

结合多分支纹理特征提取和注意力机制的肝脏肿瘤自动分割方法

邱云飞, 王月洋

辽宁工程技术大学软件学院, 辽宁 葫芦岛 125105

基金项目: 国家自然科学基金(62173171, 61404069); 辽宁省自然科学基金(2015020095); 辽宁省教育厅科学技术研究项目(LJYL051)

通信作者: 王月洋, 1063839622@qq.com 收稿/录用/修回: 2023-06-12/2023-09-21/2023-10-12

摘要

针对计算机断层扫描图像中肝脏肿瘤边界模糊、类型多样、与周围组织对比度低等特点, 以及现有网络对医学图像中的纹理信息利用不充分等问题, 提出了一种结合多分支纹理特征提取和注意力机制的肝脏肿瘤自动分割方法。首先, 设计了一个并行卷积编码器, 替换基准网络 U-Net 中的双卷积模块, 用于提取两种不同感受野下的表层特征; 接着, 提出了一个纹理特征提取网络, 将其搭建于 U-Net 的跳跃连接部分, 以提取多尺度特征下肝脏肿瘤的深层纹理信息; 最后, 在解码阶段引入了一个带有残差路径的通道注意力模块, 旨在有效捕获通道间的依赖关系, 增强肝脏肿瘤分割任务相关特征。将所提方法在 LiTS2017 和 3DIRCADb-01 肝脏肿瘤分割数据集上进行了实验论证, 结果表明, 所提方法在评价指标和可视化结果上均优于对比方法, 对于小尺寸和边界模糊的肿瘤分割具有优势, 有望为肝脏肿瘤筛检提供新的参考。

关键词

图像处理

肝脏肿瘤

U-Net

并行卷积

纹理特征

注意力机制

中图法分类号: TP391.4

文献标志码: A

Automatic Liver Tumor Segmentation Method Integrating Multi-branch Texture Feature Extraction and Attention Mechanism

QIU Yunfei, WANG Yueyang

School of Software, Liaoning Technical University, Huludao 125105, China

Abstract

To address the issues of fuzzy boundaries, diverse tumor types, low contrast with surrounding tissues in computerized tomography (CT) images, as well as insufficient utilization of texture information in existing networks for medical images, an automatic liver tumor segmentation method that combines multi-branch texture feature extraction and attention mechanism is proposed. Firstly, a parallel convolutional encoder is designed to replace the dual convolutional modules in the U-Net baseline network, aiming to extract superficial features under two different receptive fields. Secondly, a texture feature extraction network is introduced in the skip-connection part of the U-Net to extract deep texture information of liver tumors at multiple scales. Finally, a channel attention module with a residual path is incorporated in the decoding stage to effectively capture inter-channel dependencies and enhance the relevant features for liver tumor segmentation tasks. The proposed method is evaluated on the LiTS2017 and 3DIRCADb-01 liver tumor segmentation datasets. Experimental results demonstrate superiority of the proposed method in terms of evaluation metrics

Keywords

image processing;

liver tumor;

U-Net;

parallel convolution;

texture feature;

attention mechanism

and visualizations in comparison with the baseline methods, which shows advantages in segmenting small-sized and blurry boundary tumors, providing promising insights for liver tumor screening.

0 引言

随着计算机技术的不断发展与完善,计算机辅助诊断(computer-aided diagnosis, CAD)技术在医学影像领域中得到了广泛的应用^[1]。其中,肝脏肿瘤自动分割技术作为CAD核心问题之一,对于肝癌的临床诊断和治疗具有重要意义。然而,由于计算机断层扫描(computed tomography, CT)图像本身的低对比度、噪声和伪影等因素,加之人体内部器官和组织之间相似性的存在,给肝脏肿瘤自动分割的实现带来了挑战^[2]。

在过去的几十年里,众多学者致力于肝脏肿瘤自动分割技术的研究,提出了基于图像处理和机器学习的方法,如阈值分割^[3-4]、边缘检测^[5]、区域生长^[6]、基于水平集的方法^[7-8]、基于神经网络的方法^[9]等。然而,这些方法高度依赖于先验知识,难以扩展应用。随着深度学习的迅速发展,越来越多的研究者开始使用卷积神经网络(convolutional neural network, CNN)对肝脏肿瘤进行分割,并根据输入图像维度的不同将这些网络细分为2D、2.5D和3D三类。2D网络因其在设备性能需求低、实现成本代价小和计算速度快等方面的优势,在肝脏肿瘤自动分割任务中具有广泛的应用前景,并成为了当前的研究热点。

其中,U-Net^[10]是一种广泛应用于医学图像分割的CNN架构,其特殊的U形结构可以有效地提取医学图像的特征信息,在肝脏肿瘤分割领域取得了许多成功的应用。然而,它只考虑了同层级的连接,在上下层信息之间存在语义鸿沟,无法完成细小肿瘤的精准分割。为了克服这一问题,UNet++^[11]在跳跃连接部分引入了更多的跨层连接,将不同层级的特征以叠加的方式整合,在增强特征传递和重用的同时减轻了梯度消失和过拟合问题。但是这种做法也大幅度增加了模型的参数量。ResUNet++^[12]通过引入残差块、注意力机制和空洞空间金字塔池化模块来融合多尺度特征和增强肿瘤特征的表达能力,并通过条件随机场和测试时数据增强等后处理技术来精细化分割结果,提高了分割精度。但是该方法对于肿瘤的边界分割不够精确,且后处理技术会导致计算方面的额外负担,不适用于大规模数据的处理。TransUNet^[13]使用Transformer^[14]作为编码

器,结合自注意力机制和跨层连接等技术,增强了特征能力的表达和上下文信息的获取,提高了模型的建模能力和泛化性能,对于处理小尺寸和边缘模糊的肿瘤具有优势。Swin-Unet^[15]使用带有偏移窗口的分层Swin-Transformer^[16]作为编码器来提取上下文特征,并设计了具有补丁扩展层的基于对称Swin-Transformer的解码器来执行上采样操作,从而恢复特征图的空间分辨率,解决了传统U-Net中注意力机制和长距离信息传递问题。但这两种方法对训练数据的数量和质量要求较高,且Transformer结构本身的可解释性较差,因此难以在资源受限的环境中应用。

鉴于以上方法的缺陷与不足,提出了一种结合多分支纹理特征提取和注意力机制的U型网络(U-Net integrating Multi-Branch Texture Feature Extraction and Attention Mechanism, MBTFEAM-UNet)。该方法采用多种技术手段,能够从不同角度提取肝脏肿瘤的特征,充分利用CT图像中的上下文信息和纹理信息,提高分割的准确率和鲁棒性。其中,并行卷积模块能够提取不同感受野下的肿瘤表层特征;纹理特征提取网络旨在从不同尺度的特征图中提取图像中深层特征,充分利用医学图像的局部和全局纹理信息,增强分割的准确性;而注意力模块可以充分利用上下文信息,提升关键信息的权重,减少冗余信息的传递,进一步提高分割性能。

1 相关工作

1.1 纹理特征提取

纹理特征提取网络通常使用多个卷积层和池化层来逐渐降低图像分辨率,并在每个层次上提取不同的特征。在某些情况下,还可以使用循环神经网络(RNN)或注意力机制来捕捉纹理的时序信息或空间分布,被广泛应用于医学图像分析领域。对于肝脏肿瘤分割任务而言,纹理特征提取网络的设计对于分割结果的准确性和鲁棒性具有重要影响。例如,KUO等^[17]使用基于U-Net和ResNet的深度学习模型来进行肝脏肿瘤纹理特征的提取,并将这些特征输入到支持向量机(SVM)模型中进行治疗反应预测。实验结果表明,该方法可以有效地预测肝癌患者的治疗反应,在评价指标和灵敏度方面取得了优秀的表现。另外,DIAO等^[18]利用预先学习的

纹理信息生成伪标签, 提出了一种基于纹理信息的自动伪标签模块, 从而帮助深度学习模型学习更加准确的特征, 提高分割的精度, 该方法在多个数据集上均取得了优异的分割性能。这些研究表明, 纹理特征提取网络作为一种有效的特征提取模型, 能够在医学图像分析领域中发挥重要作用, 并为进一步提高肝脏肿瘤分割的准确性和效率提供了新的思路和方法。

1.2 注意力机制

注意力机制通过在训练过程中对特征权重进行调整完成模型对长距离特征依赖的建模, 达到提高分割精度的目的。常用的注意力机制分为通道注意力和空间注意力两类。其中空间注意力通过计算特征之间的相似性来联系上下文信息, 找到图像中最重要的部分进行处理, 解决局部感受野的问题。通道注意力通过全局池化操作对特征进行压缩, 捕获通道间的依赖关系, 有选择地增强或抑制不同通道。如今, 注意力作为一种即插即用的模块被应用

于深度学习的诸多领域。FU 等^[19]将空间注意力和通道注意力相结合, 提出了一种双注意力机制网络, 可以在场景分割任务中获得较好的性能。LIU 等^[20]提出了一种基于注意力机制的网络剪枝方法, 可以减少模型的参数数量并提高模型的精度。在肝脏肿瘤分割中, 注意力机制可以帮助网络更加集中地关注与肝脏肿瘤分割相关的区域, 从而提高分割的准确性和鲁棒性。

