

DOI:10.16136/j.joel.2023.03.0203

基于双解码路径 DD-UNet 的脑肿瘤图像分割算法

苏 赋^{1*}, 方 东¹, 王龙业¹, 高思萱², 张凯信¹, 敬 梁¹

(1. 西南石油大学 电气信息学院, 四川 成都 610500; 2. 西南石油大学 经济管理学院, 四川 成都 610500)

摘要:针对医学图像中病灶区域尺度不一、边界模糊和周围组织强度不均匀所导致的分割精度降低问题,提出了一种基于双解码器的脑肿瘤图像分割模型。为了增强特征的表征力,提出了高阶微分残差模块并使用不同空洞率的扩张卷积用于提取特征编码,提高了网络模型的分割性能;引入上下文语义信息感知模块(multi scale dilation, MSD),从不同的目标尺度中提取更多的精细信息,提高了对结构细节信息的捕获能力,同时减少了编解码器之间的特征差异;在空间解码路径中使用选择性聚合空间注意力模块(spatial aggregation attention module, SAAM),增加了对有效空间特征的权重比例,减少了无效的特征干扰。在脑肿瘤数据集上进行了实验验证,实验结果表明,所提算法的 Dice 系数、平均交并比、敏感性、特异性、准确率等指标分别为:93.35%、90.71%、91.15%、99.94%、96.75%。

关键词: MRI 脑肿瘤分割; 双路径 U-Net; 高阶残差; 注意力机制; 上下文语义感知

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1005-0086(2023)03-0328-09

Brain tumor image segmentation algorithm based on dual decoding path DD-UNet

SU Fu^{1*}, FANG Dong¹, WANG Longye¹, GAO Sixuan², ZHANG Kaixing¹, JING Liang¹

(1. School of Electrical Engineering and Information, Southwest Petroleum University, Chengdu, Sichuan 610500, China; 2. School of Economics and Management, Southwest Petroleum University, Chengdu, Sichuan 610500, China)

Abstract: In view of the reduction of segmentation accuracy caused by different scales of focus area, fuzzy boundary and uneven intensity of surrounding tissues in medical images, a brain tumor image segmentation model based on double decoder is proposed. In order to enhance the representation of features, a high-order differential residual module is proposed, and the expanded convolution with different void rates is used to extract the feature coding, which improves the segmentation performance of the network model; The context semantic information perception module is introduced to extract more fine information from different target scales, which improves the ability to capture structural details and reduces the feature differences between codecs. The selective spatial aggregation attention module (SAAM) is used in the spatial decoding path to increase the weight proportion of effective spatial features and reduce the interference of invalid features. Experimental verification on different brain tumor data sets is completed. The experimental results show that the Dice coefficient, average intersection union ratio, sensitivity, specificity and accuracy of the proposed algorithm are 93.35%, 90.71%, 91.15%, 99.94% and 96.75%, respectively.

Key words: MRI brain tumor segmentation; dual path U-Net; higher order residuals; attention mechanism; context semantic awareness

* E-mail: 2458048563@qq.com

收稿日期: 2022-03-27 修订日期: 2022-05-10

基金项目: 成都市国际科技合作项目(2020-GH02-00016-HZ)资助项目

0 引言

大脑是中枢神经系统最大最精密的器官。它是人们的中枢,控制着人的一切生命活动。脑肿瘤是指脑内的一些异常细胞的生长以及堆积,它是不可控的,而脑肿瘤具有高发病率和致死率,是世界上最致命的癌症之一,严重影响了人们生命安全。现在医疗机构普遍采用世界卫生组织^[1]提出的分类标准,按照脑肿瘤在其病理学上恶化程度可划分为低级别胶质瘤(low grade glioma, LGG)和高级别胶质瘤(high grade glioma, HGG),其中 HGG 通常细胞分化不良、扩散迅速、难以根除、预后较差, LGG 具有良好分化、扩散缓慢的特点,通过良好的治疗可以取得相对比较好的预后效果^[2]。

在临床诊断中,电子计算机断层扫描图像(computed tomography, CT)^[3]和磁共振成像(magnetic resonance imaging, MRI)是最常用的诊断手段,而 MRI 诊断是临床医生分析脑部时的首选,其中 MRI 是一种非侵入式的软组织成像技术,它具有高分辨率、多维度成像的特点,可对人体各部位 360°扫描成像,能够客观、具体地显示人体内的各个组织、器官的相对位置关系,更精确地提供脑肿瘤的外观、大小、位置等有用信息。因而成为脑肿瘤成像的主要手段。

随着深度学习技术的不断发展,深度学习在自然图像集中的图片分类、图片检测、图片分割等领域都取得了非常好的成绩。迄今为止,深度学习技术在医学领域的应用受到研究学者们的青睐,采用深度学习模型的分割方法一方面解决了传统的脑肿瘤分割方法中,如灰度阈值分割^[4]、小波多尺度变换^[5]和模糊核聚类^[6]等精度较差和需要先验知识的问题,另一方面解决了手工标记易错、耗时的问题。2015年, SHELHAMER 等^[7]提出了全卷积神经网络(fully convolutional networks for semantic segmentation, FCN)用于语义分割,用全卷积层替代全连接层,采用端到端的训练方式,实现了图像级分类到像素级分类的转变。RONNEBERGER 等^[8]在 FCN 的设计思想上进行改进并提出了一种新的编码器-解码器的网络结构,该网络结构呈 U 型,编码器使用下采样提取特征,再输入到解码器使用上采样还原特征尺寸大小。同时编解码之间使用了跳跃连接,结合上下文信息特征映射到解码器,将编码信息映射成二值分割掩膜。证明了特征连接的有效性,并激发了众多的研究者寻找最有效的特征连接,许多的图像分割算法^[9-12]都是基于 U-Net 进一步优化。

