

## 融合注意力机制与残差可形变卷积的肝肿瘤分割方法

杨文瀚, 廖苗\*

湖南科技大学计算机科学与工程学院, 湖南 湘潭 411201

**摘要** 手术与化疗作为肝癌的主要治疗手段需要精确提取肝脏病变区域。针对目前肝肿瘤分割方法存在的小型肿瘤丢失、肿瘤边界分割模糊、分割严重错误等问题,提出一种融合注意力机制与残差可形变卷积的肝肿瘤分割方法。以U-Net为主干网络,在编码卷积层末尾增加一条带有反卷积与激活函数的残差路径,该路径与上层跳跃连接相连,解决池化与反卷积操作中的信息损失造成的小目标分割缺失与边界模糊问题;利用可形变卷积增强模型对肿瘤边界的特征提取能力;在跳跃连接层中添加一定数量的卷积层,弥补简单跳跃连接在特征融合时造成的语义空白;通过双注意力机制,模型更加关注肿瘤特征;采用混合损失函数,该函数在保证训练稳定的情况下解决类不平衡造成的分割性能下降的问题。在肝脏肿瘤公开数据集(LITS)上进行实验,所提方法的肿瘤分割Dice系数达85.2%,分割性能优于其他对比网络,能够达到辅助医疗诊断的要求。

**关键词** 肝癌; 肿瘤分割; U-Net; 残差结构; 注意力

中图分类号 TP391.41

文献标志码 A

DOI: 10.3788/LOP221369

## Fusion of Attention Mechanism and Deformable Residual Convolution for Liver Tumor Segmentation

Yang Wenhan, Liao Miao\*

School of Computer Science and Engineering, Hunan University of Science and Technology,  
Xiangtan 411201, Hunan, China

**Abstract** Surgery and chemotherapy, as the main treatments for liver cancer, require accurate extraction for the liver lesion area. Therefore, to solve the problems of the current segmentation methods for liver tumors, such as the loss of small-sized tumors, fuzzy segmentation of tumor boundaries, and severe missegmentation, a new method for liver tumor segmentation based on the attention mechanism and deformable residual convolution is proposed. U-Net was used as the backbone network, and a residual path with deconvolution and activation function was added at the end of the encoding convolution layer to connect with the upper layer, thereby solving the problem of missing small target segmentation and fuzzy boundaries caused by information loss in pooling and deconvolution operations. Furthermore, a deformable convolution was used to enhance the model for extracting features of tumor boundaries. Several convolution layers were added to the skip connection layer to compensate for the semantic gaps caused by simple skip connections in feature fusion. The model pays more attention to tumor characteristics through the dual-attention mechanism. The mixed loss function was used to address the problem of segmentation performance degradation caused by a class imbalance under the condition of ensuring the stability of training. The experiment was carried out using the Liver Tumor Segmentation Challenge (LITS) dataset. The experimental results show that the Dice coefficient of tumor segmentation of the proposed method reaches 85.2%. Moreover, the proposed method has a better segmentation performance than other comparison networks, meeting the requirements of auxiliary medical diagnosis.

**Key words** liver cancer; tumor segmentation; U-Net; residual structure; attention

## 1 引言

肝癌以死亡率高、诊断难等特点严重威胁着人类

的生命健康。现代医学技术针对肝癌的治疗手段主要以化疗与手术切除为主<sup>[1]</sup>。两种治疗方式对治疗靶区即肿瘤区域有很大程度的依赖,精确的治疗靶区的获

收稿日期: 2022-04-20; 修回日期: 2022-05-09; 录用日期: 2022-05-25; 网络首发日期: 2022-06-10

基金项目: 国家自然科学基金(62272161)、湖南省教育厅资助科研项目(20B239)、湖南省自然科学基金(2021JJ30275)

通信作者: \*Liaomiaohi@126.com

取能为后续治疗方案的展开提供巨大帮助。计算机断层成像(CT)作为常用的医学成像技术,能直观地反映肝脏区域肿瘤的位置、大小、数量、形态等信息,能为手术与化疗提供精确定位,提高手术与化疗的成功率<sup>[2]</sup>。目前临床上肿瘤区域提取任务主要由富有经验的医师在CT图像中采取手动分割的方式实现,不仅耗时费力、复现性差,而且存在着严重的主观倾向。随着人们对自身健康的关注程度不断升高,高精度、全自动的肝脏肿瘤分割方法成为医学图像分割领域的一个研究热点。

现阶段针对CT图像中肝脏肿瘤的分割算法大致分为两类:基于传统的方法与基于深度学习的方法。传统方法诸如区域生长<sup>[3-4]</sup>、水平集<sup>[5]</sup>、图割<sup>[6]</sup>、活动轮廓模型<sup>[7-8]</sup>、支持向量机<sup>[9-10]</sup>和聚类<sup>[11-12]</sup>等难以有效适应肝脏肿瘤的形状多变、对比度低、个体差异大等情况,且该类方法通常需要人工设置参数或者提取特征,主观性强。近年来,深度学习以精度高、鲁棒性强、无需人工干预等优势被广泛运用在了医学图像分割领域。2017年Shelhamer等<sup>[13]</sup>提出的全卷积神经网络(FCN)让图像分割实现了端对端的像素级输出。在FCN基础上,Ronneberger等<sup>[14]</sup>采用编码器-解码器(ED)结构构建出U型网络(U-Net),U-Net通过跳跃连接层连通浅层与深层,紧密连接上下文信息,提高模型的信息利用率,对数据量小的医学图像数据集有较好的分割效果,成为目前医学图像分割领域的主流基准网络。在此之后,许多研究人员对U-Net进行了改进,取得了不错的成果。Milletari等<sup>[15]</sup>将U-Net中所有卷积模块之间的连接方式改为残差连接,保证网络能快速收敛的同时提高了U-Net的可扩展性。Li等<sup>[16]</sup>提出了一种瓶颈监督的U-Net模型(BS-Unet),BS-Unet通过引入膨胀卷积、稠密连接和先启(inception)模块扩大了感受野,增强了深层特征提取能力与提高了泛化性,但大量的池化与跨卷积操作会导致图像空间信息丢失。Schlemper等<sup>[17]</sup>提出融合空间注意力机制的Attention U-Net模型,该模型更加关注感兴趣区域,在图像分割任务上取得了一定成果,但是该方法缺乏对图像通道信息的考虑。Zhou等<sup>[18]</sup>对U-Net的跳跃连接层进行了重新设计,以相同的分辨率进行了密集连接,构建了一种多尺度的U-Net网络(U-Net++),以弥补简单跳跃连接造成的不同层次特征在融合时的语义空白。Li等<sup>[19]</sup>将U-Net++应用于CT图像的肝脏及肿瘤分割,并引入了深度监督机制与混合损失函数优化分割精度。Huang等<sup>[20]</sup>在U-Net++的基础上,引入了全尺度的跳跃连接和深度监督机制,提出U-Net3+网络,进一步提升了分割精度,同时降低了模型参数量,但是复杂的连接方式容易造成过拟合现象,并且没有解决数据类不平衡问题造成的分割困难问题。高飞等<sup>[21]</sup>利用加权的损失函数解决了数据类不平衡问题,有效提高了对肝脏肿瘤的分割精度。刘一鸣等<sup>[22]</sup>