2 方法

如图 1 所示, 提出的肝脏肿瘤自动分割方法分为数据预处理和肿瘤分割网络两个部分。首先, 对原始腹部 CT(计算机断层扫描)图像进行数据预处理, 以增强肝脏组织与周围器官的对比度并获取肝脏感兴趣区域(Region of Interest, ROI)。然后, 将处理后的图像输入到提出的 MBTFEAM-UNet 网络中训练, 输出肿瘤与背景的概率图, 从而完成肝脏肿瘤的分割。

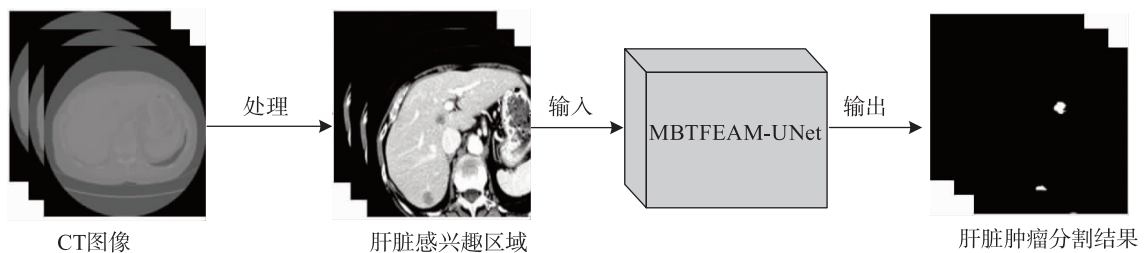


图 1 肝脏肿瘤分割流程图

Fig.1 Flow chart of liver tumor segmentation

2.1 数据预处理

为了增强肝脏及肿瘤与周围组织器官的对比度, 减少无关组织对分割结果的干扰, 对原始腹腔

CT 图像的横断面切片进行了调窗处理, 并通过固定裁剪的方式获取肝脏 ROI, 预处理流程如图 2 所示。

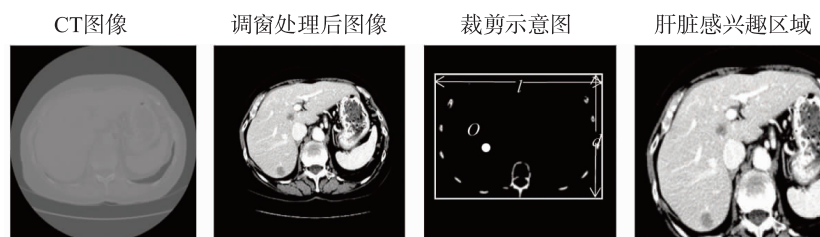


图 2 数据预处理

Fig.2 Data pre-processing

首先, 根据肿瘤掩码, 按照式(1)、式(2)计算窗宽窗位, 对原始 CT 图像进行调窗处理, 增强病灶与周围组织的对比度。

$$C(t) = \frac{M_t - m_t}{2} \tag{1}$$

$$W(t) = M_t - m_t \tag{2}$$

式中： M_i 和 m_i 分别表示肿瘤掩码的最大值和最小值， $C(t)$ 和 $W(t)$ 分别表示计算得到的窗位和窗宽大小。然后，采用阈值和形态学方法^[21]完成对 CT 图像中肋骨和脊椎的分割，并将分割结果进行行列投影，取第一个和最后一个非零像素值所在的行和列构建长为 l 宽为 d 的矩形框。最后，鉴于肝脏通常位于 CT 图像横断面切片左侧，以 $(\frac{l}{4}, \frac{d}{2})$ 为中心点 O ，取像素大小为 320×320 的方形框对调窗后的图像进行裁剪，得到肝脏 ROI。

2.2 MBTFEAM-UNet 网络

为了解决肝脏肿瘤分割任务中肿瘤边界分割不精确、细小肿瘤难以分割等问题，同时克服现有网络在提取医学图像纹理信息方面不充分的限制，以 U-Net 为基准网络，结合同步卷积编码器、纹理特征提取网络和注意力机制，提出了一个用于肝脏肿瘤分割任务的新网络——MBTFEAM-UNet。

2.2.1 整体结构

MBTFEAM-UNet 整体结构如图 3 所示。网络的具体组成如下：左半部分由 4 个并行卷积编码器 (PCE) 组成，每个编码器均执行两次不同大小卷积核的并行卷积、批归一化和 ELU 函数激活操作，用于提取不同感受野下的表层特征，同时使用 ELU

激活函数对抗噪声干扰，加速模型收敛。在每次编码完成后对得到的特征进行一次 2×2 最大池化操作，将图像尺寸缩减为输入的 $1/2$ ，增大特征提取的感受野并降低计算复杂度。将经过最后一次池化的特征图进行两次相同的并行卷积模块编码，使得模型更好地提取全局特征，同时为上采样奠定基础。网络的跳跃连接部分搭建了一个纹理特征提取网络 (MBTFE-Net)，旨在从 CT 图像中提取到多尺度的深层纹理信息。此外，采用一个残差路径使模型能够关注到训练过程中产生的微小变化，以增强肿瘤的边缘信息并减少冗余信息的传递。网络的右半部分包含 4 个解码器，均采用 2×2 双线性插值完成对输入特征的上采样，逐步将特征图恢复为原尺寸大小。解码过程中将上采样得到的深层特征与经过跳跃连接提取到的纹理特征进行拼接后输入到一个带有残差路径的通道注意力模块 (RCA)，完成对通道间依赖的建模，增强关键通道的灵敏度，使得分割结果更集中于图像中的病变区域，减少背景特征对分割结果带来的影响。将最终得到的特征图通过一个 1×1 卷积输出预测结果，并采用一个混合的损失函数衡量预测结果与真实标签之间的差异，从而指导模型参数的优化，解决医学图像中普遍存在的类不平衡问题。

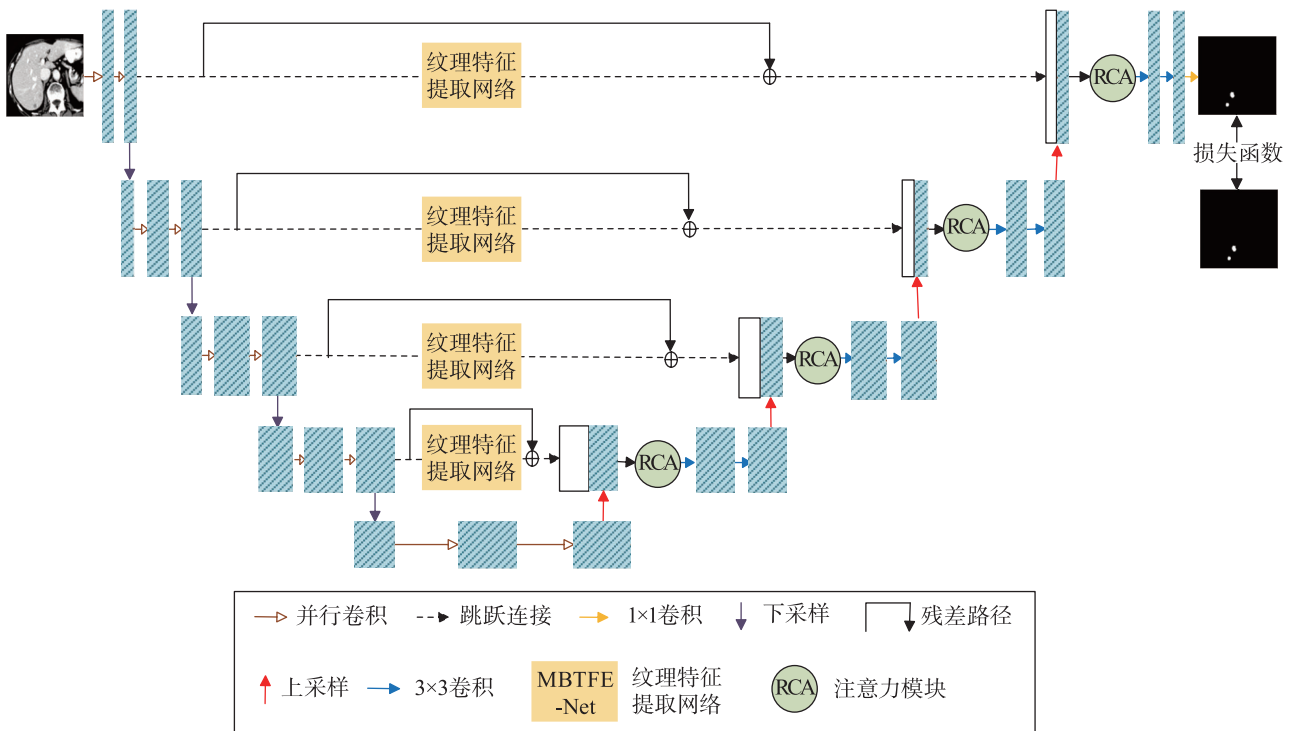


图 3 MBTFEAM-UNet 结构图
Fig.3 MBTFEAM-UNet structure diagram

2.2.2 并行卷积编码器

U-Net 中特征提取过程均由两次 3×3 卷积、批归一化和 ReLU 激活函数实现, 提取到的特征较单一, 且当学习率较大时, ReLU 激活函数易造成神经元的坏死, 导致网络学习能力受损。针对上述缺点, 设计了一个并行卷积编码器, 其结构如图 4 所示。该编码器对输入特征分别以大小为 3×3 和 5×5 的卷积核进行表层特征的提取, 随后以加操作完成特征融合, 以此实现不同感受野下表层特征的提取。除此之外, 将 ELU 作为该网络的激活函数, 相较于 ReLU 函数简单的将负值区归为 0, 该函数具有左侧软饱和性, 能够防止训练过程中由于学习

