

DOI:10.3969/j.issn.1005-202X.2026.07.013

医学影像物理

MRFA-Net: 一种基于多尺度残差特征挖掘的肺癌 CT 图像分割网络

张奔腾¹, 王尧芬¹, 胡凌燕¹, 王海玲¹, 宫晓梅²

1. 上海工程技术大学电子电气工程学院, 上海 201620; 2. 同济大学附属上海市肺科医院放疗科, 上海 200433

【摘要】医学影像中的肺部肿瘤分割对于辅助诊断和治疗规划具有重要意义。然而,由于肿瘤在形态、边界和大小上的高度异质性,精准分割仍然具有较大挑战。提出一种基于残差特征挖掘与深层特征融合的深度分割网络MRFA-Net,旨在提升肺肿瘤在CT图像中的自动分割性能。提出多尺度残差特征提取模块和深层特征聚合模块,通过多尺度上对残差特征进行提取以及深层特征融合,更有效地挖掘出具有代表性的肿瘤特征,并且在跳跃连接中引入反向注意力机制以提升模型对关键区域的感知能力。在MSD公开数据集和私有肺癌CT数据集上进行的实验结果表明MRFA-Net具备优异性能,Dice相似系数分别达到74.15%和74.77%。该方法为肺部肿瘤分割提供有效的解决方案,具备良好的临床应用潜力。

【关键词】肺部肿瘤分割;残差特征提取;深层特征聚合;反向注意力机制

【中图分类号】R318;TP391

【文献标志码】A

【文章编号】1005-202X(2026)07-0921-09

MRFA-Net: a lung cancer CT image segmentation network based on multi-scale residual feature mining

ZHANG Benteng¹, WANG Raofen¹, HU Lingyan¹, WANG Hailing¹, GONG Xiaomei²

1. School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China; 2. Department of Radiation Therapy, Shanghai Pulmonary Hospital Affiliated to Tongji University, Shanghai 200433, China

Abstract: Lung tumor segmentation from medical images is crucial for supporting diagnosis and treatment planning. However, tumors exhibit substantial heterogeneity in shape, contour and size, which makes precise segmentation still challenging. Therefore, this paper proposes a deep segmentation network named MRFA-Net (multi-scale residual feature aggregation network). Built upon residual feature mining and deep feature fusion, the network aims to improve the automatic segmentation performance of lung tumors on CT images. Specifically, two key modules are designed: the multi-scale residual feature extraction module and the deep pyramid aggregation module. By extracting residual features at multiple scales and fusing deep features, the network can capture representative tumor characteristics more effectively. Additionally, a reverse attention mechanism is integrated into skip connections to enhance the model's perception of critical regions. Experimental results on the public MSD dataset and a private lung cancer CT dataset demonstrate that MRFA-Net achieves excellent performance, with a Dice similarity coefficient of 74.15% and 74.77%, respectively. The proposed method provides an effective solution for lung tumor segmentation and holds promising prospects for clinical practice.

Keywords: lung tumor segmentation; residual feature extraction; deep pyramid aggregation; reverse attention mechanism

【收稿日期】2026-03-21

【基金项目】上海市科委科技创新行动计划(23010501700);江西省卫生健康重点科技项目(2023ZD008);申康三年行动计划肺科培育项目(SKPY2021006)

【作者简介】张奔腾,硕士研究生,研究方向:医学图像处理,E-mail: 781348902@qq.com

【通信作者】王尧芬,博士,副教授,研究方向:医学图像处理、脑机接口,E-mail: rfwangsues@163.com

前言

在癌症临床诊疗中,对肿瘤的高效标注已成为提升癌症早期诊断与精准治疗的关键,能够显著优化临床决策流程。标注肿瘤区域后,结合基因检测数据(如EGFR突变状态),可建立影像-基因组学关联模型^[1-3],指导靶向治疗。传统手动标注高度依赖医师经验,存在耗时冗长(单病例标注需30 min以上)、主观性偏差(不同医师标注结果JacCard指数差

异可达20%)及大规模筛查场景下可扩展性不足等固有缺陷。自动分割技术通过算法驱动实现标注效率的指数级提升,同时降低人工干预引入的系统误差,尤其在处理多模态、高分辨率影像数据时展现出显著优势。

近年来,深度学习技术重塑了肿瘤分割范式。卷积神经网络(CNN)凭借局部感知与层次化特征提取能力,在医学图像分析中奠定基础^[4-6]。通过堆叠卷积与池化层^[7],CNN可逐级提取从边缘纹理到器官结构的抽象特征,但其原生结构更偏向分类任务,在像素级分割中存在感受野受限与空间信息丢失问题。2015年Ronneberger等^[4]提出U-Net架构,将编码器-解码器结构与跳跃连接结合,实现多尺度特征融合与精确定位^[8]。此后围绕U-Net的改进不断推进,ResU-Net引入残差模块缓解网络退化^[9],使肝脏肿瘤分割Dice相似系数(Dice Similarity Coefficient, DSC)提升至0.92;DenseU-Net采用密集连接策略增强浅层特征复用能力,提高微小肺结节边界表达效果。注意力机制亦被广泛引入,Attention U-Net在跳跃连接中嵌入空间-通道注意力门控^[10-12],使胰腺癌分割DSC由0.78提升至0.83;PraNet通过反向注意力机制在Kvasir-SEG数据集上实现IOU为0.898,较基线提升9.6%^[13]。随着Transformer在自然语言处理领域取得突破,Vision Transformer(ViT)被引入视觉任务^[14],其长程依赖建模能力可弥补CNN局部感受野不足。Swin-U-NETR在编码阶段引入Swin Transformer提取全局上下文^[15],在BraTS2023中脑肿瘤分割DSC达0.91;TransFuse采用CNN-Transformer双分支结构,在肝脏肿瘤分割中IOU达0.78^[16]。然而,ViT仍面临计算复杂度高(如DeiT-Base达17.6 GFLOPs)及医学数据规模受限等问题^[17]。

