

DOI:10.3969/j.issn.1005-202X.2024.06.012

医学影像物理

## 基于多尺度特征融合与注意力的肝脏分割方法

冉梅子<sup>1</sup>, 胡小军<sup>2</sup>, 姜晓燕<sup>1</sup>, 范应方<sup>2</sup>, 王航<sup>1</sup>, 王海玲<sup>1</sup>, 高永彬<sup>1</sup>

1. 上海工程技术大学电子电气工程学院, 上海 201620; 2. 南方医科大学第五附属医院肝胆外科, 广东广州 510000

**【摘要】**由于CT影像对比度低、肝脏形状不规则、相邻器官边界模糊,目前基于卷积神经网络的方法在肝脏分割任务上的表现不佳,尤其是在边界识别和小目标检测方面。基于此,提出一种基于多尺度特征融合与注意力的肝脏分割方法(MFFA UNet)。首先,利用多尺度特征融合获取丰富的分割信息,同时使用空间和通道注意力机制捕获全局空间和通道间的关系。其次,通过深度监督模块充分利用中间隐藏层的输出,增强网络的学习能力,加快网络收敛速度。此外,采用一种混合损失函数,以解决类别不平衡的问题,进一步提升模型的分割效能。实验结果表明,所提出的MFFA UNet方法在公共数据集LITS上的表现超越当前主流分割网络,分割结果更接近真实值。

**【关键词】**肝脏分割;注意力机制;多尺度特征融合;深度监督;MFFA UNet

**【中图分类号】**R318

**【文献标志码】**A

**【文章编号】**1005-202X(2024)06-0739-08

### Liver segmentation method based on multi-scale feature fusion and attention

RAN Meizi<sup>1</sup>, HU Xiaojun<sup>2</sup>, JIANG Xiaoyan<sup>1</sup>, FAN Yingfang<sup>2</sup>, WANG Hang<sup>1</sup>, WANG Hailing<sup>1</sup>, GAO Yongbin<sup>1</sup>

1. School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China; 2. Department of Hepatobiliary Surgery, the Fifth Affiliated Hospital of Southern Medical University, Guangzhou 510000, China

**Abstract:** Due to the low contrast of CT images, irregular shape of the liver, and blurred boundaries with adjacent organs, the existing methods based on convolutional neural network underperform in liver segmentation tasks, especially for boundary recognition and small object detection. A novel liver segmentation method is proposed based on multi-scale feature fusion and attention, namely MFFA UNet. Multi-scale feature fusion is firstly employed to acquire abundant segmentation details, while spatial and channel attention mechanisms are utilized to capture global spatial and inter-channel relationships. Additionally, a deep supervision module fully leverages the output of intermediate hidden layers, enhancing the learning capability of the network, which in turn accelerates the network's convergence speed. Moreover, a hybrid loss function is adopted to address the issue of class imbalance, further boosting the model's segmentation efficacy. Experimental results demonstrate that the proposed MFFA UNet outperforms the prevailing segmentation networks on the public LITS dataset, producing results that are closer to the ground truth.

**Keywords:** liver segmentation; attention mechanism; multi-scale feature fusion; deep supervision; MFFA UNet

### 前言

肝癌是世界上常见的、高致死率的恶性肿瘤。如果能够在早期及时发现并采用手术治疗,能够在很大程度上提高患者的生存率<sup>[1]</sup>。目前对于肝脏图像的分割主要分为传统方法和深度学习方法。传统方法主要包括区域生长法<sup>[2-3]</sup>、图割法<sup>[4]</sup>、水平集法<sup>[5]</sup>、

阈值法<sup>[6-7]</sup>等。虽然传统的分割方法在肝脏图像分割问题上颇有成效,但受人工手动设置的影响存在一定偏差,难以满足临床应用的实际需求。基于深度学习的经典医疗影像分割网络U-Net<sup>[8]</sup>,能够在有限的数据集上表现良好,解决了医疗领域中数据较少的问题。该网络由编码器和解码器两部分组成,具有U形架构的特点,故被称为U-Net。受U形架构的启发,后续又提出许多基于U-Net的变体。包括UNet++<sup>[9]</sup>、UNet3+<sup>[10]</sup>、R2U-Net<sup>[11]</sup>、Attention U-Net<sup>[12]</sup>、Res-UNet<sup>[13]</sup>、Dense-UNet<sup>[14]</sup>、E2Net<sup>[15]</sup>等。这些方法多是在卷积神经网络的基础上进行不同改进提升模型效果,但它们忽略了卷积的局限性。卷积操作通常在局部区域内做运算,只能提取局部特征,无法捕获全局信息建立长距离的依赖关系。在自然语言处