ZHOU 等^[13]提出的 U-Net++, 使用的一种紧密连接的编解码器网络,可以集成不同尺度的特征信息; HUANG 等^[14]提出的 U-Net3+ 简化了 U-Net++ 的编码和解码路径之间的连接,从而提高了全卷积网络获取全局信息的能力以及训练效率,然而,上述方法在其结构上仍存在局限性。一方面,这些方法考虑了编码路径、解码路径或其他路径之间的特征连接,但未能合理有效地进行组合。另一方面,由于脑肿瘤具有形状不一、病灶区与非病灶区之间的灰度分布交叉重叠以及不同的组织学亚区等特点,导致肿瘤区域的轮廓、空间位置信息难以识别。

基于以上研究,本文提出了用于脑肿瘤分割的双解码(double decoder U-net, DD-UNet)网络模型,实现脑肿瘤 MRI 图像的高效分割。编码器部分采用本文提出的一种基于数值分析的高阶残差编码(higher order encoding, HOE)模块,解决了提取细节信息不连续的问题;同时,引入上下文语义信息感知模块(multi scale dilation, MSD),用于连接解码路径与编码部分,并采用了一种跨层级的特征融合连接策略。此外,在设计的双路径解码器中,上下文语义解码路径用于对语义信息进行粗粒度分割,空间路径中引入选择性聚合空间注意力模块(spatial aggregation attention module, SAAM)对粗粒度语义分割图与空间信息联合映射,最终提取到的信息更加丰富,使得分割精度更高。

1 相关工作

1.1 总体结构

本文提出了用于脑肿瘤分割的双解码模型 DD-UNet, 该网络包含一个编码器和两个解码器,如图 1 所示。编码部分:使用尺寸大小为 256×256 的 3 通道脑肿瘤图片作为输入,编码模块包含 5 个 HOE 模块,在不增加参数量的情况下,每一个卷积层采用了不同大小的空洞率来增大感受野,使用步长因子为 2 的 3×3 卷积来压缩特征图大小,同时将特征通道数增加为上一次下采样的两倍。编码器与上下文语义解码路径部分(context decoder path)通过 MSD 相连接,并采用了一种跨层级的特征融合连接策略。值得注意的是,对于不同尺寸大小、通道的特征,采用不同大小的池化层,即 2×2 、 4×4 和 8×8 (图 1 中分别用不同的虚线表示),以确保相连接特征的分辨率一致,且为聚合语义特征提供了相同数量的特征通道。解码器部分:左解码器部分为上下文语义路径,使用 U-Net 变体模型将提取到的语义特征进行粗粒度分割;右解码器部分为空间路径(spatial deco-

der path),应用 SAMM 对粗粒度语义分割图与空间信息联合映射到空间路径进行细粒度分割,使得单一信息的融合变成了有选择性地对脑肿瘤的空间位置信息进行提取、筛选、融合的过程,最后再经过一层 1×1 卷积并通过 sigmoid 归一化获得脑肿瘤预测分割图。

1.2 编码 HOE 模块

语义特征提取的连续性可以较好地提高模型的分割性能,本文以数值微分的思想,设计出了 HOE 模块^[15-17],如图 2(c)所示。本文改变了传统方法中使用多个一阶残差块堆叠的方式,而采用四阶龙格库塔方法中的比例因子自适应地去拟合输入的特征,减少了图像在编码阶段提取特征时的损失。其中 HOE 模块使用了 8 层空洞卷积,其中空洞率大小为 $r_i = 1, 1, 2, 2, 3, 3, 5, 5$ 。在深层语义特征提取中,从多个感受野学习不同目标区域的特征信息,从而捕获更多结构的细节信息。

一般地,一阶残差模块如图 2(a)所示,如式(1)所示:

$$X_{l+1} = h\Psi(X_l; \{\omega_l\}) + X_l, \quad 1 \leq l \leq N-1, \quad (1)$$

式中, $h\Psi(X_l; \{\omega_l\})$ 为残差部分, X_l 为特征输入, X_{l+1} 为特征输出。

对于更深的层,如图 2(b)所示,从上一个隐藏状态 X_L 到下一个深度 L 层的残差块,如式(2)所示:

$$X_{L+1} = h\Psi(X_L; \{\omega_L\}) + X_L, \quad 1 \leq L \leq N-1. \quad (2)$$

本文使用四阶龙格-库塔方法作为编码阶段的特征提取模块,如式(3)所示:

$$\begin{cases} X_{L+1} = X_L + \frac{1}{6}(k_0 + 2k_1 + 2k_2 + k_3) \\ k_0 = h\Psi(X_L; \{\omega_L\}) \\ k_1 = h\Psi\left(X_L + \frac{1}{2}k_0; \{\omega_L\}\right), 1 \leq L \leq N-1 \\ k_2 = h\Psi\left(X_L + \frac{1}{2}k_1; \{\omega_L\}\right) \\ k_3 = h\Psi(X_L + k_2; \{\omega_L\}), \end{cases} \quad (3)$$