提出结合空间注意力与通道注意力的双注意力机制的U-Net,该网络利用双注意力机制对跳跃连接层中特征图与上采样特征图进行更加有效的特征融合,进而提高分割精度,但是没有考虑融合特征图之间的层次与语义差异。

为了弥补上述方法存在的不足,提高肝脏肿瘤分割精度,本文提出了一种融合注意力机制与残差可形变卷积的肝脏肿瘤分割方法。所提方法的主要创新点如下。

1)重新设计U-Net结构。将U-Net中卷积层连接方式改为残差连接,避免梯度消失与梯度弥散现象;在编码区卷积层末端添加一条带有反卷积与激活函数的残差路径,并将其连接至上层跳跃连接,以解决池化与反卷积操作造成的小目标缺失与边界模糊的问题;在跳跃连接层中增加一定数量的卷积层,解决不同层次特征融合时产生的语义鸿沟。

2)引入可形变卷积,通过在标准卷积每个采样点处增加一个偏置量,卷积核可以在原位置附近区域采样,扩大了感受野。针对肿瘤边界的非刚性特点,形变卷积相比于标准卷积更能提取其中特征。

3)在跳跃连接层末端引入结合空间注意力与高效通道注意力的双注意力模块,使网络关注肝脏区域中的肿瘤特征,抑制其他无关信息传递。

4)结合Dice损失函数与Focal损失函数,提出一种混合损失函数,利用Dice函数关注分割结果与金标准相似度的特征和Focal函数能稳定回传梯度的特征,让模型快速收敛,同时解决数据集中类不平衡的问题。

## 2 所提方法内容

### 2.1 模型结构

所提方法采用U-Net结构作为主干网络。U-Net通过跳跃连接的方式将编码区的浅层特征与解码区的深层特征相连,结合了上下文信息,提高了模型的特征利用率,在医学图像分割领域中运用广泛。

为了进一步提高U-Net在肝脏肿瘤分割任务上的分割精度,对U-Net结构进行如下改进,具体结构如图1所示。本文认为池化与反卷积操作是分割任务中小型肿瘤缺失和分割边界模糊的主要原因,为此对编码区的卷积模块进行改进,提出了一种残差分支编码(RBE)模块,替换掉传统U-Net中第一层以外的双卷积模块。所有卷积层的连接方式改为残差连接。通过对经过池化-卷积-反卷积-激活操作后的高分辨率特征图与上层的跳跃连接层的特征图进行加和,利用残差的思想,让模型学习到特征图与其经过池化-卷积-反卷积-激活后的特征图之间的残差,即小目标信息与边缘信息,这些信息通过跳跃连接层进行传递。这样在特征融合阶段小目标信息与边缘信息能够更好地在分割结果中得到映射。同时在跳跃连接层中添加一定数

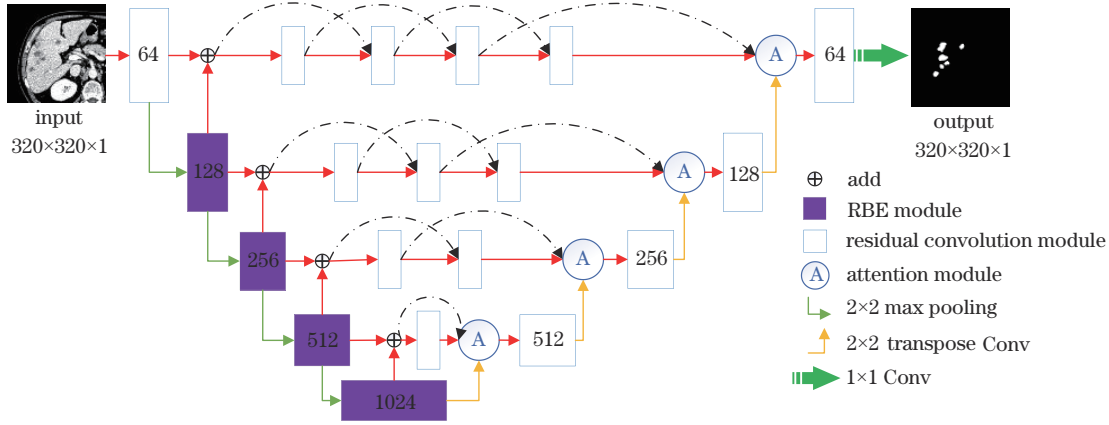


图 1 改进 U-Net 模型的结构

Fig. 1 Structure of improved U-Net model

量的残差卷积模块,以解决不同层次信息在融合时产生的语义鸿沟。使用可形变的卷积代替普通卷积,因为肝脏肿瘤的边界形状不规则且变化程度很大,具有典型的非刚性特征。传统的卷积核是规则的矩形结构,对不规则形状的特征提取效果不够理想。可形变卷积通过在传统卷积核的采样点上添加一个偏置量,让卷积核的采样点能够在原采样点的附近进行卷积运算,扩大了感受野,更便于捕获各种形状与尺度的肝脏肿瘤。引入双注意力模块,让网络模型着重关注肝脏区域的肿瘤特征而抑制其他无关特征的传递。最后在跳跃连接层中加入一定数量的卷积层,并且以残差连接的方式连接卷积层。

从图 1 可以看到,所提网络结构沿用了 U-Net 的编码器-解码器结构,从上到下一共 5 个层次,整个网络分为三个部分,即编码部分、解码部分、跳跃连接部分。网络主要由残差卷积模块构成,该模块结构如图 2 所示,由 2 个相同的卷积层构成,卷积层由 1 个  $3 \times 3$  的可形变卷积、批归一化与 ReLU 激活函数组成,连接方式为残差连接。第 2、3、4、5 层的编码模块替换为所提出的 RBE 模块,其具体结构如图 3 所示,输入特征经过一系列操作后形成两个输出特征,即输出特征 1 与输出特征 2。其中输出特征 1 输出到本层的跳跃连接层与下采样操作,输出特征 2 会与上一轮的跳跃连接特征相加。不同尺度的跳跃连接层中添加了一定数