**【收稿日期】**2023-10-26

**【基金项目】**广州市科技计划项目(202206010093);上海市科委社会发展项目“科技创新行动计划”(21DZ1204900)

**【作者简介】**冉梅子,硕士研究生,研究方向:目标检测、图像分割,  
E-mail: ranmeizi0106@163.com

**【通信作者】**姜晓燕,博士,副教授,研究方向:计算机视觉、图像分割,  
E-mail: xiaoyan.jiang@sues.edu.cn

理领域,Transformer的提出被用于序列到序列的预测进行建模,其惊人表现引起了人们广泛关注<sup>[16]</sup>。Chen等<sup>[17]</sup>提出TransUNet通过将Transformer与U-Net相结合,利用多头自注意力机制提取全局上下文输入序列,解决卷积只能在局部内运算,无法捕获全局信息的问题。同时借助U-Net的结构来恢复局部的空间信息,实现更为精确的定位和分割。因此,本文选取TransUNet作为基线网络。

由于CT影像对比度低、相邻器官边界模糊,使得分割任务变得十分困难。Jiang等<sup>[18]</sup>提出RMAU-Net在残差挤压激励(Res-SE)块中,通过残差连接来缓解梯度消失问题,同时显式地建模特征通道之间的相互依赖性和特征重新校准旨在更好地学习特征表示。Feng等<sup>[19]</sup>提出CPFNet在编码器和解码器之间设计了多个全局金字塔引导(GPG)模块,旨在通过重建跳跃连接为解码器提供不同级别的全局上下文信息。同时设计了一个尺度感知金字塔融合(SAPF)模块,动态融合多尺度上下文信息在高层特征中。Fan等<sup>[20]</sup>提出MSN-Net将高层、低层特征和多尺度语义信息相结合,获得更具代表性的特征。Xie等<sup>[21]</sup>提出MCI-Net设计一个多尺度上下文提取模块,通过组合4个级联的混合空洞卷积分支来捕捉更广

泛和更深层的特征。对于肝脏分割任务而言,无论是低层特征图的细节信息,还是高层特征图的语义信息,都至关重要。而如何将两者有效融合,是提升模型效果的关键。基于此,本文提出一种基于多尺度特征融合与注意力的肝脏分割网络(MFFA UNet)。主要贡献如下:(1)提出一种多尺度特征融合与残差注意力结合(Multi-scale Feature Fusion and Residual Attention, MFRA)模块,与原始网络中的单一跳跃连接相比,MFRA模块能够增强对的特征表示和利用,同时捕获全局空间和通道间的关系,有效提高模型捕捉边缘和细节的能力。(2)利用深度监督(Deep Supervision, DS)机制,将中间隐藏层输出汇入分割头,加速模型的收敛速度。(3)设计一种混合损失函数,解决类别不平衡问题,进一步提升模型性能。

## 1 方法

本文在基线网络TransUNet的基础上,加入MFRA模块,提升模型的学习能力。并在解码器中添加DS机制,充分利用不同层的侧输出,进一步提高模型的分割效果。网络的整体架构具体如图1所示。

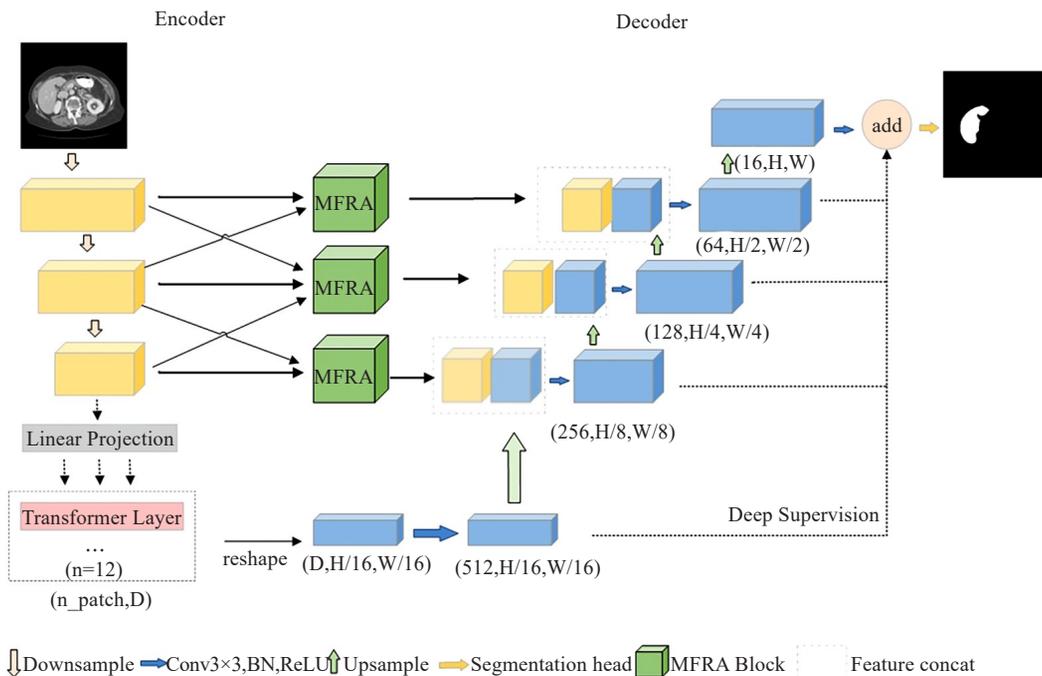


图1 MFFA UNet架构图

Figure 1 MFFA UNet structure

### 1.1 MFRA

为了解决信息丢失和边界模糊问题,本文提出MFRA模块,模块具体架构如图2所示。首先,对特

征图分别进行通道注意力和空间注意力操作,以获取一维通道注意力图和二维空间注意力图。将得到的注意力图分别与原始特征图相乘,获得带有权重

的特征图。为避免梯度消失问题,引入残差连接,将得到的结果与原始输入特征图相加。然后将不同尺度的带有权重的特征图,在通道维度上进行拼接操作。通过这种将不同尺度的特征图进行融合的方