率太大导致的神经元坏死问题。

2.2.3 纹理特征提取网络

U-Net 网络中简单的跳跃连接对于全局多尺度信息进行建模具有一定的局限性。此外, 由于经过编码器得到的特征仅能表示图像中肿瘤的表层信息, 通过跳跃连接直接传递至解码器与上采样得到的高级特征拼接时因存在语义鸿沟, 易造成特征映射模糊, 导致肿瘤边缘分割模糊、小肿瘤检测效果不佳。针对该缺点, 提出了一个纹理特征提取网络, 搭建于 U-Net 网络的跳跃连接部分, 其结构如图 5 所示。该提取器由 3 个相同的多分支纹理特征提取模块 (MBTFE) 构成, 每个模块的处理流程分

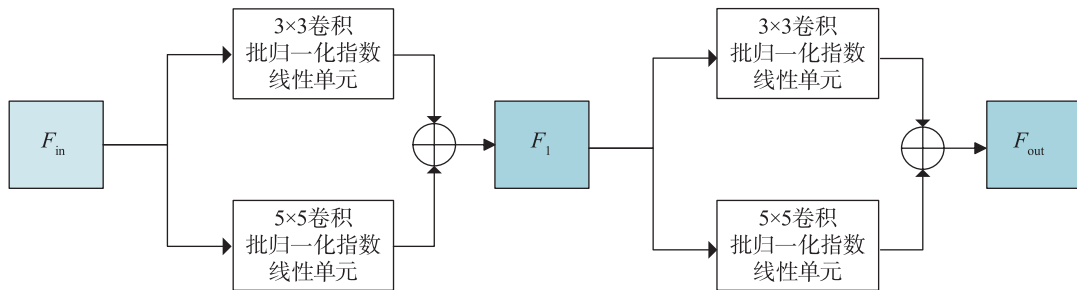


图 4 并行卷积编码器

Fig.4 Parallel convolutional encoder

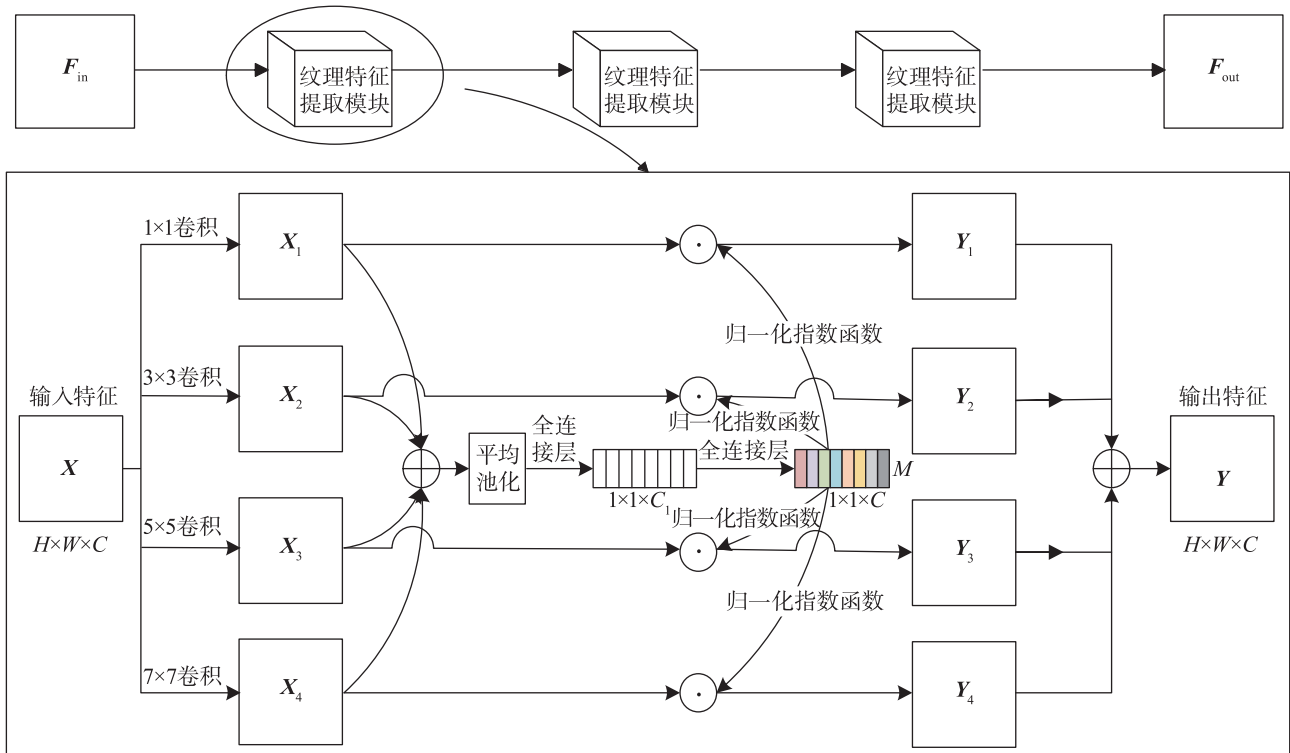


图 5 纹理特征提取网络结构图

Fig.5 Structure diagram of the texture feature extraction network

为 3 个步骤：多尺度特征提取、相关性计算和特征选取。首先，对经过编码器生成的特征图 F_{in} 使用 4 种不同尺寸 (1×1 、 3×3 、 5×5 、 7×7) 的卷积核进行特征提取，并将这些提取到的特征图相加，从而完成不同感受野下多尺度特征的融合，得到与原尺寸大小相同的特征图 \tilde{X} 。这个过程的计算公式为

$$\tilde{X} = \sum_{i=1}^4 X_i \quad (3)$$

式中： $\tilde{X}, X_i \in \mathbb{R}^{H \times W \times C}$ 。接着，采用平均池化操作从空间维度对拼接的特征图 \tilde{X} 进行压缩，获取通道间的分布关系，然后依次输入两个全连接层 (FC) 完成特征的降维和升维，生成权重特征图 M ，对其按通道方向使用 SoftMax 函数计算通道相关性，与经过不同卷积核提取到的多尺度特征分别相乘后加和到一起，生成具有通道相关性的特征图 Y 。此过程可用公式表示为

$$M = \omega_2 \left(\omega_1 \left(\frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \tilde{X}(i, j) \right) \right) \quad (4)$$

$$Y_i = X_i \odot F^{\text{SoftMax}}(M) \quad (5)$$

$$Y = \sum_{i=1}^4 Y_i \quad (6)$$

$$C_1 = \max \left(\frac{C}{r}, L \right) \quad (7)$$

式中： $Y, Y_i, \tilde{X} \in \mathbb{R}^{H \times W \times C}$ ， $M \in \mathbb{R}^{1 \times 1 \times C}$ ， $\omega_1 \in \mathbb{R}^{C_1 \times C}$ ， $\omega_2 \in \mathbb{R}^{C \times C_1}$ ， H 和 W 分别为特征图的长和宽， C 为通道数， L 和 r 分别设置为 32 和 16， \odot 为哈达玛积。

2.2.4 注意力模块

编解码结构虽然实现了特征的自动提取，但对于提取到的各类的信息都赋予了同样大小的关注度，未能实现对重要信息的聚焦和对无关信息的忽略，导致分割精度下降。对此，所提方法在原始 U-Net 网络的解码部分引入了一个带有残差路径的通道注意力模块，该模块通过计算通道间的依赖关系重新调整各通道的权重，从而实现关键信息的增强。此外，残差路径的设计不仅能够使模型在训练过程中能够关注到微小的变化，同时还能够避免梯度消失和梯度爆炸情况的发生。带有残差路径的通道注意力模块如图 6 所示。

该模块主要包括压缩、激励、相乘以及残差四个阶段。首先使用全局平均池化操作对经过上采样和跳跃连接拼接后的特征进行空间维度的压缩，将每个通道上的空间特征编码为一个全局特征。接着，加入一个 FC 层，对得到的全局特征进行降维处理，从而降低计算量。然后，使用激活函数 ReLU 增强模型的非线性能力，加速模型训练。接下来，再使用一个 FC 层将特征恢复为原始维度并使用 Sigmoid 函数激活，从而完成对每个通道重要性的预测。随后，将经过激励阶段得到的权重值与原特征图对应通道的二维矩阵相乘，重新调整输入特征的权重，完成肿瘤信息的增强，并减少背景信息对分割结果的影响。最后，将原特征图与调整通道权重后的特征图相加，获得最终的通道注意力特征图 F_c 。该模块的计算公式为