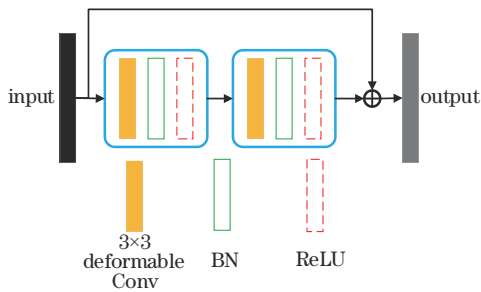


图 2 残差卷积模块的结构图

Fig. 2 Structure of residual convolution module

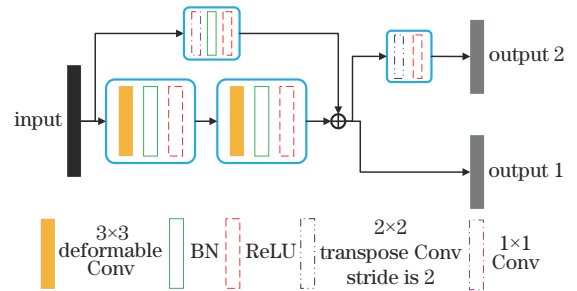


图 3 RBE 模块的结构

Fig. 3 Structure of RBE module

量的残差卷积模块,末端添加了融合双注意力机制的注意力模块。在编码部分,输入图像每向下增加一层,通道数增加为原来的 2 倍,分辨率改变为原来的  $1/2$ ;在解码部分,特征图每向上增加一层,通道数改变为原来的  $1/2$ ,分辨率扩大为原来的 2 倍。解码区末端连接一个  $1 \times 1$  大小的卷积,生成最终的分割结果图,输出的结果图的分辨率大小保持与输入图像相同。

### 2.2 可形变卷积

神经网络中卷积核的作用是提取物体特征。传统的卷积核形状是矩形的,对固定形状的目标特征提取较好,但是对形状不固定的非刚性目标则缺乏适应性,泛化能力不强。而肝脏肿瘤的形状多变,具有典型的非刚性特点,使用传统卷积核对其特征进行提取时提取能力具有一定的局限性,因此引入可形变卷积<sup>[23]</sup>来解决这种问题。可形变卷积在传统的卷积核中的每个采样点上增加一个偏移量,让每个采样点能够在原采样点的附近区域得到采样,而不局限于矩形区域。该区域可以是任意形状,这样扩大了感受野,也增强了对形变目标的特征提取能力。

图 4(a)为传统卷积核,每个圆点表示卷积核中的每个采样点;图 4(b)为可形变卷积核,图中箭头表示增加的偏移量,箭头的起始端为传统卷积核采样点,箭头末端为偏移后的新采样点,可以看到偏移后的采样点已经不再是矩形形状,变成了不规则的形状。传统

卷积中,对于给定的输入特征图  $X$ ,输出特征图  $Y$  中每个像素  $p_0$  的表达式为

$$Y(p_0) = \sum_{p_n \in R} w(p_n) \cdot X(p_0 + p_n), \quad (1)$$

式中:  $w(p_n)$  表示卷积核在  $p_n$  上的权值;  $R \in \{(-1, 1), (0, 1), \dots, (1, 1)\}$ 。可形变卷积在采样点的坐标中添加了一个坐标偏移量  $\Delta p_n$  来实现采样点的坐标偏移,表达式为

$$Y(p_0) = \sum_{p_n \in R} w(p_n) \cdot X(p_0 + p_n + \Delta p_n), \quad (2)$$

该偏移量通过卷积学习获取,可以是小数。卷积核中每个采样点拥有两个偏移值,每个偏移值代表一个轴。

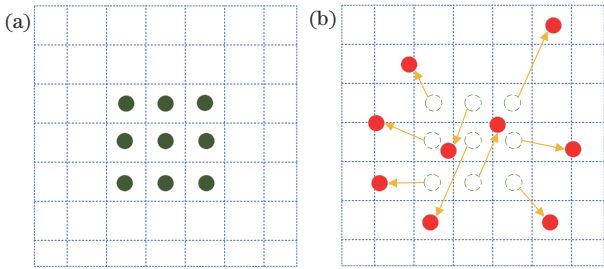


图 4 可形变卷积示意图。(a)传统卷积核;(b)可形变卷积核  
Fig. 4 Schematic of deformable convolution. (a) Traditional convolution kernel; (b) deformable convolution kernel

可形变卷积的过程:首先在原始特征图上运用一个传统卷积核进行特征提取,得到一个长宽与原特征图相同的通道数为  $2N$  的新特征图,其中 2 对应  $x$  轴与  $y$  轴,  $N$  代表卷积核中采样点的个数;然后该特征图中的偏移值作用在原传统卷积核上,形成可形变卷积核;最后使用可形变卷积核进行特征提取。由于  $\Delta p_n$  的值可能是小数,因此采用双线性插值计算  $X(p_0 + p_n + \Delta p_n)$ 。

### 2.3 双注意力机制

注意力机制来源于人类的视觉科学,当人类的视觉在关注某一片区域时对视野里其他区域的关注度会明显降低,对于肝脏肿瘤分割任务来说,就可以借助注意力机制来改善所构建的神经网络。传统 U-Net 中,为了解决特征图上采样带来的大量空间信息缺失的问题,引入了跳跃连接层,将编码区域的特征图拼接到解码区对应层级。但是这样在拼接的特征图中会有大量冗余的特征信息,冗余信息在特征融合时会对分割结果造成一定的干扰,导致错误分割。运用空间注意力与通道注意力机制对跳跃连接层中的特征信息进行处理,对重要的通道中的特征信息和分割相关区域的特征信息进行激活,抑制无关区域特征信息的激活,减少冗余信息在跳跃连接层中的传递,让跳跃连接层中传递的信息能在特征融合时精确映射,弥补上采样过程造成的特征信息缺失,从而达到提高分割精度的效果。

结合通道注意力与空间注意力机制,提出一种具有双注意力机制的结构模型,如图 5 所示。该注意力结构首先将跳跃连接层特征  $F_1$  和上采样后的特征  $F_2$  分别输入通道注意力模块,得到新特征  $F_{1a}$  和  $F_{2a}$ ;再对  $F_{1a}$  和  $F_{2a}$  分别进行  $1 \times 1$  卷积,得到  $F_{1b}$  和  $F_{2b}$ ;并将卷积结果  $F_{1b}$  和  $F_{2b}$  相加,得到特征  $F_c$ ;接着对相加结果依次进行 ReLU 激活和  $1 \times 1$  卷积操作,并将卷积后的结果输入 Sigmoid 激活函数,生成空间注意力权重图;最后将跳跃连接层传递的原始特征  $F_1$  与该权重图相乘,得到最终的输出特征图。空间注意力模块对经过通道注意力模块的特征  $F_{1a}$  与转置特征  $F_{2a}$  分别进行  $1 \times 1$  卷积然后进行加和,对相同的感兴趣区域的特征信号进行了加强,并通过一系列线性与非线性变换形成了一个空间区域的权重图,通过原特征图与该权重图相乘,实现对分割相关区域特征的激活和对无关区域的抑制。

