

# DMR-KAN: 基于多尺度区域强化的 三维肿瘤影像分割方法

方丁毅<sup>1</sup>, 程换新<sup>1\*</sup>, 骆晓玲<sup>2</sup>, 王若峥<sup>3</sup>

(1. 青岛科技大学自动化与电子工程学院, 山东青岛 266061; 2. 青岛科技大学机电工程学院, 山东青岛 266061;  
3. 新疆医科大学附属肿瘤医院放疗中心, 新疆乌鲁木齐 830022)

**摘 要:** KAN (Kolmogorov-Arnold Networks) 模型通过新的线性函数拟合方式使得图像分割的精准度得到提升。然而, 其拟合角度单一、提取标签位置信息差等问题, 导致其对标签细节特征信息的处理能力较差, 令网络精度提升受限。针对上述问题, 设计了多尺度双通道三维影像分割模型, 该模型通过整合多角度三维影像输入, 将多角度 KAN 模块与多尺度卷积加权残差通道相结合, 显著增强了网络对图像微小特征的提取能力。在网络注意力机制方面, 设计了多视角自注意力残差模块, 该模块通过多维度特征交互有效捕获标签空间位置信息, 使占比较低 (<10%) 的标签区域仍能保持优异的分割精度。模型在 BraTS2021 MRI 多模态三维脑肿瘤数据集与 LiTS2017 肝肿瘤 CT 三维数据集集中进行实验。改进模型精准度分别为 86.54% 与 88.07%; 在脑肿瘤数据集中, 增强肿瘤、全部肿瘤和肿瘤核心区域的 Dice 评价指标达到 83.67%、88.79% 和 85.28%, 相较 U-KAN 网络分别提升 3.38、2.85 和 1.62 个百分点; 在肝肿瘤数据集中, 肝脏与肿瘤区域的 Dice 评价指标达到 91.36% 与 84.77%, 分别提升了 1.69 个百分点与 1.02 个百分点。实验结果表明该模型对三维肿瘤影像分割效果提升显著。

**关键词:** 三维影像分割; 多角度; 双通道; 自注意力; 标签区域强化

中图分类号: TP391.41

文献标识码: A

文章编号: 0372-2112(2025)08-2818-12

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250379

## DMR-KAN: A 3D Medical Image Segmentation Method Based on Multi-Scale Region Enhancement

FANG Ding-yi<sup>1</sup>, CHENG Huan-xin<sup>1\*</sup>, LUO Xiao-ling<sup>2</sup>, WANG Ruo-zheng<sup>3</sup>

(1. College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao, Shandong 266061, China;  
2. College of Mechanical and Electrical Engineering, Qingdao University of Science and Technology, Qingdao, Shandong 266061, China;  
3. Cancer Hospital Affiliated to Xinjiang Medical University Radiation Therapy Center, Urumqi, Xinjiang 830022, China)

**Abstract:** The KAN (Kolmogorov-Arnold Networks) model enables the accuracy of image segmentation to be improved by a new linear function fitting method. However, the problems of single fitting angle and poor extraction of label position information lead to its poor ability to process the detailed feature information of labels, which limits the improvement of network accuracy. To address the above problems, a multi-scale dual-channel 3D image segmentation model is designed, which significantly enhances the network's ability to extract minute features from images by integrating multi-angle 3D image inputs and combining the multi-angle KAN module with multi-scale convolutional weighted residual channels. In terms of the network attention mechanism, a multi-view self-attentive residual module is designed, which effectively captures the label spatial location information through multi-dimensional feature interactions, so that the label region with a relatively low percentage (<10%) can still maintain excellent segmentation accuracy. The model is experimented on BraTS2021 MRI multi-modal 3D brain tumor dataset and LiTS2017 liver tumor CT 3D dataset. The accuracy of the improved model is 86.54% and 88.07%, respectively; in the brain tumor dataset, the Dice evaluation indexes of the enhanced tumor, all tumors, and tumor core region reach 83.67%, 88.79%, and 85.28%, which are improved by 3.38, 2.85, and 1.62 percentage points, respectively, compared with the U-KAN network; in the liver tumor dataset, the liver and tumor region's Dice evaluation index reached

91.36% and 84.77%, which were improved by 1.69 percentage points and 1.02 percentage points, respectively. The experimental results show that the model improves the effect of 3D tumor image segmentation significantly.

**Key words:** 3D image segmentation; multi-angle; dual-channel; self-attention; labeled region reinforcement

## 1 引言

随着智慧医疗的发展,医学影像分割逐渐在其中发挥重要作用.精准、快速地分割影像中的病灶能够为诊断与治疗及时提供病情信息,也为自助医疗与远程医疗提供理论支持,例如自动化无人门诊、远程微创手术等现代医疗手段<sup>[1]</sup>.因此,准确有效的分割算法对智慧医疗的发展具有重要意义.

传统分割方法大多基于二维影像,二维影像是由三维影像通过切片操作制作而来,这种操作会造成大量的原始影像信息丢失.因此,将三维影像作为网络输入数据能够有效保留影像中标签的原始信息.然而,三维影像输入造成的结果是网络参数的大幅增加,导致运算成本大幅提高<sup>[2,3]</sup>.CM-UNet<sup>[4]</sup>通过使用大卷积核来提高感受野的方法降低参数量,该模型使用多尺寸卷积核并行组成解码器通道,从而获得多尺度感受野.但是这种改进会导致原始影像中的细节特征丢失,降低标签特征的利用率.Wang等人<sup>[5,6]</sup>提出通道模型解决了这一问题,该模型通过使用不同尺寸的卷积块同时对图像特征进行提取并加权融合,使得在扩大感受野的同时兼顾了模型对细节特征的提取能力,但是这种方法的分割精度却并不理想.

注意力机制的出现为网络性能的提升提供了全新思路<sup>[7,8]</sup>.Zhang等人<sup>[9]</sup>提出了门控注意力机制,在跳跃链接部分添加加权门控融合机制,改善了网络对标签细节信息的提取能力;Varshaneya等人<sup>[10]</sup>提出通过在通道注意力机制中加入残差网络,改进其中的桥连接结构,提高了注意力机制对网络解码时的指导作用;Dao等人<sup>[11]</sup>为缓解GPU(Graphics Processing Unit)计算缓慢问题,提出了FA(Fast Attention)分块注意力机制用以减小GPU的HBM与SRAM之间的内存读写次数.虽然现有注意力机制能够有效获取解码层输出特征图像的特征信息,但是对于标签像素占比较低的医学三维影像却效果有限.

最近,KAN(Kolmogorov-Arnold Networks)模型<sup>[12]</sup>的出现为三维影像分割开辟了新的方法.KAN模型将传统MLPs(MultiLayer Perceptrons)当中的线性权重替换成B样条函数,这一改动使得模型在计算速度、精度与可解释性上更胜MLPs.通过与传统U型网络结构相结合,Moradzadeh等人<sup>[13]</sup>提出了U-KAN模型被使用在医学影像分割领域,通过独创的KAN-Layer替换传统的卷积模块与解码模块,使得模型在二维影像分割中取得了不错的效果.尽管U-KAN模型对二维影像的分割效果已经达到了一定标准,但是对于输入三维影像的

3D U-KAN模型的表现却并不理想,结合模型特点与分割原理进行分析后发现,影响分割精度的主要原因为网络特征提取模型单一,对于原始影像的细节特征利用不足,网络整体过分依赖KAN模型的拟合能力,这些问题极大地影响了网络在对三维影像进行分割时的精准度.

基于以上问题,本文提出了DMR-KAN(Dual-channel&Multi-scale Resnet KAN)模型.首先,为了增强KAN模型对三维影像的拟合能力,本文模型设计了一个三维多角度特征提取KAN模块来提升网络线性拟合部分的拟合能力;紧接着,为了解决网络特征提取方式单一导致的特征利用率差的问题,本文设计了多尺度残差卷积模块与前者组成双通道,两者并行对原始影像进行特征提取,以加强网络编码部分对原始影像特征信息的利用率,从而提高网络的分割能力;此外,本文设计了多角度自注意残差模块,通过自我强化注意力机制,生成标签对应区域的强化权重三维矩阵,以强化由每层提取出的不同尺寸特征图中肿瘤的空间特征与位置信息;最后,在损失函数部分,本文使用Dice损失与L2正则化损失函数的加权损失函数,通过利用不同深度网络输出的特征图与分割结果计算损失后加权相加,有效改善网络加深导致的深层特征图信息丢失的问题.本文模型在公开数据集BraTS2021 MRI多模态三维脑肿瘤数据集与LiTS 2017肝肿瘤CT三维数据集中进行验证.