因此,本文提出融合残差特征挖掘与深层特征聚合策略的肺肿瘤CT图像分割网络MRFA-Net(Multi-scale Residual Feature Aggregation Network)。在编码阶段构建多尺度残差特征提取(Multi-scale Residual Feature, MRF)模块,结合残差结构与双空洞率膨胀卷积,通过动态门控机制自适应融合多尺度特征:低空洞率卷积捕获局部细节,高空洞率卷积保持边界信息,实现对肺肿瘤多尺度结构的精细表征。残差路径中引入通道注意力机制以抑制冗余通道响应,增强对复杂结构区域的表达能力。瓶颈层设计深层特征聚合(Deep Pyramid Aggregation, DPA)模块,舍弃低级特征以减少干扰,并采用金字塔卷积提取不同感受野下的高语义特征,通过逐元素融合强化关键区域响应,同时结合空间注意力实现显著区域动态加权,提升细粒度特征建模能力。

1 方法

1.1 网络架构

图1为MRFA-Net的网络结构图。在网络的编码器部分,通过MRF模块逐级提取相关性强的肿瘤特征,多尺度的特征提取旨在使模型能够学习不同感受野的肿瘤特征,从而增强其对异质性肿瘤特征的捕捉能力。4个层次的特征从上到下分为 $\{f_i^j, i=1, 2, \dots, 4\}$,其中 $\{f_i^j, j=1, 2\}$ 特征为低级特征, $\{f_i^j, j=3, 4\}$ 特征为高级特征。在网络深层阶段,通过引入DPA模块并行聚合高级语义信息特征,这一阶段生成的细节信息丰富的特征图将作为上采样的输入。在每层的跳跃连接中还引入反向注意力机制(Reverse Attention, RA),通过跨域协同更充分地抑制无关区域,减少信息冗余对分割性能的影响。

1.2 MRF模块

本文提出一种基于残差连接思想的多尺度特征提取编码器结构,采用两种空洞率的膨胀卷积和动态门控融合机制,其中低空洞率卷积侧重捕获肿瘤区域的纹理细节,高空洞率卷积则有效保持目标边界信息,结合动态门控机制的筛选从而实现肺肿瘤CT图像多尺度特征的精准表征。在残差连接中结合通道注意力机制抑制低效通道中的信息,进一步增强网络对肺肿瘤CT图像中边界高度异质的表达能力,有效优化传统U型网络在处理尺寸差异显著的目标区域时因适应性不足导致的分割性能下降。

如图2所示,通过引入选择性卷积机制(Selective Kernel, SK)实现对不同感受野特征的自适应调节。具体来说,对于给定的输入特征图 $E_j \in R^{C \times H \times W}$,本文设计一组核大小分别为5和7的卷积核对其进行并行卷积。在同一特征下捕获不同感受野下的肿瘤特征,并将这些不同尺度卷积生成的特征进行连接。随后将这些分支的输出堆叠并通过软注意力机制进行加权融合,得到不同感受野的权重系数:

$$\alpha_i = F_{gp}(\Sigma(\text{Conv}_{k_i}(E_j))), i = 1, \dots, N \quad (1)$$

其中, k_i 表示不同尺度的卷积核, N 为分支数目。权重系数 α_i 由输入图像的全局平均池化向量经过非线性映射获得,体现了输入内容对不同感受野的偏好。选择卷积核可以显著提升网络对肺部肿瘤在不同空间尺度下结构表达的能力。

为了进一步提升模型对肺肿瘤区域的判别能力,实现在通道维度筛选出对肿瘤表达最具区分度的特征,在残差分支上融合使用CBR(Convolution-Batch Normalization-ReLU)块与高效通道注意力机制(Efficient Channel Attention, ECA)^[15]。CBR块由