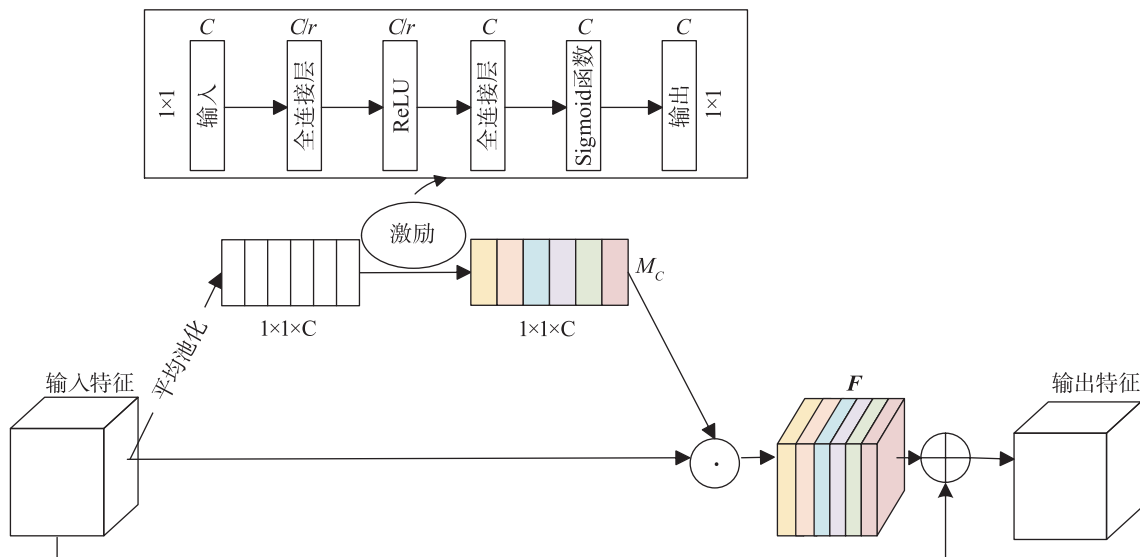


图 6 RCA 模块结构图

Fig.6 RCA module structure diagram

$$\mathbf{M}_C = \rho_2 \left(\sigma_{\text{ReLU}} \left(\rho_1 \left(\sigma_{\text{SigMoid}} \left(\frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W \mathbf{X}(i, j) \right) \right) \right) \right) \quad (8)$$

$$\mathbf{U} = \mathbf{X} \odot \mathbf{M}_C \quad (9)$$

$$\mathbf{F}_C = \mathbf{U} + \mathbf{X} \quad (10)$$

$$C_2 = \frac{C}{r} \quad (11)$$

式中: $\mathbf{M}_C \in \mathbb{R}^{1 \times 1 \times C}$, $\rho_1 \in \mathbb{R}^{C_2 \times C}$, $\rho_2 \in \mathbb{R}^{C \times C_2}$, \mathbf{X} , \mathbf{U} , $\mathbf{F}_C \in \mathbb{R}^{H \times W \times C}$.

2.3 损失函数

实验使用骰子系数作为模型预测结果与专家手动标注(下文称之为金标准)之间重叠度的评价指标,并以此为最终优化目标。然而,在训练过程中单独使用 Dice 损失容易因小目标的错误预测导致梯度发生剧烈变化,不利于网络的训练收敛。鉴于肿瘤分割的实质是二分类任务,所以引入了二分类交叉熵损失来稳定训练过程。Dice 损失 L_{Dice} 、二分类交叉熵损失 L_{BCE} 以及最终采用的加权损失 L_{total} 如下:

$$L_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N P_i T_i}{\sum_{i=1}^N P_i + \sum_{i=1}^N T_i} \quad (12)$$

$$L_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N (T_i \ln(P_i) + (1 - T_i) \ln(1 - P_i)) \quad (13)$$

$$L_{\text{total}} = \lambda (L_{\text{Dice}} + L_{\text{BCE}}) \quad (14)$$

式中: N 为图像中的像素总数, T_i 为第 i 个像素的专家分割结果(背景和肿瘤分别以 0, 1 标注), P_i 为第 i 个像素被预测为肿瘤的概率, λ 为权重因子。

3 实验结果与分析

3.1 数据集

为验证所设计网络的有效性,选取两个用于肝

脏肿瘤分割的 CT 图像公开数据集(LiTS2017 和 3DIRCADb-01)进行实验。LiTS2017 数据集包含了来自于 6 个不同医院 201 个病例的 CT 扫描图像,其中的 131 个病例由 4 位专业影像医师分别进行了肝脏和肿瘤的标注,剩余 70 个未公开标注。每个 CT 图像的切片大小为 512×512 , 层厚为 0.45 mm ~ 6.0 mm, 切片数量在 42 到 1 026 之间。3DIRCADb-01 数据集包含来自不同医院的 20 个病例的 CT 图像,其中的 15 个病例图像中包含由专业影像医师标注了大小不等的 1 到 7 个肝肿瘤。每张图像的切片大小为 512×512 , 厚度为 1.0 mm ~ 4.0 mm, 数量在 1 到 46 之间。为了提高图像对比度、减少无关组织干扰,分别对两个数据集按照 2.1 节所述方法进行预处理。预处理完成后,从 LiTS2017 和 3DIRCADb-01 数据集中分别选取了 18 399 张和 1 666 张包含肝脏区域在内的 CT 图像,为了保证训练样本的丰富性和数量,以 8:1:1 的比例随机划分为训练集、验证集和测试集。

3.2 数据扩充

由于经过预处理后的数据样本量较少,容易导致模型在训练过程中难以充分提取肿瘤特征,出现过拟合问题。为了避免这一问题的发生,对预处理后用作训练集的样本以在线增强的方式进行了数据扩充。具体而言,实验采用了随机裁剪(单类区域最大占比 0.75)、随机翻转(概率为 0.5)、边缘填充(邻近像素填充)和光度失真四种方法来对即将输入网络的训练样本及其标注进行处理,并通过设定足够的迭代次数间接实现训练样本量的增加。图 7 展示了原图和标注在经过数据增强后的图像。

3.3 评价指标

实验采用骰子系数(dice similarity coefficient,

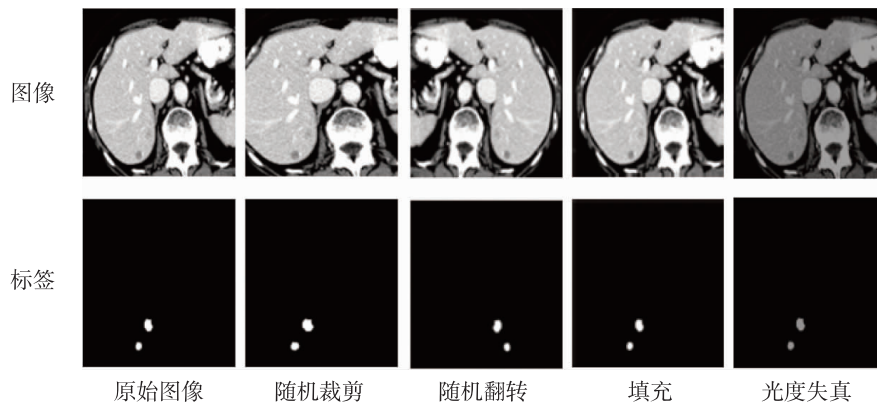


图 7 数据增强示意图

Fig.7 Schematic diagram of data augmentation

DSC)、体素重叠误差(volumetric overlap error, VOE)、体素相对误差(relative volume difference, RVD)、对称位置的平均表面距离(average symmetric surface distance, ASSD)、豪斯多夫距离(Hausdorff distance, HD)和均交并比(mean intersection over union, MIoU) 6项评价指标对模型进行性能评估^[22]。

P_{DSC} 用于度量分割结果 R 和金标准 T 之间的相似度,计算公式如下:

$$P_{DSC} = \frac{2|R \cap T|}{|R| + |T|} \quad (15)$$

式中: $|\cdot|$ 为集合中的元素个数。

P_{VOE} 用于计算分割结果和金标准之间的重叠误差程度,计算公式如下:

$$P_{VOE} = 1 - \frac{|R \cap T|}{|R \cup T|} \quad (16)$$

P_{RVD} 表示分割结果和金标准之间的差异程度,计算公式如下:

$$P_{RVD} = \frac{|R| - |T|}{|T|} \quad (17)$$

P_{ASSD} 用于计算分割结果和金标准之间对称位置的平均表面距离,计算公式如下:

$$P_{ASSD} = \frac{1}{S(R) + S(T)} \left(\sum_{s_T \in S(T)} d(s_T, S(R)) + \sum_{s_R \in S(R)} d(s_R, S(T)) \right) \quad (18)$$

式中: s 表示集合中的单个体素, $S(\cdot)$ 表示表面体素集合, $d(s, S(\cdot))$ 表示某一体素到另一体素集合中的最短距离。

P_{HD} 用来描述分割结果和金标准之间对称位置的最大表面距离,计算公式为

$$P_{HD} = \max \left\{ \max_{s_T \in S(T)} (d(s_T, S(R))), \max_{s_R \in S(R)} (d(s_R, S(T))) \right\} \quad (19)$$

P_{MIoU} 用于计算所有类别交集和并集之比的平均

值,计算公式如下:

$$P_{MIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{|R \cap T|}{|R \cup T|} \quad (20)$$

式中: k 为分割类别数目, $k+1$ 为包含背景类别总数。

3.4 实验环境与参数设置

为确保实验结果的公正性,所有实验都在配备了 Intel[®] Core[™] i7-8700 CPU@3.20 GHz 的处理器以及 6 张显卡型号为 RTX A6000(48 GB) 的硬件设备上进行。网络结构以基于 PyTorch 实现的语义分割框架 MMSegmentation 构建,并采用 Adam 优化器进行模型训练、初始学习率设置为 1×10^{-5} , 衰减率为 0.9, 输入批量大小设置为 8。模型训练迭代 30 000 次后停止。

