



· 西北大学学报 110 周年刊庆 ·

基于多尺度特征和反向注意力的肝脏肿瘤 自动分割方法

张 瑞,唐乔湛,李斯卉,宋江玲

(西北大学 医学大数据研究中心,陕西 西安 710127)

摘要 肝脏肿瘤分割旨在定位肝脏肿瘤区域,以辅助医生进行精准诊治。鉴于深度学习能自动学习医学图像中复杂的特征和结构,已成为肝脏肿瘤分割的主流方法之一。但肝脏肿瘤的大小、形态存在显著差异及边缘模糊等问题,限制了深度学习模型的分割性能。基于此,该文提出了一种融合多尺度特征和反向注意力机制的深度网络,并用于肝脏肿瘤的自动分割。具体地,基于 U-Net 模型的框架,分别设计了多尺度特征提取模块和基于深度监督的反向注意力模块,使得网络能根据分割目标的大小自适应地选择不同尺度的特征,并引导网络关注分割目标的边缘特征,进而提高网络的边缘分割能力。此外,设计了一种新的混合损失,以解决医学图像分割中的类别不平衡问题。最后,在 MICCAI2017 LiTS 挑战赛数据集的数值实验结果表明,所提方法的 Dice 系数、平均对称表面距离 ASSD 分别为 76.12% 和 3.25 mm。

关键词 肝脏肿瘤分割;多尺度特征提取;反向注意力

中图分类号: TP391.4 **DOI**: 10.16152/j.cnki.xdxbr.2023-06-007

Automatic liver tumor segmentation using multi-scale features and reverse attention

ZHANG Rui, TANG Qiaozhan, LI Sihui, SONG Jiangling

(Medical Big Data Research Center, Northwest University, Xi'an 710127, China)

Abstract The objective of liver tumor segmentation is to pinpoint the region containing the liver tumor, aiding medical professionals in delivering precise diagnosis and treatment. Owing to its capacity to automatically learn intricate features and structures from medical images, deep learning has emerged as one of the mainstream approaches to liver tumor segmentation. Nonetheless, the size, shape, and blurred edges of liver tumor exhibit significant variations, which constrain the segmentation performance of deep learning models. Consequently, we introduce a deep network that incorporates multi-scale features and a reverse attention mechanism for the automatic segmentation of liver tumor. Specifically, drawing upon the U-Net model framework, this study designs a multi-scale feature extraction module and a reverse attention module based on deep supervision. This enables the network to adaptively select features of varying scales according to the size of the seg-

收稿日期:2023-10-15

基金项目:国家自然科学基金(12071369,62006189);陕西省自然科学基金(2021JQ-430,2023-JC-QN-0028);中国博士后科学基金(2022M722580)。

第一作者:张瑞,女,教授,博士生导师,从事机器学习理论与算法、生理信号模式识别等研究,rzhang@nwu.edu.cn。

mented target and direct the network's attention towards the edge features of the segmented object, thereby enhancing the network's edge segmentation capability. Furthermore, a novel hybrid loss is devised to address the issue of class imbalance in medical image segmentation. The recent numerical experiments conducted on the MICCAI2017 LiTS Challenge dataset have demonstrated that the proposed method achieves a Dice coefficient of 76.12% and an average symmetric surface distance (ASSD) of 3.25 mm.

Keywords liver tumor segmentation; multi-scale feature extraction; reverse attention

肝脏是腹腔内最大的实质性器官,具有代谢、解毒、免疫防御等生理功能。肝脏肿瘤是发生在肝脏的异常组织增生,可分为肝脏良性肿瘤和肝脏恶性肿瘤(也称为肝癌)。据统计,我国肝癌发病率占所有癌症的9%,居癌症发病谱第4位;肝癌死亡率占所有癌症的14%,居癌症死亡谱第2位^[1]。文献[2]通过对10 966例肝癌患者的临床资料进行回顾性分析,发现早期肝癌患者经治疗,其5年生存率高达62.5%~77.2%,而中、晚期肝癌患者的5年生存率仅23.8%。因此,尽早对肝脏肿瘤进行良恶性诊断对于及时诊断患者肝脏情况、维护肝脏健康、提高肝癌生存率具有重要意义^[3]。而肝脏肿瘤的准确分割是实现这一目标的重要前提。

肝脏肿瘤分割是指在相应医学影像中对肝脏肿瘤区域进行精确定位。计算机断层扫描(computed tomography, CT)是一种能够提供高分辨率图像的影像学检查,具有扫描速度快、对人体伤害小等优点,是临床中主要用于肝脏肿瘤分割的检查手段^[4]。传统的肝脏肿瘤分割主要是通过专业的放射科医生在CT上进行手动勾勒。然而,随着患者数量的增加,且每位患者的CT包含几十张甚至几百张切片图像,逐个切片进行人工分割与标记极其耗时费力,会大幅增加医生的工作压力。同时,人工标记容易掺杂医生的主观意识,不同水平的医生对肝脏病变的认识和判断往往不同,容易出现误诊和漏诊现象。随着科学技术的发展,利用计算机视觉和人工智能进行医学图像分割的方法大量涌现,这在很大程度上减轻了医生的工作量,降低了误诊和漏诊现象。同时也一定程度缓解了偏远地区医疗资源不平衡的问题,实现医疗资源下沉。

目前,肝脏肿瘤分割算法主要分为3类:基于传统图像分割的算法、基于机器学习的分割算法和基于深度学习的分割算法。其中,传统图像分割算法主要包括阈值分割法^[5]、水平集分割法^[6]、区域生长分割法^[7]等。基于机器学习的分

割算法主要是将特征工程和分类器相结合,具体地,通过对医学图像进行特征提取(如图像的轮廓信息、图像的强度分布、图像的纹理特征、图像的频域特征等),进而结合分类器(如K-均值算法^[8]、SVM^[9]、AdaBoost^[10]等)来完成前景像素和背景像素的分类,以实现肝脏肿瘤分割。上述第1类方法在分割阶段需要进行人工交互来完成算法的初始化设置,且算法的分割效果受初始化设置的影响较大;第2类方法依赖于人工设计的特征,其特征表示能力有限,导致分割方法泛化性能差,分割结果不够精准。