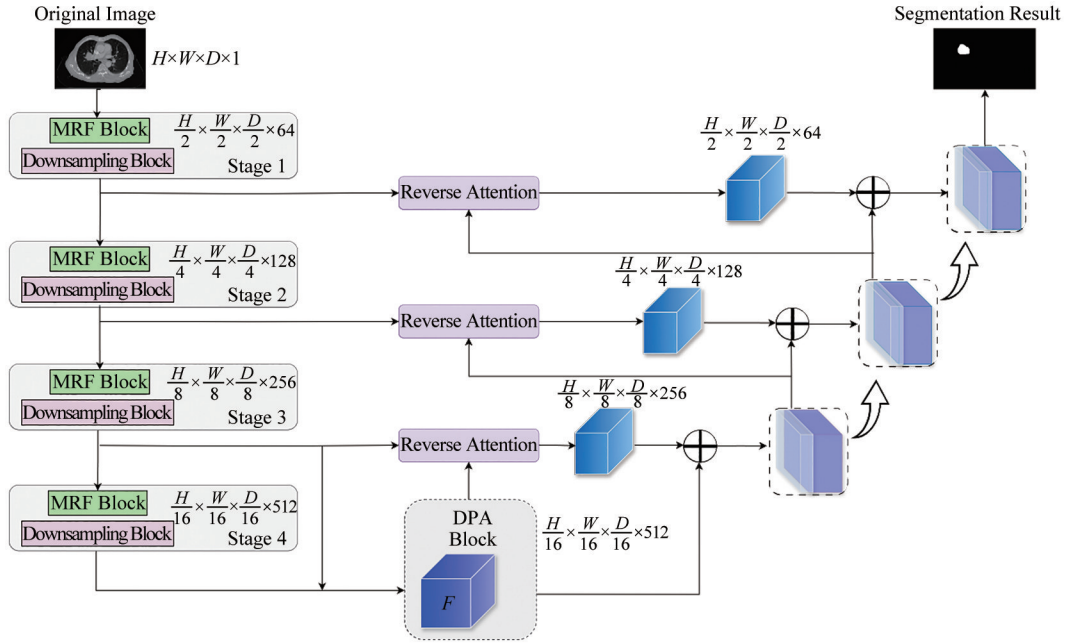


图1 MRFA-Net网络结构

Figure 1 MRFA-Net structure

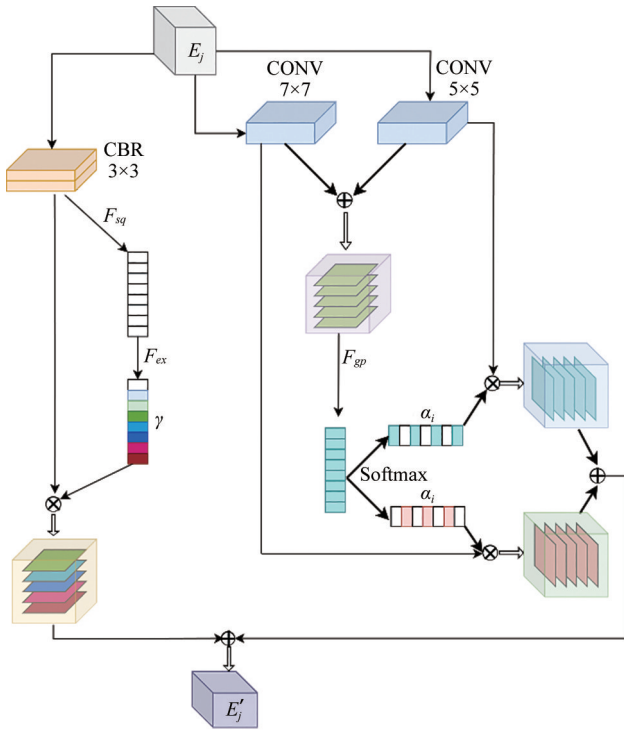


图2 多尺度特征提取模块

Figure 2 Multi-scale residual feature extraction module

两个连续的 3×3 卷积操作构成, 每层卷积均附带批归一化 (BatchNorm) 和非线性激活函数 ReLU, 通道注意力生成过程如下:

$$E_{\text{local}} = \text{ReLU}(\text{BN}(\text{CBR}_2(\text{ReLU}(\text{BN}(\text{CBR}_1(E_j)))))) \quad (2)$$

随后, 特征图 E_{local} 进入通道注意力生成模块。

首先经过 F_{sq} (Squeeze) 操作, 对特征图进行全局平均池化, 将空间维度压缩, 得到通道描述向量 z ; 接着进行 F_{ex} (Excitation) 操作, 使用核大小为 m 的一维卷积进行局部通道感知, 并通过 Sigmoid 激活函数生成通道权重系数 γ :

$$\gamma = \sigma(\text{Conv}_m(z)) \quad (3)$$

将生成的注意力权重 γ 与原始特征图 E_{local} 进行逐通道加权, 得到经过通道筛选后的特征 E_{channel} :

$$E_{\text{channel}} = E_{\text{local}} \cdot \gamma \quad (4)$$

最后, 将并行多尺度分支与残差注意力分支的输出通过残差连接进行加法融合, 获得对大小不一的肿瘤区域更精准的特征表达 E'_j :

$$E'_j = \left(\sum_{i=1}^N \alpha_i \otimes E_j \right) + E_{\text{channel}} \quad (5)$$

1.3 DPA 模块

本文在网络的瓶颈层设计一个双分支结构, 如图3所示, 该结构包含空间金字塔特征提取路径和深度特征融合路径两部分。前者通过空洞卷积组增强特征图的表达能力, 后者旨在减少信息冗余的同时将不同编码器层提取到的肿瘤特征进行充分聚合。

1.3.1 空间金字塔特征提取路径 在空间金字塔特征提取路径中, 编码器的输入首先会经过一个核大小为 1×1 的卷积对拼接后的特征进行通道个数的调整。随后设计一组核大小为 3×3 , 空洞率分别为 6、8、12 的扩张卷积, 在同一特征中捕获不同感受野下的肿瘤特征, 并将这些不同尺度卷积生成的特征进

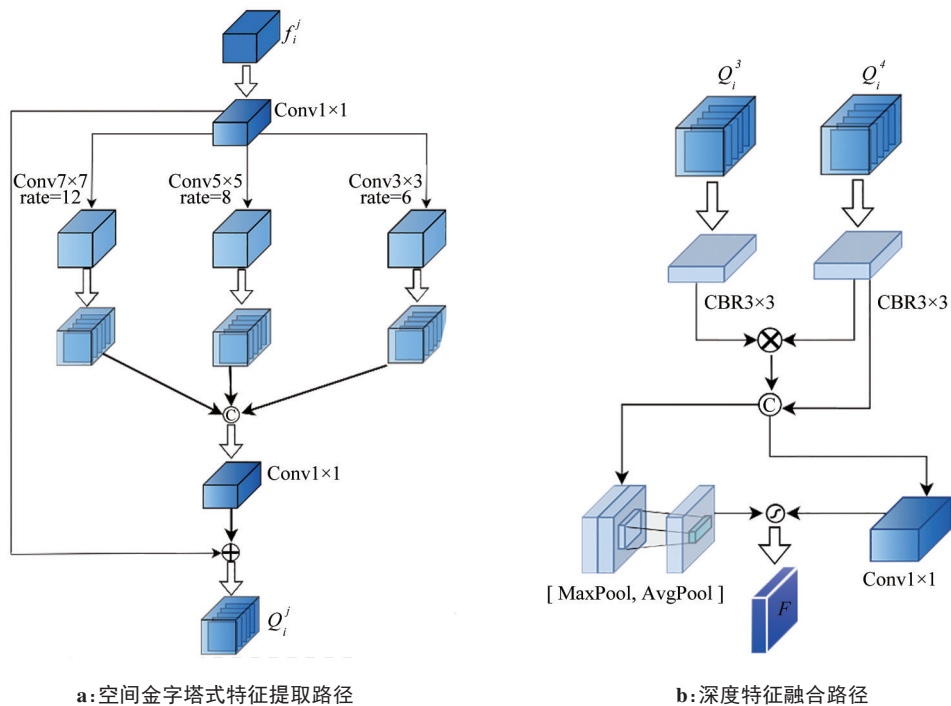


图3 深度特征聚合模块

Figure 3 Deep pyramid aggregation module

行拼接。最后,通过在降维特征中加入 f_i^j 残差分支,完成多尺度特征提取过程。设编码器输出的特征图为 $f_i^j \in R^{C_j \times H_j \times W_j}$,多尺度卷积核组的大小分别为 $\{r_1, r_2, \cdots, r_n\}$ 。提取后的特征数学表达式如下:

$$Q_i^j = \text{Conv}_{1 \times 1}(\text{Concat}(\text{Conv}_{r_1}(\text{Conv}_{1 \times 1}(f_i^j)), \text{Concat}(\text{Conv}_{r_2}(\text{Conv}_{1 \times 1}(f_i^j)), \text{Concat}(\text{Conv}_{r_3}(\text{Conv}_{1 \times 1}(f_i^j)))))) + \text{Conv}_{1 \times 1}(f_i^j) \tag{6}$$