## 2 方法

DMR-KAN模型主要包括以下几个部分:(1)三维多角度特征提取KAN模块(Multi-angle Feature Extraction KAN, MFE-KAN),作为线性拟合通道的特征提取模块,主要作用是利用其优秀的可解释的线性拟合方式对原始三维影像中的特征进行高效提取;(2)基于双通道特征提取的方法,本文设计了一个多尺度残差卷积模块(Multi-scale Residual Convolutional module, MRC),用于对原始影像信息进行多尺度卷积特征提取并与MFE-KAN组成双通道编码层;(3)多角度自注意残差模块(Multi-scale Self-attention Residual module, MSR),将该模块放在每一个双通道编码层与解码层的跳跃连接部分,用于强化提取特征图中的肿瘤位置信息及其细节特征.DMR-KAN的模型结构与概述如图1所示.在本节中,将进一步对网络结构及各模块进行详细阐述.

### 2.1 三维多角度特征提取KAN模块

在传统的三维影像特征提取网络中,往往采用单一角度或从单一方向对原始影像进行特征提取<sup>[14-16]</sup>,

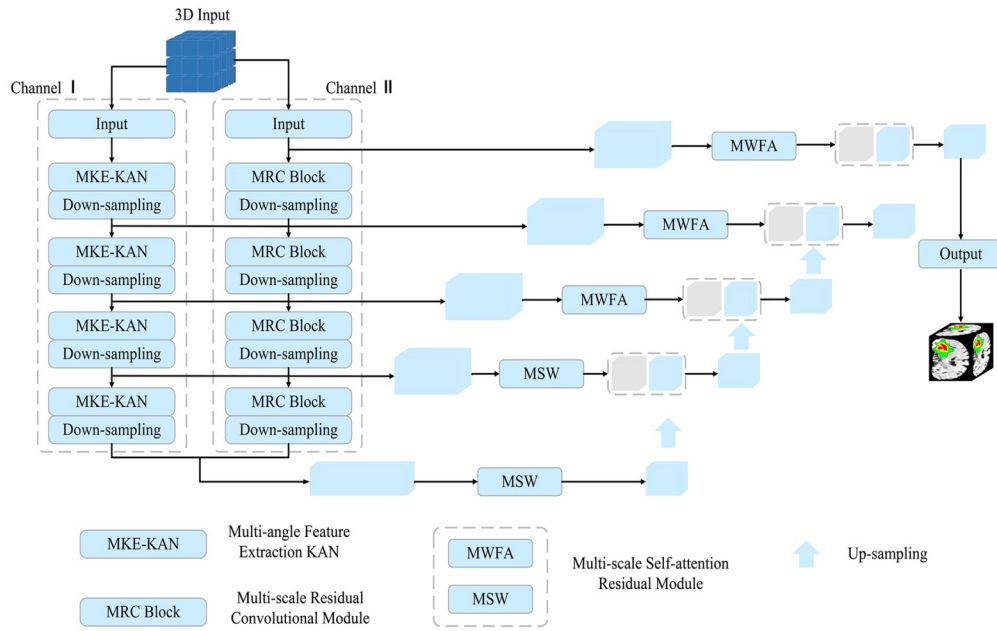


图1 DMR-KAN的模型具体结构

在现有的KAN模型相关的分割网络中也是如此. 这种单一的特征提取方式会令标签边缘的细小特征随着网络的加深、特征图像分辨率的减小而丢失, 从而使得网络对标签细节特征的提取能力减弱, 最终影响网络的分割精度. 为了解决这个问题, 本文设计了三维多角度特征提取KAN模块(MFE-KAN), 以实现三维影像的多角度特征建模. 该模块在保留了KAN模型强大的线性拟合建模与推理能力的同时, 充分利用原始特征信息对网络参数的传播进行指导, 保证了网络对细节特征的保留能力, 有效提升网络的分割精度.

如图2所示, 该模块的主干部分由多角度残差KAN模块并联构成. 在网络对特征提取时, 多个多角度残差KAN模块并行对3D特征图像进行建模, 每个模块分别从三维特征图的3个维度( $H, W, D$ )对特征图进行特征提取, 得到3个不同角度的特征矩阵, 并在最后进行自注意加权融合. 在融合模块部分, 使用由多角度残差KAN模块拟合得到的二维特征向量经过归一化与平均池化计算, 得到每一组向量的动态权重, 以完成多层特征向量的加权融合. 假定3D输入特征图 $F$ 的尺寸 $I \in \mathbb{R}^{B \times C \times H \times W \times D}$ , 其中 $B$ 为批次,  $C$ 为通道数, 在经过本模块后尺寸变为 $I' \in \mathbb{R}^{B \times 2C \times H/2 \times W/2 \times D/2}$ . 然后,  $F$ 与MRC模块的输出相结合后, 被送入下采样层与MSR模块中. 对于任意一个多角度残差KAN模块均可以被定义为以下公式:

$$\tilde{F}_n^{m_p} = \text{Norm} \left( \text{KAN} \left( \text{Flatten} \left( F_n \right) \right) \right) \quad (1)$$

$$F_n^{m_p} = DW \left( \text{Reshape} \left( \tilde{F}_n^m + \text{Norm} \left( \text{Flatten} \left( F_n \right) \right) \right) \right) \quad (2)$$

其中,  $F_n$ 代表由第 $n$ 层解码器输出的三维特征图,

$F_n^{m_p}$  ( $p \in [1, m]$ ) 表示任意多角度KAN模块编码后的二维特征向量.  $F_n$ 在沿某一方向平展后送入KAN模块, 得到经由建模后的二维特征矩阵与原始输入矩阵进行残差链接, 这种方式在有效提取特征图中的标签信息特征的同时, 最大程度地保留了原始特征信息, 确保了网络深层传播时保有足够的特征信息来指导深层网络的参数. 经过KAN模块后, 将得到的二维特征向量矩阵重新整形成原始尺寸的三维特征图并经过一个卷积核大小为3的双卷积模块. 同时, 将二维向量利用全局平均池化得到一个自注意权重, 将整形得到的三维特征图与自注意权重相乘, 最终获得单维度的自注意加权三维特征图, 多个模块输出的三维特征图将被融合成一个特征图送入下层编码器中, 该过程具体被定义为

$$F'_n = \sum_{m=1}^m \text{AvgPool} \left( \tilde{F}_n^{m_p} \right) \times F_n^{m_p} \quad (3)$$

其中,  $F'_n$ 代表经过加权多角度特征融合后的输出三维特征图.

## 2.2 多尺度残差卷积模块

传统卷积分割网络采用单通道输入模式对原始三维特征图进行特征提取<sup>[17-19]</sup>. 近期, U-Mamba<sup>[20,21]</sup>与U-KAN将输入的三维影像通过平展与标记操作, 将三维数据变为二维特征序列后作为输入进行特征提取, 这种方法在兼顾上下文信息的同时充分利用了特征图当中的全部信息对建模进行指导. 这种降维—提取—重建的建模方法虽然展现出不错的分割效果, 但是这种方法不仅未考虑到标签部分的三维空间特征信息对分割精度的影响, 甚至原始影像中大量的背景干扰信息也在很大程度上影响了二维序列信息的有效性. 而传