随着深度学习技术的发展,其强大的特征自主挖掘能力,使其在医学影像分割领域得到了广泛应用^[11]。其中,2015年Ronneberger等人提出的U-Net网络^[12],由于其结构简洁,尤其针对少量数据仍具有较好的分割性能,近年来在医学影像分割任务中受到极大关注。很多基于U-Net及其变体的模型与方法被提出用于实现肝脏肿瘤分割,包括H-DenseUNet^[13]、U-Net++^[14]、RIU-Net^[15]、MS-FANet^[16]等。

尽管基于深度学习的肝脏肿瘤自动分割方法得到了一定的发展,但仍存在局限性,具体体现在两方面:一是不同患者的肝脏肿瘤在大小、位置、形态等方面存在显著差异,易导致误检、漏检现象,已有方法通过金字塔结构、空洞卷积等提取不同尺度的特征,并将上述多尺度特征直接进行拼接或相加操作,忽略了大目标的细节信息和小目标的全局信息^[17-20];二是肝脏肿瘤的边缘模糊导致边缘分割精度低,已有方法通过引入额外的边缘分支、设计边缘损失函数以提高网络的边缘分割能力,但上述方法需要引入边缘先验知识,且边缘分割效果受边缘先验知识的影响较大^[21-25]。图1(a)~(c)是不同患者的CT切片中肝脏及肝脏肿瘤分布示意图,其中蓝色线是肝脏区域的边缘,红色线是肝脏肿瘤区域的边缘。从图1可以发现,(a)中患者的肝脏肿瘤大小各异、位置分散,少数肝脏肿瘤的边缘模糊;(b)中患者的肝脏

肿瘤边缘与肝脏边缘重合;(c)中患者的肝脏肿瘤极小。

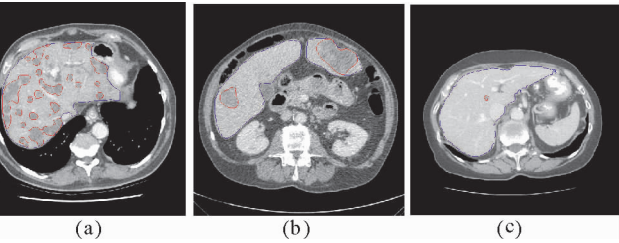


图 1 肝脏及肝脏肿瘤分布示意图

Fig. 1 Liver and liver tumor distribution diagram

基于上述分析,本文在 U-Net 网络框架下,提出了一种融合多尺度特征提取模块和基于深度监督的反向注意力模块的肝脏肿瘤自动分割模型 (multi-scale reverse attention u-net, MRA U-Net),其主要架构为编码器-解码器结构。具体地,在编码器上的多尺度特征提取模块使网络能够根据分割目标大小自适应地选择合适的尺度特征,从而缓解肝脏肿瘤大小、位置、形态存在显著性差异的问题;在编码器和解码器之间的基于深度监督的反向注意力模块能够引导网络关注编码器特征中有关分割目标的边缘特征,进一步为解码器特征补充细节特征,实现对肝脏肿瘤边缘的精细分割;此外,针对医学图像分割中的类别不平衡问题,本文设计了一种混合加权损失,使网络在训练过程中关注难分的像素点,以减少误诊、漏诊现象。

1 方法

1.1 MRA U-Net 整体结构

本文提出的肝脏肿瘤分割网络 MRA U-Net 是一个对称的编码器-解码器结构,其框架结构如图 2 所示。在编码器部分,本文设计了多尺度特征提取模块,以对 CT 图像的不同尺度特征进行提取,得到包含大目标细节信息和小目标全局信息的多尺度特征。具体地,首先,通过多尺度卷积层将输入特征沿通道维度划分成特征子集,并采用卷积操作和类残差连接提取不同特征子集上的多尺度特征。随后,采用通道注意力模块对上述特征进行融合,以适应不同大小的分割目标。解码器部分采用 U-Net 网络的解码器。而在编码器和解码器之间的跳跃连接上,本文设计了基于深度监督的反向注意力模块,通过从解码器上产生反向注意力特征图,以引导网络关注编码器特征中有关分割目标的边缘特征,进而为解码器特征补充分割目标的细节特征。最后,通过对基于深度监督的反向注意力模块输出的融合边缘信息的分割结果进行深度监督,使网络更好地学习分割目标的边缘,实现网络对肝脏肿瘤的精准定位。接下来,本文将详细介绍多尺度特征提取模块和基于深度监督的反向注意力模块的构造方式。

