<sup>1,2</sup>, LIU Bin<sup>3</sup> and ZHANG Jianxin<sup>1,2</sup>

1 School of Computer Science and Engineering, Dalian Minzu University, Dalian, Liaoning 116600, China

2 Institute of Machine Intelligence and Bio-computing, Dalian Minzu University, Dalian, Liaoning 116600, China

3 International School of Information Science and Engineering, Dalian University of Technology, Dalian, Liaoning 116620, China

**Abstract** Brain tumor is a common brain disease that heavily threatens human health, so accurate brain tumor segmentation is vital for clinic diagnosis and treatment of patients. Due to different shapes and sizes, unstable positions and fuzzy boundaries of brain tumors, it is a challenging task to achieving high-precision automatic brain tumors segmentation. Recently, U-Net has become the mainstream model for medical image segmentation due to its concise architecture and excellent performance. But it also has some problems, such as limited local receptive field, spatial information loss and insufficient use of context information. Therefore, we propose a new cascade U-Net model based on dynamic convolution and non-local attention mechanism, named CDAU-Net. Firstly, a two-stage cascade 3D U-Net architecture is proposed to reconstruct more detailed and high-resolution spatial information of brain tumors. Then, the expectation-maximization attention is added to the skip connection of CDAU-Net, and the tumor context information is better utilized by improving the network's ability to capture long-distance dependencies. Finally, the normal convolution is replaced by the dynamic convolution with local adaptive ability in CDAU-Net, which can further enhance the local feature capture ability of the network. Extensive experiments are conducted on public dataset BraTS 2019/2020 and compared with other representative methods, and experimental results show that the proposed method is effective in brain tumor segmentation. The CDAU-Net obtained the Dice values of whole tumor, tumor core and enhancing tumor segmentation on BraTS 2019/2020 verification sets are 0.897/0.903, 0.826/0.828 and 0.781/0.786, respectively, which achieves good brain tumor segmentation performance.

**Keywords** Brain tumor segmentation, U-Net, Cascade network, Dynamic convolution, Expectation-maximizing attention

## 1 引言

脑肿瘤是一种生长在大脑中的异常细胞,可分为良性

肿瘤与恶性肿瘤<sup>[1]</sup>,无论是良性或恶性脑肿瘤,都会在密闭的颅骨内产生额外的压力而伤害附近脑组织,对人类健康具有较大威胁<sup>[2]</sup>。临床实践表明,早期发现脑肿瘤对治疗较为关

基金项目:国家自然科学基金(61972062);国家民委中青年英才培养计划项目;辽宁省应用基础研究计划项目

This work was supported by the National Natural Science Foundation of China(61972062), Young and Middle-aged Talents Program of National Civil Affairs Commission and Applied Basic Research Project of Liaoning Province.

通信作者:张建新(jxzhang0411@163.com)

键。磁共振成像(Magnetic Resonance Imaging, MRI)是目前最常用的一类脑肿瘤检测手段,可产生不同类型组织对比图像来为脑肿瘤诊疗提供有效信息。由于医生手动分割和分析MRI中脑肿瘤是一项艰巨而耗时的任务<sup>[3]</sup>,就产生了对计算机辅助脑肿瘤自动分割的迫切需求。然而,脑肿瘤存在着形态各异、大小不一、边界不清晰无规律等特点,导致自动化高精度分割脑肿瘤仍然是当前医疗影像分析中的难点任务。

近年来,深度学习被有效应用于脑肿瘤分割任务中并逐渐成为领域主流方法。2014年,Long等<sup>[4]</sup>将卷积神经网络(Convolutional Neural Network, CNN)中的全连接层替换为卷积层而构造出全卷积网络(Fully Convolutional Network, FCN),解决了图像分割中的空间不连续问题并有效实现图像像素级分类,成为深度学习图像语义分割的奠基性工作。针对医学图像分割任务,考虑到FCN对特征图恢复不够精细、缺乏空间一致性等问题,Ronneberger等<sup>[5]</sup>提出了FCN的变体网络U-Net,该网络由对称的编码器与解码器构成,采用跳跃连接对下采样和上采样特征进行拼接来保留更多的维度信息和空间信息,因其简洁的网络架构和优异的分割性能,其迅速成为包括脑肿瘤分割在内的一系列医学图像分割任务的主流方法。考虑到二维U-Net对三维MRI脑肿瘤图像空间上下文信息捕获的局限性,研究者提出了三维U-Net<sup>[6]</sup>并大幅提升了二维网络模型的脑肿瘤分割性能。此外,针对基于三维U-Net在小尺度和不规则脑肿瘤区域的高分辨率特征表示能力不足问题等,研究者将自编码器<sup>[7]</sup>、注意力机制<sup>[8-10]</sup>和级联架构<sup>[11-16]</sup>等引入网络中,极大促进了三维U-Net脑肿瘤分割网络模型的发展。其中,在脑肿瘤分割级联网络上,Myronenko<sup>[11]</sup>将自编码器(Variational Auto Encoder, VAE)分支级联到三维U-Net中,有效提升了脑肿瘤分割精度,该方法在2018年脑肿瘤分割挑战赛(Brain Tumor Segmentation, BraTS)上获得第一名;随后,Jiang等<sup>[12]</sup>和Buatois等<sup>[13]</sup>使用级联三维U-Net分割脑肿瘤,在第一阶段获得粗略预测分割结果,在第二阶段将粗略分割结果和三维脑肿瘤MRI图像拼接起来作为输入预测出更精细的分割结果,其中Jiang等<sup>[12]</sup>的方法获得了BraTS2019挑战赛第一名;此外,Li等<sup>[14]</sup>、Liu