法,可以增强特征图的表示获取更丰富的分割信息,有效解决肝脏分割任务中因边界模糊而导致的错分割问题。

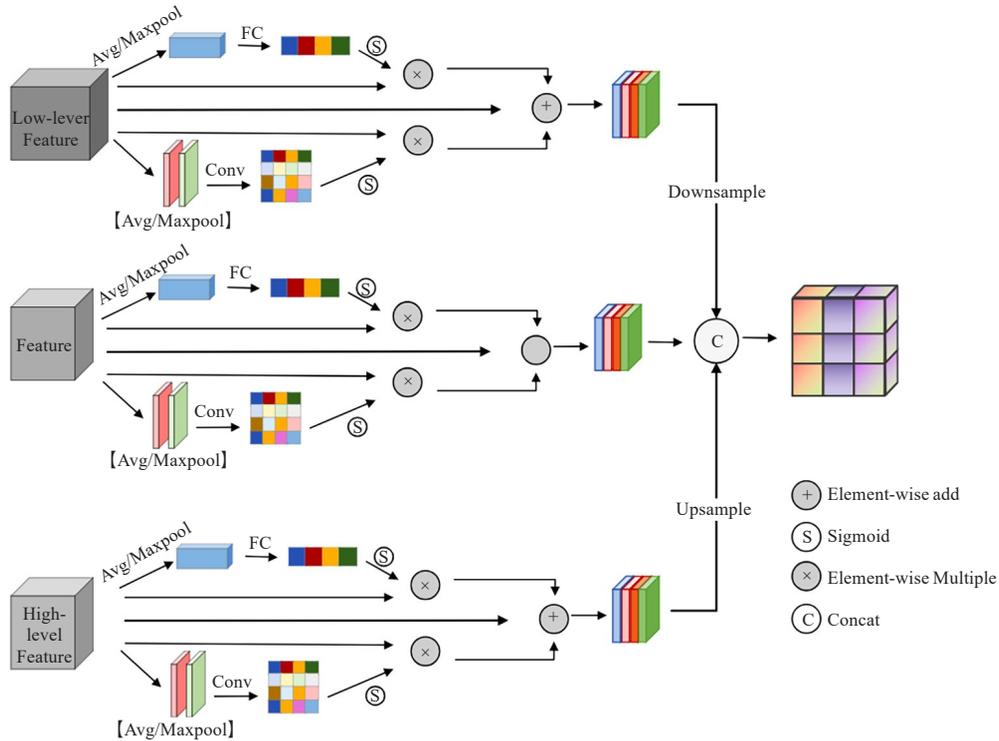


图2 多尺度特征融合与残差注意力结合模块

Figure 2 Multi-scale feature fusion and residual attention module

通道注意力模块主要关注特征间的通道关系,以识别输入图像中“什么”是重要的。空间注意力模块则侧重特征间的空间关系,着重关注输入图像中信息“在哪里”<sup>[22-24]</sup>。使用通道注意力和空间注意力机制,能够使网络侧重关注感兴趣的部分,抑制无关噪声干扰。通道和空间注意力公式描述如下:

$$M_c(F) = \sigma(\text{MLP}(\text{Avgpool}(F)) + \text{MLP}(\text{Maxpool}(F))) \\ = \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \quad (1)$$

$$M_s(F) = \sigma(f^{7 \times 7}([\text{Avgpool}(F); \text{Maxpool}(F)])) \\ = \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])) \quad (2)$$

其中, $\sigma$ 代表激活函数 Sigmoid, $f^{7 \times 7}$ 代表卷积核大小为 $7 \times 7$ 的卷积操作, $W_i$ 表示参数矩阵,Maxpool和Avgpool代表最大池化和平均池化操作,MLP代表多层感知机, $M_c(F)$ 、 $M_s(F)$ 分别表示通道注意力图和空间注意力图。

为了避免梯度消失问题,在通道和空间注意力计算中,引入残差连接。公式描述如下:

$$F^* = M_c(F) \times F + M_s(F) \times F + F \quad (3)$$

其中, $F$ 表示输入特征图, $F^*$ 表示经由通道和空间注意力操作后获得的带权重特征图。

随后通过在通道维度上进行拼接,来实现不同尺度的特征融合。经过MFRA模块,最终输出的结果为:

$$F_{\text{out}} = \text{Concat}(F_{\text{low}}^*, F^*, F_{\text{high}}^*) \quad (4)$$