式中, ω_L 为权值矩阵, X_L 为特征输入, X_{L+1} 为最终

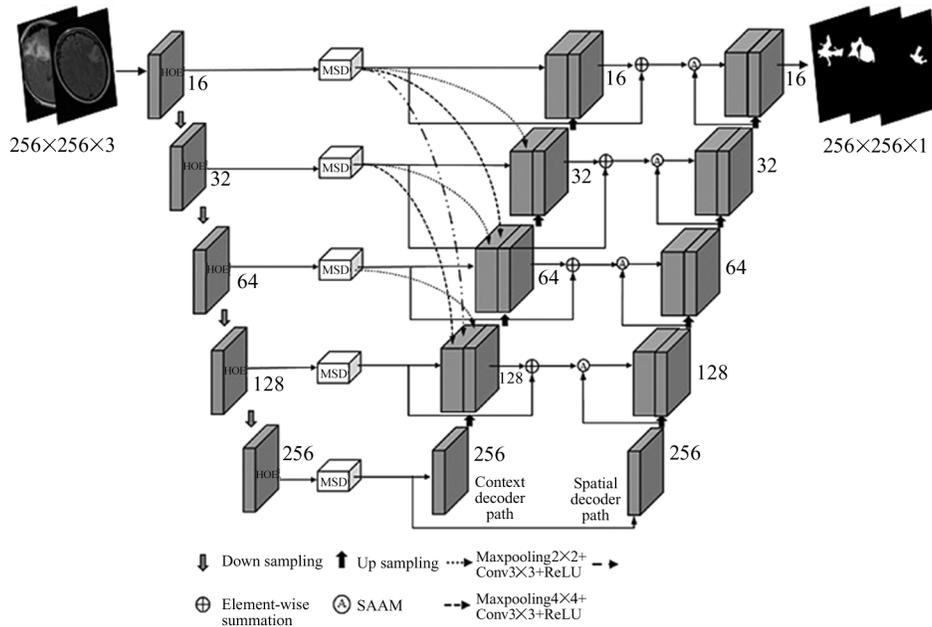


图 1 DD-UNet 网络结构

Fig. 1 DD-UNet network structure

的特征输出。

1.3 多尺度上下文语义信息感知模块

多尺度语义表达可以有效增强特征表征能力,减少编解码器特征信息之间的差异性。本文在编码器与上下文解码路径中引入了 MSD,其结构如图 3 所示。其设计方法是将几个带有空洞率的卷积层采

用级联的方式连接,随着空洞率以斐波那契序列规律增加而产生锯齿波的形式^[18-21],能更好地满足对不同尺度大小的物体进行特征提取,同时有助于梯度在低层级和高层级的传递。

这种设计的目的是由一系列卷积操作产生不同尺度大小的感受野,其最终大小可以完全覆盖一个

正方形区域,没有任何缺失的“孔”,并从更广泛的像素范围中获取信息。假设有 N 层卷积层,其中卷积核大小为 $k \times k$,空洞率大小为 $[r_1, \dots, r_i, \dots, r_n]$,对于每个卷积层 i ,其对应的扩张卷积的核大小为 $k_{di} \times k_{di}$,其中 $k_{di} = k + (k - 1) \times (r_i - 1)$, r_i 是空洞率。接下来,将扩张卷积层内两个非零值之间的最大距离定义为如式(4)所示:

$$m_i = \max[m_{i+1} - 2r_i, m_{i+1} - 2(m_{i+1} - r_i), r_i], \quad (4)$$

式中, $m_n = r_n$, $m_2 \leq k$ 。

该模块采用了一种空洞率以斐波那契序列规律的方式增长,而不是让空洞率以传统的指数序列规律增长,斐波那契序列被定义为如式(5)所示:

$$r_i = \frac{1}{\sqrt{5}} \times \left[\left(\frac{1 + \sqrt{5}}{2} \right)^i - \left(\frac{1 - \sqrt{5}}{2} \right)^i \right]. \quad (5)$$

与通常采用的指数序列相比,它的扩张速率较小,因此可以抑制网格效应。本文创建了一个8层的扩张卷积网络,其中以 $r_i = 1, 1, 2, 3, 5, 8, 13, 21$ 的扩张速率增加。假设输入的特征为 $X_0 \in R^{H \times W \times C}$,融合的特征输出图为 $X_{MSD} \in R^{H \times W \times C}$,总体的特征融合公式化如下:

$$\begin{aligned} \Pi_C = & X_{1 \times 1}^{r=1}(X_0) \oplus X_{3 \times 3}^{r=1}(X_1) \oplus X_{5 \times 5}^{r=2}(X_2) \oplus \\ & X_{3 \times 3}^{r=3}(X_3) \oplus X_{5 \times 5}^{r=5}(X_4) \oplus X_{3 \times 3}^{r=8}(X_5) \oplus \\ & X_{3 \times 3}^{r=13}(X_6) \oplus X_{3 \times 3}^{r=21}(X_7), \end{aligned} \quad (6)$$

$$X_{MSD} = \sigma(f(m, \Pi_C)) + \alpha, \quad (7)$$

式中, X_i 表示第 i 层的特征输入 ($i \geq 0$), Π_C 表示不同尺度特征通道之和, $X_{k \times k}^{r_i}(\cdot)$ 表示输入的特征图通过卷积核大小 $k \times k$ 且空洞率为 r_i 的卷积操作, m 表