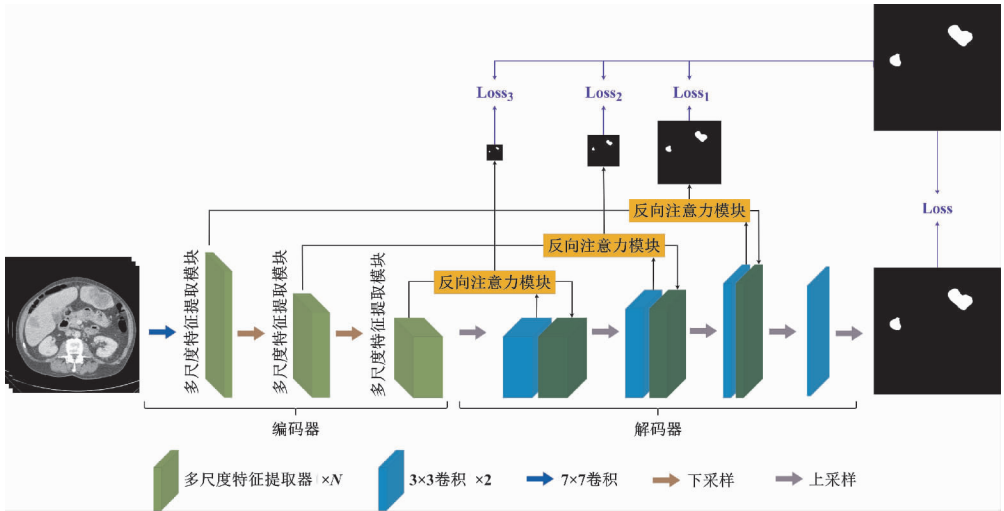


图 2 MRA U-Net 示意图

Fig. 2 MRA U-Net schematic

1.2 多尺度特征提取模块

肝脏肿瘤的大小、位置、形态等存在显著差异,大部分多尺度特征提取方法在扩大感受野的同时忽略了大目标的细节信息和小目标的全局信

息。为了提高网络对不同大小目标的分割能力,本文在编码器上设计了多尺度特征提取模块,如图 3 所示,该模块共包括 N 个多尺度特征提取器,每个特征提取器是由 6 个卷积层和 1 个注意

力层构成的残差结构。

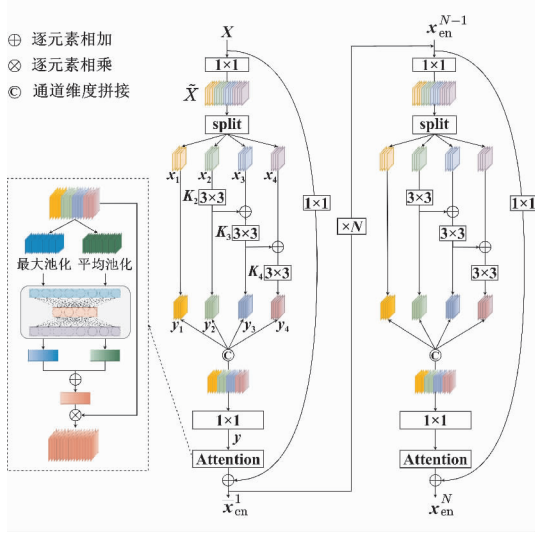


图3 多尺度特征提取模块示意图

Fig. 3 Multi-scale convolution diagram

每个特征提取器进行特征提取的过程主要分为以下4个步骤。

1) 对输入特征 $X \in \mathbf{R}^{H \times W \times C}$ 进行维度变换,即 X 经过 1×1 卷积层后得到特征 $\tilde{X} \in \mathbf{R}^{H \times W \times C_1}$ 。进而沿通道维度将 $\tilde{X}$ 平均划分成 s 份,得到 s 个特征子集 $x_h \in \mathbf{R}^{H \times W \times C_1/s}$ ($h = 1, 2, \dots, s$),

$$x_h = \text{split}(\text{Conv}^{1 \times 1}(X)) \quad (h = 1, 2, \dots, s) \quad (1)$$

式中:split 表示划分特征子集的操作; $\text{Conv}^{1 \times 1}$ 表示 1×1 卷积操作。

2) 特征子集 x_1 直接作为输出 y_1 , 特征子集 x_2 经过 3×3 卷积操作 K_2 后得到输出特征 y_2 , 其余每个特征子集 x_h ($2 < h \leq s$) 和 y_{h-1} 以类残差连接的方式相加后经过 3×3 卷积操作 K_h 得到输出特征 y_h , 即

$$y_h = \begin{cases} x_h, & h = 1; \\ K_h(x_h), & h = 2; \\ K_h(x_h + y_{h-1}), & 2 < h \leq s. \end{cases} \quad (2)$$

由于特征经过 3×3 卷积操作后会得到一个更大的感受野,因此 y_h 具有比 y_{h-1} 更大的感受野,即每个 y_h 代表一种尺度下的图像特征。

3) 将所有 $y_h \in \mathbf{R}^{H \times W \times C_1/s}$ 沿通道维度进行拼接后经过 1×1 卷积操作得到多尺度特征 $y \in \mathbf{R}^{H \times W \times C_2}$,

$$y = \text{Conv}^{1 \times 1}(\text{concat}(y_1, y_2, \dots, y_s)) \quad (3)$$

进而采用通道注意力模块(如图3中的子图所示)学习特征 y 每个通道的注意力分数,并据此得到

加权多尺度特征 $\text{Att}(y)$,

$$\text{Att}(y) = y \times \sigma(\text{MLP}(\text{MaxPool}(y)) + \text{MLP}(\text{AvgPool}(y))) \quad (4)$$

式中:concat 表示特征拼接操作;MaxPool 和 AvgPool 分别表示最大池化和平均池化;MLP 表示共享的全连接层; σ 表示 Sigmoid 激活函数。

4) 将输入特征 X 经过 1×1 卷积操作后进行维度调整,使其能够与加权多尺度特征 $\text{Att}(y)$ 进行融合,得到第一层多尺度特征提取器的最终输出特征 x_{en}^1 。上述过程可表示为

$$x_{\text{en}}^1 = \text{ReLU}(\text{Conv}^{1 \times 1}(X) + \text{Att}(y)) \quad (5)$$

式中:ReLU 表示 ReLU 激活函数。

由于多尺度特征提取模块中包含 N 个多尺度特征提取器,重复上述操作,则第 i 层多尺度特征提取器的输出特征为 x_{en}^i ,多尺度特征提取模块的最终输出为 x_{en}^N 。

1.3 基于深度监督的反向注意力模块

在传统的 U-Net 编码器-解码器结构中,解码器特征中包含高级的语义信息,使网络能够在图像中分辨出分割目标,而编码器特征中包含丰富的边缘、纹理等信息,这些特征通过跳跃连接的方式输入到解码器中,从而为解码器特征补充分割目标的位置信息,实现对分割目标的定位。但这种简单的跳跃连接方式使编码器无法反向利用解码器中的高级语义信息,从而导致编码器难以利用分割目标的整体信息对其局部细节进行微调,使得定位不够精确。因此,本文设计了一个基于深度监督的反向注意力模块(见图4),利用解码器特征产生反向注意力特征图,进而引导网络关注编码器特征中有关分割目标的边缘特征,以完成对于病灶区域局部细节的定位。