其中,Concat表示通道拼接, $F_{\text{low}}^*$ 表示浅层特征, $F_{\text{high}}^*$ 表示深层特征。

## 1.2 深度监督

网络不同阶段的输出,对最终分割结果都有影响。Wu等<sup>[25]</sup>设计4种方式的监督路径,以在模型优化过程中保留丰富的多尺度深层特征。实验证明多路径监督能够更好地保存信息,进一步提升网络的分割效果。受其影响,本文在网络的解码器中使用深度监督策略。具体操作如图3所示,先经过 $1 \times 1$ 大小的卷积,对特征通道数进行压缩。然后对得到的结果进行上采样操作,将低分辨率的特征恢复到和

上层特征相同的分辨率,然后进行对应元素相加操作。重复上述操作,将最终得到的结果汇入分割头。充分利用中间层的侧输出,向模型训练提供更丰富的信息,最终提升网络分割精度。

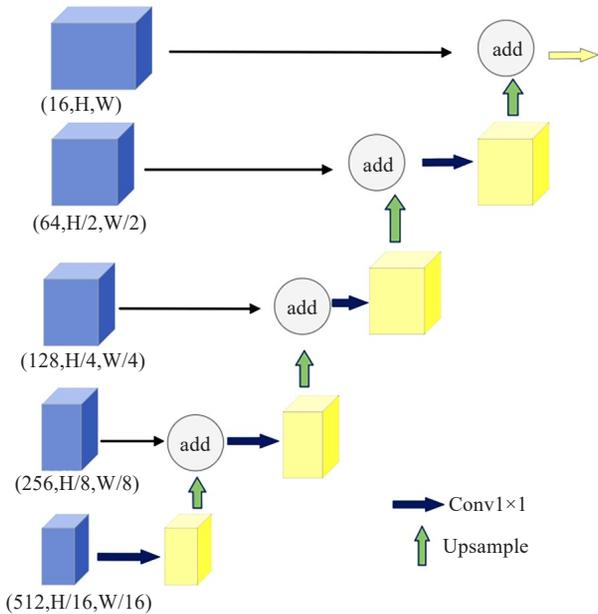


图3 深度监督  
Figure 3 Deep supervision

### 1.3 混合损失

二元交叉熵损失函数(BceLoss)是图像分割任务中常用的损失函数,它平等地计算每个像素点的损失。公式如下:

$$BceLoss = -\frac{1}{N} \sum_i^N (y_i \log(p_i) + (1 - y_i) \log(1 - p_i)) \quad (5)$$

DiceLoss 是医学影像分割中经常使用的损失函数。可以缓解样本中前背景不平衡带来的负面影响<sup>[26]</sup>。它在训练中更侧重对前景区域的挖掘,缓解正负样本不平衡问题,但会使模型的训练过程不稳定<sup>[27]</sup>。公式如下:

$$DiceLoss = 1 - \frac{2 \sum_{i=1}^N y_i p_i}{\sum_{i=1}^N y_i + \sum_{i=1}^N p_i} \quad (6)$$

其中,  $N$  表示像素点的总个数,  $y_i$  表示像素  $i$  的真实值,  $p_i$  表示像素  $i$  的预测值。

在肝脏分割任务中,有的切片内目标区域占比较小,大面积背景可能影响分割结果。针对上述问题,本文提出一种混合损失函数,将 BceLoss 和 DiceLoss 相结合。公式如下:

$$Loss = \alpha DiceLoss + (1 - \alpha) BceLoss \quad (7)$$

其中,  $\alpha$  是控制 DiceLoss 和 BceLoss 权重的超参数,取值为  $[0, 1]$ , 通过以 0.2 为间隔进行一系列消融实验

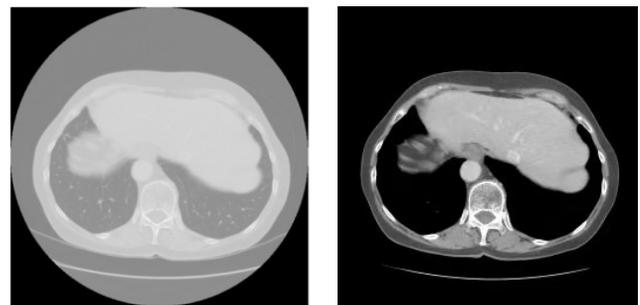
后,对比实验结果中的各项评价指标确定  $\alpha$  的最优值为 0.6。

## 2 实验与结果

### 2.1 数据集预处理

本实验使用的数据集是肝脏分割挑战赛数据集 LITS2017。LITS 数据集由 201 例患者的 3D 腹部 CT 扫描数据组成,其中已公布肝脏和肿瘤标签的数据为 131 例,另外 70 例用做测试尚未公布标签,由训练有素的放射科医生或肿瘤科医生手工标注。病人的切片数量为 74~987,平面分辨率为 0.55~1.00 mm,切片厚度为 0.45~6.00 mm<sup>[28]</sup>。按照 6:2:2 的比例,将数据划分为训练集、验证集和测试集。将所有的数据沿着 z 轴方向进行无间隔切片处理,共获得 19 163 张二维的 CT 切片。