1.3.2 深度特征融合路径 图4中不同颜色表示特征响应强度,颜色越暖表示该位置的特征激活越强,颜色越冷表示激活越弱。浅层特征(E1、E2)主要响应肺部轮廓与边缘信息,而深层特征(E3、E4)则更集中于肿瘤区域并包含更丰富的语义表达,但同时存在空间细节弱化的问题。为避免浅层纹理干扰并降低

计算负担,本文在融合策略中仅选取编码器后两层深层特征进行聚合。首先,对相对浅层特征 Q_i^3 与深层特征 Q_i^4 进行逐元素相乘,以增强跨层语义关联;随后将乘积结果与 Q_i^4 在通道维度拼接,实现语义与结构信息的融合。在此基础上,引入空间注意力机制以强化肿瘤相关判别特征^[18-22]。对于输入特征图 $F \in R^{H \times W \times C}$,分别进行全局最大池化与全局平均池化,获得空间响应特征;将两者在通道维度拼接后,经卷积生成尺寸为 $H \times W \times 1$ 的空间特征图,并通过Sigmoid函数归一化得到空间注意力权重矩阵 M_s 。经过逐元运算后部分特征融合路径的输出 F 如下:

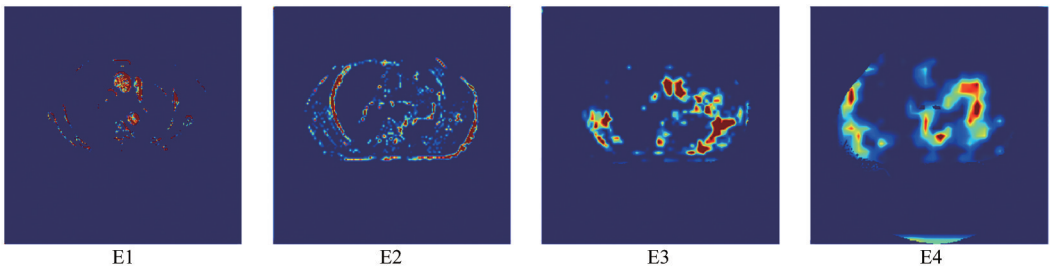
$$F = M_s \otimes \text{Concat}((\text{CBR}(Q_i^3) \otimes \text{CBR}(Q_i^4)), \text{CBR}(Q_i^4)) \tag{7}$$


图4 来自不同深度编码器层特征图的可视化结果

Figure 4 Visualization of feature maps from multi-deep encoder layers

E1、E2中的浅层特征主要包含肺部的轮廓和边缘信息;而E3、E4的深层特征更强调病灶区域并包含复杂的语义信息

1.4 反向注意力模块

如图5所示,反向注意力模块旨在引导网络关注初始预测中如边缘等被忽略的细节,从而精炼特征表示^[19]。首先,输入的初始预测图经Sigmoid激活得到注意力图 M 。为挖掘缺失的边界信息,笔者对其执行逆转操作,得到反向注意力特征图 $1-M$ 。该操作有效抑制已准确定位的高置信度目标内部区域,并增强低置信度区域(如肿瘤边缘)的权重^[20-23]。最后,将 $1-M$ 与输入特征图 F 进行元素级相乘,得到精炼特征图 F' :

$$F' = F \otimes (1 - M) \quad (8)$$

其中, $\otimes$ 表示沿空间维度的元素级相乘。该机制迫使模型在后续处理中发掘更多难以识别的细节特征,有效提升最终的分割精度。

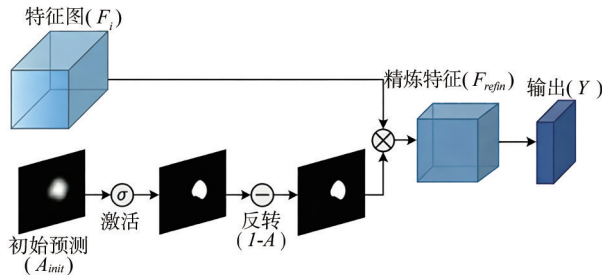


图5 反向注意力模块

Figure 5 Reverse attention module

2 实验

2.1 数据集

本研究采用 Medical Segmentation Decathlon (MSD)第6子任务(肺部肿瘤分割)公开数据集,包含63例非小细胞肺癌(Non-Small Cell Lung Cancer, NSCLC)患者的三维CT及对应肿瘤标注,图像分辨率为 512×512 像素,标注由放射科医师完成。按照既定方案将数据划分为训练集43例(537张)、验证集8例(101张)和测试集12例(157张)。训练前对所有体数据统一进行窗宽/窗位截断、归一化与重采样处理,以满足网络输入要求^[24-26]。

为验证模型的泛化性能,进一步采用同济大学合作单位提供的私有数据集,并经过医院伦理委员会批准(审批号:L20-316Z)。共85例经病理证实的NSCLC患者胸部CT(DICOM格式, 512×512 像素,层厚1~3 mm)。病灶标注由两名具有10年以上经验的副主任医师双盲完成,并经专家组审核确认。采用分层随机抽样将数据划分为训练集51例(711张)、验证集17例(196张)和测试集17例(175张)。

2.2 实验环境

在本文的实验中,模型使用Adam优化器进行训练,初始学习率设置为0.001,批大小为4,epoch为100。本文在NVIDIA GeForce RTX4070super(8 GB)图形处理单元(GPU)上训练所有模型,并应用相同的设置和训练策略。此外,训练过程中采用余弦退火学习率调度,确保学习率能够平稳地降至最小值0.00001。