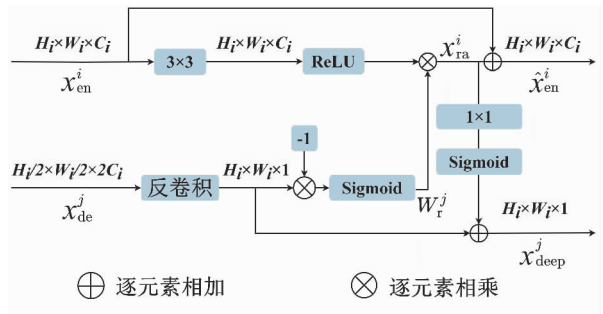


图4 基于深度监督的反向注意力模块示意图

Fig. 4 Schematic diagram of reverse attention module based on deep supervision

将在第 i 层编码器和第 j 层解码器之间跳跃连接上的编码器特征和解码器特征分别记为 x_{en}^i

$\in \mathbf{R}^{H_i \times W_i \times C_i}$ 和 $x_{de}^j \in \mathbf{R}^{H_i/2 \times W_i/2 \times 2C_i}$, 基于深度监督的反向注意力模块引导网络关注目标的边缘特征主要分为 4 个步骤。

1) 通过对解码器特征 x_{de}^j 进行反卷积得到分割结果, 并对其进行取反操作后, 采用 Sigmoid 激活函数得到反向注意力特征图 $W_r^j \in \mathbf{R}^{H_i \times W_i \times 1}$, 即

$$W_r^j = \sigma(-1 \times \text{Deconv}(x_{de}^j)) \quad (6)$$

式中: Deconv 表示反卷积操作; σ 表示 Sigmoid 激活函数。

2) 对编码器特征 x_{en}^i 通过 3×3 卷积以及 ReLU 激活函数后与反向注意力特征图 W_r^j 相乘得到加权特征 $x_{ra}^i \in \mathbf{R}^{H_i \times W_i \times C_i}$, 用以关注编码器特征中除当前分割结果中显著目标之外的边缘特征, 实现对当前分割结果的细节特征补充。将 x_{ra}^i 和原始的编码器特征 x_{en}^i 相加, 得到融合有效边缘信息的输出特征 x_{en}^i 。

$$\hat{x}_{en}^i = x_{ra}^i + x_{en}^i = \text{ReLU}(\text{Conv}^{3 \times 3}(x_{en}^i)) \times W_r^j + x_{en}^i \quad (7)$$

式中: $\text{Conv}^{3 \times 3}$ 表示 3×3 卷积操作; ReLU 表示 ReLU 激活函数。

3) 采用 1×1 卷积、Sigmoid 激活函数对 x_{ra}^i 进行操作得到边缘分割结果, 进一步和解码器特征 x_{de}^j 通过反卷积得到的分割结果相加, 得到包含边缘信息的分割结果 $x_{deep}^j \in \mathbf{R}^{H_i \times W_i \times 1}$ 。

$$x_{deep}^j = \sigma(\text{Conv}^{1 \times 1}(x_{ra}^i)) + \text{Deconv}(x_{de}^j) \quad (8)$$

式中: $\text{Conv}^{1 \times 1}$ 表示 1×1 卷积操作。

4) 为了监督边缘信息的学习, 本文将基于深度监督的反向注意力模块输出的分割结果 x_{deep}^j ($j = 1, 2, \dots, M$) 分别上采样至原始输入图像大小, 再与原始输入图像的真实标签 y_{gt} 计算损失, 最后将所有损失加权求和得到用于网络深度监督过程的损失, 具体表示为

$$\text{Loss}_{\text{deep}} = \sum_{j=1}^M a_j \text{Loss}(x_{deep}^j, y_{gt}) \quad (9)$$

式中: a_j ($j = 1, 2, \dots, M$) 表示深度监督不同阶段损失的权重。

1.4 损失函数

针对肝脏肿瘤分割中常常出现的类别不平衡问题, 本文设计了一种混合加权损失函数, 使网络在训练过程中关注难分的像素点, 以减少误诊、漏诊现象。

$$\text{Loss} = w_1 \text{loss}_{\text{ce}} + w_2 \text{loss}_{\text{tversky}} \quad (10)$$

式中: w_i ($i = 1, 2$) 表示各损失函数的权重; loss_{ce}

表示交叉熵损失 (cross-entropy loss, CE)^[26]; $\text{loss}_{\text{tversky}}$ 表示 Tversky 损失^[27]。具体表示为

$$\text{loss}_{\text{ce}}(p, g) = -\frac{1}{N} \sum_{i=1}^N \sum_s p_{is} \log g_{is} \quad (11)$$

$$\text{loss}_{\text{tversky}}(p, g) = \frac{\sum_{i,s} p_{is} g_{is}}{\sum_{i,s} p_{is} g_{is} + \alpha \sum_{i,s} \bar{p}_{is} g_{is} + \beta \sum_{i,s} p_{is} \bar{g}_{is}} \quad (12)$$

式中: p 表示网络的分割结果; g 表示真实标签; p_{is} ($\bar{p}_{is}$) 表示 p 中位置 i 的像素属于 (不属于) 类别 s 的概率; g_{is} ($\bar{g}_{is}$) 表示 g 中位置 i 的像素属于 (不属于) 类别 s 。

结合深度监督部分的损失, 本文在网络训练过程中使用的总损失为

$$\text{Loss}_{\text{total}} = \text{Loss} + \text{Loss}_{\text{deep}} \quad (13)$$