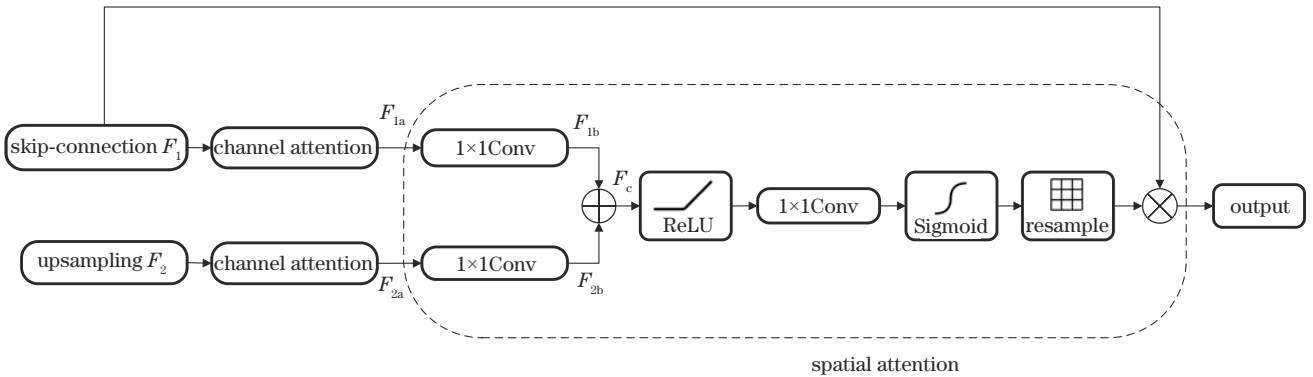


图 5 双注意力结构模型  
Fig. 5 Dual attentional structure model

通道注意力模块的具体结构如图 6 所示。该结构是一种高效的通道注意力模型,首先通过全局平均池化得到集中的特征;然后采用一维卷积生成通道的共

享权重,根据通道的维度建立非线性映射,确定一维卷积中卷积核的大小;最后经过 Sigmoid 函数激活,生成通道带状权重矩阵。通过输入特征  $T_1$  与权重矩阵相

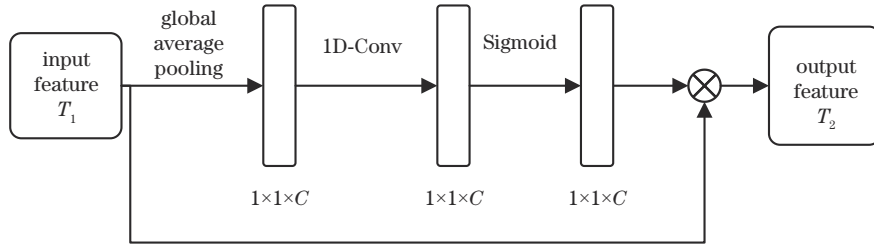


图 6 通道注意力结构

Fig. 6 Structure of channel attention

乘,最后得到输出特征  $T_2$ 。其中通道数  $C$  与一维卷积的卷积核大小  $k$  之间的关系为

$$C = 2^{(\gamma \times k - q)}, \quad (3)$$

$$k = \left\lceil \frac{\log_2 C}{\gamma} + \frac{q}{\gamma} \right\rceil_{\text{odd}}, \quad (4)$$

式中:  $\gamma$  和  $q$  表示常数项系数,分别设置为 2 和 1。通过这样的形式将通道注意力机制与空间注意力机制有效融合在一起,既能有效利用特征图的空间信息,也能充分考虑到不同通道的像素点的类别权重,能提高对上下文信息的利用率,从而保证更加精确的分割结果。

#### 2.4 混合损失函数

Dice 损失函数是医学图像分割中常用的损失函数,它能指导网络通过不断的学习让分割结果不断逼近金标准。Dice 损失函数的公式为

$$l_{\text{Dice}} = 1 - \frac{2 \times \sum_{i=1}^N G_i P_i}{\sum_{i=1}^N G_i + \sum_{i=1}^N P_i}, \quad (5)$$

式中:  $G_i$  为 CT 图像中第  $i$  个像素的专家手动分割结果(背景标记为 0,目标标记为 1);  $P_i$  为网络预测第  $i$  个像素属于目标的概率;  $N$  为 CT 图像中的像素数目。Dice 损失函数训练过程中如果预测值与真实值的差距变化较大,会出现损失值函数曲线混乱,难以观察出收敛信息,甚至无法收敛的情况。二元交叉熵(BCE)损失函数也可以运用在二分类语义分割中,它主要关注每个像素属于某个类别的概率,其值是背景区域和目标区域中所有像素交叉熵的均值,表达式为

$$l_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [G_i \log P_i + (1 - G_i) \log(1 - P_i)]. \quad (6)$$

该损失函数具有能稳定回传梯度信息、快速收敛的特点,能有效弥补 Dice 函数训练过程中梯度变化剧烈的缺陷。但是一旦出现目标区域像素占比远小于背景像素的这种不平衡情况时,BCE 损失函数的预测会更加趋向于背景。Focal 损失函数能很好地解决这个问题,表达式为

$$l_{\text{Focal}} = -\frac{1}{N} \times$$

$$\sum_{i=1}^N [(1 - P_i)^{\gamma_1} \times G_i \log P_i + P_i^{\gamma_1} \times (1 - G_i) \log(1 - P_i)], \quad (7)$$

式中:  $\gamma_1$  为超参数。Focal 损失函数是在 BCE 损失函数的基础上进行改进的,通过引入权重因子  $(1 - P_i)^{\gamma_1}$ ,改善正负差距过大引起的类不平衡问题。本文结合 Focal 函数的优点与 Dice 函数的优点,提出一种混合损失函数,表达式为

$$l_{\text{mix}} = l_{\text{Focal}} - \alpha \ln l_{\text{Dice}}, \quad (8)$$

式中:  $\alpha$  为权重参数,经多次实验设定该值为 0.01。

### 3 实验内容

#### 3.1 数据集

采用由 Codalab 网站提供的肝脏肿瘤挑战赛(LITS)公共数据集<sup>[24]</sup>。该数据集包含 131 个序列病例的训练集与 70 个病例的测试集,训练集数据由来自不同机构的 4 名放射科医师对肝脏与肝脏肿瘤区域进行手动标注,而测试集数据没有标注。结果只能在网站上提交后查看。所有图像分辨率为  $512 \times 512$ ,切片数量为 42~1026,片层间距为 0.45~6.0 mm,训练集中序列号为 32、34、38、41、47、87、89、91、105、106、114、115、119 的病例 CT 图像中不包含肝脏肿瘤,切片总数为 58638 张,测试集第 59 号序列的扫描方向与其他测试集数据不同,切片数量为 34880 张。采集的数据来自世界各地不同机构,图像的扫描方向、质量、空间分辨率也存在明显差异,这种差异在测试集与训练集的 CT 图像中更为明显,虽然增加了模型分割难度,但是该数据集是目前肝脏肿瘤分割研究中使用最广泛的数据集。图 7 展示了 LITS 数据集中的部分 CT 图像。从 LITS 训练集中的 131 个病例中随机选取包含肝脏器官的 7183 幅 CT 图像作为实验数据集,其中训练集和验证集的比例设置为 8:2。