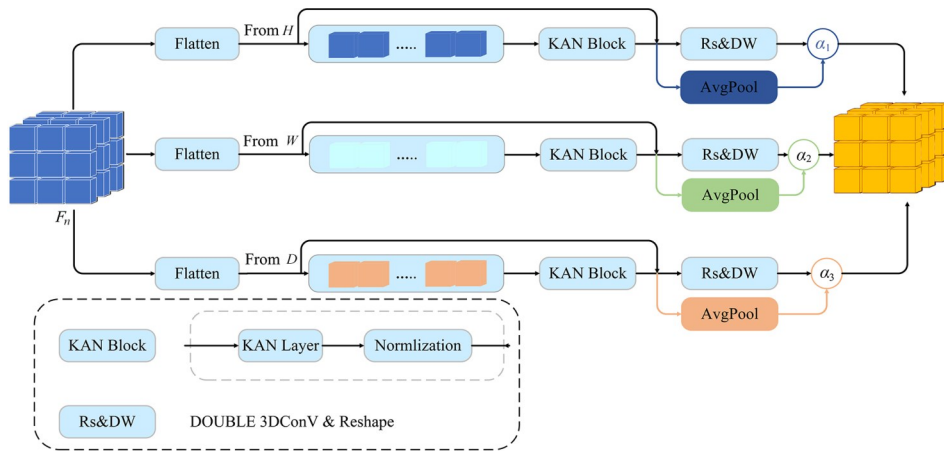


图2 三维多角度特征提取KAN模块结构

统的卷积方法与上述方法相比,卷积操作能够有效提取特征像素块周围的空间特征信息并在核区域进行特征融合,使得卷积后每个体素中包含对应卷积核大小的空间特征信息的同时,卷积操作方式有效压缩了背景标签的体素占比,在很大程度上增强了网络的鲁棒性. 基于上述理论,本文设计了一个多尺度残差卷积模块(Multi-scale Residual Convolution module-Block, MRC-Block)用于空间特征提取,同时设计了3个动态权重对从不同卷积核得到的特征图进行动态加权,之后将各卷积通道得到的特征信息进行点乘后对MFE-KAN得到的特征图像进行指导,从而使每层解码器输出的特征图像获取更多特征信息. MRC-Block 具体结构见图3.

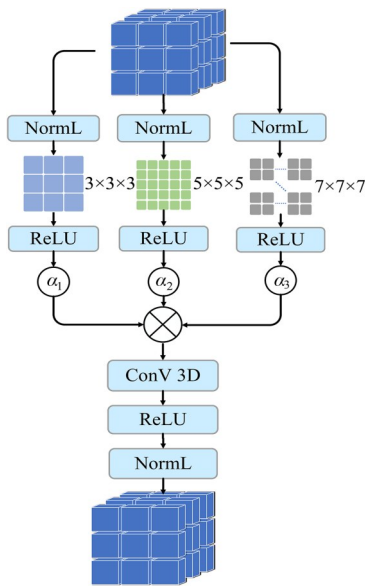


图3 多尺度残差卷积模块

当三维特征图输入进本模块时,3个不同尺寸的卷积模块同时对三维影像并行处理,得到3个不同感受野的特征图后,将三者通过加权融合得到多尺度融合特

征图. 在进行特征图融合时,选择使用矩阵点乘的方式对三维特征矩阵进行特征融合,这种融合方式使得标签区域与背景区域的特征值差距拉大,这种差距会在后续的激活函数与归一化操作当中进一步扩大,使得特征图包含的标签边缘信息能够有效传递到下层编码器当中,具体流程可由以下公式表示:

$$F_{n_{\text{conv}(1,2,3)}} = \text{ReLU}(\text{Conv}^{(3,5,7)}(\text{NL}(F_n))) \quad (4)$$

$$\overline{F}_n^T = \alpha_1 F_{n_{\text{conv}1}} \otimes \alpha_2 F_{n_{\text{conv}2}} \otimes \alpha_3 F_{n_{\text{conv}3}} \quad (5)$$

$$\overline{F}_n = \text{NL}(\text{ReLU}(\text{Conv}^{3 \times 3 \times 3}(\overline{F}_n^T))) \quad (6)$$

其中,  $F_n$  为逐层解码器的输入特征图;  $\alpha_n$  作为超参数,为通道融合的自适应加权权重,其初始值为1且与网络的精准度相关. 在网络训练前期,网络的精准度往往取决于网络对肿瘤标签整体的识别程度,此时大卷积核带来的大感受野体素能够使得网络更快地识别出肿瘤整体的大致区域,以加快前期的网络精度的提升速度. 所以在训练前期,  $\alpha_3$  将会增大以引导网络快速确定肿瘤大致区域,而  $\alpha_1$  则会下降以降低权重. 为了弥补  $\alpha_1$  下降可能导致的细节特征丢失过重从而影响MFE-KAN的特征提取能力,此时  $\alpha_2$  会适当上升进行补偿,从而使卷积网络能够使用自身获取的肿瘤区域信息对MFE-KAN进行指导的同时,不会因细节信息丢失过重而产生负反馈.

在网络训练后期,肿瘤标签的整体区域已被大致划分出来,此时精度的提升则取决于网络对标签边缘细节分割能力的强弱,对标签边缘细节的分割能力越强,网络最后的分割精度就越高. 此时小卷积核对标签边缘的特征提取能力可以有效提升网络对标签边缘的分割能力. 依照上述原理,随着网络精度的逐渐提升,  $\alpha_3$  会逐渐下降而  $\alpha_1$  则逐渐增加以达到细节特征提取的要求. 而此时大小为5的卷积核在细节特征提取方面则会弱于大小为3的卷积核,所以此时令  $\alpha_2$  逐渐下降以增加  $\alpha_1$  在三者中的权重占比. 根据网络训练时的精准

度,动态调整  $\alpha_n$  的相应数值,有效匹配网络在训练不同阶段的需求.各参数具体自调节曲线如图4所示.

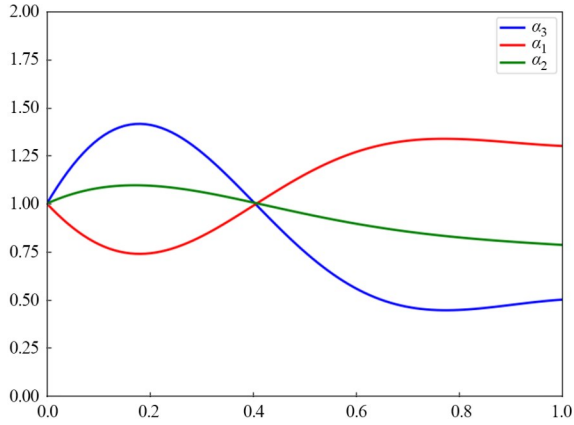
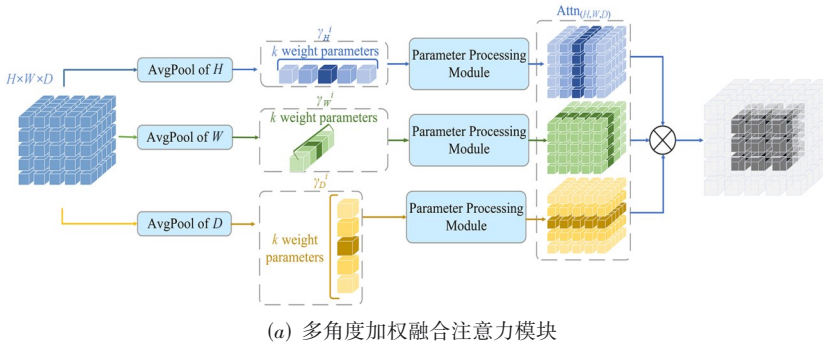
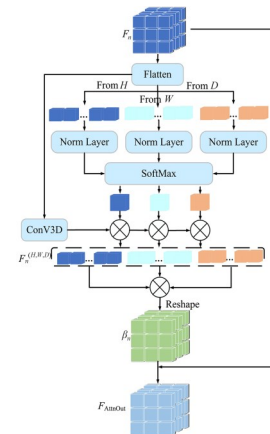


图4 自调节权重随精度变化曲线



(a) 多角度加权融合注意力模块



(b) 多尺度自注意权重模块

图5 多角度加权融合注意力模块和多尺度自注意权重模块结构

### 2.3.1 多角度加权融合注意力模块

传统U形网络中,解码器中的上采样模块采用线性插值的方法使特征图逐渐恢复至原始尺寸<sup>[26,27]</sup>.然而,无论是双线性插值还是三线性插值,其本质方法是根据相邻体素的特征值与应恢复的尺寸要求计算插入数值的个数与具体值,这种方法在深层网络的小尺寸特征图中更占优势.但是在浅层网络中,肿瘤的不规则性与边缘无序性,线性插值往往会导致原始特征的丢失,从而造成精度下降.为了解决浅层网络中大尺寸特征图的恢复问题,本文设计了一个多角度加权融合注意力模块(Multi-angle Weighted Fusion Attention module, MWFA).该模块的主要功能是缓解大尺寸特征图中肿瘤区域占比过低、肿瘤边缘形状复杂无序而导致的特征图恢复效果差的问题,在图5(a)中对模块结构进行了展示.首先,该模块接收来自浅层编码器的特征图像  $F_N^{C \times H \times W \times D}$  并将图像分别沿3个不同的角度  $H$ 、 $W$ 、 $D$

### 2.3 多角度自注意残差模块

随着网络层数的加深,原始三维影像会因为卷积与下采样等操作令尺寸逐渐减小,使得每一个体素的感受野增大,从而导致特征图中标签的细节信息模糊、缺失<sup>[22,23]</sup>.此时,在U型网络的跳跃链接中加入注意力机制<sup>[24,25]</sup>成为了当下的最优解决方法.合适的注意力机制能够充分利用编码器输出图像的特征信息,大幅改善因为上采样中线性插值操作导致的特征信息恢复偏差问题.在以往设计的注意力机制中,对于同一网络中不同深度的跳跃链接层往往采用同一个注意力模块,但是同一种注意力模块并不能适合每一种尺寸的特征图.基于此问题,受通道注意力与区域加权思想的启发,本文提出了多角度自注意残差模块,针对不同深度、不同尺寸的特征图设计了不同的注意力模块.具体结构如图5所示.