由于直接得到二维 CT 切片数据存在对比度低,相邻器官边界模糊等问题,直接分割效果较差,因此对数据先进行预处理操作。首先使用 HU 值截断,根据人体组织学将 HU 设置为  $[-250, 250]$ ,其包含了 99% 的肝脏区域。对得到的结果进一步进行归一化处理,减小无关组织的影响降低分割难度。将图片大小调整为  $256 \times 256$ ,加快模型的训练速度。经过预处理后得到的 CT 切片如图 4 所示,可以看出经过预处理的图片轮廓更清晰,对比度也更明显。



a: 原始 2D CT 图像      b: 预处理后的 2D CT 图像

图4 CT 图像预处理前后对比

Figure 4 Comparison of CT image before and after preprocessing

另外,为了提高模型的泛化能力和鲁棒性,实验中对所有图片进行简单的数据增强处理,包含随机旋转、翻转、平移操作。

### 2.2 实验环境和参数设置

本文用于比较的方法(U-Net、UNet++、ResUNet、TransUNet)均已重新训练,实验在 CPU (Intel E5-2678)、GPU (Nvidia GeForce RTX 3080Ti)、Ubuntu18.04 系统、PyCharm 平台上进行。使用 SGD 作为优化器,初设学习率为 0.01,权重衰减设置为

0.000 1, 批次大小设定为 8, 训练轮数统一设置为 100。

### 2.3 实验评价指标

本文使用多个指标评估网络性能, 包括 Dice 相似系数 (Dice Similarity Coefficient, DSC)、豪斯多夫距离 (Hausdorff Distance, HD)、交并比 (IoU)、精确率 (Precision)、召回率 (Recall)。DSC 是一种评估相似度系数, 通常用于计算两个样本的相似度或者重叠度, 取值为 0~1, 其数值越高表示模型效果越好。HD 度量空间中两个子集之间的距离。对分割出的边界较为敏感, 其数值越低表示模型效果越好<sup>[29]</sup>。IoU、Precision、Recall 是语义分割常用指标, 其数值越高表示模型效果越好。各评价指标的计算公式如下:

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

$$IoU = \frac{TP}{TP + FP + FN} \quad (9)$$

$$HD(X, Y) = \max(hd(X, Y), hd(Y, X)) \quad (10)$$

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

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

其中, TP 表示真正例, 即样本的真实类别是正例, 模型预测也为正样本的个数。FP 表示假正例, 即样本的真实类别是负例, 但模型预测为正样本的个数。FN 表示假负例, 即样本的真实类别为正例, 但模型预测为负样本个数。

### 2.4 实验结果与分析

**2.4.1 消融实验** 为了验证所提出的模型 MFFA UNet 在肝脏分割上的效果, 本文在 LITS 数据集上进行消融实验。以 TransUNet 作为基线网络, 分别验证添加 MFRA、DS 模块后的肝脏分割效果。得到的消融实验结果如表 1 所示, 与基线网络 TransUNet 相比, 本文模型 MFFA UNet 对于肝脏区域的分割效果明显更佳。

由表 1 所示, 在网络中加入 MFRA 模块后, 各个评价指标皆有所提升。实验结果表明将低层特征的细节信息和高层特征的语义信息融合, 利用两者之间的优势互补, 增强特征的代表。并通过注意力机

制关注目标区域, 减少无关信息的影响, 明显改善模型分割效果。其次在解码器中添加 DS 模块, 保存不同尺度的中间层特征信息, 对网络性能也有一定提升。最终实验结果表明, 本文所提出的 MFFA UNet 模型在肝脏分割任务上表现更优越, 各个指标均优于基线网络。

图 5 为消融实验结果对比图。其中, 图 5a 是原始 CT 图像的 2D 切片数据, 图 5b 是专家手工分割的标签, 图 5c 是基线网络 TransUNet 分割的结果, 图 5d~f 分别表示 TransUNet+MFRA、TransUNet+DS、MFFA UNet 对肝脏区域分割的结果。图 5c 的 TransUNet 分割出的结果中, 存在将周边器官错误识别为目标区域的问题。在识别形状不规则的边界区域时, 出现边界分割不光滑情况。针对上述问题, 本文通过多尺度特征融合的方法, 将对分割任务同样重要的低层细节信息和高层语义信息相融合, 增强对特征的代表, 缓解边缘区域分割有误的问题。引入空间和通道注意力机制, 使得网络侧重点关注目标区域, 抑制无关信息影响, 有效缓解了错误分割问题。由图 5 可以看出 MFFA UNet 分割的结果更接近专家手动分割的结果。相对于基线网络 TransUNet, 错误分割情况明显改善, 边缘区域也更加平滑。