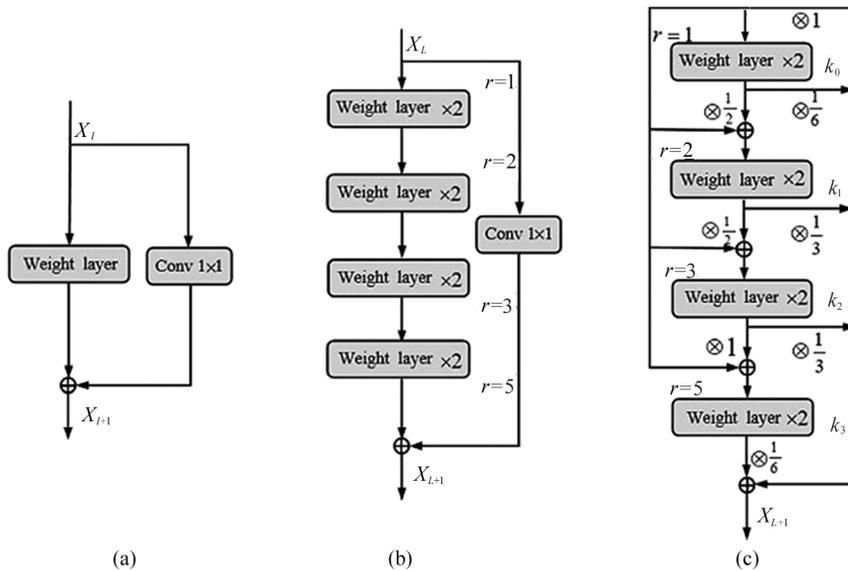


图 2 残差编码块:(a) 传统一阶残差编码块;(b) 一阶 8 层残差编码块;(c) HOE 编码块

Fig. 2 Residual coding block: (a) Traditional first-order residual coding block; (b) First-order 8-layer residual coding block; (c) HOE coding block

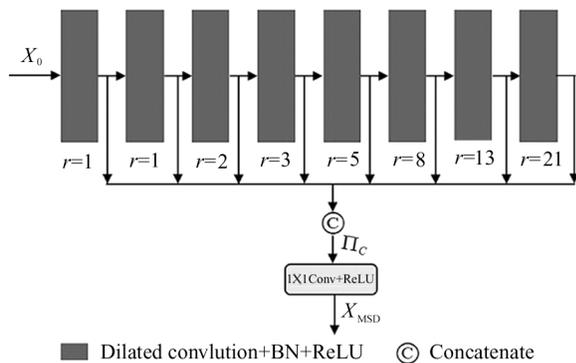


图 3 多尺度上下文信息感知模块
Fig. 3 Multiscale context information awareness module

示卷积权重, $f(\cdot)$ 表示卷积运算, σ 表示 ReLU 激活函数, α 表示卷积偏置。

1.4 选择性全局聚合空间注意力模块

本文提出了用于选择性聚合空间信息特征的注意力模块 SAAM,如图 4 所示。该结构延续了传统的 CBAM 注意力机制^[22]的设计思想。本文的 SAAM 分为两部分:第一部分对输入的特征图应用通道注意力机制,将特征图 $M_{in} \in R^{H \times W \times C}$ 输入到通道注意力机制进行关键信息的提取,提取到的通道注意图为 $f_c \in R^{H \times W \times C}$,再对生成的注意图进行选择空间特征提取,与输入的原特征图 M_{in} 进行元素相乘,生成的空间注意图为 $F_{s1} \in R^{H \times W \times C}$;第二部分

是对输入的特征 $M_{in} \in R^{H \times W \times C}$ 应用空间注意力机制进行最大池化操作处理,得到 $f_{\max-pool} \in R^{H \times W \times 1}$, 之后将其输入到 1×1 conv, 进行通道参数压缩, 然后经过一个 sigmoid 激活函数并与特征输入进行元素乘积运算得到 $F_{s1} \in R^{H \times W \times C}$ 。最后将第一部分和第二部分所提取到的空间特征采用叠加的方式聚合得到 $F_{s2} \in R^{H \times W \times C}$ 。简而言之, 聚合的二维空间注意力图提取过程公式化如式(8)–(10)所示:

$$F_{s1} = f_{s1}(f_c(M_{in}) \otimes M_{in}) \otimes M_{in}, \quad (8)$$

生成空间注意力图, 该操作的实现是首先将输入的特

$$F_{s1} = f_{s1}(M_{in}) \otimes M_{in}, \quad (9)$$

$$F_{s2} = F_{s1} + F_{s2}, \quad (10)$$

式中, $f_c(\cdot)$ 、 $f_{s1}(\cdot)$ 以及 $f_{s1}(\cdot)$ 分别表示 Channel-Spatial-Attention 注意力模块中的通道注意力生成、空间注意力生成和 Spatial-Attention 模块的空间注意力生成, F_{s1} 表示 Spatial-Channel-Attention 注意力模块的特征生成图, F_{s2} 表示 SAAM 空间注意力模块最终提取的注意力特征图。