3.5 结果与分析

3.5.1 激活函数对分割性能的影响

ReLU、LeakyReLU 和 ELU (exponential linear unit) 是 CNN 中常用的 3 种激活函数,其取值曲线如图 8 所示。为了探讨不同激活函数对分割精度的影响,分别将其作为 MBTFEAM-UNet 的激活函数进行实验,按照式(15)~式(20)计算得到的评价指标数值如表 1 所示。从中可以发现,不同的激活

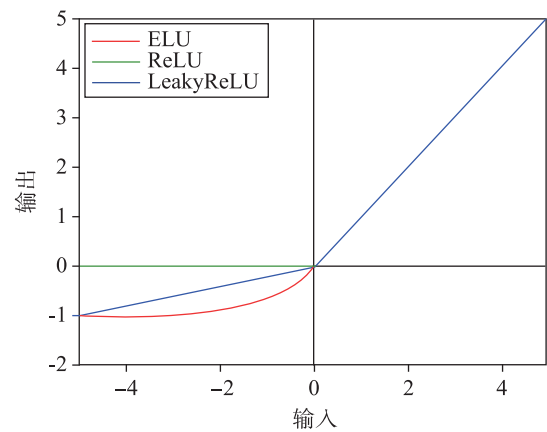


图 8 不同激活函数曲线

Fig.8 Curves of different activation functions

表 1 不同激活函数在 LiTS2017 数据集上的分割性能对比结果

Tab.1 Comparison results of segmentation performance of different activation functions on LiTS2017 dataset

激活函数	DSC /%	VOE /%	RVD /%	ASSD /mm	HD /mm	MIoU /%
ReLU	81.37	31.41	-7.75	6.69	10.68	84.19
LeakyReLU	82.97	29.31	-7.35	6.12	9.97	85.22
ELU	83.95	27.66	-6.41	2.54	8.05	86.03

注: 加黑字体为最优结果。

函数对于实验精度有一定的影响。当使用 ReLU 作为激活函数时, 模型的分割精度远不如使用 LeakyReLU 和 ELU 时取得的精度高, 这是因为当输入值为负值时, ReLU 函数使得梯度为 0, 导致神经元坏死, 从而权重无法继续更新。而 LeakyReLU 和 ELU 函数对负值输入有一定的坡度, 在一定程度上减少了静默神经元的出现, 允许其基于梯度的学习, 保证了模型参数的更新。相较于 LeakyReLU, ELU 采用一种平滑的指数函数对输入的负值进行处理, 具有左侧软饱和性, 更能减少噪声对于分割精度的干扰, 在实验中取得了最优结果。

3.5.2 MBTFE 模块数量对分割精度的影响

为了在尽可能地减少模型参数的同时保证分割精度的提升, 对提出的纹理特征提取网络中的 MBTFE 模块数量设计在 LiTS2017 数据集上进行了实验, 得到的分割精度及模型参数量大小如表 2 所示。可以看到, 分割精度的提升幅度在很大程度上取决于 MBTFE 模块数量的多少。当 MBTFE 模块

数量较少时, 网络因深层纹理特征提取不充分导致最终的肿瘤分割效果不佳。随着模块数量的增加, 图像的深度纹理信息得以充分的提取, 得到的评价指标也愈来愈好, 但与此同时模型的参数量也在逐步增大。MBTFE 模块数量从 1 到 3 时, 评价指标变化明显, DSC 值分别有 4% 左右的提升, 此时认为分割效果提升的重要性大于参数量的增加带来的计算开销。但当 MBTFE 模块数量从 3 增加到 4 时, 评价指标 DSC 值仅有 0.34% 的提升, 而参数量增加了 29.42 M, 模型计算成本的增长远大于其性能的提升, 所以最终设计的纹理特征提取网络由 3 个 MBTFE 模块组成。

3.5.3 损失函数权重对分割性能的影响

实验以 0.5 的增长步幅探讨了损失函数权重因子 λ 的取值在 LiTS2017 数据集上对分割性能的影响, 按照式(15)~式(20)计算得到的评价指标数值如表 3 所示。实验结果表明: 当权重因子 λ 取值为 1.5 时, 模型在各项指标上达到了最优值, 此时分割性能最佳。

表 2 不同 MBTFE 模块数量条件下在 LiTS2017 数据集上的分割性能对比结果

Tab.2 Comparison results of segmentation performance under different MBTFE module quantity condition on the LiTS2017 dataset

MBTFE 模块数量	DSC /%	VOE /%	RVD /%	ASSD /mm	HD /mm	MIoU /%	Params /百万
1	75.89	38.48	-10.20	6.15	13.68	81.31	42.81
2	80.05	33.03	-8.48	4.87	11.92	83.97	72.23
3	83.95	27.66	-6.41	2.54	8.05	86.03	101.64
4	84.29	27.06	-6.11	1.95	7.79	86.28	131.06

注: 加粗字体为最优结果。

表 3 不同损失函数权重取值条件下在 LiTS2017 数据集上的分割性能对比结果

Tab.3 Comparison results of segmentation performance under different loss function weight values condition on the LiTS2017 dataset

λ 取值	DSC /%	VOE /%	RVD /%	ASSD /mm	HD /mm	MIoU /%
0.5	83.03	29.07	-7.12	5.92	9.73	85.37
1.0	83.75	28.04	-6.73	3.76	8.28	85.89
1.5	83.95	27.66	-6.41	2.54	8.05	86.03
2.0	83.19	28.85	-7.02	4.15	9.30	85.48
2.5	82.70	29.59	-7.76	6.52	10.07	85.13

注: 加粗字体为最优结果。

3.5.4 中间层特征分析

为了证实所提方法的有效性, 将 U-Net 和 MBTFEAM-UNet 的中间层特征进行可视化分析, 得到的各阶段特征图如图 9 所示。从图 9(a)中可以看到, 在编码路径中, 由于并行卷积编码器使用不同

尺寸的卷积核扩大了原编码器的感受野, 使得编码各阶段在局部纹理特征提取的同时兼顾了全局纹理信息的利用, 从而实现了对不同尺寸肿瘤表层特征的有效提取。而 U-Net 随着网络层次的加深因纹理信息匮乏而造成了小尺寸肿瘤信息的遗失。从图 9

(b)中可以看到, 由于 MBTFE-Net 对图像表层信息进行了更深层次的提取以及 RCA 模块在提升肿瘤所在位置权重的同时关注信息细微变化, 使得所提方法在解码各阶段得到的肿瘤信息更加清晰、具

体, 从而实现了不同大小肿瘤及其边缘的精准分割。而原网络仅能分割出部分大体积肿瘤, 且存在由于对上下文信息利用不充分而在最终的分割结果中误将背景像素判别为肿瘤的情况。

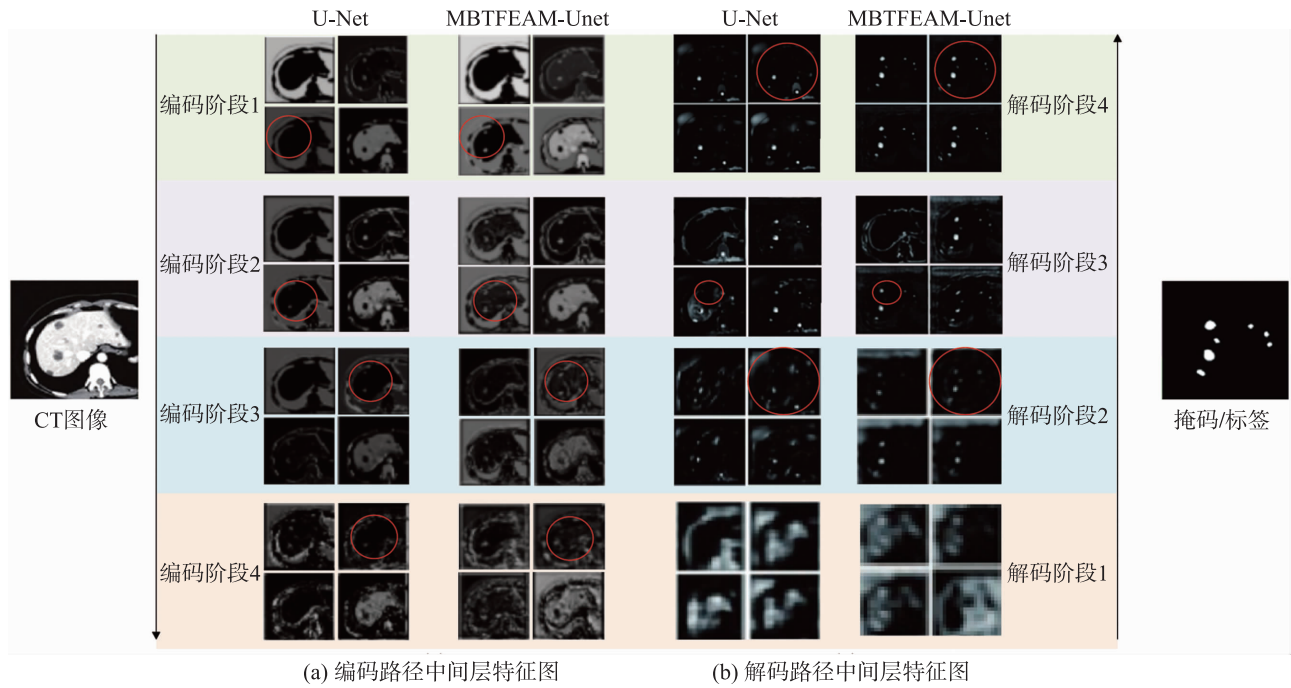


图9 中间层特征图

Fig.9 Middle layer feature map

3.5.5 对比实验

为了全面客观分析所提方法在分割性能上的表现, 选用基准网络 U-Net 及其医学分割领域主流改进方法 U-Net ++、ResUNet ++、TransUnet、Swin-Unet 进行对比实验, 在 LiTS2017 和 3DIRCADb-01 数据集上按照式(15)~式(20)计算得到的评价指标结果如表 4 和表 5 所示。根据表 4 所示的结果, 所提方法在 LiTS2017 数据集中取得的各项评价指