2.3 评价指标

为全面评估所提方法的分割性能,本文采用6项指标:DSC、交并比(IOU)、召回率(Recall)、95%豪斯多夫距离(HD95)、相对体积差(RVD)和平均表面距离(ASD)^[27-29]。上述指标从区域重叠、分类性能、表面距离等多个维度综合反映模型的分割效果:

$$DSC = \frac{2 \times |P \cap G|}{|P| + |G|} \quad (11)$$

$$IOU = \frac{|P \cap G|}{|P \cup G|} \quad (12)$$

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

$$HD95(P, G) = \max \left\{ 95th \min_{p \in \partial P} d(p, g), 95th \min_{g \in \partial G} d(g, p) \right\} \quad (14)$$

$$ASD(P, G) = \frac{1}{|\partial P| + |\partial G|} \left(\sum_{p \in \partial P} \min_{g \in \partial G} d(p, g) + \sum_{g \in \partial G} \min_{p \in \partial P} d(g, p) \right) \quad (15)$$

$$RVD = \frac{|P| - |G|}{|G|} \quad (16)$$

其中, P 是预测区域, G 是真实区域。TP(True Positive)是真阳性像素数,FN(False Negative)是假阴性像素数。Recall越高,表示漏检越少。 ∂P 、 ∂G 分别为预测和真实边界, $d(p, g)$ 是欧几里得距离。 $RVD > 0$ 表示过分割, $RVD < 0$ 表示欠分割。绝对值越小越好。

3 结果与讨论

3.1 对比实验分析

为全面评估MRFA-Net在肺肿瘤分割任务中的表现,本文从定量指标与可视化效果两个层面对比分析不同模型的分割性能。表1展示包括AttentionU-Net、MobileV3Net、VGG16-U-Net、TransU-Net、SAM2-U-Net、DeepLabv3+Res50Net以及本文模型(MRFA-Net)在MSD第6任务数据集上的评估结果,图6则对应展示各模型在两个典型病例中的分割可视化对比效果。

从表1可见,本文模型在DSC(74.15%)、Recall(81.75%)、ASD(15.430 1 mm)和RVD(39.753 6%)

指标上均优于其他对比方法,展现出良好的泛化性能。在 DSC 方面,本文模型优于排名第二的 Attention U-Net (71.55%),并远高于 VGG16-U-Net (57.92%)等传统 CNN 架构,显示出在模型对肿瘤区域的更高拟合能力。IOU 同样显著提升说明本文方法在分割区域一致性方面更具优势;在边界精度方

面,本文方法在 ASD 上取得 15.430 1 mm 的最小值,优于 Attention U-Net (20.976 1 mm)与 DeepLabv3+ Res50 (19.657 9 mm),说明其对肿瘤复杂边缘结构具有更好的建模能力;HD95 也达到相对较低的 41.879 1 mm,在考虑到 MSD 数据集肿瘤形态多样性的前提下,具备较好的临床适应性。

表 1 MSD 第 6 任务数据集上的肿瘤分割结果比较
Table 1 Comparison of tumor segmentation results on the Task 6 dataset of MSD

模型	DSC/%	IOU/%	Recall/%	HD95/mm	ASD/mm	RVD/%
Attention U-Net	71.55	58.40	81.29	62.654 2	20.976 1	56.602 7
MobileV3Net	66.82	52.12	74.66	46.131 4	16.092 6	63.014 0
VGG16-U-Net	57.92	45.93	62.53	49.208 1	18.234 5	65.878 3
TransU-Net	69.72	63.84	77.72	34.721 6	18.689 4	60.500 3
SAM2-U-Net	65.65	56.20	76.44	34.168 4	16.946 2	78.748 6
DeepLabv3+Res50Net	67.47	60.71	79.94	60.187 4	19.657 9	52.184 7
本文模型	74.15	63.45	81.75	41.879 1	15.430 1	39.753 6

图 6 进一步验证本文方法在实际分割中的优势。第一组样本中,MobileV3Net 与 VGG16-U-Net 存在明显伪影、区域破碎或偏移现象;SAM2-U-Net 与 TransU-Net 虽能覆盖主肿瘤区域,但边界细节不足。相比之下,本文模型能够准确恢复病灶位置与结构,区域连贯、轮廓清晰,最接近 Ground Truth。第二组样本肿瘤面积较小,

对小目标识别能力要求更高。MobileV3Net 与 VGG16-U-Net 出现大量虚假小区域或空洞,Attention U-Net 与 DeepLabv3+Res50Net 虽保留主区域,但存在边界收缩或偏移问题。本文方法不仅准确定位小目标,还显著减少背景误检与冗余边界预测,体现出较强的小目标敏感性与精细建模能力。

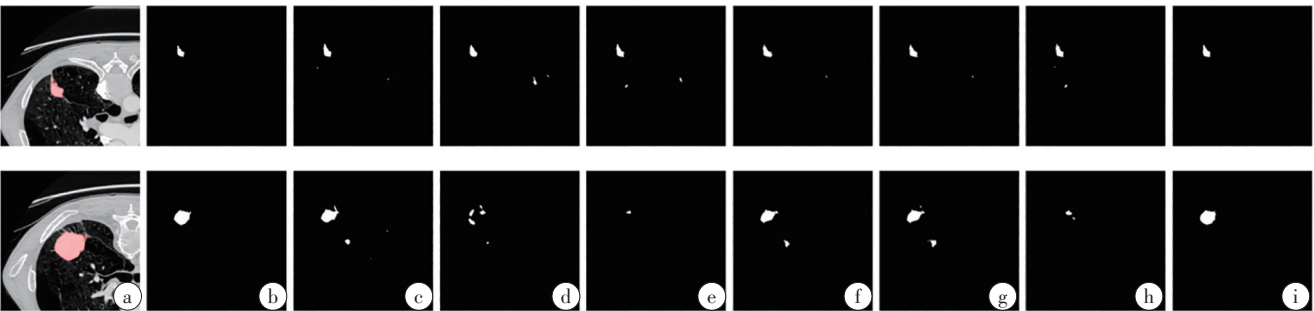


图 6 在 MSD 第 6 任务数据集上的对比图
Figure 6 Comparison results on the Task 6 dataset of MSD

第 1、2 行图像分别为第一、二组样本;a: 原始 CT 图像;b: 真实标注;c: Attention U-Net;d: MobileV3Net;e: VGG16-U-Net;f: TransU-Net;g: SAM2-U-Net;h: DeepLabv3;i: 本文模型

为验证模型在不同数据分布下的鲁棒性与泛化能力,本文在私有肺癌数据集上进一步评估。表 2 给出了 6 项指标对比,定量分析表明,本文方法在 DSC (74.77%)、IOU (62.33%)和 ASD (8.427 1 mm)等关键