2 数值实验

2.1 数据集及预处理

本文的实验验证采用 MICCAI2017 LiTS 挑战赛数据集^[28], 该数据集包含从 7 个临床站点收集的 201 名患者的 CT 扫描图, 其中训练集包含 131 名患者的 CT 扫描图 (标签可见), 在线测试集包含 70 名患者的 CT 扫描图 (标签不可见)。训练集中 CT 的标注是由具有 3 年肿瘤成像经验的放射科医生使用 ITK-SNAP 软件完成, 包含肝脏肿瘤、肝脏、背景 3 类标签。由于不同临床站点使用的扫描设备跟扫描协议不同, 所有 CT 的层内分辨率范围为 $0.56 \sim 1.0 \text{ mm}$, 层厚范围为 $0.45 \sim 6.0 \text{ mm}$, 轴向切片的大小为 512×512 , 切片数范围为 $42 \sim 1\,026$ 。考虑到该挑战赛不再接收在线测试集上分割结果的提交, 本文仅使用标签可见的 131 名患者的 CT 扫描图作为全部数据集。

本文对于数据集的预处理包含重采样、调窗、归一化等步骤。具体地, 首先统计所有 CT 的空间采样率, 得到空间采样率的中位数为 $[0.76, 0.76, 1.0]$, 并将所有 CT 重采样至 $[0.76, 0.76, 1.0]$; 其次, 通过调窗操作将所有 CT 的强度值截断到 $[-200, 200]$ HU 范围内, 以增强 CT 的对比度, 减少无关组织器官对肝脏肿瘤分割结果的影响; 最后, 对所有 CT 进行归一化, 并按照轴位进行切片, 将相邻的 3 张 CT 切片和中间 CT 切片对应的标签分别作为输入图像和真实标签, 使网络学习相邻切片间的信息。数据集中包含肝脏的切

片有 19 156 张,包含肿瘤的切片有 7 190 张,数据预处理流程如图 5 所示。

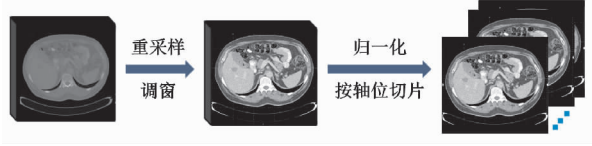


图 5 CT 预处理流程图

Fig. 5 CT preprocessing flow chart

2.2 实验环境及参数设置

本文实验使用 Adam 进行优化,初始学习率为 0.000 1,学习率的更新公式为

$$l_r = l_r \times \left(1 - \frac{i}{t_i}\right)^{0.9},$$

式中: i 表示当前迭代次数; t_i 表示全部迭代次数。数据集的批量大小设置为 8,训练轮数为 150 轮,本文采用翻转、旋转、对比度增强等方式对数据集进行数据增强,并将图像大小调整至 256×256 ,混合损失函数中 w_1 为 0.3, w_2 为 0.7, Tversky 损失函数中 α 为 0.1, β 为 0.9。本文所有表格中显示的实验结果均为 5 折交叉验证结果的平均值,所有实验均在 Linux Ubuntu 18.04.4 LTS 平台上借助显存为 12 GB 的 NVIDIA GeForce GTX 2080Ti、Python 3.7 以及 Pytorch 1.7.1 框架完成的。

2.3 实验结果与分析

2.3.1 评价指标

为了验证所提分割网络的性能,本文采用的模型性能评价指标有 Dice 系数 (dice coefficient, Dice)^[29]、体积重叠误差 (volumetric overlap error, VOE)^[30]、相对体积差 (relative volume difference, RVD)^[31]、平均对称表面距离 (average symmetric surface distance, ASSD)^[31]、最大对称表面距离 (maximum symmetric surface distance, MSD)^[32]。若网络的分割结果和真实标签分别为 A 和 B ,则各评价指标的计算公式如下。

1) Dice 用于计算网络的分割结果和真实标签之间的相似性,其值越大,分割效果越好,其数学表达式为

$$\text{Dice} = \frac{2|A \cap B|}{|A| + |B|} \quad (14)$$

2) VOE 用于计算网络的分割结果和真实标签的体素之间的体积重叠误差,值越接近 0,分割效果越好,其公式为

$$\text{VOE} = 1 - \frac{|A \cap B|}{|A \cup B|} \quad (15)$$

3) RVD 计算网络的分割结果和真实标签的体素之间的相对体积重叠误差,值越接近 0,分割效果越好,定义为

$$\text{RVD} = \frac{|B| - |A|}{|A|} \quad (16)$$

4) ASSD 计算网络的分割结果边界 $S(A)$ 上的体素到真实标签边界 $S(B)$ 以及真实标签边界 $S(B)$ 上的体素到网络的分割结果边界 $S(A)$ 的所有距离的平均值,值越接近 0,分割效果越好,表示为

$$\text{ASSD} = \frac{\left(\sum_{s_A \in S(A)} d(s_A, S(B))\right) + \left(\sum_{s_B \in S(B)} d(s_B, S(A))\right)}{|S(A)| + |S(B)|} \quad (17)$$

式中: d 表示欧几里得距离。

5) MSD 将 ASSD 中计算距离平均值替换为计算最大距离,值越接近 0,分割效果越好,具体公式为

$$\text{MSD} = \max\left(\max_{s_A \in S(A)} d(s_A, S(B)), \max_{s_B \in S(B)} d(s_B, S(A))\right) \quad (18)$$

2.3.2 消融实验

本文通过设计消融实验,以验证多尺度特征提取模块、基于深度监督的反向注意力模块以及混合损失函数的有效性。表 1 展示了消融实验结果,其中最优结果使用黑色加粗字体进行标记。

多尺度特征提取模块的有效性。通过比较表 1 中“基线模型 (U-Net)”和“基线模型 + 多尺度特征提取模块”的分割性能,发现多尺度特征提取模块的增加将 Dice 提高了 7.9%,将 VOE、ASSD、MSD 分别降低了 7.28%、7.67 mm、29.77 mm,同时使 RVD 更接近 0,提升了网络的分割性能。一方面,多尺度特征提取模块从不同的特征子集上提取了丰富的多尺度特征,能够同时包含大目标的细节信息和小目标的全局信息;另一方面,采用通道注意力模块对多尺度特征进行通道加权,使网络能够根据分割目标的大小自适应地选择合适的多尺度特征。上述结果说明本文所设计的多尺度特征提取模块是有效的。