标均优于对比网络。相较于原始网络, 所提方法在 DSC 和 MIoU 指标上分别提升了 14.80% 和 9.83%, VOE 和 RVD 指标分别降低了 19.49% 和 9.66%, 证明预测结果和金标准之间的重叠程度更好, 体素之间的差异更小, 分割性能更佳; 相较于采用密集连接思想的 U-Net ++, 所提方法在各项指标上也有显著的提升, 这是因为 MBTFE-Net 的多分支纹理特征提取模块可以从不同尺度充分提取肿瘤图像中的深

表 4 不同方法在 LiTS2017 数据集上的分割性能对比结果

Tab.4 Comparison of segmentation performance of different methods on the LiTS2017 dataset

方法	DSC /%	VOE /%	RVD /%	ASSD /mm	HD /mm	MIoU /%
U-Net	69.15	47.15	-15.07	11.19	22.73	76.20
U-Net ++	73.35	42.32	-14.88	9.64	19.83	78.48
ResUNet ++	76.45	38.96	-13.14	8.82	18.19	79.81
TransUnet	77.75	35.32	-11.51	6.55	16.80	80.65
Swin-Unet	79.20	33.45	-9.20	5.04	12.43	82.56
MBTFEAM-UNet	83.95	27.66	-5.41	2.54	8.05	86.03

注: 加粗字体为最优结果。

表5 不同方法在 3DIRACADb-01 数据集上的分割性能对比结果

Tab.5 Comparison of segmentation performance among different methods on the 3DIRACADb-01 dataset

方法	DSC/%	VOE/%	RVD/%	ASSD/mm	HD/mm	MIoU/%
U-Net	78.42	35.50	-13.49	14.82	16.32	82.08
U-Net + +	82.80	29.35	-7.54	8.17	10.93	85.20
ResUNet + +	85.44	25.41	-6.37	6.72	8.67	87.18
TransUnet	73.90	41.39	-20.43	16.74	19.15	79.75
Swin-Unet	76.26	32.25	-17.33	14.68	17.16	80.98
MBTFEAM-U-Net	89.66	18.75	-4.71	1.06	5.57	90.54

注: 加粗字体为最优结果。

层纹理信息, 而 RCA 模块在对通道依赖进行建模的同时, 也可以使模型更加注重肿瘤的边缘信息, 从而使得肿瘤边缘更加清晰; 相较于 Transformer 架构方法 TransUnet 和 Swin-Unet, 所提方法在 DSC 和 MIoU 指标上也有较大幅度的提升, 并且取得的 ASSD、HD 指标更低, 进一步证明了所提方法对于肝脏肿瘤分割任务的有效性。

由表 5 可知, 在 3DIRACADb-01 数据集中, 由于训练样本数据量过少, TransUnet 和 Swin-Unet 在各项指标上的表现远不及其他网络, 这表明了 Transformer 架构在应对数据样本较少的分割任务时存在一定的弊端; U-Net 网络的肿瘤分割性能虽然在评价指标 DSC 上取得了 78.42% 的结果, 但其他指标表明该网络的预测结果与金标准之间的差异较明显; ResUNet++ 在各项评价指标上均达到次优

值, 且由于残差块和注意力机制的引入使得对称位置的平均表面距离和豪斯多夫距离进一步缩减, 分割结果和金标准之间的重叠程度较高; 所提方法在 3DIRACADb-01 数据集中的各项评价指标均达到最优值, 相比于次优值, DSC、MIoU 指标提高了 4.22% 和 3.36%, ASSD 和 HD 分别下降了 5.66 mm 和 3.1 mm, 评价指标 DSC 相比于骨干网络 U-Net 提升了 11.24%, 在很大程度上减少了漏检情况的发生。

3.5.6 结果可视化分析

为了主观评估各方法的分割性能, 对 LiTS2017 和 3DIRACADb-01 的测试集图像进行了分割, 以更加直观地展示所提方法在肝脏肿瘤分割任务中的有效性。图 10 为 LiTS2017 数据集中不同方法的分割结果, 可以看到, U-Net、U-Net++、ResUNet++、

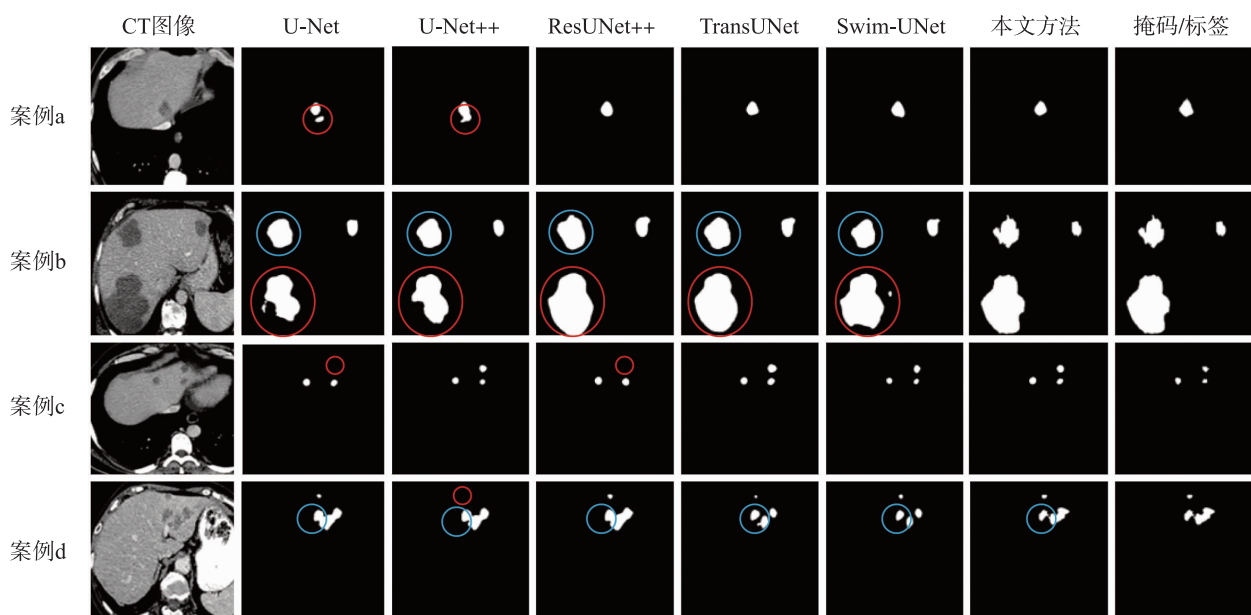


图 10 不同方法在 LiTS2017 数据集上的分割结果

Fig.10 Segmentation results of different methods on the LiTS2017 dataset

TransUNet 以及 Swin-Unet 均能大致定位到肿瘤所在位置。但这些方法在处理边界模糊, 尺寸较小的肿瘤时容易出现误检、漏检以及边界分割不精确等情况。例如: 对于案例 a 中的图像, U-Net 和 U-Net++ 的肿瘤分割边界不精确; 在案例 b 中对比方法均未能实现对大体积肿瘤的准确分割, 并且 Swin-Unet 出现了将背景误判为肿瘤的情况; 在案例 c 中, U-Net 和 ResUNet++ 出现了漏检情况; 在案例 d 中, 只有所提方法能够精确地分割出肿瘤边界, 且 U-Net++ 未能实现对图像上方小尺寸肿瘤的检测。

图 11 为不同方法在 3DIRACADb-01 上的分割结果, 从图中可以看出, 所提方法在边缘模糊以及

细小肿瘤的分割上具有绝对的优势。例如: 在案例 a 中 Swin-Unet 误把背景分割成了肿瘤, TransUNet 和 U-Net 对于肿瘤的边缘分割不够精确; 在案例 b 中, 只有 ResUNet++ 和 MBTFEAM-UNet 精准分割了肿瘤, 其余 4 种方法均存在将背景误判为肿瘤的情况; 在案例 c 的 CT 图像中, 肿瘤大小不一、类型复杂多样, 虽然用 ResUNet++ 方法取得了与金标准相近的结果, 但相比于 MBTFEAM-UNet, 在分割的视觉效果上仍处于劣势; 在案例 d 中, 所有方法都分割出了肿瘤的大致位置与形状, 但 MBTFEAM-UNet 分割出的肿瘤边界与金标准最为相近。综上所述, MBTFEAM-UNet 网络输出的分割图像具有肿瘤分割边界清晰、细节纹理丰富等优势。