等<sup>[15]</sup>和Yang等<sup>[16]</sup>同样设计了由粗略到精细的多阶段级联三维U-Net方法并充分利用多尺度信息。以上方法有效表明了级联网络在脑肿瘤分割中的优越性能,也使其成为该领域研究的一个重要分支。

本文借鉴三维级联网络在脑肿瘤分割任务上的突出效果,考虑U-Net网络中存在局部感受野有限、空间信息丢失和未能充分利用上下文信息等问题,着重在三维级联网络基础上提升网络全局特征关注表达和自适应调整感受野能力。首先,从加强脑肿瘤全局上下文信息与扩大感受野着手,将一种轻量级的非局部注意力期望最大化(Expectation-maximization Attention, EMA)模块<sup>[17]</sup>引入级联网络中;此外,针对使用普通卷积难以有效捕获不同形状大小的肿瘤特征问题,尝试在网络中引入基于注意力的动态卷积<sup>[18]</sup>(Dynamic Convolution),由输入动态决定卷积核来更好地适应脑肿瘤不同图像差异巨大的情况,可在不加深或加宽网络的情况下进一步提升网络性能。综上,本文构建了一种融合EMA和动态卷积的三维级联U-Net脑肿瘤分割新模型——级联动态注意力U-Net(Cascade Dynamic Attention U-Net, CDAU-Net),主要贡献可概况如下:(1)以三维级联脑肿瘤分割网络架构为基础,从增强网络的肿瘤全局上下文信息表达能力以及对不同尺度肿瘤的感受野自适应调整角度着手,提出了一种新的三维级联脑肿瘤分割网络。(2)CDAU-Net在第二阶段网络第四层横向连接中引入一种轻量级非局部注意力期望最大化注意力模块以提高网络捕获长距离依赖特征的能力,同时将第二阶段网络中所有普通卷积替换为动态卷积来自适应匹配局部特征感受野,实现在整体上提升网络的脑肿瘤分割性能。(3)在BraTS 2019和BraTS 2020数据集上对CDAU-Net方法进行了验证,消融实验结果和对比实验结果均证明了所提方法的有效性,其能够获得与领域先进方法具有一定竞争力的性能。

## 2 网络框架及算法原理

### 2.1 完整网络框架

本文提出的三维级联动态注意力U-Net脑肿瘤分割网络架构如图1所示。

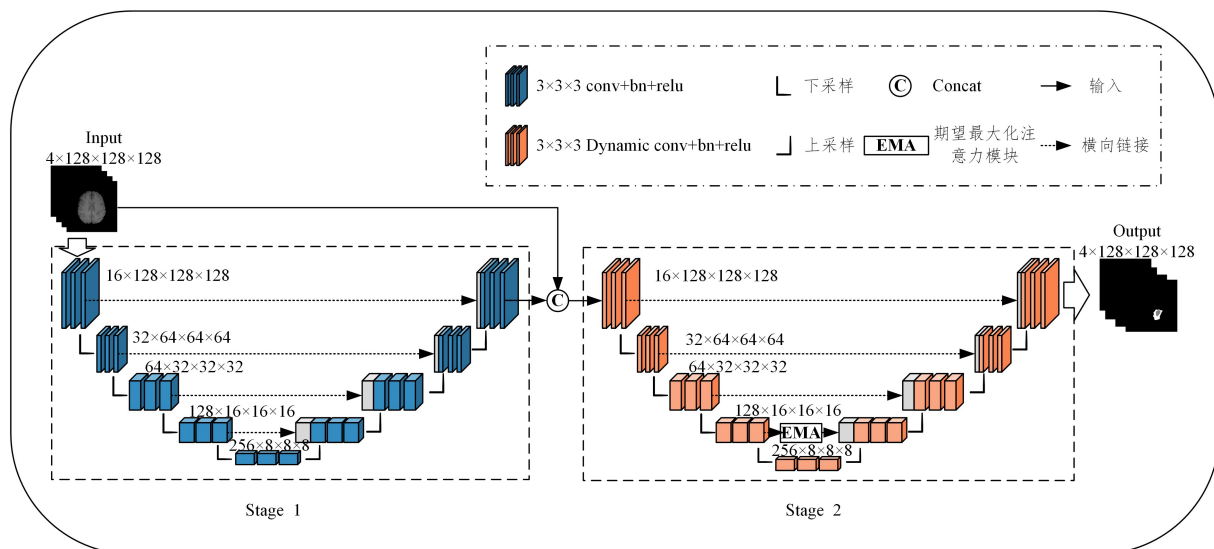


图1 基于级联动态注意力U-Net的脑肿瘤分割网络架构

Fig. 1 Architecture of cascade dynamic attention-based U-Net for brain tumor segmentation