#### 3.2 数据预处理

为了提高 CT 图像中肿瘤区域像素占比与图像的对比度,便于特征提取,对训练集与测试集的 CT 图像进行了裁剪与窗口和窗宽的设定。在保证肝脏区域完

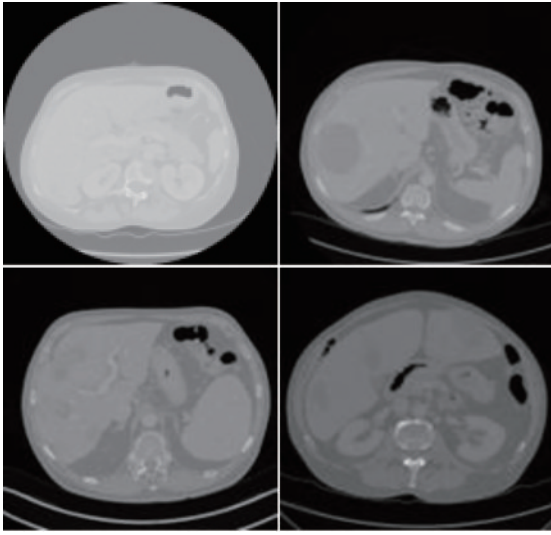


图7 LITS数据集的部分图像

Fig. 7 Partial images of LITS dataset

全显示的情况下将图像由  $512 \times 512$  分辨率大小裁剪为  $320 \times 320$ 。原始 CT 图像中肿瘤区域占比较小,且与其他组织器官缺乏良好的灰度对比,直接分割往往精度较低。为了降低肿瘤分割的复杂性,提高分割精度,首先对原始 CT 图像进行裁剪、窗位和窗宽调整等预处理。根据解剖学先验,肝脏通常位于人体腹腔右部。在 CT 断层扫描中,肝脏区域主要出现在横切面视角成像图像的腹腔左侧,如图 8(a)所示。为了获取包含完整肝脏的感兴趣区域,去除不相关组织器官的干扰,需对 CT 图像进行裁剪。具体操作包括:首先采用阈值和数学形态学方法<sup>[25]</sup>对 CT 图像中的肋骨和脊椎进行分割,得到如图 8(b)所示结果;然后对分割结果进行行和列投影,且分别取第一个以及最后一个不

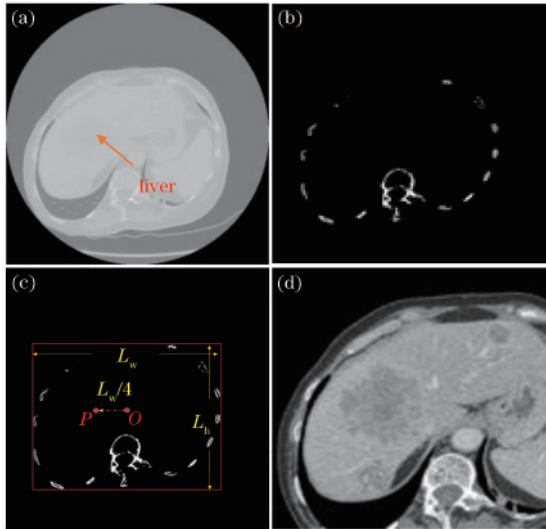


图8 数据预处理。(a)原始CT图像;(b)肋骨和脊椎分割结果;(c)裁剪示意图;(d)预处理图像

Fig. 8 Data pre-processing. (a) Original CT image; (b) segmentation result of ribs and spine; (c) cropping diagram; (d) pre-processed image

为零的像素所在的行和列,构建如图 8(c)所示长为  $L_w$ 、宽为  $L_h$  的矩形框,定位腹腔区域;考虑到肝脏区域通常位于 CT 切片腹腔左侧,将矩形框中心点  $O$  向左平移  $L_w/4$  得到点  $P$ ,并以点  $P$  为中心取大小为  $320 \times 320$  的方形区域作为最终的裁剪结果;最后对裁剪区域图像的窗位与窗宽进行调整。根据放射医学先验知识,可显示的人体组织 CT 值范围一般为  $-1000 \sim 1000$  Hu,而肝、脾、肾、肌肉等器官软组织的 CT 值范围通常为  $40 \sim 70$  Hu,为了增加肝脏与毗邻组织的对比度,将 CT 图像的窗位和窗宽分别设置为 60 Hu 和 170 Hu,此时图像的 CT 值截断范围为  $-25 \sim 145$  Hu。从图 8(d)可以看出,预处理后的 CT 图像中去除了大量不相关的组织与器官,肿瘤像素占比增大且肿瘤与周围组织的对比度显著提高。所有数据在输入模型之前都进行了归一化处理。

### 3.3 评价指标

为了定量评价网络分割性能,使用 5 种常用的评价指标,包括 Dice 系数、体积重叠误差 (VOE)、相对体积差异 (RVD)、平均表面距离 (ASD) 和最大对称表面距离 (MSSD)。

Dice 系数用于计算分割结果  $A$  与金标准  $B$  之间的重叠指数,计算公式为

$$P_{\text{Dice}} = \frac{2|A \cap B|}{|A| + |B|} \times 100\%, \quad (9)$$

式中:  $|\cdot|$  表示求给定区域的像素数目;  $P_{\text{Dice}}$  取值范围为  $[0, 1]$ , 0 和 1 分别表示分割结果与金标准之间完全不重叠或完全重叠。VOE 与 Dice 系数相反,值越小表明算法分割性能越好,计算公式为

$$P_{\text{VOE}} = 1 - \frac{|A \cap B|}{|A \cup B|} \times 100\%。 \quad (10)$$

RVD 用于衡量分割结果与金标准的相对体积差,表达式为

$$P_{\text{RVD}} = \frac{|B| - |A|}{|A|} \times 100\%。 \quad (11)$$

RVD 正或负值表明算法倾向于欠分割或过分割,但其值大小并不能直接反映分割性能的优劣,需结合其他指标进行综合评价。ASD 和 MSSD 分别表示分割结果与金标准之间的平均表面距离和最大对称表面距离,表达式分别为

$$P_{\text{ASD}} = \frac{1}{|S(A)| + |S(B)|} \left\{ \sum_{I_a \in S(A)} D[I_a, S(B)] + \sum_{I_b \in S(B)} D[I_b, S(A)] \right\}, \quad (12)$$

$$P_{\text{MSSD}} =$$

$$\max \left\{ \max_{I_a \in S(A)} D[I_a, S(B)], \max_{I_b \in S(B)} D[I_b, S(A)] \right\}, \quad (13)$$