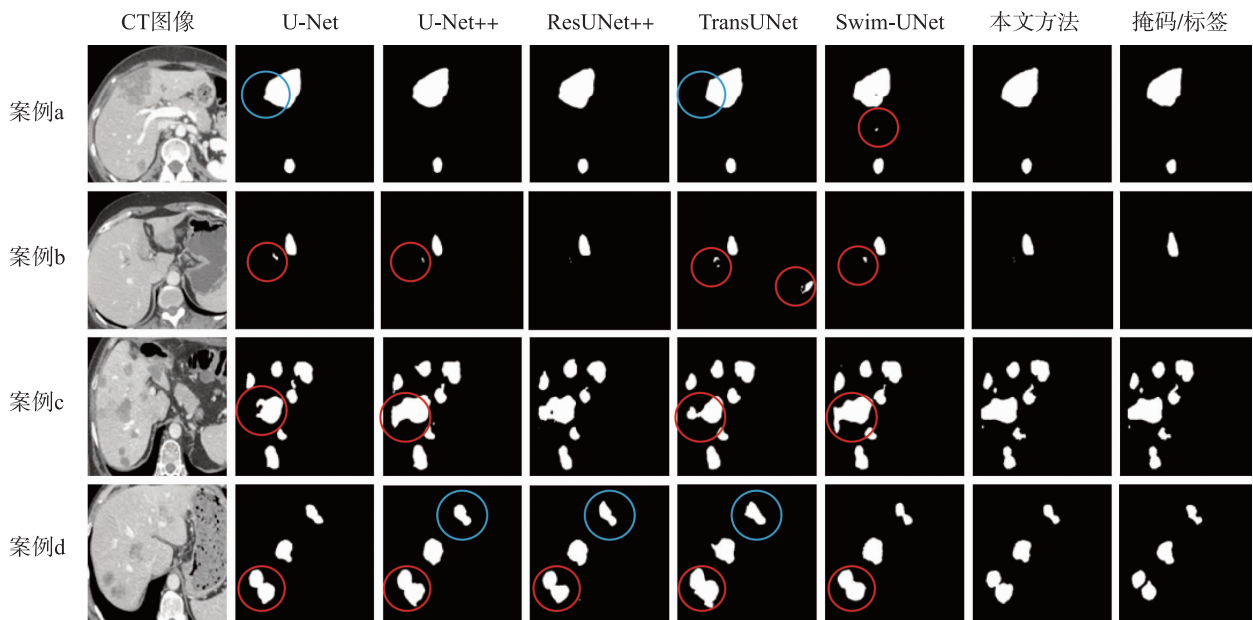


图 11 不同方法在 3DIRACADb-01 数据集上的分割结果

Fig.11 Segmentation results of different methods on 3DIRACADb-01 dataset

3.5.7 消融实验

如方法所述, MBTFEAM-UNet 在骨干网络 U-Net 中引入了并行卷积编码器(PCE)、纹理特征提取网络(MBTFE-Net)以及带有残差路径的通道注意力模块(RCA)。为了验证不同模块对分割性能提升的有效性, 以从无到有的形式与初始网络进行组合对 MBTFEAM-UNet 进行消融实验, 在肝脏肿瘤分割数据集 LiTS2017 和 3DIRCADb-01 上根据式(15)~式(20)计算得到的评价指标结果如表 6 和表 7 所示。结果表明: 将原网络中的双卷积编码器用并行卷积编码器替换后, 模型的分割性能有明显

提升, ASSD 和 HD 指标数值大幅度降低, 表明预测结果与金标准之间的重叠程度较好, 且肿瘤边缘分割清晰; 在跳跃连接处添加 MBTFE-Net 时, 网络性能改进最大, 评价指标 DSC 在两个数据集上相较于初始网络分别提升了 12.32% 和 8.06%。因此, MBTFE-Net 的引入对于分割性能的提升起着至关重要的作用; RCA 模块的加入使得评价指标 DSC 在两个数据集中相较于原始网络分别提升了 3.45% 和 4.35%, 这是因为 RCA 模块能够在模型训练过程中完成对通道依赖的建模, 增强肿瘤的边缘信息, 从而提升模型的分割能力; 原始网络与 PCE 和 MBT-

FE-Net 两种模块的组合在分割结果的各项指标中取得了次优值, 其均交并在两个数据集上分别取得了 84.94% 和 89.98% 的结果, 显然, 通过逐步添加模块, 分割性能相较于原始网络得到了极大的提升; 所提方法在两个数据集中的各项评价指标均达到了最优值, 其 DSC 指标高达 83.95% 和 89.66%, 相较于原始网络分别提高了 14.80% 和 11.24%, 充分呈现了该方法在肝脏肿瘤分割性能方面的卓越表

现。此外, 对 RCA 模块中提到的残差路径进行了消融实验, 结果显示, 残差路径的引入虽然仅对 DSC 指标产生了微小提升, 但在 ASSD 和 HD 两项指标上相较于未加残差路径的 RCA 模块均出现了约 1% 的下降。该数据表明, 残差路径的引入能够使所提方法对肿瘤的边缘分割更加精确。综上所述, 3 种模块的添加都积极的促进了分割性能的提升。

表 6 不同模块在 LiTS2017 数据集上的消融实验结果

Tab.6 Ablation experimental results of different modules on the LiTS2017 dataset

组合策略			DSC /%	VOE /%	RVD /%	ASSD /mm	HD /mm	MIoU /%
PCE	MBTFE-Net	RCA						
			69.15	47.15	-15.07	11.19	22.73	76.20
✓			73.79	40.35	-12.87	6.55	14.17	79.63
	✓		81.47	31.27	-7.46	4.78	10.56	84.20
		✓	72.60	43.01	-13.48	7.96	15.30	78.25
✓	✓		82.48	29.82	-6.10	3.86	9.27	84.94
✓		✓	74.71	39.10	-12.25	5.64	13.53	80.28
	✓	✓	81.65	31.01	-7.36	4.42	9.83	84.34
✓	✓	✓(没有残差路径)	83.60	28.18	-5.83	3.51	9.06	85.77
✓	✓	✓	83.95	27.66	-5.41	2.54	8.05	86.03

注: 加粗字体为最优结果。

表 7 不同模块在 3DIRACADb-01 数据集上的消融实验结果

Tab.7 Ablation experimental results of different modules on the 3DIRACADb-01 dataset

组合策略			DSC /%	VOE /%	RVD /%	ASSD /mm	HD /mm	MIoU /%
PCE	MBTFE-Net	RCA						
			78.42	35.50	-13.49	14.82	16.32	82.08
✓			84.62	26.81	-8.24	8.05	10.11	86.51
	✓		86.48	23.26	-7.26	5.10	9.25	87.77
		✓	82.77	29.40	-9.79	9.37	12.89	85.17
✓	✓		88.98	19.86	-5.59	2.30	7.24	89.98
✓		✓	85.45	25.71	-7.86	7.78	9.82	86.55
	✓	✓	87.96	20.89	-6.64	4.07	8.58	88.93
✓	✓	✓(没有残差路径)	89.22	19.32	-5.21	2.15	6.97	90.01
✓	✓	✓	89.66	18.75	-4.71	1.06	5.57	90.54

注: 加粗字体为最优结果。

3.5.8 跨器官分割能力评估

为了进一步验证所提方法在肿瘤分割任务中的有效性, 对肾脏肿瘤分割公开数据集 KiTS2019 类比 2.1 节所述预处理方法进行调窗裁剪处理, 并从

中选取了包含病灶在内的 5 662 张横断面切片对 MBTFEAM-UNet 进行对比实验, 最终取得的评价指标和分割效果分别如表 8 和图 12 所示。

实验结果表明, 所提方法无论从客观评价指标

还是主观分割结果可视化中均优于现有方法，在各项评价指标中均取得了最优值。可视化结果显示所提方法能够实现对小尺寸肿瘤的精准分割，且分割

出的肿瘤边界更加清晰。由此可见，所提方法具有一定的通用性和泛化能力，对于其他医学图分割任务具有一定的参考价值。

表 8 不同方法在 KiTS2019 数据集上的分割性能对比结果
Tab.8 Comparison of segmentation performance of different methods on the KiTS2019 dataset

方法	DSC /%	VOE /%	RVD /%	ASSD /mm	HD /mm	MIoU /%
U-Net	83.96	26.04	-10.21	17.58	11.04	84.89
U-Net++	87.17	22.74	-5.19	14.53	6.94	88.29
ResUNet++	88.71	19.19	-4.54	13.41	5.91	89.55
TransUnet	88.02	19.98	-5.25	13.92	6.78	89.16
Swin-Unet	90.58	16.26	-3.86	10.34	5.02	90.77
MBTFEAM-UNet	93.80	11.67	-2.51	8.43	3.71	94.00

