

基于堆叠树形聚合结构空洞卷积的肝脏肿瘤分割

高飞^{1*}, 闫宾¹, 陈健¹, 乔凯¹, 宁培钢², 史大鹏²¹战略支援部队信息工程大学信息工程学院, 河南 郑州 450001;²河南省人民医院影像科, 河南 郑州 450002

摘要 为了克服传统肝脏肿瘤分割网络下采样带来的细节信息丢失问题,同时提取丰富的多尺度信息,提出了一种基于堆叠树形聚合结构空洞卷积的肝脏肿瘤分割算法。首先,在编码器网络中提出了残差密集模块;然后,在编码器-解码器网络中加入树形聚合结构的空洞卷积模块,有效消除了普通空洞卷积带来的棋盘伪影现象,提高了分割精度。最后,用加权的损失函数解决了图像中前景和背景不平衡的问题。实验结果表明,本算法在电子计算机断层扫描图像数据集上的 Dice 相似性系数、像素正确率和交并比分别为 0.8026、0.7974 和 0.7317。

关键词 图像处理; 残差网络; 密集连接; 空洞卷积; 肝脏肿瘤

中图分类号 TP391

文献标志码 A

doi: 10.3788/AOS202141.1810002

Liver Tumor Segmentation Based on Dilated Convolution of Stacked Tree Aggregation Structure

Gao Fei^{1*}, Yan Bin¹, Chen Jian¹, Qiao Kai¹, Ning Peigang², Shi Dapeng²¹College of Information System Engineering, PLA Strategic Support Force Information Engineering University, Zhengzhou, Henan 450001, China;²Department of Radiology, Henan Provincial People's Hospital, Zhengzhou, Henan 450002, China

Abstract In order to overcome the loss of detail information caused by down sampling of traditional liver tumor segmentation networks and extract rich multi-scale information at the same time, this paper proposes an algorithm of liver tumor segmentation based on dilated convolution of stacked tree aggregation structure. First, a residual dense module is proposed in the encoder network. Then, a dilated convolution module of stacked tree aggregation structure is added to the encoder-decoder network, which can effectively eliminate the checkerboard artifacts caused by ordinary dilated convolution and improve the segmentation accuracy. Finally, a weighted loss function is used to solve the problem of the imbalance between the foreground and the background in the image. The experimental results show that the Dice similarity coefficient, pixel accuracy rate and intersection ratio of the algorithm on the computer tomography image data set are 0.8026, 0.7974 and 0.7317, respectively.

Key words image processing; residual network; dense connection; dilated convolution; liver tumor

OCIS codes 100.3008; 100.4996; 200.4260

1 引言

电子计算机断层扫描(CT)图像中肿瘤的自动分割对于肝癌的准确诊断和指导治疗具有重要的临床应用价值。CT 图像中肝脏肿瘤形状、大小、灰度的高度变异性,导致肝脏肿瘤可能与其他器官和血

管相邻,肿瘤内部图像灰度不均匀,肿瘤与邻近正常组织的对比度较低,继而产生一系列伪影和弱边界、噪声等不良现象,使肿瘤区域与正常肝组织之间的边界比较模糊,增加了肿瘤分割的难度,因此,CT 图像中的肝脏肿瘤自动分割是一个具有挑战性的问题。

传统 CT 图像的肝脏肿瘤分割算法主要有阈值

收稿日期: 2021-03-04; 修回日期: 2021-03-28; 录用日期: 2021-04-07

基金项目: 国家重点研发计划(2018YFC0114500)

通信作者: *gfflyfly@163.com

法、区域增长法、水平集理论、图割和机器学习法。Moltz 等^[1]提出了一种基于灰度阈值法的肝脏肿瘤分割算法,根据图像的灰度值设定阈值进行二分类,并加入了形态学操作。Lu 等^[2]用蒙特卡罗方法模拟图像的灰度值分布,以选取种子点和建立区域合并准则,但这种分割算法对肿瘤边界比较模糊的图像效果欠佳。Casciaro 等^[3]通过对肿瘤轮廓进行曲线演化构建能量函数,包括当前的曲线形状和分割目标,不断优化能量函数使曲线形状与分割目标轮廓一致,并在此基础上提出了一种基于梯度矢量流和图割的肝脏肿瘤分割算法。Li 等^[4]提出了一种融合图像梯度、区域竞争和先验信息的统一水平集模型,用于 CT 图像中肝脏肿瘤的分割,并利用无监督模糊聚类方法估计肿瘤的概率分布。Zareei 等^[5]在主动活动轮廓的优化中使用了梯度矢量流和遗传算法,减轻了分割算法对初始轮廓的依赖,并结合肝脏位置及其先验知识,使初始曲线接近指定的区域边界,降低了局部最小效应。Shimizu 等^[6]提出了一种用于肝肿瘤分割的 AdaBoost 分类器,通过提取图像的灰度均值、纹理特征训练多个弱 AdaBoost 分类器,进而构建一个强分类器。Conze 等^[7]提出了一种结合随机森林和层次多尺度树的肿瘤分割算法。Vorontsov 等^[8]将一种支持向量机分类器用于 CT 图像中的肝肿瘤分割,并通过一种全方位可变形曲面模型对结果进行细化。上述算法均存在较大的局限,如阈值法、区域增长法对肿瘤和周边正常组织灰度接近的情况分割效果较差;大部分传统分割算法需要用户手工设置种子点或肿瘤的初始位置等先验知识,难以实现肝脏的全自动分割,且基于人工设计的特征和先验信息的有效性对分割精度的影响较大。