式中:  $S(\cdot)$  表示指定区域的表面像素;  $D[p, S(\cdot)]$  表示像

素  $p$  到  $S(\bullet)$  的欧氏距离。ASD 和 MSSD 单位均是 mm, 且值越小表明算法分割性能越好。

### 3.4 实验环境与参数设置

所提网络模型是在基于 Python 语言的 PyTorch 框架下实现的。硬件配置: AMD Ryzen 5900X 处理器, 32 GB 内存和单张 GeForce RTX 3080 显卡。采用 Adam 优化算法。学习率为  $10^{-4}$ , 一阶矩的指数衰减率为 0.9, 二阶矩的指数衰减率为 0.999, 用于维持数值稳定的常数为  $1 \times 10^{-8}$ , 权重衰减为  $10^{-6}$ 。为了避免过拟合, 网络训练过程中还采用了 Dropout 策略, 其值设置为 0.5。epoch 和 batchsize 分别设置为 100 和 6。在 Codalab 网站上线上提交网络模型的测试结果。

## 4 结果与分析

### 4.1 概率图可视化对比

为了验证分割方法的有效性, 对不同网络末端经  $1 \times 1$  卷积后输出的分割结果概率图进行可视化对比, 对比结果如图 9 所示。对比了 U-Net<sup>[14]</sup>、Attention

U-Net (A-Unet)<sup>[17]</sup>、U-Net++<sup>[18]</sup>、U-Net3+<sup>[20]</sup>、MDCC-Unet<sup>[26]</sup> 与所提网络共 6 种网络。可以看到: 传统的 U-Net 存在严重的欠分割问题, 对形状较大的肿瘤分割效果较好, 而对形状较小的肿瘤都未能检出; 引入空间注意力机制的 Attention-Unet 明显改善了小型肿瘤检出率低的问题, 但是对肿瘤的边界分割仍然不够精确; U-Net3+ 是在 U-Net++ 基础上提出的多尺度密集连接的结构, 但是在这张图像的分割中, 出现了较多概率值在 0.5 左右的模糊分割区域, 且误分割现象严重, 分割效果不如 U-Net++ 好, 可以看出单纯增加网络连接的复杂程度在某些情况下不一定带来更好的分割效果; 综合来看, MDCC-Unet 与所提方法的分割效果较好, 更加贴近于金标准, 但是 MDCC-Unet 在肿瘤边缘区域的像素概率值较低, 对于肿瘤边界的判定不如所提方法精确。由此可以说明, 所提方法能够提高小肿瘤检出率并且获得更加精确的肿瘤分割边界, 进而提高对整体肝脏肿瘤的分割精度。

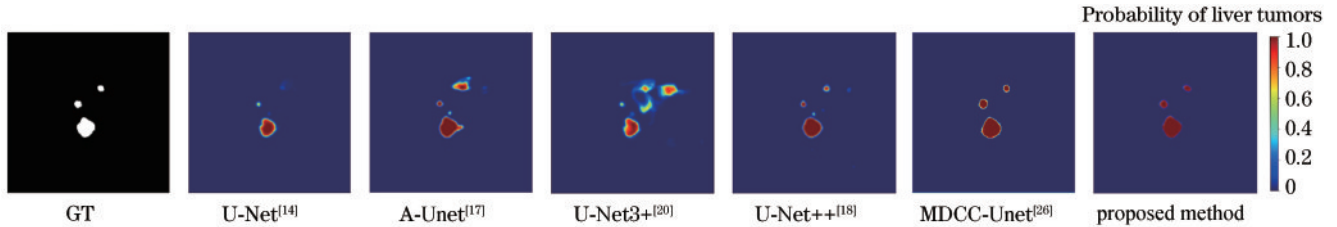


图 9 不同方法的概率图可视化对比

Fig. 9 Visual comparison of probability graphs of different methods

### 4.2 消融实验

为了验证不同模块对分割精度的影响, 以 U-Net 为基准进行了大量消融实验, 损失函数使用相同参数的混合损失函数。实验结果如表 1 所示, 可以看出: U-Net 对肝脏肿瘤分割的 Dice 值为 71.8%, 单独添加双注意力模块 (dual attention)、新的跳跃连接结构 (new skip-connection)、可形变卷积 (deformable Conv)、RBE 四个模块后, 分割精度得到一定的增加。其中 RBE 模块对分割精度提升最高, Dice 系数达

82.0%, VOE 指数下降到 36.7%; 可形变卷积模块对 RVD 评价指标与对分割边界评价指标 ASD 与 MSSD 的提升明显, RVD 上升到 -12.4%, ASD 与 MSSD 指标分别减小到 1.56 mm 与 7.92 mm。在模块的组合过程中, 新增的模块都能有效地改善肿瘤分割精度, 整体指标呈现上升趋势, 所有模块加入后分割效果达到最佳。

### 4.3 不同方法对比

由于 LITS 测试集没有提供标签, 测试集的分割

表 1 消融实验对比结果

Table 1 Comparison results of ablation experiments

Module	Dice / %	VOE / %	RVD / %	ASD / mm	MSSD / mm
U-Net	71.8	48.5	-25.3	2.44	9.13
U-Net+new skip-connection	73.1	47.2	-22.1	2.01	9.01
U-Net+dual attention	79.3	37.2	-11.5	1.69	8.13
U-Net+deformable Conv	76.8	40.1	-12.4	1.56	7.92
U-Net+RBE	82.0	36.7	-7.3	1.34	7.32
U-Net+deformable Conv+dual attention	81.9	38.8	-8.7	1.42	7.54
U-Net+RBE+deformable Conv+dual attention	83.8	35.3	-6.2	1.29	7.21
U-Net+all modules	85.2	32.4	-3.2	0.98	6.61

结果无法与金标准(GT)进行定性对比。因此选取所提方法在验证集上的分割结果与其他网络进行定性对比。图 10 展示了所提方法和其他不同网络以及金标准之间的部分分割结果对比。其中第 1 列为原始 CT 图像,第 2~7 列为 U-Net<sup>[14]</sup>、Attention U-Net<sup>[17]</sup>、U-Net++<sup>[18]</sup>、U-Net3+<sup>[20]</sup>、MDCC-Unet<sup>[26]</sup>与所提方法的分割结果,最后一列为医生手动分割的金标准。

从图 10 可以看出:传统的 U-Net 在肿瘤形状规则

并且像素占比较大的情况下分割效果才较好,对小型肿瘤检出程度不高,欠分割现象比较严重;A-Unet、U-Net++、U-Net3+ 相比于 U-Net 分割效果有一定提升,但是仍然存在着小型肿瘤缺失、边界模糊、分割错误等现象;MDCC-Unet 与所提方法的分割效果明显优于前几种对比网络,没有出现小肿瘤丢失、明显的错误分割现象,相比而言,所提方法在肿瘤边界的判定上比 MDCC-Unet 接近金标准,整体分割效果更优。