此外, 为了验证所提出的损失函数对模型分割结果的影响, 本文分别使用 BceLoss、DiceLoss 和提出的混合损失函数来训练模型。损失函数的消融实验结果如表 2 所示。相较于仅使用单个损失函数来训练网络, 本文所提出的混合损失函数效果明显更佳。通过将 DiceLoss 与 BceLoss 结合, 即解决类别不平衡问题, 同时也让模型收敛得更好。

对于混合损失函数中的超参数  $\alpha$ , 本文采用 0.2 的步长, 在 0~1 的范围内进行多次实验。实验结果可视化如图 6 所示。通过比较各项评价指标, 确定了  $\alpha$  的最优值为 0.6。

**2.4.2 对比实验** 为了验证本文提出的方法在肝脏分割任务上的有效性, 和现有的语义分割模型 U-Net、UNet++、ResUNet、TransUNet 在数据集 LITS 上进行了对比实验。结果如表 3 所示, 与经典的医学影像分割网络 U-Net 相比, 本文方法在多个评价指标上皆有提升。与其他分割方法相比, 本文在 U 型架构的基础上, 重新设计中间跳跃连接。通过多尺度特征融合的方法, 将浅层特征与深层特征相结合, 有效提升模型分割效果。并利用通道注意力和空间注意力模块, 自适应调整各通道间的特征响应值, 关注感兴趣区域抑制无关信息, 缓解错误分割问题。实验结果证明本文方法的有效性和优越性, 在肝脏分割任务中明显胜于其他方法。

表 1 MFFA UNet 在数据集 LITS 上的消融实验

Table 1 MFFA UNet ablation experiments on LITS dataset

模型	DSC/%	HD/mm	IoU/%	精确率/%	召回率/%
TransUNet	92.9	6.10	89.5	92.6	93.9
TransUNet+MFRA	93.9	5.35	90.6	93.3	94.6
TransUNet+DS	93.3	5.61	89.8	93.1	94.0
MFFA UNet	94.8	5.01	91.5	94.5	95.2

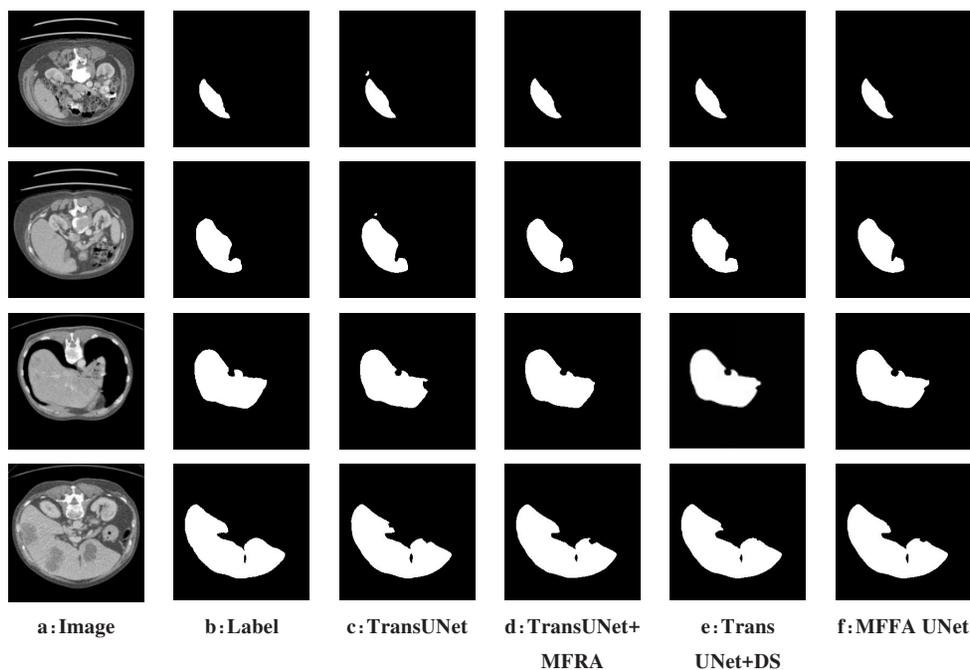


图5 MFFA UNet消融实验分割结果

Figure 5 Segmentation results in MFFA UNet ablation experiments

第1~4行图像代表不同CT切片的实验结果

表2 MFFA UNet的损失函数消融实验

Table 2 Loss function in MFFA UNet ablation experiment

损失函数	DSC/%	HD/mm	IoU/%	精确率/%	召回率/%
BceLoss	91.3	6.43	89.2	92.0	93.2
DiceLoss	92.6	6.20	89.5	92.2	93.1
Proposl Loss	94.8	5.01	91.5	94.5	95.2

表3 不同模型在数据集LITS上的对比实验

Table 3 Comparative experiments of different models on LITS dataset

模型	DSC/%	IoU/%	精确率/%	召回率/%
U-Net	91.5	87.9	92.2	92.5
UNet++	93.2	89.8	93.1	93.7
ResUNet	93.7	90.1	93.5	94.0
TransUNet	92.9	89.5	92.6	93.9
MFFA UNet	94.8	91.5	94.5	95.2