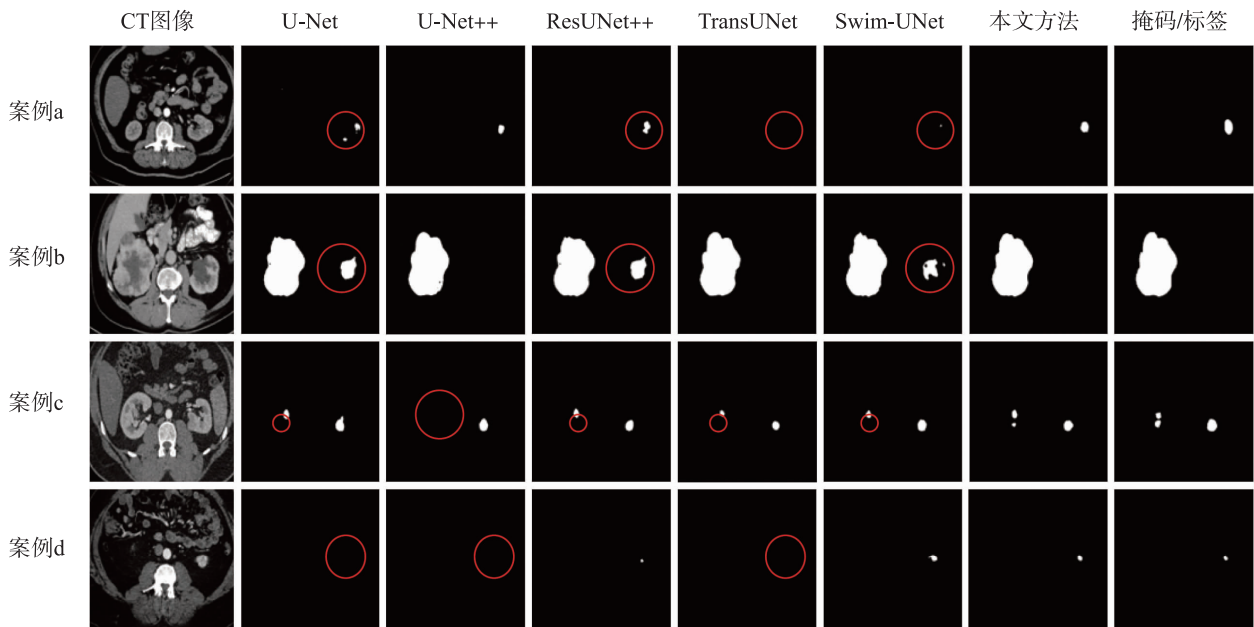


图 12 不同方法在 KiTS2019 数据集上的分割结果
Fig.12 Segmentation results of different methods on KiTS2019 dataset

4 结论

针对 CT 图像中肝脏肿瘤边界模糊、类型多样、与周围组织对比度低等特点以及现有网络对医学图像中的纹理信息利用不充分等问题，在原始 U-Net 网络的基础上，提出了一种结合纹理特征提取和注意力机制的网络模型——MBTFEAM-UNet。该方法设计了一个并行卷积编码器替换掉 U-Net 网络中的双卷积模块，用于提取不同感受野下的图像表层信息。此外，在 U-Net 网络的跳跃连接部分搭建了一

个纹理特征提取网络，其中的每个模块均采用不同大小的卷积核扩大感受野，使得网络能够捕获更多的深层图像纹理信息，以应对 CT 图像中肿瘤形状和大小的多变性。同时，在解码阶段引入带有残差路径的通道注意力模块来学习通道间的关联性，提升模型对肿瘤边界的分割能力。将所提方法在肝脏肿瘤分割数据集 LiTS2017 和 3DIRACADb-01 上完成了实验，并在肾脏肿瘤分割数据集 KiTS19 上进行了跨器官分割性能评估。实验结果表明，所提方法能够有效地解决现有网络对医学图像纹理信息提

取不充分的问题, 在边界模糊以及细小肿瘤的分割中表现出了明显优势, 能够在医疗诊断中起到辅助作用, 有望为肿瘤筛检工作提供新的参考。此外, 所提方法也为医学图像分割的其他相关任务提供了思路。

未来的工作方向包括进一步提高肝脏肿瘤分割的准确度和效率, 探索更好的纹理特征提取网络和注意力机制, 以及在临床医学中更广泛地应用深度

学习技术。

数据可用性声明

支撑本研究的科学数据已在中国科学院科学数据银行 (Science Data Bank) ScienceDB 平台公开发布, 访问地址为 <https://www.doi.org/10.57760/sciencedb.11107> 或 <http://cstr.cn/31253.11.sciencedb.11107>。

参考文献

- [1] 吕明慧, 周帅, 朱强. 基于深度学习乳腺超声计算机辅助诊断系统研究进展[J]. 中国医学影像技术, 2020, 36(11): 1722 – 1725.
LYU M H, ZHOU S, ZHU Q. Research progresses of breast ultrasound computer aided diagnosis systems based on deep learning [J]. Chinese Journal of Medical Imaging Technology, 2020, 36(11): 1722 – 1725.
- [2] 高飞, 闫镔, 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): 81 – 92.
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): 81 – 92.
- [3] MOLTZ J H, BORNEMANN L, KUHNIGK J M, et al. Advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in CT scans[J]. IEEE Journal of Selected Topics in Signal Processing, 2009, 3(1): 122 – 134.
- [4] MAHDY L N, EZZAT K A, TORAD M, et al. Automatic segmentation system for liver tumors based on the multilevel thresholding and electromagnetism optimization algorithm[J]. International Journal of Imaging Systems and Technology, 2020, 30(12): 1256 – 1275.
- [5] CASCIARO S, FRANCHINI R, MASSOPTIER L, et al. Fully automatic segmentations of liver and hepatic tumors from 3-D computed tomography abdominal images: Comparative evaluation of two automatic methods[J]. IEEE Sensors Journal, 2011, 12(3): 464 – 473.
- [6] ABD-ELAZIZ O F, SAYED M S, ABDULLAH M I. Liver tumors segmentation from abdominal CT images using region growing and morphological processing[C]//2014 International Conference on Engineering and Technology. Piscataway, USA: IEEE, 2014: 1 – 6.
- [7] XU L, ZHU Y, ZHANG Y, et al. Liver segmentation based on region growing and level set active contour model with new signed pressure force function [J/OL]. Optik, 2020 [2023 – 01 – 20]. <https://www.sciencedirect.com/science/article/pii/S0030402619316031>. DOI: 10.1016/j.ijleo.2019.163705.
- [8] LI C, WANG X, EBERL S, et al. A likelihood and local constraint level set model for liver tumor segmentation from CT volumes [J]. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2967 – 2977.
- [9] LI W. Automatic segmentation of liver tumor in CT images with deep convolutional neural networks[J]. Journal of Computer and Communications, 2015, 3(11): 146 – 148.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin, Germany: Springer, 2015: 234 – 241.
- [11] ZHOU Z, RAHMAN Siddiquee M M, TAJBAKHS N, et al. Unet++: A nested u-net architecture for medical image segmentation[C]//4th International Workshop of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision. Berlin, Germany: Springer, 2018: 3 – 11.
- [12] JHA D, SMEDSRUD P H, JOHANSEN D, et al. A comprehensive study on colorectal polyp segmentation with ResUNet++, conditional random field and test-time augmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(6): 2029 – 2040.
- [13] CHEN J, LU Y, YU Q, et al. Transunet: Transformers make strong encoders for medical image segmentation[EB/OL]. (2021 – 03 – 20) [2023 – 02 – 06]. <https://arxiv.org/abs/2102.04306>.

- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J/OL]. Advances in neural information processing systems, 2017(2017-10-20)[2023-01-26]. <https://dl.acm.org/doi/10.5555/3295222.3295349>. DOI: 10.48550/arXiv.1706.03762.
- [15] CAO H, WANG Y, CHEN J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation[C]//2022 Workshops: Computer Vision. Berlin, Germany: Springer, 2023: 205-218.
- [16] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//IEEE/CVF International Conference on Computer Vision. Piscataway, USA: IEEE, 2021: 10012-10022.
- [17] KUO C L, CHENG S C, LIN C L, et al. Texture-based treatment prediction by automatic liver tumor segmentation on computed tomography[C]//2017 International Conference on Computer, Information and Telecommunication Systems. Piscataway, USA: IEEE, 2017: 128-132.
- [18] DIAO Z, JIANG H, ZHOU Y. Leverage prior texture information in deep learning-based liver tumor segmentation: A plug-and-play Texture-Based Auto Pseudo Label module[J/OL]. Computerized Medical Imaging and Graphics, 2023[2023-01-30]. <https://www.sciencedirect.com/science/article/pii/S0895611123000356?via9030ihub>. DOI: 10.1016/j.compmedimag.2023.102217.
- [19] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2019: 3146-3154.
- [20] LIU Z, SUN M, ZHOU T, et al. Rethinking the value of network pruning[EB/OL]. (2019-03-05)[2023-04-18]. <https://arxiv.org/abs/1810.05270>.
- [21] 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.
- [22] 周芳芳, 王一达, 宋阳, 等. 基于2.5D级联卷积神经网络的CT图像胰腺分割方法[J]. 中国医学物理学杂志, 2020, 37(6): 786-791.
ZHOU F F, WANG Y D, SONG Y, et al. Segmentation of pancreas in CT images based on 2.5D cascaded convolutional neural network[J]. Chinese Journal of Medical Physics, 2020, 37(6): 786-791.

作者简介

邱云飞(1976-), 男, 博士, 教授。研究领域为数据挖掘与智能信息处理等。

王月洋(1998-), 男, 硕士生。研究领域为计算机视觉和医学图像处理。