指标上均优于对比方法;RVD 最低,为 32.756 1%,说明体积估计更为稳定。

结合图 7 可视化结果可进一步印证上述结论。第一组样本中,Attention U-Net 与 MobileV3Net 存在

表 2 私有数据集上的肿瘤分割结果比较

Table 2 Comparison of tumor segmentation results on the private dataset

模型	DSC/%	IOU/%	Recall/%	HD95/mm	ASD/mm	RVD/%
Attention U-Net	69.03	62.04	80.49	42.514 1	9.734 4	43.557 2
MobileV3Net	63.57	51.50	65.28	52.748 9	14.016 4	47.552 3
VGG16-U-Net	71.90	59.61	74.90	44.466 4	10.484 9	38.771 5
TransU-Net	72.11	60.87	74.23	44.151 7	11.787 7	40.963 4
SAM2-U-Net	71.87	60.93	80.71	46.693 5	10.645 9	52.063 4
DeepLabv3+Res50Net	72.71	61.11	73.66	30.399 0	9.044 1	34.585 9
本文模型	74.77	62.33	76.18	36.402 2	8.427 1	32.756 1

漏检或伪分割现象；VGG16-U-Net、TransU-Net 与 SAM2-U-Net 虽覆盖主区域，但边缘拟合仍存在锯齿或偏移。本文模型在整体形状与边界贴合度上更接近 Ground Truth，与其在 HD95 和 ASD 指标上的优势一致。第二组样本病灶面积较大且形态复杂，MobileV3Net 与 Attention U-Net 再次出现边界收缩；TransU-Net 与 SAM2-U-Net 虽有所改善，但细节恢复仍不足。相比之下，本文模型既保留复杂结构，又获得更加平滑且真实的分割轮廓。

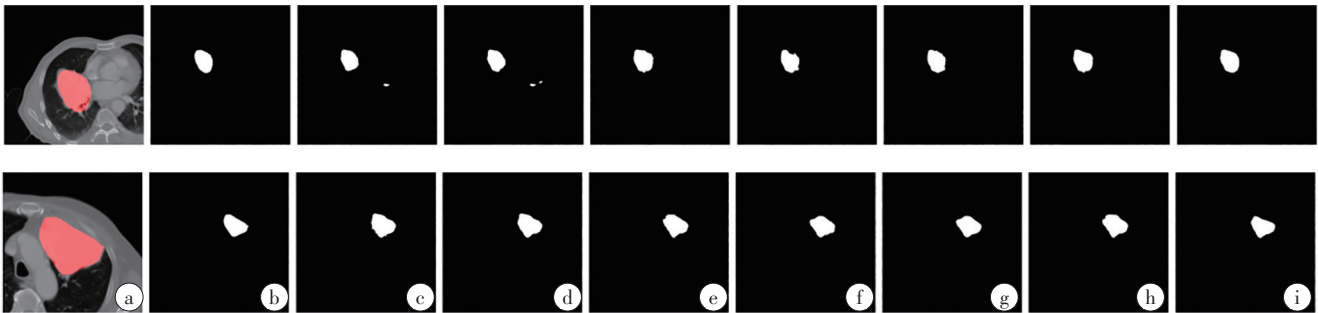


图 7 在私有数据集上的对比图

Figure 7 Comparison results on the private dataset

第 1、2 行图像分别为第一、二组样本；a: 原始 CT 图像；b: 真实标注；c: Attention U-Net；d: MobileV3Net；e: VGG16-U-Net；f: TransU-Net；g: SAM2-U-Net；h: DeepLabv3；i: 本文模型

为进一步直观展示性能分布情况，图 8 给出了私有数据集上 DSC 与 HD95 的病例级分布箱线图。可以观察到，在 DSC 指标上，MRFA-Net 的分布整体向上偏移，中位数明显提高，且四分位区间相对收窄，说明性能提升并非由少数异常样本驱动，而是在多数病例上实现一致性改善。与此同时，在 HD95 指标上，MRFA-Net 的中位数明显下降，且分布范围缩小，表明边界误差得到有效控制。DSC 提升与 HD95 降低的协同趋势进一步说明模型不仅提高了区域重叠精度，同时改善了边界定位能力。

3.2 消融实验分析

3.2.1 针对模块有效性的研究 为了证明本文模块具有实际效果，本研究采用逐次增量添加模块的方法

进行实验。首先评估基线模型在原始肺癌 CT 图像上的肿瘤分割性能，在此基础上依次在基线模型上添加 MRF 和 DPA 模块，验证两个模块的有效性，并进一步探讨模块的参数数量对性能影响的评估。

如表 3 所示，Attention U-Net 作为基线网络，在私有数据集上的 DSC 为 69.03%，HD95 为 42.514 1 mm，ASD 为 14.285 7 mm。为了证明 MRFA-Net 模型的有效性，首先将 RA 坐标注意力机制引入到基线网络的跳跃连接中，这一操作使得模型的分割能力得到较为明显的改善，DSC 得分提升至 69.75%。在实验的基础上分别继续添加 MRF 模块和 DPA 模块，DSC 继续提高到 72.28% 和 72.97%，HD95 和 ASD 也收敛到更值得信赖的得分水准。