为重建更精细的高分辨率空间信息,在网络主体上采用级联架构,将两个具有 4 层上下采样层的三维 U-Net 级联形成两阶段网络,来实现对脑肿瘤区域的从粗糙到精细的分割结果。其中,第一阶段使用三维 U-Net 对脑肿瘤进行粗略分割,网络输入的多模态脑肿瘤尺寸为  $4 \times 128 \times 128 \times 128$ ,输出尺寸同样为  $4 \times 128 \times 128 \times 128$ ;第二阶段网络输入为粗略分割结果与多模态三维脑肿瘤 MRI 的聚合结果,其尺寸为  $8 \times 128 \times 128 \times 128$ ,通过学习聚合信息使粗略分割结果指导网络从三维脑肿瘤 MRI 学习更精细的分割结果,最后的输出尺寸仍为  $4 \times 128 \times 128 \times 128$ 。由于三维 U-Net 在横向连接上使用的是常规的直接连接方式,在描述脑肿瘤图像的长距离特征依赖关系上存在局限性,为此引入轻量级非局部注意力期望最大化注意力 EMA 模块,考虑到计算代价只在第二阶段中第四层横向链接位置添加 EMA 模块,该模块输入特征尺寸为  $128 \times 16 \times 16 \times 16$ ,在增加全局信息描述的同时也可扩大网络感受野。此外,将第二阶段三维 U-Net 中编解码模块中的普通卷积替换为动态卷积,可根据肿瘤尺寸自适应调整网络感受野,从而加强肿瘤图像的局部特征提取能力。EMA 模块和动态卷积的具体情况将在下文中进行介绍。此外,为能更好地进行网络训练,将第一阶段的输出和第二阶段的输出分别与真值图像计算损失值,然后合并为最终损失值共同监督网络。

### 2.1.1 期望最大化注意力

全局信息对于脑肿瘤分割较为重要,引入非局部自注意力可以增强肿瘤图像长距离特征依赖,但常规非局部自注意力需要逐点计算空间注意力和生成注意力图,在与三维卷积结合时会导致模型计算代价急剧增加。为此,使用轻量级非局部的期望最大化注意力 EMA 模块<sup>[17]</sup>,其是通过一组重建特征来计算非局部注意力,由于重建基的体素比输入特征图的体素少得多,可以大幅降低非局部注意力的计算代价。EMA 模块具体结构如图 2 所示,其使用  $1 \times 1 \times 1$  的卷积改变特征图  $X$  通道数,并将其输入到虚线框中的 EM 算法以计算非局部注意力,为了避免过拟合,将输入特征图以残差的方式与重构特征图求和。

在模块具体优化计算上,EMA 模块使用期望最大化(Expectation-maximization, EM)算法迭代出一组紧凑的基,进而在这组基上计算注意力。给定输入特征图  $X \in R^{C \times D \times H \times W}$  和初始基  $\mu \in R^{C \times K}$ ,其中  $C$  是通道数, $D \times H \times W$  是输入特征图大小, $K$  是基的个数,基  $\mu$  和非局部自注意力  $Z \in R^{D \times H \times W \times K}$  分别学习参数和潜在变量。EM 算法的目标是寻找最优的可学习参数使数据似然最大化。EM 实现时主要分期望(Expectation, E)步和最大化(Maximization, M)步, E 步估计  $Z$  的期望, M 步通过最大化似然更新,具体过程为:

$$Z^{(t)} = \text{softmax}(X^T (\mu^{(t-1)})) \quad (1)$$

$$\mu_k^{(t)} = \frac{\sum_{n=1}^V (Z_{nk}^{(t)} X_n)}{\sum_{n=1}^V Z_{nk}^{(t)}} \quad (2)$$

其中,  $t$  为迭代次数,  $V = D \times W \times H$ ,  $Z_{nk}$  为位置  $n$  处第  $k$  个通道的注意力向量。文中  $t$  设置为 3,即 E 步和 M 步交替执行 3 次。最后,通过收敛的  $\mu$  和归一化注意力  $Z$  重构特征映射  $\tilde{X}$ ,表示如下:

$$\tilde{X} = \mu Z^T \quad (3)$$

其中,  $\tilde{X}$  为重构后的特征映射。在网络中嵌入 EMA 模块可有效聚合长距离上下文信息,并实现在保持类间差异的同时缩小类内部差异,从而提升级联脑肿瘤分割网络的分割性能。

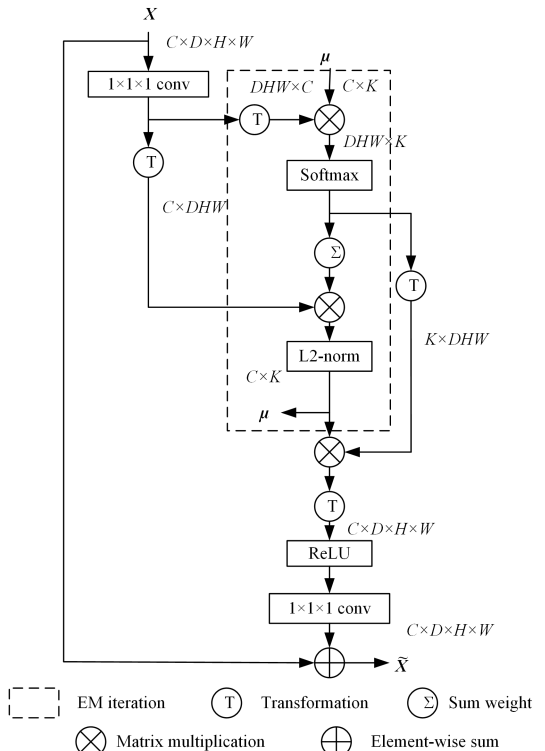


图 2 期望最大化注意力模块结构图

Fig. 2 Structure of EMA module

### 2.1.2 动态卷积

深度神经网络使用普通卷积时,网络训练完成后所有卷积核参数会固定并且卷积会共享卷积核参数,一般需要通过加宽或加深网络来学习更多的参数以产生更好的性能。为在不增加网络深度或宽度的情况下提升模型性能,Chen 等<sup>[18]</sup>提出了动态卷积的概念,使卷积层能够根据输入图像的特性来自适应调整感受野能力。