基于深度监督的反向注意力模块的有效性。在肝脏肿瘤大小、位置、形态存在显著差异的情况下,仅通过 Dice 反映的网络分割结果和真实标签的相似程度无法很好地度量分割效果,需要通过距离度量来定量分析网络对于分割目标边缘的分

割能力。通过比较表 1 中“基线模型(U-Net)”和“基线模型 + 基于深度监督的反向注意力模块”的分割性能,发现基于深度监督的反向注意力模块的加入使 ASSD、MSD 降低了 7.38 mm、9.88 mm,同时使 Dice 增加了 6.71%,从而说明基于深度监督的反向注意力模块的嵌入对于肝脏肿瘤边缘分割精度的提升是有意义的。

损失函数的有效性。为了验证本文所设计的混合加权损失函数的有效性,采用不同的损失函数设置对 MRA U-Net 进行优化并评估网络的分割性能。具体地,本文设置的情况有“CE + Tversky”“CE + Dice”“CE”“Tversky”,通过对比可以发现,采用“CE + Tversky”对网络进行优化时,Dice 高达 76.12%,验证了混合加权损失函数的有效性。其中,Tversky 损失通过对假阳性样本和

假阴性样本加权,降低网络的误诊率和漏诊率,进而提升网络的分割性能。

2.3.3 超参数分析

为了进一步探讨本文使用的混合损失函数中交叉熵损失和 Tversky 损失的权重设置对模型效果的影响,本文采用不同的损失权重设置进行实验以对比网络的分割性能,实验结果如图 6 中的直方图所示,可以发现,当 $w_1 = 0.3, w_2 = 0.7$ 时,网络的分割性能最佳。此时 Tversky 损失作为混合损失函数的主要部分,使网络在前期训练时重点关注难分的像素,随着 Tversky 损失的下降,交叉熵损失引导网络学习分割结果中全体像素的分布与真实标签中全体像素分布的差异,提升网络的整体分割性能。

表 1 不同模块和不同损失在 LiTS 数据集上的消融实验

Tab. 1 Ablation experiments with different modules and different losses on LiTS dataset

模型	损失函数	Dice/%	VOE/%	RVD/%	ASSD/mm	MSD/mm
基线模型(U-Net)	CE + Tversky	64.78	48.33	0.185	12.36	80.23
基线模型 + 多尺度特征提取模块	CE + Tversky	72.68	41.05	-0.060	4.69	50.46
基线模型 + 基于深度监督的反向注意力模块	CE + Tversky	71.49	42.49	0.092	4.98	70.35
MRA U-Net	CE + Tversky	76.12	35.63	-0.031	3.25	48.54
MRA U-Net	CE + Dice	75.01	39.55	0.042	3.96	50.68
MRA U-Net	CE	73.96	40.74	0.094	4.33	63.18
MRA U-Net	Tversky	74.68	40.33	-0.101	3.75	52.64

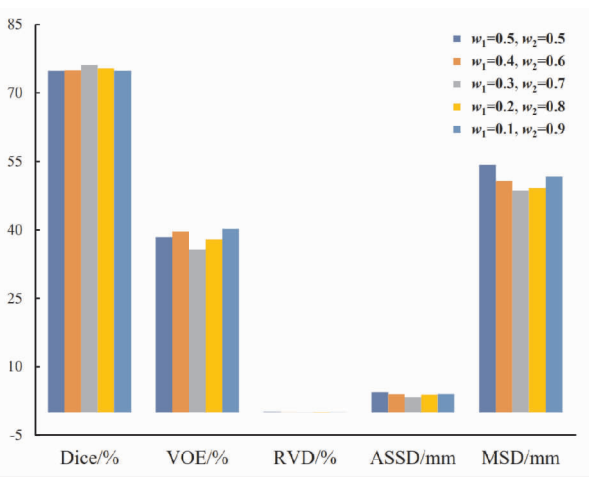


图 6 MRA U-Net 在不同混合损失函数权重下的分割性能

Fig. 6 Segmentation performance of MRA U-Net under different hybrid loss function weights

2.3.4 对比实验

为了验证本文所提方法的优越性,将本文所提方法和 U-Net、Attention U-Net、UNet++、TransUNet、RIU-Net、MS-FANet 的分割性能进行了比较。如表 2 所示,本文所提方法相较次优方法在 Dice 上提高了 1.92%,同时将 VOE、ASSD、MSD 降低了 1.07%、2.74 mm、1.24 mm,并且本文所提方法的 RVD 更接近 0。根据上述量化结果可知,本文所提方法在分割结果和真实标签的相似度、分割目标的边缘分割效果等方面表现良好,同时反映出多尺度特征提取模块和基于深度监督的反向注意力模块相结合的有效性。

2.3.5 分割结果可视化

2.3.4 小节从定量的角度分析了本文所提方法的有效性,本小节通过对不同方法的分割结果

进行可视化,以从定性的角度分析本文所提方法与其他方法的性能差异。图 7 展示了采用不同方法对不同患者 CT 切片进行分割的可视化结果。根据图 7 中第 1 行的可视化结果能够发现,U-Net 着重关注与对比度相关的特征,其将与肝脏肿瘤灰度值相似的组织误分割为肝脏肿瘤,同时 Attention U-Net 也存在上述问题,但其误分割现象

比 U-Net 少。从图 7 中第 3 行的可视化结果可以看出,本文所提的 MRA U-Net 对于小目标肝脏肿瘤的分割能力较强,能够有效避免漏诊现象的出现。通过图 7 中第 2、4、5 行的可视化结果能够发现,本文所提的 MRA U-Net 分割的肝脏肿瘤边缘比其他方法的分割结果更精确,验证了基于深度监督的反向注意力模块对网络边缘分割能力的提升。