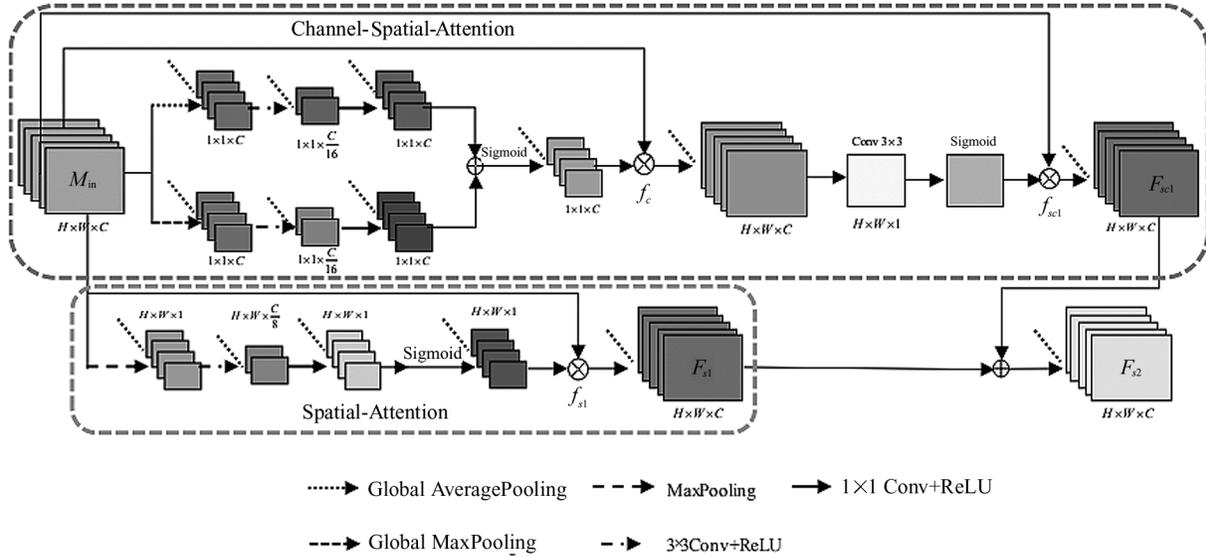


图 4 选择性全局聚合空间注意力机制模块

Fig. 4 Selective global spatial aggregation attention mechanism module

2 实验结果与分析

2.1 实验环境及参数设置

实验所用的操作系统为 Windows 64 位, 内存 16 GB, 处理器为 Zen 5 3600 6 - Core Processor 3.60 GHz 六核, 显卡为 Nvidia GeForce GTX 3060, 显存 12 G。本文所设计的模型方法使用深度学习 Tensorflow 2.5 框架和 Python 3.8 进行实验, 网络训练过程中使用的优化器为 Adam 优化器, 初始学习率 0.001, 批数量 (batch size) 设置为 16, 训练迭代次数为 150 轮, 并且使用在线数据增强。

2.2 实验数据集

本文实验使用的数据集来自癌症基因组图谱数据库 (the cancer genome atlas, TCGA), 从 TCGA 低级别胶质瘤样本中筛选了 110 例患者, 对应于 110 个肿瘤序列图片组, 每个病例都包含 4 个不同的 MRI 序列, 分别是 T1、T1ce、T2 和 Flair 序列, 每个序列患者切片数量 15–90 不等, 实验数据有效图片一共

3 929 张, 其中每张图像大小为 256×256 。另外, 将增强后的数据集随机抽取 80% 的图像作为训练集, 20% 的图像作为测试集, 且对图片进行了颅骨分离, 使得网络在训练的时候, 能更集中训练大脑内部区域。

经过统计得出, 本文使用的 TCGA 数据集正负样本比例失衡, 其比约为 3 : 2, 所以本文使用的损失函数 (loss function) 采用 ABRAHAM 等^[23] 提出的 (focal tversky loss, FTL), 如式(11)–(12)所示, 解决了正负样本数量不平衡, 从而影响分割精度的问题。

$$TI_C =$$

$$\frac{\sum_{i=1}^N p_{ic} g_{ic} + \epsilon}{\sum_{i=1}^N p_{ic} g_{ic} + \delta \sum_{i=1}^N p_{\bar{i}c} g_{\bar{i}c} + \beta \sum_{i=1}^N p_{ic} g_{\bar{i}c} + \epsilon}, \quad (11)$$

$$FTL_C = \sum_C (1 - TI_C)^{\frac{1}{\gamma}}, \quad (12)$$

式中, ϵ 是为了防止分母为零而引入的一个正常数,

p_{ic} 表示病变类别 c 在第 i 个像素的标准值, g_{ic} 表示病变类别 c 在第 i 个像素的标准值, $g_{\bar{c}}$ 表示为非病变类别 c 在第 i 个像素的预测值。通过调整超参数 δ 和 β 来提高召回率以及通过调整 γ 值来增强对感兴趣区域(ROI)的关注度,从而增强了对边界轮廓信息模糊以及病灶周围对比度低等的识别能力。其中, $\delta = 0.7, \beta = 0.3, \gamma = 0.75$ 。