划分为  $K$  个区域 ( $K \in \text{GCF}(H, W, D)$ ),之后模块会对各个区域进行平均池化操作,得到3组二维权重参数矩阵.接下来,将权重参数矩阵归一化并恢复成与输入特征图尺寸相同的参数矩阵,将3组矩阵按同一角度逐元素相乘,再将其归一化并经过 Sigmoid 激活函数得到最终的权重矩阵后,与初始输入矩阵相乘,放大特征图中的肿瘤部分特征值来指导解码器输出的特征图.以  $H$  方向为例,具体操作由以下公式表示:

$$\gamma_H^i = \text{AvgPool} \left( F_N^{C \times \frac{H}{K} \times W \times D} \right) \quad (7)$$

$$\gamma_H = \text{Concat}(\gamma_H^1, \gamma_H^2, \dots, \gamma_H^K) \quad (8)$$

其中,  $\gamma_H^i$  表示原始特征图  $F_N^{C \times H \times W \times D}$  按照  $H$  角度依次分为  $K$  个区域后,经由平均池化得到的二维权重参数矩阵.将所得的  $K$  个二维权重参数矩阵进行拼接得到初始三维权重矩阵  $\gamma_H$ ,之后对初始三维权重矩阵进行归

一化,强化每一个区域的特征值用以反映区域内标签的具体位置信息.在初始三维权重矩阵的尺寸复原后,将3个不同角度下的权重矩阵逐元素相乘,经过数据处理模块后得到一个空间注意力权重矩阵,具体流程如下:

$$\text{Attn}_{(H,W,D)}^{C \times H \times W \times D} = \text{Sigmoid} \left( \text{Norml} \left( \text{Reshape} \left( \gamma_{(H,W,D)} \right) \right) \right) \quad (9)$$

$$\begin{aligned} \text{Attn}'^{C \times H \times W \times D} \\ = \text{Attn}_H^{C \times H \times W \times D} \otimes \text{Attn}_W^{C \times H \times W \times D} \otimes \text{Attn}_D^{C \times H \times W \times D} \end{aligned} \quad (10)$$

$$\text{Attn}^{C \times H \times W \times D} = \text{Sigmoid} \left( \text{Norml} \left( \text{Attn}'^{C \times H \times W \times D} \right) \right) \quad (11)$$

其中,  $\text{Attn}_{(H,W,D)}^{C \times H \times W \times D}$  代表每个角度下的三维权重矩阵,将三者进行逐元素相乘并进行特征数据处理,得到空间注意力权重矩阵  $\text{Attn}^{C \times H \times W \times D}$ .

### 2.3.2 多尺度自注意权重模块

对于深层网络,多次下采样操作会不可避免地令特征图尺寸大幅减小(不足原始影像10%),此时极易引起肿瘤细节信息模糊、缺失等问题.为解决此问题,本文设计了一个多尺度自注意权重模块(Multi-scale Self-attention Weight module, MSW),该模块充分利用了小尺寸特征图中的全部信息,并以此为基础生成基于不同尺度下的权重矩阵对标签区域进行加权强化,这种方法有效解决了传统注意力机制对深层小尺寸特征图处理能力有限的问题,模型具体结构如图5(b)所示.

模型的输入来自深层网络编码得来的小尺寸特征图,将特征图并行送入不同角度( $H, W, D$ )的自注意残差模块,该模块包括一个token&平展层、归一化层、Softmax函数层以及一个二维卷积残差链接.以角度 $H$ 为例,三维特征图送入自注意残差模块后被平展成一组二维特征值序列,同时记录下对应特征值原始位置以便后期尺寸还原.之后,二维特征值序列经过归一化层与Softmax函数,此时二维特征值序列变为一组与初始特征信息具有高度关联的二维权重矩阵.将二维权重矩阵与二维卷积残差链接中得到的原始二维序列对应相乘,这种操作不仅能够放大标签位置的上下文信息,还能充分保留标签的原始特征信息,有效避免信息干扰与特征过拟合问题.最后,将从不同角度得到的二维权重矩阵相乘得到多角度二维权重矩阵.将多角度二维权重矩阵整形并与原始特征图逐元素相乘得到最后的输出,输出的结果用以指导上采样时的特征恢复.该模块有效改善因线性插值采样与小尺寸特征图带来的语义信息丢失和恢复误差,具体可由下述公式表示:

$$f'_n = \text{NL}_{(H,W,D)} \left( \text{token} \left( F_n \right) \right) \quad (12)$$

$$f_n^{(H,W,D)} = \text{Softmax} \left( f'_n \right) * \text{Conv} \left( \text{token} \left( F_n \right) \right) \quad (13)$$

上式表示的是输入的三维特征图以某一角度转换

为二维权重矩阵的过程,3个不同角度的二维权重均按照上述公式得到.其中,  $f_n^{(H,W,D)}$  代表这3组不同的二维权重序列.之后,将3组权重序列相乘并整形成原尺寸三维自注意权重矩阵并与原始特征矩阵相乘,得到模型输出,具体流程如下:

$$\beta_n^{H \times W \times D} = \text{Reshape} \left( \prod \left( f_n^H, f_n^W, f_n^D \right) \right) \quad (14)$$

$$F_{\text{AttnOut}}^{B \times C \times W \times H \times D} = F_n \otimes \beta_n^{H \times W \times D} \quad (15)$$

其中,  $\beta_n^{H \times W \times D}$  代表三维自注意权重矩阵.

## 3 实验结果

### 3.1 数据处理及参数设置

本次实验采用BraTs 2021 MRI脑肿瘤数据集与LiTS 2017肝肿瘤CT三维影像数据集作为实验数据集.脑肿瘤数据集中包含1251张 $240 \times 240 \times 155$ 的脑肿瘤核磁共振影像,数据集中的标签包含flair、t1ce、t1和t2这4种模态的三维影像;肝肿瘤数据集中包含131例肝肿瘤三维CT影像,数据集中标签包含背景区域、肝脏区域与肿瘤区域,3种区域分别以0、1和2在label中进行标记.

数据集预处理中,首先对脑肿瘤数据集进行处理,将图像进行反转与随机旋转等扩增操作,扩大数据集容量的同时加强网络对各种形态肿瘤的分割能力,增强网络的抗扰能力.其次,MRI数据集存在背景区域过大,标签体素占比过低等类不平衡问题,因此在数据集部分采取对数据集边缘的黑边进行修剪,将图像裁剪为 $160 \times 160 \times 128$ ,提升有效区域在图像中的占比,有效缓解了部分标签的类不平衡导致的网络训练中止、精度不提升等问题.

在肝肿瘤数据集预处理中,首先对全部数据进行数据清理,剔除残缺与像素不匹配的影像5张;其次对影像进行中心切割操作,在减小影像大小的同时有效增加肿瘤标签区域占比;然而,数据集中各影像尺寸中深度(即 $D$ 维度)不一( $D \in [75, 598]$ ),因此在裁剪过程中,通过标签数据判断标签区域所在深度范围并裁剪,之后再通过裁剪或补零的方式将深度调整至最近的2的幂次数,以此保证网络在上采样时尺寸对齐,最终将影像大小裁剪至 $(448, 448, 2^n)$ .