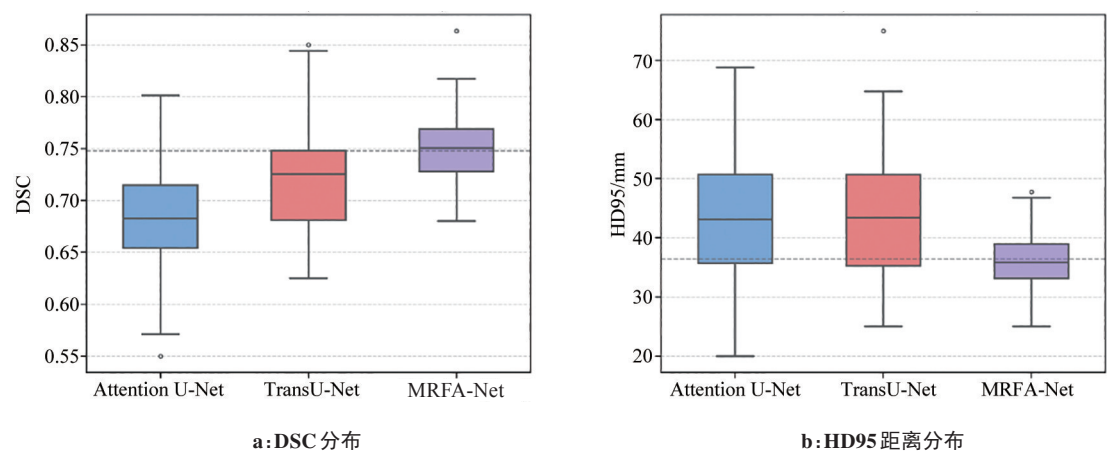


图 8 私有肺癌数据集上的分割性能分布箱线图

Figure 8 Boxplots of segmentation performance on the private lung cancer dataset

表 3 私有数据集上的消融研究比较

Table 3 Comparison of ablation studies on the private dataset

模型	DSC/%	HD95/mm	ASD/mm
Attention U-Net	69.03	42.514 1	14.285 7
Attention U-Net+RA	69.75	52.239 0	12.347 5
Attention U-Net+RA+MRF	72.28	44.866 4	9.734 4
Attention U-Net+RA+DPA	72.97	41.753 2	10.813 3
本文模型	74.77	36.402 2	8.427 1

3.2.2 模块内部参数配置的消融研究 除了验证核心模块的有效性外,本文进一步探究MRF模块中的卷积核尺寸(Kernel Size)以及DPA模块中的空洞率(Dilation Rate)对分割性能的影响,以确定最佳的感受野配置。

MRF模块旨在通过多尺度卷积提取残差特征。笔者对比不同卷积核组合的性能,结果如表4所示。当使用较小的卷积核组合(3×3+5×5)时,由于感受野受限,模型难以覆盖大体积肿瘤的上下文,DSC仅为73.54%。相反,盲目扩大卷积核至(7×7+9×9)虽然增加了感受野,却引入了无关背景噪声并导致参数量激增至3.42 M,性能反而未见提升(74.05%)。相比之下,本文模型采用的(5×5+7×7)组合在参数量(3.04 M)与性能(74.77%)之间取得了最佳平衡,有效兼顾了局部细节与全局轮廓的感知。

3.2.3 参数量与分割性能的权衡分析 此外,本文评估了模型参数量对分割性能的影响,以验证所提模块的参数效率。如图9所示,基础模型(Baseline)的参数量约为19.1 M,其DSC为69.03%。在引入反向注意力机制(Attention U-Net+RA)后,参数量跃升至21.7 M,DSC提高至70.75%。

从图中折线的斜率变化可以观察到,在引入本文

表 4 MRF 模块中不同卷积核尺寸组合的性能对比

Table 4 Performance comparison of different kernel size combinations in the MRF module

卷积核尺寸	Parameters/M	DSC/%	HD95/mm
3×3+5×5	2.85	73.54	38.12
5×5+7×7(本文模型)	3.04	74.77	36.40
7×7+9×9	3.42	74.05	37.85

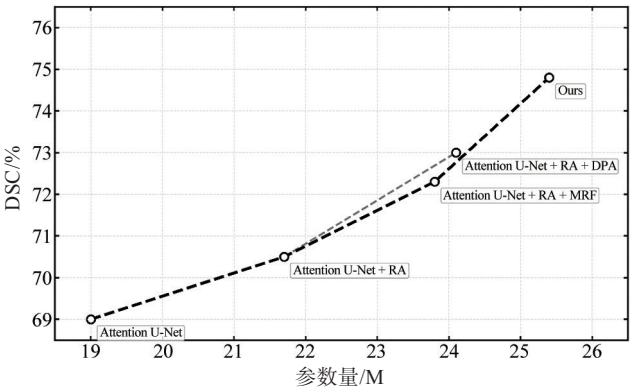


图 9 参数提升对性能增长的影响趋势

Figure 9 Effects of number of parameters on performance improvement

设计的两个核心模块后,性能增益显著。特别是引入深层特征聚合模块(Attention U-Net+RA+DPA)后,模型参数量从21.7 M仅增加至24.2 M,但DSC大幅提升至72.97%。最终,本文提出的完整模型(MRFA-Net)在参数量仅为25.4 M的情况下,取得了74.77%的DSC。MRFA-Net在未明显增加计算负担的前提下,实现了性能的跨越式提升,展现出优越的精度-效率权衡能力(Accuracy-Efficiency Trade-off),充分验证了选择性残差特征提取与深层特征融合机制的高效性。

4 结论

为满足临床中对肺部肿瘤分割在精度与鲁棒性方面的多样化需求,本文提出一种基于残差特征挖掘与深度特征融合的肺癌CT分割模型MRFA-Net。在编码器阶段设计MRF模块,通过空洞卷积与门控融合机制实现多尺度上下文信息的提取,并利用残差连接中的注意力机制抑制冗余通道,从而提升网络对肺肿瘤复杂边界与细节的建模能力。此外,本文进一步提出DPA模块,用于在解码阶段深度整合高级肿瘤特征信息。该模块舍弃浅层特征干扰,结合不同感受野的卷积操作以及逐像素增强方式,显著提高模型对肺部肿瘤边缘和形态的感知能力。通过联合MRF与DPA两个模块,MRFA-Net能够有效缓解小体积或结构相似区域的误分割问题,并提升对边界模糊和形态多样肿瘤的建模能力,从而显著提高肺肿瘤CT图像的分割精度。在MSD公开数据集和私有肺癌CT数据集上进行的实验结果表明MRFA-Net具备优异的性能,DSC分别达到74.15%和74.77%。该方法为肺部肿瘤分割提供有效的解决方案,具备良好的临床应用潜力。