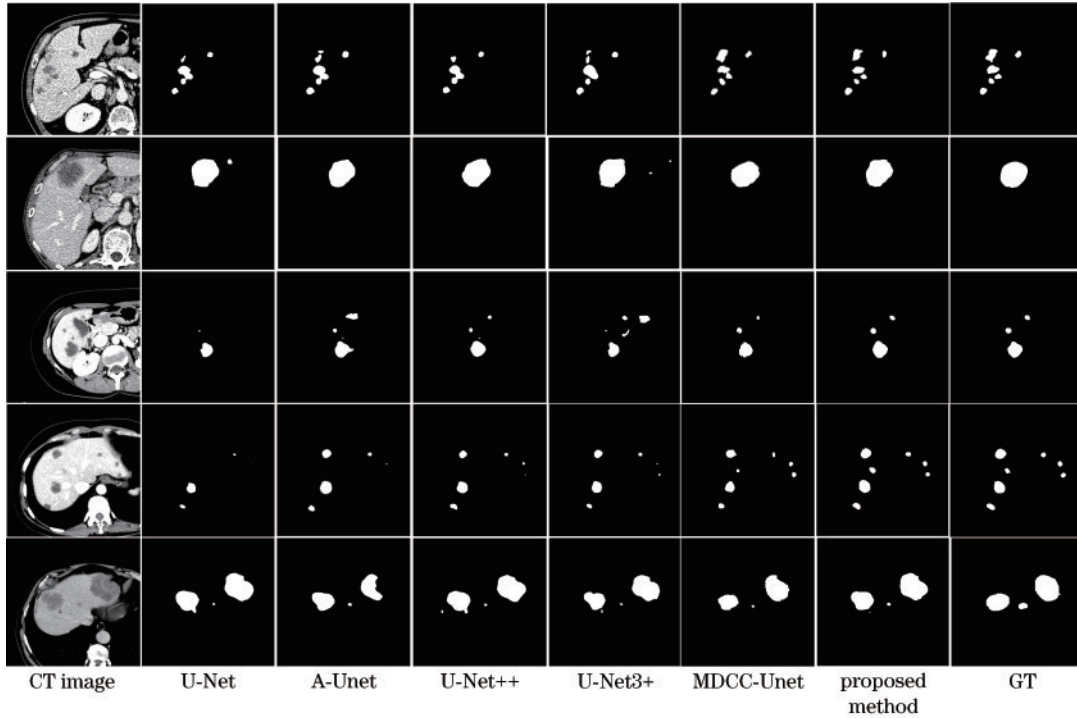


图 10 不同网络的分割结果

Fig. 10 Segmentation results of different networks

采用 LITS 测试集对不同方法进行测试,并且使用 Dice、VOE、RVD、ASD、MSSD 五种评价指标对分割结果进行定量对比,结果如表 2 所示。从表 2 可以看出:所提方法相对于传统的 U-Net 模型,分割效果提升显著,其中衡量整体分割精度的 Dice 系数指标从 71.8% 提升到了 85.2%,相比于第二名的 MDCC-

Unet 依然有 1.3 个百分点的提升;对于衡量分割边界准确度的 ASD 与 MSSD 两个评价指标,所提方法相比于 U-Net,从 2.42 mm 与 9.13 mm 下降到了 0.98 mm 与 6.61 mm,ASD 指标相对于第二名 MDCC-Unet 下降了 0.24 mm,MSSD 指标相比于第二名 U-Net++ 下降了 0.51 mm,其他各项指标也均优于对比网络,且提升效果显著。由此可说明,可形变卷积在肝脏肿瘤分割任务上比传统卷积能更有效地提取相关特征,双注意力机制的引入、RBE 模块的引入、在跳跃连接层中添加残差卷积模块的方案能有效提高对肿瘤边界的分割精度,进而使 U-Net 模型在肝脏肿瘤分割任务中获取更高的分割精度。

表 2 不同方法在 LITS 测试集上的分割性能对比

Table 2 Performance comparison of different methods on LITS test set

Method	Dice / %	VOE / %	RVD / %	ASD / mm	MSSD / mm
U-Net <sup>[14]</sup>	71.8	48.5	-25.3	2.42	9.13
BS-Unet <sup>[15]</sup>	75.1	44.0	-23.4	1.73	8.54
A-Unet <sup>[17]</sup>	76.2	40.6	-14.4	1.42	7.92
U-Net++ <sup>[18]</sup>	80.8	37.2	-10.7	1.26	7.12
U-Net3+ <sup>[20]</sup>	81.2	36.9	-11.2	1.31	7.41
MDCC-Unet <sup>[26]</sup>	83.9	34.1	-6.6	1.22	7.53
proposed method	85.2	32.2	-3.2	0.98	6.61

此外,所提方法还与近几年来一些在 LITS 数据集上进行实验且具有代表性的肝脏肿瘤方法进行了定量的比较,定量实验的对比结果如表 3 所示,表 3 中所有实验数据均来自于对应的原始文献,其中“—”表示相应文献中没有给出对应的值。从表 3 可以看到,所提方法的各项评价指标也是优于其他对比方法的。可以说明,所提方法在近年来相同领域的研究成果中是先进的。



表 3 在 LITS 数据集上与其他方法的比较  
Table 3 Performance comparison with other methods  
on LITS dataset

Method	Dice / %	VOE / %	RVD / %	ASD / mm	MSSD / mm
RA-UNet <sup>[27]</sup>	59.5	38.9	-15.2	1.289	6.775
Method in Ref. [28]	67.0	45.0	4.0	6.660	57.930
Method in Ref. [22]	67.6	34.1	-6.4	—	—
Method in Ref. [29]	73.69	37.80	-15.78	—	—
MA-Net <sup>[30]</sup>	74.9	21.0	-18.0	—	—
CUResNets <sup>[31]</sup>	75.2	37.9	-15.9	—	—
Method in Ref. [32]	83.32	11.62	-15.98	—	—
Proposed method	85.2	32.2	-3.2	0.98	6.61

## 5 结 论

针对肝脏肿瘤分割中存在的小型肿瘤检出率低、分割边界模糊、分割精度不高的问题,提出了一种融合注意力机制与残差可形变卷积的肝肿瘤分割方法。以 U-Net 网络作为主干,用残差连接代替 U-Net 模块间的连接方式,避免了梯度弥散与消失的现象;引入 RBE 模块,通过在卷积层末端添加一条带有反卷积与激活函数的残差路径,将其与上层跳跃连接相连,弥补池化与反卷积操作造成的边缘信息与小目标信息丢失的缺陷;在跳跃连接层中添加一定数量卷积层,增强其中信息的深层特征表达,弥补简单连接造成的语义空白;引入结合空间与通道注意力机制的双注意力模块,提高对肿瘤信息的关注度,抑制其他无关信息;利用可形变卷积,更好地提取不规则的肿瘤边界特征。实验结果表明,所提方法的分割精度指标 Dice 系数达 85.2%。相比于 U-Net,所提方法显著提高了小型肿瘤检出率和对肿瘤边界的分割精度,各项评价指标有较大提升,同时优于其他对比网络,能够基本满足临床需求,起到帮助医生进行肝脏疾病诊断等医疗辅助的作用。

## 参 考 文 献

- [1] Parikh N D, Pillai A. Recent advances in hepatocellular carcinoma treatment[J]. *Clinical Gastroenterology and Hepatology*, 2021, 19(10): 2020-2024.
- [2] Liu L P, Wang L, Xu D, et al. CT image segmentation method of liver tumor based on artificial intelligence enabled medical imaging[J]. *Mathematical Problems in Engineering*, 2021, 2021: 9919507.
- [3] Baāzaoui A, Barhoumi W, Ahmed A, et al. Semi-automated segmentation of single and multiple tumors in liver CT images using entropy-based fuzzy region growing

[J]. *Innovation and Research in BioMedical Engineering*, 2017, 2(38): 98-108.