最后对2种数据集进行分类,按照6:3:1的比例将其分为训练集、验证集和测试集对网络进行训练与验证.基于以上数据集,本次实验模型使用Python对模型进行构建,在NVIDIA RTX 4070 TI Super上进行训练与调试.训练时,模型的学习率设置为0.001,每次训练图像为16张,批次为2,迭代次数为100,优化器选择Adam优化器,用验证集和测试集验证与评估模型性能.

### 3.2 损失函数

在医学影像分割领域中, Dice Loss 是常用于评价网络分割效果的一种损失函数, 尤其是对多标签分割任务的图像分割效果显著, 其公式如下:

$$\text{DiceLoss} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (16)$$

其中,  $X$  为实际标签像素点,  $Y$  为对应标签的预测像素点. 但是 U 型网络特征导致的不对称监督造成了部分细节特征损失, 使得 Dice 损失函数在反向传递时在特征保留方面受到一定限制. 为解决这个问题, 本文除了设计上述的模块来缓解细节特征损失外, 在 Dice 损失函数的基础上, 引入逐层 L2 正则化损失函数<sup>[28]</sup>与加权相加 Dice, 得到新的损失函数作为本文模型训练时的损失函数.

首先提取编码器每一层输出特征图并对其进行随机通道选择, 保证得到的特征图尺寸与输出特征图经过最大池化后的特征图尺寸一致, 并将两者进行 L2 正则化计算得到逐层 L2 正则化损失函数, 具体描述如下:

$$L_{\text{ELL}} = \frac{1}{N \cdot I} \sum_{n=1}^{N-1} \sum_{i=1}^I \left\| \text{SRC}(F_n^i) - \text{MaxPool}(F_{\text{out}}) \right\|^2 \quad (17)$$

其中,  $F_n^i$  代表解码器每一层输出的特征图 ( $E_1^1, E_1^2, \dots, E_1^n$ ), SRC 表示对  $F_n^i$  在通道维度进行随机选择操作, 得到的特征图像与经过最大池化所得到的图像尺寸一致,  $F_{\text{out}}$  代表网络最终的输出特征图,  $L_{\text{ELL}}$  则表示上述两者经由 L2 正则化计算后得到的损失函数.

此外, 为了充分利用浅层特征信息从而解决信息利用率问题, 在损失函数中增加不同深度特征图的正则化损失函数来平衡由于反复卷积所带来的细节特征信息丢失问题. 将输出特征图浅通道特征与深通道特征分开后进行 L2 正则化计算, 这种计算方式能够有效地将浅层特征信息用于深层特征信息分割指导上, 从而提升浅层信息利用率并有效解决由于不对称监督学习所带来的上下文信息不对称的问题, 具体表示如下:

$$L_{\text{CQR}} = \frac{1}{N \cdot I} \sum_{n=1}^N \sum_{i=1}^I \left\| \hat{F}_n^i - \overline{F}_n^i \right\|^2 \quad (18)$$

其中,  $F_n^i$  代表第  $n$  层的每一张输出特征图,  $\hat{F}_n^i$  代表深通道特征,  $\overline{F}_n^i$  代表浅通道特征,  $L_{\text{CQR}}$  表示两者经由 L2 正则化计算后得到的具有浅层特征指导信息的损失函数. 通过这种方式能有效解决信息利用率和解决信息不对称的问题.

基于上述公式, 得到最终加权正则化损失函数具体表示如下:

$$L_{\text{DWL}} = \text{DiceLoss} + \alpha_1 L_{\text{ELL}} + \alpha_2 L_{\text{CQR}} \quad (19)$$

其中,  $\alpha_1$  与  $\alpha_2$  均为可调节权重, 可随反向传播效果进行调节.  $L_{\text{DWL}}$  表示经由 DiceLoss 与前两者加权得到的最终损失函数.

### 3.3 实验结果

#### 3.3.1 对比试验

2 组数据集中共包含 5 种标签, 分别为脑肿瘤中的肿瘤区域 (WT)、肿瘤核心 (TC) 以及增强肿瘤 (ET) 和肝肿瘤中的肝脏 (Liver) 和肿瘤区域 (Tumor). 因此, 在对网络模型的分割性能进行评价时, 不仅要评价网络分割的精准度进行对比, 还应该计算数据集中每种标签的 Dice 分数以反映网络对肿瘤各部分的分割能力, 具体对比实验数据见表 1, 加粗数据为本文结果.

实验数据表明, 本文提出的 DMR-KAN 在精准度与各标签的 Dice 分数 2 种评价指标中均有显著提升. 在脑肿瘤数据集中, 网络精准度相较于传统注意力机制网络 Attention-Unet 与最新双通道网络 DE-Unet 分别提升了 8.75 个百分点与 6.05 个百分点. 其次, 实验中将本文网络与较新的 Swin-Unet、U-KAN、SegMamba 进行对比, 不仅在精准度上相较于新网络有一定提升, 而且在各标签的 Dice 分数上也有显著提升, 以分割难度最大的肿瘤核心区域 (Tumor Core, TC) 的 Dice 分数为例, 本文网络相较于三者分别提升了 4.89、1.62 和 0.52 个百分点, 而与本文的基础网络 (base model)——U-KAN 网络相比, 精准度上提升了 3.25 个百分点, 各标签的 Dice 分

表 1 DMR-KAN 与其他模型的对比实验数据

单位: %

模型	BraTs 2021				LiTS 2017		
	精准度	Dice 分数			精准度	Dice 分数	
		ET	WT	TC		Liver	Tumor
nn-unet(2021) <sup>[29]</sup>	77.67	77.09	79.41	76.33	76.66	80.82	72.49
Attention-Unet(2021) <sup>[30]</sup>	77.79	77.64	78.91	76.13	78.56	82.96	74.16
Trans-Unet(2024) <sup>[31]</sup>	78.83	78.06	79.67	76.57	80.74	84.79	76.68
DE-Unet(2023) <sup>[32]</sup>	80.49	78.94	79.91	77.17	82.67	86.17	79.17
Swin-Unet(2023) <sup>[33]</sup>	82.55	79.21	83.06	80.39	85.88	88.84	82.91
U-KAN(2024)	83.29	80.29	85.94	83.66	86.71	89.67	83.75
SegMamba(2024) <sup>[34]</sup>	84.37	80.67	86.90	84.76	87.35	90.73	83.96
<b>DMR-KAN</b>	<b>86.54</b>	<b>83.67</b>	<b>88.79</b>	<b>85.28</b>	<b>88.07</b>	<b>91.36</b>	<b>84.77</b>

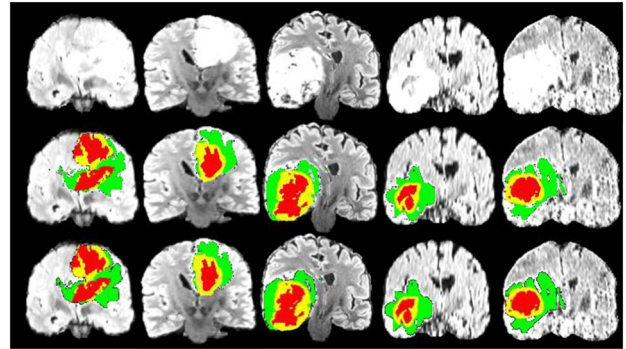
数分别提升了3.38、2.85和1.62个百分点。

在肝肿瘤数据集中,尽管数据集容量仅有126张,但本文网络依旧表现出了良好的分割效果。首先在精准度方面,本文网络相较于U-KAN与SegMamba,分别提升了1.36个百分点和0.72个百分点,与Swin-UNET相比更是提升了2.19个百分点,这反映出本文网络对于小数据集的学习能力与泛化能力的优越性。其次对于各标签Dice分数,本文模型对肝脏与肝肿瘤的分割精度在对比模型中达到了最高,分别是91.36%、84.77%。相比于U-KAN模型提升了1.69个百分点和1.02个百分点。