2.3 评价指标

为了评估网络模型的分割效果,本文通过 5 个分割评估指标进行了比较: Dice 系数(Dice coefficient, Dice)、平均交并比(mean intersection over union, mIoU)、敏感性(sensitivity, Se)、特异性(specificity, Sp)、准确率(accuracy, Acc)。如式(13)–(17)所示:

$$M_{mIoU} = \frac{T_P}{F_N + T_P + F_P}, \quad (13)$$

$$D_{Dice} = \frac{2T_P}{F_P + 2T_P + F_N}, \quad (14)$$

$$S_{Se} = \frac{T_P}{T_P + F_N}, \quad (15)$$

$$S_{Sp} = \frac{T_N}{T_N + F_P}, \quad (16)$$

为病变类别 c 在第 i 个像素的预测值, $p_{\bar{c}}$ 表示非病

$$A_{Acc} = \frac{T_P + T_N}{T_P + F_P + T_N + F_N}, \quad (17)$$

式中, T_P (真阳性)表示能正确分割的脑肿瘤病灶区的像素数目, F_P (假阳性)表示被错误分割为脑肿瘤病灶区的背景像素值, T_N (真阴性)表示被正确分割的背景区域的像素数目, F_N (假阴性)表示错误标记为背景的脑肿瘤像素数目。

2.4 实验结果分析

2.4.1 HOE 块对分割结果的影响

为了验证本文提出的四阶残差编码块对病灶区的分割效率,使用 U-Net 模型在 TCGA 数据集上进行实验,用模型的参数量和 5 个常用指标进行了实验对比,如表 1 所示,其中 M 代表可训练参数量(表 2、表 3 同指)。对比图 2(b)深度为 8 层的一阶残差模块,四阶残差编码块的参数量不仅没有增加,并且在 $mIoU$ 、 $Dice$ 、 Se 、 Acc 指标上分别提升了 1.77%、2.79%、1.98%、1.13%。所以本文推断出,高阶残差在提取特征信息时更具连续性,有效提升了脑肿瘤分割精度,证明了本文提出的四阶残差编码块的有效性。

表 1 高、低阶残差编码块对比实验

Tab. 1 Comparison experiment of high and low-order residual coding blocks

Methods	$mIoU$	$Dice$	Se	Sp	Acc	$M(\times 10^6)$
U-Net base	0.8604	0.8652	0.8458	0.9989	0.9304	7.55
U-Net(Res)	0.8746	0.8757	0.8695	0.9990	0.9318	10.63
U-Net(HOE)	0.8923	0.9036	0.8893	0.9992	0.9431	10.63

2.4.2 消融实验

为了验证添加的各个模块对脑肿瘤的分割性能,从基础模型 U-Net base 到最后的 DD-UNet 共进行了 5 次消融实验,其中 baseline 仅由一个编码器和两个解码器构成,在此基础上添加不同的模块,分别凸显了 HOE、MSD 和 SAAM 模块对脑肿瘤分割精度的影响。

实验结果如表 2 所示,在 baseline 基础上加入 HOE 模块后, $Dice$ 和 $mIoU$ 的指标分别提升了 2.95% 和 3.23%;然后添加 MSD 后, $Dice$ 的指标提升了 1.97%;最后添加选择性 SAAM, $mIoU$ 和 $Dice$ 的指标分别提升了 1.44% 和 1.02%,相比基础模型 U-Net base,本文方法在 $mIoU$ 、 $Dice$ 、 Se 、 Acc 指标上分别提升了 4.67%、6.83%、6.57%、3.11%。其中 HOE、MSD 对识别目标区域边缘轮廓信息以及不同

尺度大小的图片较为敏感, SAAM 可以精确定位病灶区的空间位置信息,证明了所添模块的有效性。

2.4.3 算法对比实验结果分析

为了进一步证明本文方法的有效性,本文将提出的 DD-UNet 与 FCN-8、U-Net、DeepLabV3+、U-Net3+、AttentionU-Net、DoubleU-Net、U²Net 共 7 种目前表现较好的算法模型进行分割性能对比,结果如表 3 所示,可以看出本文基于 DD-UNet 模型的分割算法在 $mIoU$ 、 $Dice$ 、 Se 、 Sp 、 Acc 指标上分别达到 90.71%、93.35%、91.15%、99.94%、96.15%。虽然本文方法比 FCN-8、U-Net、DeepLabV3+、U-Net3+、AttentionU-Net 参数多,但其分割性能远超过这几种方法。根据表 3 的实验结果可知, DoubleU-Net 和 U²Net 的参数量均超过本文方法,尤其

是 U²Net 方法的参数量比本文高出近 2400 万参数,但与本文算法相比,本文算法在 *mIoU*、*Dice*、*Se*、*Acc* 指标上比 U²Net 分别提升了 2.75%、1.24%、1.11%、1.12%,说明了本文方法的优越性,与其他方法相比,本文在参数量和精度之间找到了一个平衡点,使其获取更优的分割结果。