【参考文献】

- [1] 岳文怡, 盛复庚. 基于磁共振成像的乳腺癌影像基因组学的研究现状和进展[J]. 临床放射学杂志, 2024, 43(6): 991-994.
Yue WY, Sheng FG. Research status and progress of breast cancer radiogenomics based on magnetic resonance imaging[J]. Journal of Clinical Radiology, 2024, 43(6): 991-994.
- [2] 周力凯, 江雨洋, 冯亚春, 等. 基于多尺度区域与类不确定性理论的局部阈值分割方法[J]. 计算机应用, 2020, 40(增刊2): 66-72.
Zhou LK, Jiang YY, Feng YC, et al. Local image thresholding method based on multiscale region and class uncertainty theory[J]. Journal of Computer Applications, 2020, 40(S2): 66-72.
- [3] 邵浩然, 张伟, 马宇张. 基于ResUNet和Transformer的CT肝脏肿瘤图像分割[J]. 中国医学物理学杂志, 2025, 42(11): 1455-1461.
Shao HC, Zhang W, Ma YZ. CT liver tumor image segmentation based on ResUNet and Transformer[J]. Chinese Journal of Medical Physics, 2025, 42(11): 1455-1461.
- [4] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [5] Xiao X, Lian S, Luo ZM, et al. Weighted Res-UNet for high-quality retina vessel segmentation[C]//2018 9th International Conference on Information Technology in Medicine and Education (ITME). Piscataway, NJ, USA: IEEE, 2018: 327-331.
- [6] Dong RS, Pan XQ, Li FY. DenseU-Net-based semantic segmentation of small objects in urban remote sensing images[J]. IEEE Access, 2019, 7: 65347-65356.
- [7] Ding QY, Pan Y, Liu JX, et al. AMEF-Net: towards an attention and multi-level enhancement fusion for medical image classification in Parkinson's aided diagnosis[J]. IET Computer Vision, 2024, 19(1): e12324.
- [8] Fan DP, Ji GP, Zhou T, et al. PraNet: parallel reverse attention network for polyp segmentation[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2020. Cham: Springer International Publishing, 2020: 263-273.
- [9] Isensee F, Jaeger PF, Kohl SAA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nat Methods, 2021, 18(2): 203-211.
- [10] Shi LX, He WY, Wang GD. A dual-path fusion network with edge feature enhancement for medical image segmentation [J]. Mathematics, 2026, 14(1): 55.
- [11] Hatamizadeh A, Nath V, Tang YC, et al. Swin UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images [C]//Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Cham: Springer International Publishing, 2022: 272-284.
- [12] Zhang YD, Liu HY, Hu Q. TransFuse: fusing transformers and CNNs for medical image segmentation[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2021. Cham: Springer International Publishing, 2021: 14-24.
- [13] Yuan L, Chen YP, Wang T, et al. Tokens-to-token ViT: training vision transformers from scratch on ImageNet [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2021: 538-547.
- [14] Yang L, Zhang JW, Wang T, et al. Multi-scale camouflaged feature mining and fusion network for liver tumor segmentation[J]. Eng Appl Artif Intell, 2025, 148: 110398.
- [15] Su J, Liu Z, Zhang J, et al. DV-Net: accurate liver vessel segmentation via dense connection model with D-BCE loss function[J]. Knowl Based Syst, 2021, 232: 107471.
- [16] Shit S, Paetzold JC, Sekuboyina A, et al. cDice-a novel topology-preserving loss function for tubular structure segmentation[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 16555-16564.
- [17] 李翠锦, 瞿中. 基于卷积神经网络的跨层融合边缘检测算法[J]. 计算机应用研究, 2021, 38(7): 2183-2187.
Li CJ, Qu Z. Cross fusion edge detection algorithm based on CNN[J]. Application Research of Computers, 2021, 38(7): 2183-2187.
- [18] Fu XH, Bi L, Kumar A, et al. Multimodal spatial attention module for targeting multimodal PET-CT lung tumor segmentation[J]. IEEE J Biomed Health Inform, 2021, 25(9): 3507-3516.
- [19] Zhang YJ, Chen X. Lightweight semantic segmentation algorithm based on MobileNetV3 network[C]//2020 International Conference on Intelligent Computing, Automation and Systems (ICICAS). Piscataway, NJ, USA: IEEE, 2020: 429-433.
- [20] Qamar S, Fazil M, Ahmad P, et al. UNet with self-adaptive Mamba-like attention and causal-resonance learning for medical image segmentation[J]. Sci Rep, 2025, 16(1): 135.
- [21] Wang B, Shi H, Zhao ZT, et al. PSCT-Net: a parallel symmetric CNN-transformer hybrid network for medical image segmentation[J]. Med Eng Phys, 2025, 146: 104442.
- [22] Peng HX, Xue C, Shao YY, et al. Semantic segmentation of litchi branches using DeepLabV3+ model [J]. IEEE Access, 2020, 8: 164546-164555.
- [23] 李柯, 刘文忠, 秦镜洵. 改进MambaUNet网络对肝脏肿瘤CT图像的轻量化级联分割[J]. 中国医学物理学杂志, 2025, 42(8): 1068-1078.
Li K, Liu WZ, Qin JT. Improved MambaUNet for lightweight cascaded segmentation in liver tumor CT image [J]. Chinese Journal of Medical Physics, 2025, 42(8): 1068-1078.
- [24] Tagnamas J, Ramadan H, Yahyaouy A, et al. CS-Net: combined ConvNeXt-Swin-Unet for accurate medical image segmentation[J]. J Supercomput, 2026, 82(5): 276.
- [25] Yang YW, Huang HG, Zhang GD, et al. DHR-Net: an ultra-lightweight U-Net based on efficient convolutional attention for medical image segmentation[J]. Displays, 2026, 93: 103425.
- [26] Lei LH, Li W. Transformer-based multi-task model for lung tumor segmentation and classification in CT images[J]. J Radiat Res Appl Sci, 2025, 18(3): 101657.
- [27] Yi C, Jiang SF, Xiong LL, et al. D-S-Net: an efficient dual-stage strategy for high-precision segmentation of gross tumor volumes in lung cancer CT images[J]. BMC Cancer, 2025, 25(1): 1387.
- [28] Dwivedi P, Barage S, Singh R, et al. PET-based radiomic analysis in multicentre lung cancer study and impact of feature domain harmonization[J]. Phys Eng Sci Med, 2025, 48(4): 1841-1851.
- [29] 付宏建, 白宏阳, 郭宏伟, 等. 融合多注意力机制的光学遥感图像目标检测方法[J]. 光子学报, 2022, 51(12): 304-312.
Fu HJ, Bai HY, Guo HW, et al. Object detection method of optical remote sensing image with multi-attention mechanism [J]. Acta Photonica Sinica, 2022, 51(12): 304-312.

(编辑:薛泽玲)