通过数据分析可以得出,与单通道输入的新网络相比(U-KAN等),本文提出的三维多角度特征提取KAN模块与多尺度残差卷积模块组成的双通道输入,有效提升了网络特征提取能力,很大程度上改善了KAN模型特征提取模型单一,对于原始图像的细节特征利用不足等问题。而与新型多角度特征提取网络SegMamba相比,本文模型在精准度与各项Dice分数上依旧具有一定优势,分析原因为本文所使用的注意力机制有效改善了网络上采样恢复过程中的部分问题。此外,从各标签Dice分数也可以看出,本文提出的针对不同深度的注意力机制,有效改善了网络在跳跃链接中对解码特征图像信息利用率低、线性插值的局限性、肿瘤标签定位不准确、肿瘤边缘特征不规律、肿瘤细节信息提取不足等问题带来的肿瘤核心部分分割效果差的问题,很大程度上改善了网络对标签细节部分的提取能力不足的问题,大幅提升了网络的精准度与分割能力。

上述对比实验验证了本方法的性能优势,为进一步展示分割效果,本文选择了标签种类更多且模态更为复杂的脑肿瘤影像可视化分割结果与特征效果图,如图6与图7所示。在脑肿瘤分割结果图中,第一行为原始影像切片,第二行为标签影像对应位置切片,第三行为分割效果图对应位置切片,切片位置选择包含标签信息更多的肿瘤中心位置,这样更有利于展示网络的分割能力。具体来说,网络对肿瘤整体的分割程度与原标签相比,整体形状与标签边缘的曲线走向相似。而且对于部分边缘细节部分,例如第一列中绿色标签上方的小突出,本文网络能够有效识别并进行边缘标注。但对于边缘区域复杂的标签部分,例如第五列中标签右下部分,这种边缘走向复杂且无规则的区域网络的分割能力有所欠缺,但是依旧能够对边缘部分的走向进行准确预测。

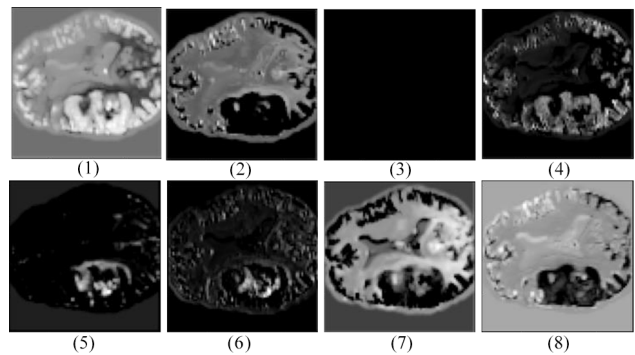
为了直观展现网络分割的具体过程,特征效果图选择了浅层网络输出的特征图,这样既能够直观地表现出网络对各标签的分割过程,又能展现出网络在分割时对不同模态特征图的处理方式。从图7可以看出,网络在对特征提取时,不同通道根据不同模态的特征



注:第一行为原始影像切片,第二行为标签影像,第三行为分割效果图。

图6 DMR-KAN网络分割脑肿瘤效果

图所展现出的独特特征对肿瘤区域进行分割,从特征图中可以清晰观察到,浅层网络对不同模态(T1、T2、Flair)的响应具有显著差异性。例如,T2加权像的高信号区域在特定通道中被强化为亮色特征,而Flair序列的病灶边缘则在其他通道呈现环状激活模式。但是在这些特征图中发现,部分通道会出现无任何特征的全黑特征图,通过对网络进行分析发现,这种情况出现的原因与输出通道数的超参数设置有关,网络最终输出图像的通道数为3,且在计算损失函数时并未计算标签为0的区域,这使得网络在编码过程中将除肿瘤外的部分均识别为背景,所以会出现全黑特征图的情况。



注:T2标签分类如图中(2)和(4),T1标签如图中(5)和(6),Flair标签如图中(1)和(4),其余图像以不同增强方式区分各标签。

图7 DMR-KAN网络分割脑肿瘤特征可视化

综合上述所有实验的数据与分析,本文提出的改进网络与U-KAN网络相比,在精准度与细节分割能力都有显著提升。

### 3.3.2 消融实验

为了验证本文提出模块的对网络的精准度与分割能力提升的有效性与合理性,本文设计了相关的消融实验。对本文提出的3个模块分别进行了不同类型的验证实验。首先,本文对3种改进模块进行了消融实验,将3种模型分别添加到原始模型中得出相关评价指标并进行对比,具体实验的验证模型与评价指标见表2,

其中加粗数据为最优结果。

通过实验对比,本文提出的3种模块与初始网络相比,在2组数据集中精准度分别提升了3.25个百分点和1.36个百分点.在脑肿瘤数据集中,各标签的Dice分数提升了3.38、2.85和1.62个百分点.其中,当只添加MFE模块时,在精度与各标签的Dice分数上提升并不明显,但是在添加了卷积双通道输入后,虽然网络精准度提升了0.16个百分点,但是ET与TC的Dice分数却有所下降.根据原理分析,卷积网络对占比较大标签的细节部分处理能力要优于小标签,这一问题在肝肿瘤的实验数据中也有体现.就小标签而言,随着卷积编码层数深入,特征图尺寸下采样变小导致细节信息丢失严重,此时合适的注意力机制对于小占比标签的细节恢复起到了尤为重要的作用.因此在添加了注意力机制之后,在2组数据集中不仅网络精度分别提升了0.77个百分点和0.28个百分点,而且脑肿瘤数据集中3种标签与肝肿瘤中的2种标签的Dice分数均有大幅提升.与基础网络相比各项对比指标提升明显,有效说明了本文设计的注意力模块有效改善了初始网络存在的部分问题,尤其是网络对于占比极小的标签(如TC和Tumor)分割能力差所导致的整体精准度的问题,大幅提升了网络的精准度与分割能力.

表2 基于原始KAN网络添加本文改进模块后的评价指标数值  
单位: %

改进模块			BraTs 2021			LiTS 2017			
MFE-KAN	MRC	MSR	精准度	Dice分数			精准度	Dice分数	
				ET	WT	TC		Liver	Tumor
U-KAN (base model)			83.29	80.29	85.94	83.66	86.71	89.67	83.75
√			85.61	82.83	87.95	84.71	87.55	90.44	84.26
√	√		85.77	82.59	88.19	84.65	87.79	90.88	83.81
√	√	√	<b>86.54</b>	<b>83.67</b>	<b>88.79</b>	<b>85.28</b>	<b>88.07</b>	<b>91.36</b>	<b>84.77</b>

其次,为了验证本文提出的2种分层注意力机制对网络的影响,在消融实验部分,本文对2种注意力模块的使用场景与可调节参数 $K$ 的数值设置分别进行了消融实验,通过在不同深度的跳跃链接中添加注意力机制与本文网络进行消融实验,对比相关改动网络的精准度与各标签Dice分数,从而验证分层注意力机制设计的有效性与必要性.除了验证注意力机制的有效性外,在网络训练的过程中发现,在双通道改进网络中添加本文的新型注意力机制后,网络的训练速度也有所提升.因此在对比试验部分,本文添加了一个新型注意力模块对网络迭代次数影响的对比试验.具体网络对比模型与评价指标见表3~表5,其中表3和表5中加粗数据为最优结果.

表3 2种注意力模块使用在不同深度下的网络性能指标 单位: %

模型	BraTs 2021				LiTS 2017		
	精准度	Dice分数			精准度	Dice分数	
		ET	WT	TC		Liver	Tumor
仅使用MWFA	85.69	82.75	87.35	84.24	87.53	90.47	84.33
仅使用MRC+MSW	85.16	82.16	87.07	83.85	86.91	89.87	83.26
DMR-KAN	<b>86.54</b>	<b>83.67</b>	<b>88.79</b>	<b>85.28</b>	<b>88.07</b>	<b>91.36</b>	<b>84.77</b>

表4 添加注意力模块对网络训练及相关参数的影响

指标	MFE-KAN (BraTs)	DMR-KAN (BraTs)	MFE-KAN (LiTS)	DMR-KAN (LiTS)
最优精度/%	85.61	86.54	87.55	88.07
训练代数/ epoch	100	80	350	200
训练时间/h	23.34	16.57	17.5	10.0
参数量/M	3.29	3.85	3.29	3.85
运算次数/ GFLOPS	276.81	316.05	27.64	32.34