在具体实现上,动态卷积使用一组  $K$  个并行卷积核而不是每层使用一个卷积核,其卷积核根据输入图像的关注度动态聚合多个并行卷积核,并以非线性方式聚合来获得更强的特征表示能力。同时,动态卷积使用压缩和激励模块 SENet<sup>[19]</sup> 计算卷积核注意力,通过使用全局平均池化压缩输入的全局空间信息,然后使用两个全连接层和一个中间的非线性层结构对特征进行降维。其中,对第一个全连接层进行降维,第二个全连接层将维度减少至  $K$ ,最后经过 softmax 为  $K$  个卷积核生成归一化的注意力权重  $\pi_k$ 。为促进注意力  $\pi_k$  的学习,  $\pi_k$  被约束为  $\sum_k \pi_k(x) = 1$ 。与 SENet 不同的是,动态卷积将注意力加在并行的卷积上,因此增加的注意力计算成本很低。动态卷积的具体结构如图 3 所示。

由于 MRI 脑肿瘤图像是三维图像,本文将原始二维动态卷积扩展到三维,进一步将级联网络中第二阶段中所有普通卷积替换为三维动态卷积,使网络对不同输入脑肿瘤图像自适应建立卷积核参数,从而获取更好的脑肿瘤分割效果。

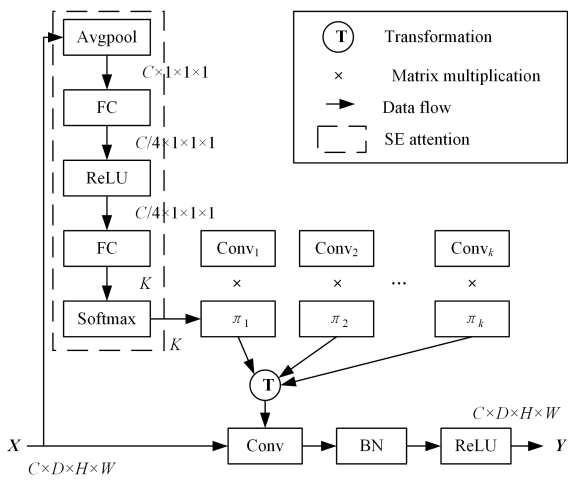


图3 动态卷积结构图

Fig. 3 Structure of dynamic convolution

## 2.2 损失函数

损失函数是指导脑肿瘤分割网络训练的重要因素。考虑到脑肿瘤分割任务中存在着严重的类不平衡问题,在损失函数上结合采用交叉熵损失函数 $L_{CE}$ 和Dice损失函数 $L_{DC}$ 。其中,交叉熵损失减少了预测结果与真值图像的误差,Dice损失函数是医疗影像分割中常用的损失函数,很好地解决了类不平衡问题,其计算公式分别为:

$$L_{CE} = - \sum_{i=0}^N y_i \log \hat{y}_i \quad (4)$$

$$L_{DC} = 1 - \frac{2 \sum_{i=0}^N y_i \hat{y}_i}{\sum_{i=0}^N (y_i + \hat{y}_i)} \quad (5)$$

其中, $y$ 表示真值图像, $\hat{y}$ 表示预测的分割结果, $N$ 为样本标签总数。

本文方法在使用损失函数时分为两阶段进行。其中,网络第一阶段的输出与真值图像计算使用交叉熵损失函数,网络第二阶段的输出与真值图像使用交叉熵损失函数和Dice损失函数的结合。此后,再将两阶段损失函数按照比例相加,完整的损失函数 $Loss$ 可表示为:

$$Loss = 0.5 \times L_{CE1} + 0.5 \times (L_{CE2} + L_{DC2}) \quad (6)$$

其中, $L_{CE1}$ 和 $L_{CE2}$ 分别表示脑肿瘤分割网络第一阶段和第二阶段交叉熵损失函数, $L_{DC2}$ 表示网络第二阶段Dice损失函数。

## 3 实验和结果

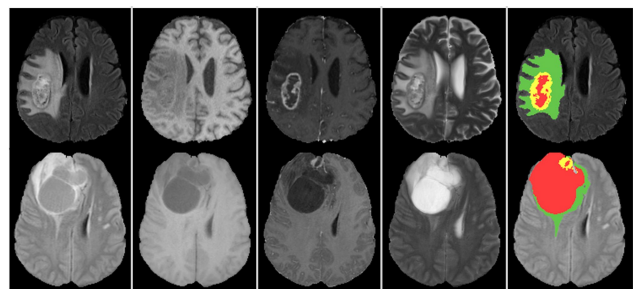
### 3.1 实验环境与配置

实验硬件环境为Intel(R) Xeon(R) Gold 5222 CPU @ 3.80 GHZ和1个Nvidia RTX3090 GPU显卡(24 GB),在代码实施上采用Pytorch深度学习框架,模型训练使用Adam优化器,初始学习率为0.001,动量为0.95,权重衰减系数为 $1 \times 10^{-5}$ 。鉴于三维脑肿瘤分割模型需要消耗较大的内存,受硬件配置资源限制,训练中batch-size大小设置为3,模型训练迭代次数设置为550。

### 3.2 数据集与预处理

使用公开的MRI脑肿瘤分割数据集BraTS 2019和BraTS 2020<sup>[20]</sup>来验证本文方法的有效性。BraTS 2019数据集包含训练集335个病例和验证集125个病例,其中训练集