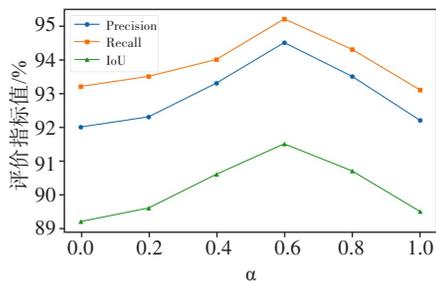


图6 超参数 $\alpha$ 消融实验

Figure 6 Ablation experiment on hyperparameter  $\alpha$

图7为不同模型在LITS数据集上分割结果对比图,可以看出多个模型都能分割出肝脏区域,但存在一定的欠分割和过分割现象。尤其是肝脏形状不规则、与周边器官距离较近、边界模糊时,U-Net、UNet++错误地将背景区域分割成目标肝脏。此外U-Net还存在对微小目标识别能力较差问题。而ResUNet和TransUNet分割出的肝脏结果,存在边缘

轮廓粗糙,边界有误等问题。与其他模型相比,本文方法MFFA UNet得到的分割结果更接近真实标签,在形状方面没有出现较大误差,边缘轮廓也较为精确,可以达到预期分割效果。

### 3 结论

本文提出一种基于多尺度特征融合与注意力(MFFA UNet)的方法,目的是从CT影像中精确分割出肝脏区域。通过低层特征的细节信息与高层特征的语义信息结合,提高特征的表示和利用。利用空间和通道注意力机制,使网络更加关注目标区域,抑制无关信息干扰。添加深度监督,充分利用不同层侧输出,并使用混合损失函数,进一步提升模型性能。通过实验证明,本文提出的方法在肝脏分割任务上优于其他语义分割网络。但本文提出的方法是

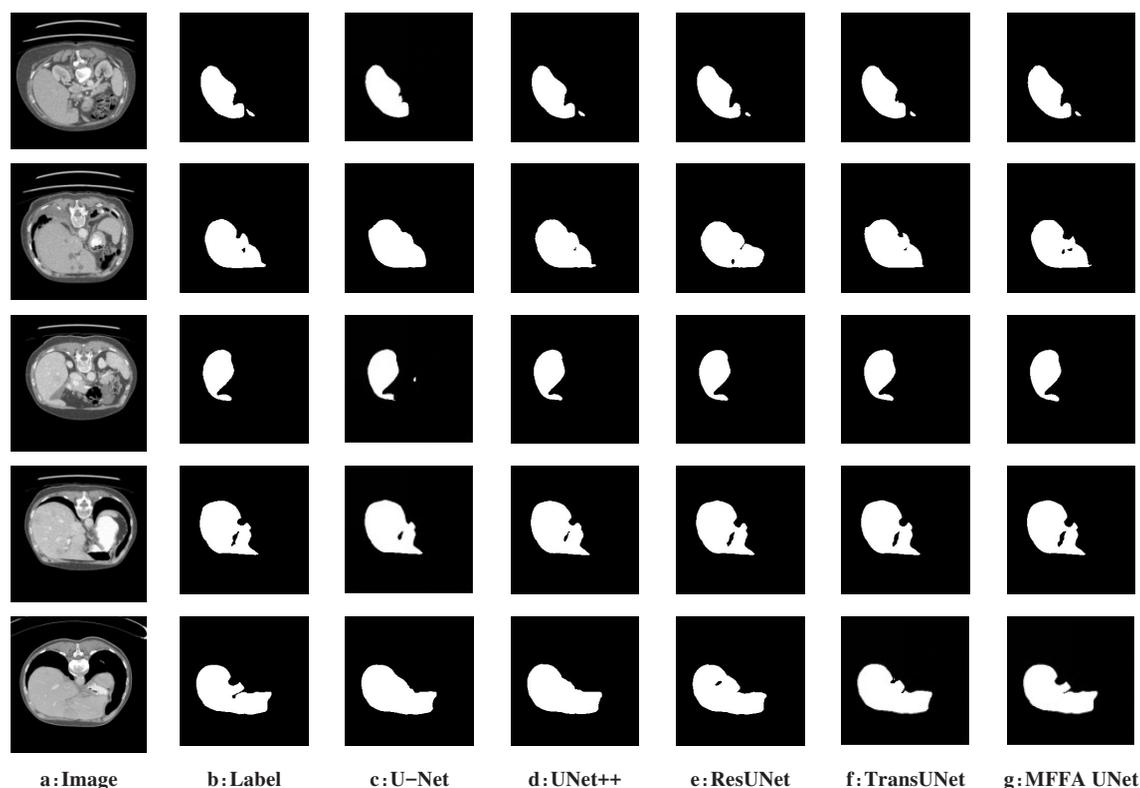


图7 不同模型肝脏分割对比图

Figure 7 Comparison of liver segmentation of different models

第1~5行图像代表不同CT切片的实验结果

基于2D切片数据的,只考虑了二维切片内的信息,忽略了切片间的关系,导致上下文信息丢失。因此,下一步工作重心是三维领域,将切片间的信息利用起来,以实现更为精确的分割。