表 2 不同方法在 LiTS 数据集上的分割性能

Tab. 2 Segmentation performance of different methods on LiTS dataset

方法	Dice/%	VOE/%	RVD/%	ASSD/mm	MSD/mm
U-Net	64. 78	48. 33	0. 185	12. 36	80. 23
Attention U-Net	71. 79	42. 04	0. 121	4. 92	89. 08
UNet ++	71. 90	42. 00	-0. 109	5. 02	85. 66
TransUNet	72. 01	41. 66	0. 058	3. 43	54. 76
RIU-Net	73. 79	39. 46	0. 160	4. 55	56. 51
MS-FANet	74. 20	36. 70	-0. 107	5. 99	49. 78
MRA U-Net	76. 12	35. 63	-0. 031	3. 25	48. 54

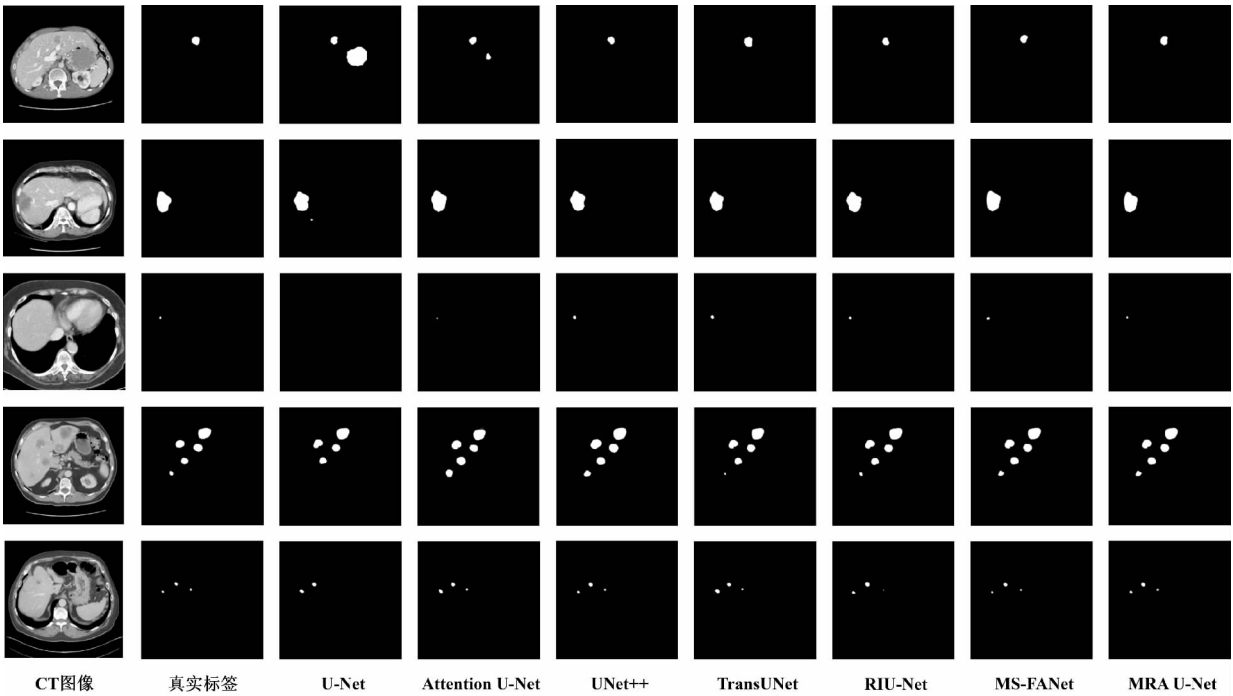


图 7 不同方法的分割结果可视化

Fig. 7 Visualization of segmentation results by different methods

3 结语

本文提出了一种融合多尺度特征提取模块和基于深度监督的反向注意力模块的肝脏肿瘤自动分割模型。首先,设计了一种新的多尺度特征提

取模块,并据此构造编码器用以提取 CT 图像的多尺度特征,通过通道注意力机制使网络能够根据分割目标的大小自适应地选择合适的多尺度特征;其次,在编码器-解码器的跳跃连接上嵌入基于深度监督的反向注意力模块,通过从解码器特征中产生反向注意力,以引导网络关注编码器特

征中有关分割目标的边缘特征,为解码器特征补充分割目标的细节信息并采用深度监督缓解梯度消失、促进网络的训练。此外,本文还设计了一种混合损失函数,以解决图像分割中存在的类别不平衡问题。经数值实验证明,本文所提方法在肝脏肿瘤的自动分割中性能表现良好。