表5 不同 $K$ 值下MSR对网络分割性能的影响 单位: %

参数 ( $K \in \text{GCF}(H, W, D)$ )	BraTs 2021				LiTS 2017		
	精准度	Dice分数			精准度	Dice分数	
		ET	WT	TC		Liver	Tumor
$K=2$	85.27	81.59	87.47	84.65	85.91	87.49	81.71
$K=4$	85.67	82.21	86.90	84.70	87.16	89.77	83.36
$K=8$	<b>86.54</b>	<b>83.67</b>	<b>88.79</b>	<b>85.28</b>	<b>88.07</b>	<b>91.36</b>	<b>84.77</b>
$K=16$	84.31	81.67	85.71	83.48	75.61	81.29	70.37

实验中的评价指标反映出,针对不同深度的跳跃链接添加相应的注意力机制(MWFA与MSW),有效提升了网络对图像中肿瘤各标签的分割能力.在脑肿瘤数据集中,本文网络相较于使用单一注意力模块在精准度上平均提升了0.85个百分点与1.38个百分点;在标签TC部分中的Dice分数提升了1.04个百分点与1.43个百分点.值得注意的是,相较于没有添加注意力机制的双通道网络(MFE-KAN+MRC),添加MSW后评价指标整体有所下降,经分析得出,原因是MSW的自注意机制更适用于深层的小尺寸特征图对上采样特征图进行加权与监督,而在浅层中使用该模块,会导致大量的无用信息(背景标签)影响自注意权重,从而使来自编码器的特征图像在跳跃链接中失去指导意义,导致跳跃链接环节对上采样特征图产生错误监督的问题,使得添加注意力机制的网络在分割能力上与初始网络相比较差.

在肝肿瘤数据集中,网络性能情况与脑肿瘤中的表现一致.本文网络相较于单一注意力网络均有提升,而在没有添加注意力机制的双通道网络(MFE-KAN+MRC)中添加MSW后,相较于MFE-KAN+MWFA的单通道组合评价指标整体略有下降,根据模块原理与现

有数据分析原因发现,两种数据集中标签占比相似(预处理后脑肿瘤中占比 5.56%,肝肿瘤中占比 6.5%),标签的三维形态具有相似的特征信息,例如两者均为正常组织包裹肿瘤区域.因此可以得知在肝肿瘤数据集中的数据结果和下降原因与脑肿瘤一致.同时,从表 4 中可以发现,相较于原模型,添加注意力机制后模型参数量与计算量分别提升了 19.48% 与 14.49%,但同时也不难发现,本文网络在脑肿瘤数据集中训练至 80 个 epoch 时达到模型参数最优解且训练时间相较于原模型降低 29.74%,而在肝肿瘤数据集中训练 200 个 epoch 达到最优精度且训练时间降低了 42.86%.

通过比较不难得出,本文提出的新型注意力机制有效减少了模型的训练时间,经过对模型原理进行分析得出,产生这一结果的原因是因为在解码层部分,上采样图像得到了更为有效的特征指导权重矩阵,令上采样中尺寸恢复环节的数据插入更为合理且符合原始图像的特征信息分布.这一改动使得网络第一代的精度得到提升,且在反向传播时,无训练参数的注意力机制不会影响反向传播的计算速度.两者综合影响,使得网络的训练速度得到提升.

而添加注意力机制后可以发现,网络在肝肿瘤数据集中的表现要优于脑肿瘤数据集,这种现象的出现与数据集之间的不同有一定关系.LiTS2017 经过预处理后仅剩 126 张,属于小样本数据集,因此单次训练所需代数远大于脑肿瘤数据集.但是通过达到最优模型所需时间不难看出,网络整体训练时间仅为 10 h 且精度达到 88.07%,证明本文模型在标签占比极低的小数

据集上也展现出了良好的分割效果.

本实验环节在最佳训练代数后继续对模型进行了训练,而此时出现精度下降现象,这种情况的产生,极有可能是模型因过度训练而产生的过拟合现象使得网络精度开始下降.在合适的训练次数及时停止训练即可解决此问题.综合以上结果与分析,本文提出的注意力模块对网络的精度与训练参数的反向传播速度具有一定的优化作用.

至于注意力模块中超参数  $K$  的设置,从表 5 中不难看出,直接影响网络对肿瘤整体的分割能力,这一提升具体反映在网络的精准度上.首先对于  $K$  值的选择,在前文中已经提到, $K$  应为三维影像尺寸的公因数.

对于本文而言, $K$  值的选择则在 2、4、8、16 中产生.而由于注意力模块中的平均池化操作将  $K$  大小的三维特征池化成一个二维权重数值,因此当  $K$  值选取不合适时,极易导致因  $K$  值选择过大使得部分权重空间中由于标签区域占比低,影响区域信息的增强效果,最终导致网络整体精度降低,甚至网络分割表现弱于无注意力机制的网络,即 MFE-KAN+MRC;或因  $K$  值选择过小使得标签边缘区域权重不够,不足以令边缘信息得到强化,从而影响网络分割性能.试验结果表明,当  $K=8$  时,精准度达到最高的 86.54% 与 88.07%,且各标签的 Dice 分数指标表现良好,根据此实验结果,本文将超参数  $K$  的值设置为 8.最后,为了验证本文提出的多尺度卷积通道对网络的性能提升效果,实验中将本文模型与多组双通道卷积模型进行了对比,具体实验数据如表 6 所示,其中加粗数据为最优结果.

表 6 不同形态下的多尺度卷积模块对分割性能的影响

单位: %

模型				BraTs 2021			LiTS 2017			
	3*3*3	5*5*5	7*7*7	精准度	Dice 分数			精准度	Dice 分数	
					ET	WT	TC		Liver	Tumor
MRC	√	√		85.63	82.11	87.46	84.49	86.91	89.79	84.33
	√		√	84.19	81.34	86.28	83.79	85.45	88.24	83.53
	√	√	√	<b>86.54</b>	<b>83.67</b>	<b>88.79</b>	<b>85.28</b>	<b>88.07</b>	<b>91.36</b>	<b>84.77</b>
	无权重			77.63	76.47	78.28	76.65	79.29	82.26	76.34
DE-Net				80.49	78.94	79.91	77.17	82.67	86.17	79.17

通过实验数据可以发现,本文提出的多尺度残差卷积模块与目前新型双通道卷积模块(DE-Net)对比中具有一定领先,在 2 种数据集中的精准度与各标签 Dice 分数上均领先较多,证明了本模型的创新性与可行性.在横向对比实验的数据中可以看出,添加了动态调节权重的精度与各标签的 Dice 分数相较于无权重模块有明显提升,且前文中提及的 3 种卷积核间的联动与互补有效增加了本模块对数据中三维细节特征的提取能力.根据上述分析结果,本文为多尺度卷积块设计的动态权重有效改善了网络整体的分割性能.

## 4 结论

在本文中,我们提出了适用于医学三维肿瘤影像分割的 DMR-KAN 网络模型.在编码器部分,使用三维多角度特征提取 KAN 模块与多尺度残差卷积模块并行的双通道编码器对原始特征图进行编码,实现了标签特征的高效利用,解决了网络对特征信息提取方式单一、细节信息提取能力弱等问题,提升了网络对三维影像整体的分割能力.此外,根据不同深度下编码器输出特征图的不同尺寸,针对性地设计了多角度加权融合

注意力模块与多尺度自注意权重模块. 有效提升了浅层网络对标签位置的定位加权能力与深层网络对标签位置信息的提取能力,改善了网络对占比极小标签(如本文数据集中的TC标签)的分割能力,同时提升网络编码层输出特征图对解码层上采样图像的指导能力. 在2种数据集中的实验表明,DMR-KAN对三维肿瘤影像的分割具有良好效果,同时对小样本数据集与肿瘤检查中常见模态的数据集分割任务具有出色的泛化性,尤其对不同模态数据集中微小标签的识别效果显著,未来将会从网络整体结构、多尺度卷积通道、多模态输入等方面对网络进行优化,以提升泛化性与精度.