深度神经网络在医学图像分割领域的应用越来越广泛,如 U 型网络 (UNet)^[9]、语义分割网络 (SegNet)^[10] 和 DeepLabv3^[11]、全连接密集连接网络 (FC-DenseNet)^[12]。Christ 等^[13]提出了一种在 CT 图像分割中使用的级联全卷积网络 (CFCN) 和密集三维条件随机场 (3D CRF) 模型。Chlebus 等^[14]对初始 FCN 框架进行了改进,并用 3D CRF 对最终的预测结果进行细化,提升了肝脏肿瘤的分割精度。FCN 的卷积核是固定的,为了扩大感受野,需要进行下采样,但同时会丢失很多细节信息,影响最终的分割精度。Han^[15]开发了一个多层的 2.5 D 卷积神经网络 (CNN) 模型,利用了 UNet 的跳层连接融合高、低层特征和残差网络 (ResNet) 的直接连接对图像进行编码。黄泳嘉等^[16]将改进的

UNet 用于肝部图像的分割。Zhou 等^[17]提出了一种基于 UNet++ 架构的分割模型,在网络中采用跳跃连接对特征进行叠加以提高分割精度,同时引入残差结构解决了梯度消失问题。Li 等^[18]提出了一种基于 UNet 和密集连接网络 (DenseNet) 的 CT 图像肝脏病变分割算法,通过融合层混合片内和片间特征优化模型,取得了不错的分割效果。上述大部分基于 UNet 的模型通过跳层连接融合高、低层特征,以恢复细节信息,分割效果较好,但特征融合方式比较简单。为了进一步提高分割精度,人们提出了多种融合上下文信息的分割算法。Gu 等^[19]利用密集卷积块和多核池化块提取更丰富的空间特征,提出了一个融合上下文信息的分割网络,提高了肿瘤的分割精度。艾玲梅等^[20]在深度 CNN 中引入注意力机制,实现了医学图像的分割。Chen 等^[21]通过注意力机制融合不同层次的特征。张文秀等^[22]用残差卷积层代替 UNet 中的原始卷积层,并在解码部分引入注意力机制,提高了图像的分割效果。肝脏肿瘤的尺寸从几毫米到十几厘米不等,因此,多尺度信息的利用对于分割网络也是至关重要的。Chen 等^[23]提出了空洞卷积的概念,在保持卷积核大小不变的情况下,利用具有一定空洞率的卷积获得更大的感受野,保留了更多的细节信息,有利于多尺度信息的提取。郑婷月等^[24]在 FCN 中引入残差块和空洞卷积,进一步提升了分割精度。吴晨玥等^[25]用空洞卷积代替了残差网络中的原始卷积,从而提高了分割精度。冯博文等^[26]将改进的空洞卷积模块用于核磁图像的分割。但空洞卷积作为稀疏计算会引入棋盘伪影,如 DeepLabv3 中基于空间金字塔池化 (ASPP) 的空洞卷积,会导致网络在肿瘤图像的边缘分割精度较低。

针对上述问题,本文提出了一种基于堆叠树形聚合结构空洞卷积的肝脏肿瘤分割算法。在编码器网络中加入了残差密集模块 (RDB),在解决梯度消失问题的同时进行特征重用。为了更好地恢复细节信息,通过跳层连接融合高层特征和低层特征。在编码器-解码器网络中加入树形聚合结构的空洞卷积 (TASD) 模块,从而扩大感受野,提取丰富的多尺度信息。此外,树形聚合结构可以有效消除普通空洞卷积带来的棋盘伪影现象,提高了肝脏肿瘤的分割精度。

2 网络模型与算法

2.1 残差密集模块

深度学习中的有监督学习是用有标签的训练集

数据训练一个深度网络,使深度网络能够逐层将隐含在数据中的特征抽象出来,并输出接近目标和任务的特征表达,然后通过类似于 Softmax 的分类器输出最终结果。深度网络的层数较多,可提取到更抽象的表达,但误差在反向传播过程中,梯度会逐渐消失或变大,产生梯度弥散或梯度爆炸问题,使低层深度网络参数不能进行有效的更新。残差网络^[27]通过构造恒等映射解决梯度消失问题,可表示为

$$y_l = h(x_l) + F(x_l, \omega_l), \quad (1)$$

$$x_{l+1} = f(y_l), \quad (2)$$

式中, x_l 和 x_{l+1} 分别为第 l 层的输入和输出, ω_l 为第 l 层的权重信息, $h(x_l)$ 为恒等映射, $F(\cdot)$ 为残差处理模块,可学习输入与输出的残差, $f(\cdot)$ 为激活函数。第 L 层学习到的特征可表示为