中高级胶质瘤(High-grade Glioma, HGG)有259个病例进行了,低级胶质瘤(Low-grade Glioma, LGG)有76个病例; BraTS 2020数据集在BraTS 2019数据集的基础上对HGG病例扩充,包含训练集369个病例和验证集125个病例,其中训练集中高级胶质瘤HGG有293个病例,低级胶质瘤LGG有76个病例。每个病例包含Flair, T1, T2和T1ce这4种模态,每张MR图像大小为 $240\text{mm} \times 240\text{mm} \times 155\text{mm}$ ,同时训练集还提供由专业医师手动分割的脑肿瘤结果。为确保脑肿瘤分割结果的公平性,验证集不公开分割真值图像(Ground Truth)标签,需要将分割结果提交至在线评估平台来评测模型分割性能。脑肿瘤影像数据标签包含4类,分别为健康部位(标签0)、坏死区域(标签1)、水肿区域(标签2)和增强肿瘤(标签4),可形成的肿瘤病灶区域包括全部肿瘤(Whole Tumor, WT)、肿瘤核心(Tumor Core, TC)和增强肿瘤(Enhancing Tumor, ET)这3类。每类肿瘤病灶区域包含一至多个标签,具体来说:WT包含标签1, 2, 4, TC包含标签1, 4, ET包含标签4。图4给出了2个代表性的MRI脑肿瘤图像样本示例,其中上图为HGG样本,下图为LGG样本,另外图中红色代表坏死区域,绿色代表水肿区域,黄色代表增强肿瘤。



(a) Flair (b) T1 (c) T1ce (d) T2 (e) Ground Truth

图4 4种模态与真值图像的脑肿瘤MRI

Fig. 4 MRI images in different modes and ground truth

在脑肿瘤图像数据的预处理上,首先使用Z-score方法对数据集标准化,即

$$z' = \frac{z - \mu}{\delta} \quad (7)$$

其中, $z$ 是输入图像, $z'$ 是归一化图像, $\mu$ 是输入图像平均值, $\delta$ 是输入图像标准差。由于脑肿瘤图像中存在大量无用的背景信息,其后将多模态三维脑肿瘤MRI图像随机裁剪为 $4 \times 128 \times 128 \times 128$ 的尺寸作为网络输入;此外,使用随机旋转、翻转和强度变换等策略对图像数据进行增强。

### 3.3 评价指标

采用医疗影像分割中常用的Dice相似系数(Dice Similarity Coefficient, DSC)和Hausdorff距离作为模型评价指标。Dice相似系数用来评估分割结果与真值图像的相似度,取值范围在0~1之间,具体为:

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

其中, $TP$ ,  $FP$ ,  $TN$ 和 $FN$ 分别表示被正确预测为脑肿瘤的体素个数、被预测为肿瘤的非肿瘤体素个数、被预测为非肿瘤的体素个数和未被检测到的肿瘤体素个数。

Hausdorff指标用来评估分割结果边界和真值图像边界之间的距离,表示为:

$$HD(T, P) = \max\left(\sup_{t \in T} \inf_{p \in P} d(t, p), \sup_{p \in P} \inf_{t \in T} d(t, p)\right) \quad (9)$$

其中,  $t$  和  $p$  分别表示在真值图像区域  $T$  和预测区域  $P$  上的点,  $d(t, p)$  是计算点  $t$  和点  $p$  之间的距离函数。为了消除离群集中非常小的子集的影响, 本文将 Hausdorff 最终结果值乘以 95%, 即 Hausdorff95。

### 3.4 实验和结果

为验证所提出的 CDAU-Net 脑肿瘤分割模型的有效性, 首先在 BraTS 2020 数据集上实施消融实验, 再与其他具有代表性的脑肿瘤分割网络进行结果比较; 同时, 为验证 CDAU-Net 的泛化性能, 还给出了在 BraTS 2019 数据集上与其他模型的对比实验结果; 最后, 给出了比较的可视化分割结果来进一步说明所提模型的效果。

#### 3.4.1 消融实验结果

为充分验证模型的消融结果, 消融实验在 BraTS 2020 训练集和验证集上实施, 使用三维 U-Net 作为基线方法, 在此基础上将三维 U-Net 修改为三维级联 U-Net, 进而在第二阶段分别添加 EMA 和动态卷积模块进行结果验证。为方便比较, 用 CU-Net 表示三维级联 U-Net 模型, 用 +EMA 和 +DConv 分别表示添加期望最大化注意力和动态卷积结果。

在 BraTS 2020 训练集消融实验上, 将该数据集中脑肿瘤图像按 8:2 的比例随机划分成新的训练集和验证集。同时, 综合训练时间与结果稳定性两方面的因素, 每个模型使用两次训练结果的平均值, 具体消融实验结果如表 1 所列。

表 1 BraTS 2020 训练集的消融实验

Table 1 Ablation experiments on BraTS 2020 training dataset

分割方法	DSC			
	ET	WT	TC	平均值
U-Net(基线方法)	0.784	0.900	0.797	0.827
CU-Net	0.788	0.897	0.816	0.834
CU-Net +EMA	0.803	0.898	0.815	0.839
CU-Net +DConv	0.797	0.905	0.840	0.847
CDAU-Net	0.813	0.905	0.836	0.851