## 【参考文献】

- [1] 曹毛毛, 陈万青. GLOBOCAN 2020 全球癌症统计数据解读[J]. 中国医学前沿杂志(电子版), 2021, 13(3): 63-69.  
Cao MM, Chen WQ. Interpretation on the global cancer statistics of GLOBOCAN 2020[J]. Chinese Journal of the Frontiers of Medical Science (Electronic Version), 2021, 13(3): 63-69.
- [2] 黄鹏, 郑洪, 梁超. 图像分割方法综述[J]. 武汉大学学报(理学版), 2020, 66(6): 519-531.  
Huang P, Zheng Q, Liang C. Overview of image segmentation methods [J]. Journal of Wuhan University (Natural Science Edition), 2020, 66 (6): 519-531.
- [3] Gambino O, Vitabile S, Lo Re G, et al. Automatic volumetric liver segmentation using texture based region growing [C]//2010 International Conference on Complex, Intelligent and Software Intensive Systems. Piscataway, NJ, USA: IEEE, 2010: 146-152.
- [4] Chen XJ, Pan LJ. A survey of graph cuts/graph search based medical image segmentation[J]. IEEE Rev Biomed Eng, 2018, 11: 112-124.
- [5] Wu YF, He CJ. Indirectly regularized variational level set model for image segmentation[J]. Neurocomputing, 2016, 171: 194-208.
- [6] Khairandish MO, Sharma M, Jain V, et al. A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images[J]. IRBM, 2022, 43(4): 290-299.
- [7] Pare S, Kumar A, Singh GK, et al. Image segmentation using multilevel thresholding: a research review[J]. Iran J Sci Technol Trans Electr Eng, 2020, 44(1): 1-29.
- [8] 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.
- [9] Zhou ZW, Rahman Siddiquee MM, 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. Cham: Springer International Publishing, 2018: 3-11.
- [10] Huang HM, Lin LF, Tong RF, 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). Piscataway, NJ, USA: IEEE, 2020: 1055-1059.
- [11] Alom MZ, Yakopcic C, Hasan M, et al. Recurrent residual U-Net for medical image segmentation[J]. J Med Imaging (Bellingham), 2019, 6(1): 014006.
- [12] Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: learning where to look for the pancreas[EB/OL]. (2018-05-20). <https://arxiv.org/abs/1804.03999>.
- [13] 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.
- [14] Cai SJ, Tian YX, Lui H, et al. Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network[J]. Quant Imaging Med Surg, 2020, 10(6): 1275-1285.
- [15] Tang YB, Tang YX, Zhu YY, et al. E2net: an edge enhanced network for accurate liver and tumor segmentation on CT scans[C]//Medical Image Computing and Computer Assisted Intervention - MICCAI 2020. Cham: Springer International Publishing, 2020: 512-522.
- [16] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000-6010.
- [17] Chen JN, Lu YY, Yu QH, et al. TransUNet: transformers make strong encoders for medical image segmentation[J]. (2021-02-08). <https://>

- arxiv.org/abs/2102.04306.
- [18] Jiang LF, Ou JJ, Liu RH, et al. RMAU-Net: residual multi-scale attention U-Net for liver and tumor segmentation in CT images[J]. *Comput Biol Med*, 2023, 158: 106838.
- [19] Feng SL, Zhao HM, Shi F, et al. CPFNet: context pyramid fusion network for medical image segmentation [J]. *IEEE Trans Med Imaging*, 2020, 39(10): 3008-3018.
- [20] Fan TL, Wang GL, Wang X, et al. MSN-Net: a multi-scale context nested U-Net for liver segmentation[J]. *Signal Image Video Process*, 2021, 15(6): 1089-1097.
- [21] Xie XW, Pan XP, Shao F, et al. MCI-Net: multi-scale context integrated network for liver CT image segmentation[J]. *Comput Electr Eng*, 2022, 101: 108085.
- [22] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 7132-7141.
- [23] Woo S, Park J, Lee JY, et al. CBAM: convolutional block attention module [C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [24] Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 13708-13717.
- [25] Wu YC, Xia Y, Song Y, et al. Vessel-Net: retinal vessel segmentation under multi-path supervision [C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. Cham: Springer International Publishing, 2019: 264-272.
- [26] Zhao RJ, Qian BY, Zhang XL, et al. Rethinking dice loss for medical image segmentation [C]//2020 IEEE International Conference on Data Mining (ICDM). Piscataway, NJ, USA: IEEE, 2020: 851-860.
- [27] Yeung M, Sala E, Schönlieb CB, et al. Unified focal loss: generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation[J]. *Comput Med Imaging Graph*, 2022, 95: 102026.
- [28] Bilic P, Christ P, Li HB, et al. The liver tumor segmentation benchmark (LiTS)[J]. *Med Image Anal*, 2023, 84: 102680.
- [29] Karimi D, Salcudean SE. Reducing the hausdorff distance in medical image segmentation with convolutional neural networks[J]. *IEEE Trans Med Imaging*, 2020, 39(2): 499-513.

(编辑:薛泽玲)