### 参考文献

- [1] 郑光远, 刘峡壁, 韩光辉. 医学影像计算机辅助检测与诊断系统综述[J]. 软件学报, 2018, 29(5): 1471-1514.  
ZHENG G Y, LIU X B, HAN G H. Survey on medical image computer aided detection and diagnosis systems[J]. Journal of Software, 2018, 29(5): 1471-1514. (in Chinese)
- [2] 陈正阳, 陈平华, 陈建平. 一种改进的Unet医学影像分割模型[J]. 计算机仿真, 2024, 41(6): 216-222.  
CHEN Z Y, CHEN P H, CHEN J P. An improved unet medical image segmentation model[J]. Computer Simulation, 2024, 41(6): 216-222. (in Chinese)
- [3] 刘金平, 吴娟娟, 张荣, 等. 基于结构重参数化与多尺度深度监督的COVID-19胸部CT图像自动分割[J]. 电子学报, 2023, 51(5): 1163-1171.  
LIU J P, WU J J, ZHANG R, et al. Toward automated segmentation of COVID-19 chest CT images based on structural reparameterization and multi-scale deep supervision[J]. Acta Electronica Sinica, 2023, 51(5): 1163-1171. (in Chinese)
- [4] CUI M T, LI K, CHEN J Y, et al. CM-unet: A novel remote sensing image segmentation method based on improved U-net[C]//IEEE Access. Piscataway: IEEE, 2023: 56994-57005.
- [5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [6] WANG M H, YUAN P, QIU S, et al. Dual-encoder UNet-based narrowband uncooled infrared imaging denoising network[J]. Sensors, 2025, 25(5): 1476.
- [7] GUO M H, XU T X, LIU J J, et al. Attention mechanisms in computer vision: A survey[J]. Computational Visual Media, 2022, 8(3): 331-368.
- [8] TAY Y, DEGHANI M, BAHRI D, et al. Efficient transformers: A survey[J]. ACM Computing Surveys, 2022, 55(6): 1-28.
- [9] 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.
- [10] VARSHANEYA V, BALASUBRAMANIAN S, DARSCHAN G. Res-SE-Net: Boosting performance of ResNets by enhancing bridge connections[M]//Machine Learning Algorithms and Applications, Hoboken: Wiley, 2021: 61-75.
- [11] DAO T, FU D Y, ERMON S, et al. FlashAttention: Fast and memory-efficient exact attention with IO-awareness [EB/OL]. (2022-06-23) [2025-05-10]. <https://arXiv.org/abs/2205.14135>.
- [12] LIU Z M, WANG Y X, VAIDYA S, et al. KAN: Kolmogorov-Arnold networks[EB/OL]. (2025-02-09) [2025-05-10]. <https://arXiv.org/abs/2404.19756>.
- [13] MORADZADEH A, WAWRZYNIAK L, MACKLIN M, et al. UKAN: Unbound Kolmogorov-Arnold network accompanied with accelerated library[EB/OL]. (2024-10-01) [2025-05-10]. <https://arXiv.org/abs/2408.11200>.
- [14] LOU A G, GUAN S Y, LOEW M H. DC-UNet: Rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation[C]//Medical Imaging 2021: Image Processing. California: SPIE, 2021: 98.
- [15] LI C X, LIU X Y, LI W Y, et al. U-KAN makes strong backbone for medical image segmentation and generation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(5): 4652-4660.
- [16] MENG W J, LIU S J, WANG H J. AFC-Unet: Attention-fused full-scale CNN-transformer unet for medical image segmentation[J]. Biomedical Signal Processing and Control, 2025, 99: 106839-106849.
- [17] FANG C L, WU K G. KANDU-Net: A dual-channel U-Net with KAN for medical image segmentation[EB/OL]. (2024-09-30) [2025-05-10]. <https://arXiv.org/abs/2409.20414>.
- [18] AZAD R, NIGGEMEIER L, HÜTTEMANN M, et al. Beyond self-attention: Deformable large kernel attention for medical image segmentation[C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2024: 1276-1286.
- [19] WANG C, WANG L, WANG N Q, et al. CFATransUnet: Channel-wise cross fusion attention and transformer for 2D medical image segmentation[J]. Computers in Biology and Medicine, 2024, 168: 107803.
- [20] REN J T, HOCHREUTER K, KALLEHAUGE J F, et al. UMamba adjustment: Advancing GTV segmentation for head and neck cancer in MRI-guided RT with UMamba and NnU-net ResEnc planner[C]//Head and Neck Tumor Segmentation for MR-Guided Applications. New York:

ACM, 2025: 123-135.

- [21] LIU J R, YANG H, ZHOU H Y, et al. Swin-UMamba: Mamba-based UNet with ImageNet-based pretraining[C]// Medical Image Computing and Computer Assisted Intervention-MICCAI 2024. New York: ACM, 2024: 615-625.
- [22] ATEK S, MEHIDI I, JABRI D, et al. SwinT-Unet: Hybrid architecture for medical image segmentation based on swin transformer block and dual-scale information[C]// 2022 7th International Conference on Image and Signal Processing and Their Applications. Piscataway: IEEE, 2022: 1-6.
- [23] LIU A L, JIA D Q, SUN K C, et al. LM-UNet: Whole-body pet-CT lesion segmentation with dual-modality-based annotations driven by Latent mamba U-Net[C]// Medical Image Computing and Computer Assisted Intervention-MICCAI 2024. Cham: Springer, 2024: 405-414.
- [24] ZHANG R H, ZOU R Z, ZHAO Y, et al. BA-Net: Bridge attention in deep neural networks[J]. Expert Systems with Applications, 2024,292: 128525-128538.
- [25] CHEN J D, SHEN X Y, ZHAO Y, et al. Attention gate and dilation U-shaped network (GDUNet): An efficient breast ultrasound image segmentation network with multi-scale information extraction[J]. Quantitative Imaging in Medicine and Surgery, 2024, 14(2): 2034-2048.
- [26] 严武军, 李建昌, 叶金霞. 轻量化GEC-3DUnet 骶髂关节影像自动分割[J]. 计算机技术与发展, 2025, 35(2): 33-40. YAN W J, LI J C, YE J X. Lightweight GEC-3DUnet based automatic segmentation of sacroiliac joint medical imaging[J]. Computer Technology and Development, 2025, 35(2): 33-40. (in Chinese)
- [27] ZHONG H, WU C. T-UNet: Triplet UNet for change detection in high-resolution remote sensing images[J]. Geospatial Information Science, 2025,28(2):437-454.
- [28] KIM D, YOO Y, PARK S, et al. SelfReg: Self-supervised contrastive regularization for domain generalization[C]// 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2022: 9599-9608.
- [29] ISENSEE F, JAEGER P F, KOHL S A A, et al. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation[J]. Nature Methods, 2021, 18(2): 203-211.
- [30] TONG X Z, WEI J Y, SUN B, et al. ASCU-Net: Attention gate, spatial and channel attention U-Net for skin lesion segmentation[J]. Diagnostics, 2021, 11(3): 501.
- [31] CHEN J N, MEI J R, LI X H, et al. TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers[J]. Medical Image Analysis, 2024, 97: 103280.
- [32] KARIMI A, FAEZ K, NAZARI S. DEU-Net: Dual-encoder U-Net for automated skin lesion segmentation[J]. IEEE Access, 2023, 11: 134804-134821.
- [33] CAO H, WANG Y Y, CHEN J, et al. Swin-Unet: Unet-like pure Transformer for medical image segmentation [M]//Computer Vision-ECCV 2022 Workshops. Cham: Springer Nature Switzerland, 2023: 205-218.
- [34] XING Z H, YE T, YANG Y J, et al. SegMamba: Long-range sequential modeling mamba for 3D medical image segmentation[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2024. Cham: Springer, 2024: 578-588.

## 作者简介



方丁毅 男, 1999年6月生, 黑龙江嫩江人. 青岛科技大学自动化与电子工程学院硕士研究生. 主要研究方向为人工智能与机器视觉.  
E-mail: 2682758035@qq.com



程焕新 男, 1966年9月生, 安徽安庆人. 青岛科技大学自动化与电子工程学院教授, 硕士生导师. 主要研究方向为机器视觉、智能控制系统、深度学习理论及应用等.  
E-mail: 2718228346@qq.com



骆晓玲 女, 1966年5月生, 山东菏泽人. 青岛科技大学机电工程学院教授、硕士生导师. 主要研究方向为过程装备及自动化的优化设计、深度学习参数优化与过程控制.  
E-mail: luoxiaoling@qust.edu.cn



王若峰 女, 1964年1月生, 辽宁营口人. 博士后、教授、主任医师(二级)、博士生导师, 国务院政府特殊津贴专家. 主要研究方向为肿瘤放疗基础及临床.  
E-mail: wrz8526@vip.163.com