由表 1 中结果可见, 基线模型三维 U-Net 在增强肿瘤 ET、全部肿瘤 WT 和肿瘤核心 TC 上获得的 DSC 值分别是 0.784, 0.900 和 0.797, 当采用基础级联模型 CU-Net 时相应值变为 0.788, 0.897 和 0.816, 尽管较基线模型在 WT 结果上略微有所下降, 但在 ET 和 TC 结果上均有所提升, 特别是在 TC 上取得了 1.9% 的精度提升, 验证了级联架构的有效性。进一步, 当在级联模型中单独加入 EMA 模块或动态卷积模块时, 相比基础级联方法整体性能均有所提升。特别地, 加入 EMA 模块时 ET 值可提升 1.5%, 加入动态卷积时 ET 和 TC 分别可提升 0.9% 和 2.4%, 有效表明了这两个模块在小尺度脑肿瘤分割上的效果。最后, 当把两个模块同时加入级联架构上形成 CDAU-Net 后, 在 ET, WT 和 TC 上分别可获得 DSC 值为 0.813, 0.905 和 0.836, 较级联基础模型 CU-Net 分别获得 2.5%, 0.5% 和 2% 的性能提升, 较基线方法 U-Net 更是分别有 2.9%, 0.5% 和 3.9% 的性能提升, 表明所提出 CDAU-Net 模型在脑肿瘤分割上的良好效果, 其原因在于加入模块可以更好地捕获肿瘤全局上下文信息和自适应调整对肿瘤区域的感受程度。

为进一步保证消融实验结果的合理性, 另外还在 BraTS 2020 验证集上进行消融实验, 使用 BraTS 2020 训练集全部数据训练模型, 然后使用训练好的模型对 BraTS 2020 验证集中 125 例脑肿瘤实施分割, 最后将分割结果提交至 BraTS 在线评估平台, 由平台返回最终指标评测结果, 具体实验结果如表 2 所列。

表 2 BraTS 2020 验证集上的消融实验

Table 2 Ablation experiments on BraTS 2020 validation dataset

分割方法	DSC			
	ET	WT	TC	平均值
U-Net(基线方法)	0.769	0.894	0.799	0.821
CU-Net	0.773	0.897	0.808	0.826
CU-Net +EMA	0.787	0.899	0.815	0.834
CU-Net +DConv	0.778	0.895	0.833	0.835
CDAU-Net	0.786	0.903	0.828	0.839

从表 2 可知, 基线方法三维 U-Net 在 ET, WT 和 TC 上的 DSC 结果为 0.769, 0.894 和 0.799, 级联网络后在 3 个指标上分别可提升 0.4%, 0.3%, 0.9%。进一步, 在级联网络之上分别添加 EMA 模块和动态卷积时, 与三维 U-Net 相比在 WT 指标上差距很小, 但在 ET 结果上分别提升 1.8% 和 0.9%, 在 TC 结果上分别提升 1.6% 和 3.4%, 证明了所引入的两个模块在脑肿瘤分割上的效果。最后, 当同时引入级联、EMA 模块和动态卷积形成 CDAU-Net 后, 在 ET, WT 和 TC 分别可获得 0.786, 0.903 和 0.828 性能结果, 与三维 U-Net 相比分别提高 1.7%, 0.9% 和 2.9%, 进一步证明了级联架构与两个模块结合的有效性。另外, 从表中数据可以发现, 虽然最终结果并未达到每个指标的最优, 但是综合 3 个指标的平均结果性能确是最优的。

#### 3.4.2 与代表性方法对比结果

为进一步验证所提出方法的有效性和竞争力, 在 BraTS 2019 和 BraTS 2020 验证集上与其他代表性脑肿瘤分割方法进行对比, 结果如表 3 和表 4 所列。

由表 3 可见, CDAU-Net 在 BraTS 2019 验证集上 DSC 可获得的 ET, WT 和 TC 结果分别为 0.781, 0.897 和 0.826, 三者平均值为 0.834。在进行对比的方法中, Zhang 等<sup>[21]</sup> 使用三维 U-Net++ 网络并改进损失函数, Xia 等<sup>[22]</sup> 将跨阶段局部网络结构引入三维 U-Net 中构建出新的于脑肿瘤分割模型, 二者分别获得平均 DSC 值 0.799 和 0.788; Liu 等<sup>[8]</sup> 提出一种上下文感知网络(Context Aware Network, CANet), 该方法从特征图中捕获具有上下文的高维特征并提出上下文引导的注意条件随机场。与 Zhang 等和 Xia 等的方法相比, CDAU-Net 具有较为明显的优势, 与 Liu 等方法相比, 在 ET, WT 和 TC 上分别提升 2.2%, 1.2% 和降低 2.5%, 整体上仍优于 CANet 模型。与 Buatois 等<sup>[13]</sup> 和 Li 等<sup>[14]</sup> 提出的由粗到细的多阶段级联网络相比, 本文方法在平均值上分别提升 3.3% 和 1.1%, 一定程度证明了其有效性。此外, 还与 Wang 等<sup>[23]</sup> 提出的 TransBTS 方法进行了对比, 该方法将视觉中近期广受关注的 Transformer 引入三维脑肿瘤分割中, 从结果上看二者在平均值上比较接近。从 Hausdorff 距离上看, CDAU-Net 方法在 ET, WT 和 TC 上分别为 3.39, 5.91 和 7.77, 平均 Hausdorff 距离为 5.69, 在表中排名第三。整体来说, 结合 DSC 和 Hausdorff 距离的结果, 本文方法与其他代表性方法相比具有良好的竞争性。

表3 BraTS 2019 验证集上与代表性方法对比结果

Table 3 Comparison results with representative methods on BraTS 2019 validation set