不同分割算法在 TCGA 数据集上分割脑肿瘤的可视化结果图,如图 5 所示。从图中可以看出 DD-UNet 网络与其他网络相比,该网络在不同尺度大小、形状不规则、轮廓边界模糊以及病灶区域对比度低等多种情况下,均取得了很好的分割效果。其分割结果最接近真实标签,与表 3 的实验数据相吻合。

表 2 改进方案对分割结果的影响

Tab. 2 The impact of the improved scheme on the segmentation results

Methods	<i>mIoU</i>	<i>Dice</i>	<i>Se</i>	<i>Sp</i>	<i>Acc</i>	<i>M</i> ($\times 10^6$)
UNet base	0.8604	0.8652	0.8458	0.9989	0.9304	7.55
baseline	0.8722	0.8741	0.8584	0.9979	0.9332	8.03
baseline + HOE	0.9045	0.9036	0.9093	0.9994	0.9593	11.77
baseline + HOE+MSD	0.8927	0.9233	0.9116	0.9991	0.9561	15.86
baseline + HOE+MSD+SAAM	0.9071	0.9335	0.9115	0.9994	0.9615	21.27

表 3 不同脑肿瘤分割算法对比实验

Tab. 3 Comparative experiment of different brain tumor segmentation algorithms

Methods	<i>mIoU</i>	<i>Dice</i>	<i>Se</i>	<i>Sp</i>	<i>Acc</i>	<i>M</i> ($\times 10^6$)
FCN-8	0.8516	0.8108	0.8167	0.9986	0.8972	0.71
U-Net	0.8604	0.8652	0.8458	0.9989	0.9304	7.55
DeepLabV3+ ^[24]	0.8732	0.8893	0.8948	0.9989	0.9394	16.34
U-Net3+	0.8406	0.9199	0.9224	0.9989	0.9535	6.32
AttentionU-Net ^[25]	0.8549	0.9208	0.8928	0.9994	0.9548	13.62
DoubleU-Net ^[26]	0.8478	0.9172	0.9026	0.9990	0.9538	29.59
U ² Net ^[27]	0.8796	0.9211	0.9004	0.9993	0.9503	44.83
DD-UNet(ours)	0.9071	0.9335	0.9115	0.9994	0.9615	21.27

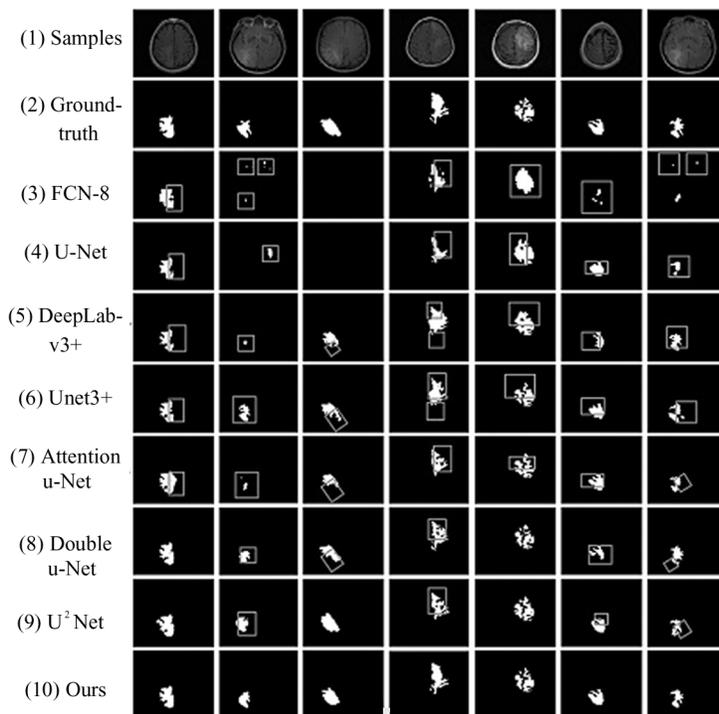


图 5 不同算法脑肿瘤分割可视化结果

Fig. 5 Visualization results of brain tumor segmentation with different algorithms

3 结 论

为了优化脑肿瘤分割,本文提出了 DD-UNet 算法。编码器采用 HOE 模块,再利用空洞卷积的技术优势,对脑肿瘤病灶的轮廓信息、空间位置信息的提取能力大幅度提升。模型使用了两个解码器,分别为上下文路径解码器和空间路径解码器,它们分别被设计用来对特征图进行粗粒度分割和细粒度分割。在空间路径解码器中采用选择性 SAAM,加强了有效信息的获取,减少了无效信息的干扰。最后对比其他基于 U-Net 改进的方法,本文方法的分割性能最优。然而,它也有一定的局限性,该实验是基于二维数据进行分割的,并没有充分利用三维的空间信息,忽略了切片之间相关性,阻碍了分割性能,后续将进一步探索并研究更高效的脑肿瘤分割算法。

参考文献:

- [1] YANG X J, JIANG T. Interpretation of "World Health organization classification of tumors of the central nervous system (2016)"[J]. Chinese Journal of Neuropsychiatric Diseases, 2016, 42(6): 321-329.
杨学军,江涛. 解读《世界卫生组织中枢神经系统肿瘤分类(2016年)》[J]. 中国神经精神疾病杂志, 2016, 42(6): 321-329.
- [2] 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.
- [3] WANG S Z, WANG X D. Advances in the application of computed tomography in surgery[J]. Journal of Shanghai Jiaotong University: Medical Edition, 2017, 37(6): 877-880.
王舒泽,王旭东. 电子计算机断层扫描技术在手术中的应用进展[J]. 上海交通大学学报:医学版, 2017, 37(6): 877-880.
- [4] SANG L, QIU M, WANG L, et al. Brain tumor MRI image segmentation based on statistical thresholding method[J]. Journal of Biomedical Engineering Research, 2010, 29(4): 237-239.
桑林琼,邱明国,王莉,等. 基于统计阈值的脑肿瘤 MRI 图像的分割方法[J]. 生物医学工程研究, 2010, 29(4): 237-239.
- [5] WANG H, YANG C M, YUAN Z H, et al. Tumor image fusion based on wavelet multiscale decomposition[J]. Journal of Northeastern University (Natural Science Edition), 2005, 26(4): 238-241.
王宏,杨春梅,袁正华,等. 基于小波多尺度分解的肿瘤图像融合[J]. 东北大学学报(自然科学版), 2005, 26(4): 238-241.
- [6] ZHANG T D, LV X Q, REN X Y, et al. Brain tumor image segmentation method based on spatial fuzzy kernel clustering[J]. Control Engineering, 2017, 24(10): 2107-2111.
张腾达,吕晓琪,任晓颖,等. 基于空间模糊核聚类的脑肿瘤图像分割方法[J]. 控制工程, 2017, 24(10): 2107-2111.
- [7] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [8] 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, October 05-09, 2015, Munich, Germany. Cham: Springer, 2015: 234-241.
- [9] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [10] CHAURASIA A, CULURCIELLO E. LinkNet: Exploiting encoder representations for efficient semantic segmentation[C]// 2017 IEEE Visual Communications and Image Processing (VCIP), December 10-13, 2017, St. Petersburg, FL, USA. New York: IEEE, 2017: 1-4.
- [11] CHEN L C, PAPANDREOU G, KOKKINOS L, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [12] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, November 30-December 2, 2018, Xi'an, China. New York: IEEE, 2017: 2881-2890.
- [13] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKHSH N, et al. UNet++: A nested U-Net architecture for medical image segmentation[C]// Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, September 20, 2018, Granada, Spain. Cham: Springer, 2018: 3-11.
- [14] HUANG H, LIN L, TONG R, et al. UNet 3+: A full-scale connected UNet for medical image segmentation[C]//

- ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 4-8, 2020, Barcelona, Spain. New York: IEEE, 2020: 1055-1059.
- [15] WEINAN E. A proposal on machine learning via dynamical systems[J]. *Communications in Mathematics and Statistics*, 2017, 1(5): 1-11.
- [16] LU Y, ZHONG A, LI Q, et al. Beyond finite layer neural networks: Bridging deep architectures and numerical differential equations[C]// *International Conference on Machine Learning*, July 10-15, 2018, Stockholm, Sweden. PMLR, 2018: 3276-3285.
- [17] HE X, MO Z, WANG P, et al. ODE-inspired network design for single image super-resolution[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 1732-1741.
- [18] ANTHIMOPOULOS M, CHRISTODOULIDIS S, EBNER L, et al. Semantic segmentation of pathological lung tissue with dilated fully convolutional networks[J]. *IEEE Journal of Biomedical and Health Informatics*, 2018, 23(2): 714-722.
- [19] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. (2015-11-23)[2022-03-27]. <https://arxiv.org/abs/1511.07122>.
- [20] WANG P, CHEN P, YUAN Y, et al. Understanding convolution for semantic segmentation[C]// *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, March 12-15, 2018, Lake Tahoe, NV, USA. New York: IEEE, 2018: 1451-1460.
- [21] CHEN L C, PAPANDEOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation [EB/OL]. (2017-06-17) [2022-03-28]. <https://arXiv.org/abs/1706.05587>.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]// *Proceedings of the 15th European Conference on Computer Vision (ECCV)*, September 8-14, 2018, Munich, Germany. Berlin: Springer-Verlag, 2018: 3-19.
- [23] ABRAHAM N, KHAN N M. A novel focal tversky loss function with improved attention U-Net for lesion segmentation[C]// *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, April 8-11, 2019, Venice, Italy. New York: IEEE, 2019: 683-687.
- [24] CHEN L C, ZHU Y, PAPANDEOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [M]// *Computer Vision-ECCV 2018*. Cham: Springer International Publishing, 2018: 833-851.
- [25] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: learning where to look for the pancreas[EB/OL]. (2018-04-13) [2022-03-27]. <https://arxiv.org/abs/1804.03999v2>.
- [26] JHA D, RIEGLER M A, JOHANSEN D, et al. Double U-Net: A deep convolutional neural network for medical image segmentation[C]// *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, July 28-30, 2020, Rochester, MN, USA. New York: IEEE, 2020: 558-564.
- [27] 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.

作者简介:

苏 赋 (1979—), 女, 工学博士, 副教授, 主要研究领域为人工智能、医学图像处理、信号与信息处理。