- [4] Xia Y Q, Qiao S H, Ye Q Q. Automatic liver segmentation method based on deep learning and region growing algorithm[J]. *International Journal of Performability Engineering*, 2020, 16(12): 1900-1909.
- [5] Zhang Y, Wu J, Jiang B X, et al. Deep learning and unsupervised fuzzy C-means based level-set segmentation for liver tumor[C]//2020 IEEE 17th International Symposium on Biomedical Imaging, April 4-7, 2020, Iowa City, IA, USA. New York: IEEE Press, 2020: 1193-1196.
- [6] Yang Z, Zhao Y Q, Liao M, et al. Semi-automatic liver tumor segmentation with adaptive region growing and graph cuts[J]. *Biomedical Signal Processing and Control*, 2021, 68: 102670.
- [7] Amina T, Lakhdar L, Hakim B, et al. Improved active contour model through automatic initialisation: liver segmentation[C]//2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering, May 25-27, 2021, Tripoli, Libya. New York: IEEE Press, 2021: 771-775.
- [8] Shao X X, Lin X M, Shang T T. Liver CT image segmentation algorithm research based on CV model[C]//2017 3rd IEEE International Conference on Computer and Communications, December 13-16, 2017, Chengdu, China. New York: IEEE Press, 2017: 1889-1892.
- [9] Priya V, Biju V G. SVM based liver tumor classification from computerized tomography images[J]. *International Journal of Advanced Engineering and Nano Technology*, 2015, 2(6): 31-36.
- [10] Devi R M, Seenivasagam V. Automatic segmentation and classification of liver tumor from CT image using feature difference and SVM based classifier-soft computing technique[J]. *Soft Computing*, 2020, 24(24): 18591-18598.
- [11] Ranjbarzadeh R, Saadi S B. Automated liver and tumor segmentation based on concave and convex points using fuzzy c-means and mean shift clustering[J]. *Measurement*, 2020, 150: 107086.
- [12] Rela M, Nagaraja S, Ramana P. Liver tumor segmentation using superpixel based fast fuzzy C means clustering[J]. *International Journal of Advanced Computer Science and Applications*, 2020, 11(11): 380-387.
- [13] 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.
- [14] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [15] Milletari F, Navab N, Ahmadi S A. V-net: fully

- convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV), October 25-28, 2016, Stanford, CA, USA. New York: IEEE Press, 2016: 565-571.
- [16] Li S, Tso G K F, He K J. Bottleneck feature supervised U-Net for pixel-wise liver and tumor segmentation[J]. Expert Systems with Applications, 2020, 145: 113131.
- [17] Schlemper J, Oktay O, Schaap M, et al. Attention gated networks: learning to leverage salient regions in medical images[J]. Medical Image Analysis, 2019, 53: 197-207.
- [18] Zhou Z W, Siddiquee M M R, Tajbakhsh N, et al. UNet++: a nested U-Net architecture for medical image segmentation[M]//Stoyanov D, Taylor Z, Carneiro G, et al. Deep learning in medical image analysis and multimodal learning for clinical decision support. Lecture notes in computer science. Cham: Springer, 2018, 11045: 3-11.
- [19] Li C, Tan Y S, Chen W, et al. ANU-Net: attention-based nested U-Net to exploit full resolution features for medical image segmentation[J]. Computers & Graphics, 2020, 90: 11-20.
- [20] Huang H M, Lin L F, Tong R F, et al. UNet 3+: a full-scale connected UNet for medical image segmentation [C]//2020 IEEE International Conference on Acoustics, Speech and Signal Processing, May 4-8, 2020, Barcelona, Spain. New York: IEEE Press, 2020: 1055-1059.
- [21] 高飞, 闫斌, 陈健, 等. 基于堆叠树形聚合结构空洞卷积的肝脏肿瘤分割[J]. 光学学报, 2021, 41(18): 1810002.  
Gao F, Yan B, Chen J, et al. Liver tumor segmentation based on dilated convolution of stacked tree aggregation structure[J]. Acta Optica Sinica, 2021, 41(18): 1810002.
- [22] 刘一鸣, 肖志勇. 基于特征融合的肝脏肿瘤自动分割方法[J]. 激光与光电子学进展, 2021, 58(14): 1417001.  
Liu Y M, Xiao Z Y. Automatic segmentation algorithm of liver tumor based on feature fusion[J]. Laser & Optoelectronics Progress, 2021, 58(14): 1417001.
- [23] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 764-773.
- [24] Bilic P, Christ P F, Vorontsov E, et al. The liver tumor segmentation benchmark (LiTS)[EB/OL]. (2019-01-03)[2021-10-23]. <https://arxiv.org/abs/1901.04056>.
- [25] Selver M A, Kocaoğlu A, Demir G K, et al. Patient oriented and robust automatic liver segmentation for pre-evaluation of liver transplantation[J]. Computers in Biology and Medicine, 2008, 38(7): 765-784.
- [26] Xu C C, Hu D, Zhang Y D, et al. Study on the segmentation method of multi-phase CT liver tumor based on dual channel U-nets[J]. Journal of Physics: Conference Series, 2021, 1828(1): 012043.
- [27] Jin Q G, Meng Z P, Sun C M, et al. RA-UNet: a hybrid deep attention-aware network to extract liver and tumor in CT scans[J]. Frontiers in Bioengineering and Biotechnology, 2020, 8: 605132.
- [28] Han X. Automatic liver lesion segmentation using a deep convolutional neural network method[EB/OL]. (2017-04-24)[2021-10-23]. <https://arxiv.org/abs/1704.07239>.
- [29] Tran S T, Cheng C H, Liu D G. A multiple layer U-net, U<sup>n</sup>-net, for liver and liver tumor segmentation in CT [J]. IEEE Access, 2021, 9: 3752-3764.
- [30] Fan T L, Wang G L, Li Y, et al. MA-net: a multi-scale attention network for liver and tumor segmentation[J]. IEEE Access, 2020, 8: 179656-179665.
- [31] Xi X F, Wang L, Sheng V S, et al. Cascade U-ResNets for simultaneous liver and lesion segmentation[J]. IEEE Access, 2020, 8: 68944-68952.
- [32] 黄泳嘉, 史再峰, 王仲琦, 等. 基于混合损失函数的改进型 U-Net 肝部医学影像分割方法[J]. 激光与光电子学进展, 2020, 57(22): 221003.  
Huang Y J, Shi Z F, Wang Z Q, et al. Improved U-Net based on mixed loss function for liver medical image segmentation[J]. Laser & Optoelectronics Progress, 2020, 57(22): 221003.