分割方法	DSC				Hausdorff95			
	ET	WT	TC	平均值	ET	WT	TC	平均值
Zhang 等 <sup>[21]</sup>	0.709	0.873	0.814	0.799	12.30	15.45	12.47	13.40
Xia 等 <sup>[22]</sup>	0.707	0.878	0.779	0.788	—	—	—	—
Liu 等 <sup>[8]</sup>	0.759	0.885	0.851	0.832	4.80	5.89	6.56	5.75
Buatois 等 <sup>[13]</sup>	0.750	0.854	0.800	0.801	3.08	7.04	5.96	5.36
Li 等 <sup>[14]</sup>	0.771	0.886	0.813	0.823	6.03	6.23	7.41	6.55
Wang 等 <sup>[23]</sup>	0.789	0.900	0.819	0.836	3.73	5.64	6.04	5.14
CDAU-Net(本文方法)	0.781	0.897	0.826	0.834	3.39	5.91	7.77	5.69

在表4中的BraTS 2020验证集消融实验结果上, Akbar等<sup>[9]</sup>在三维U-Net中使用空洞卷积并在跳跃连接中添加注意力的改进三维U-Net, 所提出方法较之有明显的性能优势。Zhang等<sup>[10]</sup>提出将EMA模块添加到DMF-Net模型横向链接中, 与之相比, CDAU-Net在ET上领先其1.6%, 但在TC上落后其1.1%, 在平均值上有0.4%的性能提升。在级联网络对比上, Liu等<sup>[15]</sup>提出一种级联三个子分支和一个主分支的分割网络, 其中子分支捕获不同脑肿瘤特征, 主分支使用空间通道融合块聚合多模态特征; Lyu等<sup>[7]</sup>使用两阶段VAE级联网络并在网络中添加注意力门。相比Liu等和Lyu等的级

联脑肿瘤分割网络, CDAU-Net方法在平均值上领先Liu等的方法2.3%, 领先Lyu等方法0.1%。此外, CDAU-Net方法与Jiang等<sup>[24]</sup>在三维U-Net中添加Swin Transformer的SwinBTS方法对比, 在ETWT和TC上分别领先1.2%, 1.2%和2.5%。以上结果有效表明了所提方法在脑肿瘤分割上的效果。从Hausdorff距离结果上看, CDAU-Net方法在WT上可达到4.57, 但是在ET和TC上略有不足, 原因在于CDAU-Net未使用更多后处理过程, 会对个别不含有ET的LGG做出不准确预测, 这种假阳性导致Hausdorff距离变为默认上限最大值, 显著影响了模型评估结果。

表4 BraTS 2020 验证集上与代表性方法对比结果

Table 4 Comparison results with representative methods on BraTS 2020 validation set

分割方法	DSC				Hausdorff95			
	ET	WT	TC	平均值	ET	WT	TC	平均值
Akbar 等 <sup>[9]</sup>	0.729	0.886	0.802	0.806	31.97	10.26	13.58	18.60
Zhang 等 <sup>[10]</sup>	0.770	0.896	0.839	0.835	32.40	7.70	11.70	17.26
Liu 等 <sup>[15]</sup>	0.764	0.882	0.801	0.816	21.39	6.49	6.68	11.52
Lyu 等 <sup>[7]</sup>	0.790	0.897	0.829	0.838	24.14	6.17	7.04	12.45
Jiang 等 <sup>[24]</sup>	0.774	0.891	0.803	0.823	26.84	15.78	8.56	17.06
CDAU-Net(本文方法)	0.786	0.903	0.828	0.839	35.03	4.57	14.58	18.06

### 3.4.3 结果可视化

为更加直观地显示脑肿瘤分割效果, 还对BraTS 2020训练集中的数据结果进行可视化, 具体如图5所示。图5中选取3个代表性病例进行展示, 每一行从左到右分别演示了Flair、真值图像(Ground Truth)、三维U-Net和CDAU-Net的分割结果, 其中真值图像、三维U-Net和CDAU-Net的

分割结果覆盖在Flair图像上。此外, 为了更直观地展示脑肿瘤分割结果, 并将肿瘤区域放大, 同时使用蓝色线条标出CDAU-Net优于U-Net的分割结果位置。通过可视化图像能看出, 所提方法可以很好地分割增强肿瘤、全部肿瘤和肿瘤核心, 结果优于三维U-Net, 但与真值图像相比还有一定差距, 分割性能还有进一步提升的空间。

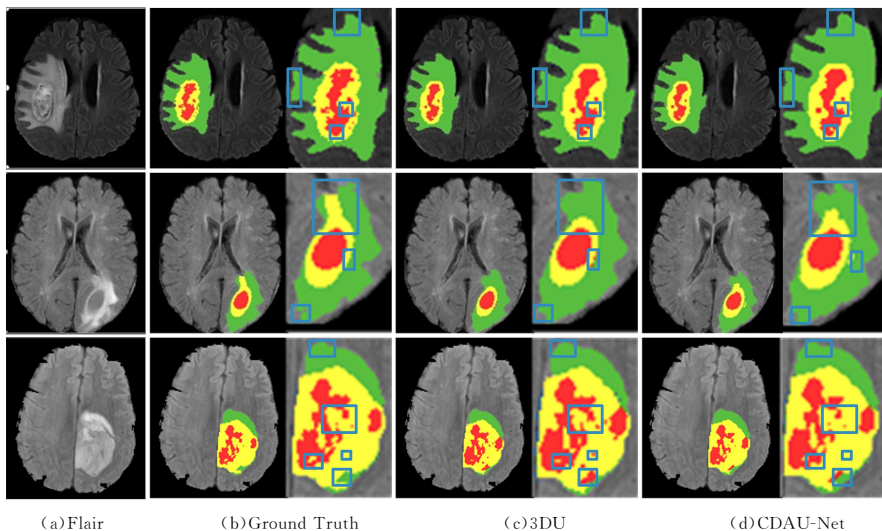


图5 BraTS 2020 训练集分割结果示例(电子版为彩图)