参考文献

- [1] 曹毛毛,李贺,孙殿钦,等. 全球肝癌 2020 年流行病学现状[J]. 中华肿瘤防治杂志, 2022, 29(5): 322-328.
CAO M M, LI H, SUN D Q, et al. Global epidemiology of liver cancer in 2020 [J]. Chinese Journal of Cancer Prevention and Treatment, 2022, 29(5): 322-328.
- [2] 夏永祥,张峰,李相成,等. 原发性肝癌 10 966 例外科治疗分析[J]. 中华外科杂志, 2021, 59(1): 6-17.
XIA Y X, ZHANG F, LI X C, et al. Surgical treatment of primary liver cancer: A report of 10 966 cases [J]. Chinese Journal of Surgery, 2021, 59(1): 6-17.
- [3] 郑永昌,毛一雷. 早诊早治,提高我国肝癌 5 年生存率[J]. 肝胆胰外科杂志, 2022, 34(6): 321-324.
ZHENG Y C, MAO Y L. Improve the 5-year survival rate of hepatocellular carcinoma in China via early diagnosis and treatment [J]. Journal of Hepatopancreatobiliary Surgery, 2022, 34(6): 321-324.
- [4] QURESHI I, YAN J, ABBAS Q, et al. Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends [J]. Information Fusion, 2023, 90: 316-352.
- [5] CIECHOLEWSKI M, OGIELA M R. Automatic segmentation of single and multiple neoplastic hepatic lesions in CT images [C] // International Work-conference on the Interplay between Natural and Artificial Computation. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007: 63-71.
- [6] YANG X P, DO YANG J, HWANG H P, et al. Segmentation of liver and vessels from CT images and classification of liver segments for preoperative liver surgical planning in living donor liver transplantation [J]. Computer Methods and Programs in Biomedicine, 2018, 158: 41-52.
- [7] 宋红,王勇,黄小川,等. 基于动态自适应区域生长的肝脏 CT 图像肿瘤分割算法[J]. 北京理工大学学报, 2014, 34(1): 72-76.
SONG H, WANG Y, HUANG X C, et al. A dynamic adaptive region growing segmentation algorithm for tumor of liver CT images [J]. Transactions of Beijing Institute of Technology, 2014, 34(1): 72-76.
- [8] MASSOPTIER L, CASCIARO S. A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans [J]. European Radiology, 2008, 18(8): 1658-1665.
- [9] VORONTOV E, ABI-JAOUDEH N, KADOURY S. Metastatic liver tumor segmentation using texture-based omni-directional deformable surface models [C] // Abdominal Imaging. Computational and Clinical Applications; 6th International Workshop, ABDI 2014. Cambridge, MA, USA: Springer International Publishing, 2014: 74-83.
- [10] LI Y, HARA S, SHIMURA K. A machine learning approach for locating boundaries of liver tumors in CT images [C] // 18th International Conference on Pattern Recognition (ICPR'06). Hong Kong, China: IEEE, 2006, 1: 400-403.
- [11] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE, 2015: 3431-3440.
- [12] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C] // Medical Image Computing and Computer-Assisted Intervention. MICCAI 2015; 18th International Conference. Munich, Germany: Springer International Publishing, 2015: 234-241.
- [13] LI X, CHEN H, QI X, et al. H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes [J]. IEEE Transactions on Medical Imaging, 2018, 37(12): 2663-2674.
- [14] ZHOU Z, SIDDIQUEE M M R, TAJBAKHS N, et al. UNet ++: Redesigning skip connections to exploit multiscale features in image segmentation [J]. IEEE Transactions on Medical Imaging, 2020, 39(6): 1856-1867.
- [15] LV P, WANG J, WANG H. 2.5 D lightweight RIU-Net for automatic liver and tumor segmentation from CT [J]. Biomedical Signal Processing and Control, 2022, 75: 103567.
- [16] CHEN Y, ZHENG C, ZHANG W, et al. MS-FANet: Multi-scale feature attention network for liver tumor segmentation [J]. Computers in Biology and Medicine, 2023, 163: 107208.
- [17] GAO S H, CHENG M M, ZHAO K, et al. Res2Net: A new multi-scale backbone architecture [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662.

[18] LI B, XIAO C, WANG L, et al. Dense nested attention network for infrared small target detection [J]. IEEE Transactions on Image Processing, 2022, 32: 1745-1758.

[19] WANG K, ZHANG X, ZHANG X, et al. EANet: Iterative edge attention network for medical image segmentation[J]. Pattern Recognition, 2022, 127: 108636.

[20] FU Z, ZHANG R. ω -net: Dual supervised medical image segmentation with multi-dimensional self-attention and diversely-connected multi-scale convolution [J]. Neurocomputing, 2022, 500: 177-190.

[21] CHEN S, TAN X, WANG B, et al. Reverse attention for salient object detection [C] // Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 234-250.

[22] LOU A, GUAN S, LOEW M. Caranet: Context axial reverse attention network for segmentation of small medical objects [J]. Journal of Medical Imaging, 2023, 10(1): 014005-014005.

[23] DI S, ZHAO Y Q, LIAO M, et al. TD-Net: A hybrid end-to-end network for automatic liver tumor segmentation from CT images[J]. IEEE Journal of Biomedical and Health Informatics, 2023, 27(3): 1163-1172.

[24] GAO H, LYU M, ZHAO X, et al. Contour-aware network with class-wise convolutions for 3D abdominal multi-organ segmentation [J]. Medical Image Analysis, 2023, 87: 102838.

[25] ZHOU Z, BIAN Y, PAN S, et al. A dual branch and fine-grained enhancement network for pancreatic tumor segmentation in contrast enhanced CT images[J]. Biomedical Signal Processing and Control, 2023, 82: 104516.

[26] SUDRE C H, LI W, VERCAUTEREN T, et al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations [C] // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017. Quebec City, Canada: Springer International Publishing, 2017: 240-248.

[27] SALEHI S S M, ERDOGMUS D, GHOLIPOUR A. Tversky loss function for image segmentation using 3D fully convolutional deep networks [C] // International Workshop on Machine Learning in Medical Imaging. Cham: Springer International Publishing, 2017: 379-387.

[28] BILIC P, CHRIST P, LI H B, et al. The liver tumor segmentation benchmark (lits) [J]. Medical Image Analysis, 2023, 84: 102680.

[29] 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). Stanford, CA, USA: IEEE, 2016: 565-571.

[30] JACCARD P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura [J]. Bull Soc Vaudoise Sci Nat, 1901, 37: 547-579.

[31] HEIMANN T, VAN GINNEKEN B, STYNER M A, et al. Comparison and evaluation of methods for liver segmentation from CT datasets [J]. IEEE Transactions on Medical Imaging, 2009, 28(8): 1251-1265.

[32] HUTTENLOCHER D P, KLANDERMAN G A, RUCKLIDGE W J. Comparing images using the Hausdorff distance [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(9): 850-863.

(编辑 张欢)

作者简介:



张瑞,博士,教授,博士生导师,“陕西省特支计划”科技创新领军人才,陕西省医学数据分析挖掘与智能应用创新团队负责人,微系统可靠性分析与仿真陕西省高校工程研究中心主任,西北大学医学大数据研究中心主任,西北大学数学学院院长。分别获得西安交通大学应用数学专业理学博士学位与新加坡南洋理工大学电子电气工程专业工学博士学位。先后赴美国伊利诺伊大学数学系、美国哈佛医学院及麻省总医院访问交流。目前从事机器学习理论与算法、医学数据分析与处理、神经计算建模等方向的教学与科研工作。已在国内外重要杂志和国际学术会议上发表学术论文 80 余篇,主持国家级省部级等科研项目 9 项,以第一完成人获陕西省科学技术奖二等奖 1 项。