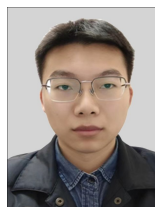
Fig. 5 Example segmentation results on BraTS 2020 training dataset

**结束语** 本文提出了一种新的三维级联动态注意力U-Net脑肿瘤分割网络DAU-Net, 其以级联三维U-Net架构为

基础,将期望最大化注意力模块和动态卷积同时引入,有效提高了网络对肿瘤全局信息和局部信息的捕获能力。在公开的 BraTS 2019 和 BraTS 2020 数据集上的消融实验结果表明了所引入模块在脑肿瘤分割上的有效性,与领域代表性方法的比较结果进一步证明了所提方法的竞争力。在今后的工作中,除在级联架构下探索更高特征表达能力的卷积和自注意力模块外,还将考虑引入更先进的数据增强策略和后处理方法,从而进一步提高脑肿瘤分割的准确度。

## 参 考 文 献

- [1] MENZE B H, JAKAB A, BAUER S, et al. The multimodal brain tumor image segmentation benchmark (BRATS) [J]. IEEE Transactions on Medical Imaging, 2014, 34(10):1993-2024.
- [2] BAUER S, WIEST R, NOLTE L P, et al. A survey of MRI-based medical image analysis for brain tumor studies[J]. Physics in Medicine & Biology, 2013, 58(13):R97.
- [3] PEREIRA S, PINTO A, ALVES V, et al. Brain tumor segmentation using convolutional neural networks in MRI images[J]. IEEE Transactions on Medical Imaging, 2016, 35(5):1240-1251.
- [4] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015:3431-3440.
- [5] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015:234-241.
- [6] ÇICEK Ö, ABDULKADIR A, LIENKAMP S S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]// International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer, 2016:424-432.
- [7] LYU C, SHU H. A two-stage cascade model with variational autoencoders and attention gates for MRI brain tumor segmentation[C]// International MICCAI Brainlesion Workshop. Cham: Springer, 2020:435-447.
- [8] LIU Z, TONG L, CHEN L, et al. CANet: Context aware network for brain glioma segmentation[J]. IEEE Transactions on Medical Imaging, 2021, 40(7):1763-1777.
- [9] AKBAR A S, FATICHAH C, SUCIATI N. Single level UNet3D with multipath residual attention block for brain tumor segmentation[J]. Journal of King Saud University—Computer and Information Sciences, 2022.
- [10] ZHANG J, JIANG Z, LIU D, et al. 3D asymmetric expectation-maximization attention network for brain tumor segmentation [J]. NMR in Biomedicine, 2022, 35(5):e4657.
- [11] MYRONENKO A. 3D MRI brain tumor segmentation using autoencoder regularization[C]// International MICCAI Brainlesion Workshop. Cham: Springer, 2018:311-320.
- [12] JIANG Z, DING C, LIU M, et al. Two-stage cascaded u-net: 1st place solution to brats challenge 2019 segmentation task[C]// International MICCAI Brainlesion Workshop. Cham: Springer, 2019:231-241.
- [13] BUATOIS T, PUYBAREAU É, TOCHON G, et al. Two stages CNN-based segmentation of gliomas, uncertainty quantification and prediction of overall patient survival [C]// International MICCAI Brainlesion Workshop. Cham: Springer, 2019:167-178.
- [14] LI X, LUO G, WANG K. Multi-step cascaded networks for brain tumor segmentation[C]// International MICCAI Brainlesion Workshop. Cham: Springer, 2019:163-173.
- [15] LIU C, DING W, LI L, et al. Brain tumor segmentation network using attention-based fusion and spatial relationship constraint [C]// International MICCAI Brainlesion Workshop. Cham: Springer, 2020:219-229.
- [16] YANG S, GUO D, WANG L, et al. Cascaded Coarse-to-Fine Neural Network for Brain Tumor Segmentation[C]// International MICCAI Brainlesion Workshop. Cham, Springer, 2020:458-469.
- [17] LI X, ZHONG Z, WU J, et al. Expectation-maximization attention networks for semantic segmentation[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019:9167-9176.
- [18] CHEN Y, DAI X, LIU M, et al. Dynamic convolution: Attention over convolution kernels[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020:11030-11039.
- [19] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:7132-7141.
- [20] BAKAS S, REYES M, JAKAB A, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge[J]. arXiv:1811.02629, 2018.
- [21] ZHANG X Y, WANG B, AN W C, et al. Glioma Segmentation Network Based on 3DU-Net++ with Fusion Loss Function [J]. Computer Science, 2021, 48(9):187-193.
- [22] XIA F, SHAO H, DENG X. Cross-stage deep-learning-based MRI fused images of human brain tumor segmentation[J]. Journal of Image and Graphics, 2022, 27(3):873-884.
- [23] WANG W, CHEN C, DING M, et al. Transbts: Multimodal brain tumor segmentation using transformer[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2021:109-119.
- [24] 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.



**CHEN Bonian**, born in 1997, postgraduate, is a member of China Computer Federation. His main research interests include computer vision and medical image analysis.



**ZHANG Jianxin**, born in 1981, Ph. D., professor, master supervisor, is a senior member of China Computer Federation. His main research interests include computer vision and intelligent medical data processing.