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, \omega_i), \quad (3)$$

式中, x_i 和 ω_i 分别为对应层的输入和权重信息。

若损失函数为 ϵ , 则误差反向传播时梯度的反向传播过程可表示为

$$\frac{\partial \epsilon}{\partial x_l} = \frac{\partial \epsilon}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_l} = \frac{\partial \epsilon}{\partial x_L} \cdot \left[1 + \frac{\partial}{\partial x_L} \sum_{i=l}^{L-1} F(x_i, \omega_i) \right], \quad (4)$$

式中, $\frac{\partial \epsilon}{\partial x_l}$ 为损失函数 ϵ 关于 x_l 的梯度信息。恒等映射机制可以保证梯度不会消失,从而抑制梯度衰减现象。

密集连接网络^[28]的特点是在每一个密集块中通过级联方式合并每一层的特征并向前推进,第 L 层有 L 个输入,即一个 L 层的网络共有 $L(L+1)/2$ 个连接。对于每一层,所有前层的特征都是其输入。密集连接在减少梯度消失的同时强调了特征继承和特征重用,并将网络的宽度控制在一个较窄的水平上,很好地解决了过拟合等问题。假设输入图像为 X , 网络结构有 L 层, 则密集连接可表示为

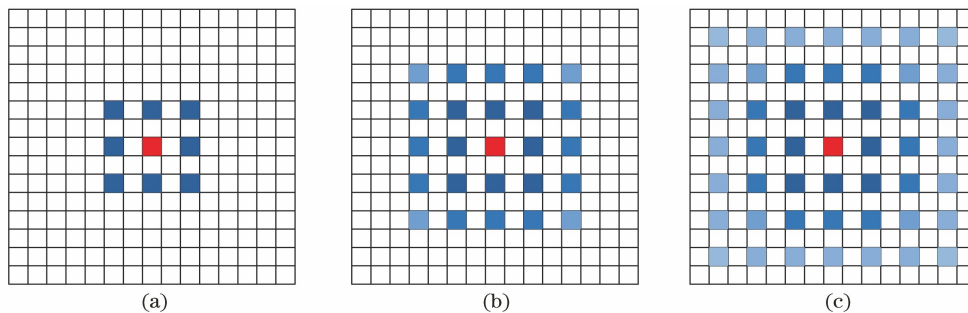


图 2 依次进行 $r=2$ 的空洞卷积操作。(a)第 1 次;(b)第 2 次;(c)第 3 次

Fig. 2 Dilated convolution operation of $r=2$ is performed in sequence. (a) First time; (b) second time; (c) third time

$$x_L = H_L(x_0, \dots, x_{L-1}), \quad (5)$$

式中, $H_L(\cdot)$ 为非线性转换函数,包括批归一化(BN)、修正线性单元(ReLU)和 3×3 卷积三个连续的操作。第 L 层的输入为 x_0, \dots, x_{L-1} 的特征信息,网络每一层的输入都是之前所有层输出的级联合并,该层所学习的特征图也会被直接传递到后面层作为输入。

受残差网络和密集连接网络的启发,提出了一种 RDB,其结构如图 1 所示。其中,每一层都是 BN+ReLU+ 3×3 卷积三个连续的操作。RDB 在保留密集连接网络优点的同时,又吸收了残差网络恒等映射的优点。设 X 为输入图像,则密集连接网络的输出 $F(X)$ 可表示为

$$F(X) = H(X) - X, \quad (6)$$

式中, $H(X)$ 为预期输出。

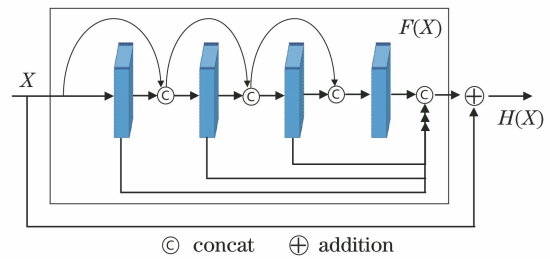


图 1 RDB 的结构

Fig. 1 Structure of the RDB

2.2 树形聚合结构的空洞卷积模块

下采样操作会降低图像的分辨率,导致细节信息的丢失,而在图像尺寸不变的情况下,空洞卷积可在不增加参数量或计算量的情况下引入丰富的多尺度信息,从而控制深度卷积神经网络中计算特征的分辨率。对像素点进行扩张率 $r=2$ 的空洞卷积,等同于对该像素点进行 5×5 卷积,依次对正中心像素点进行 $r=2, r=2, r=2$ 的空洞卷积,具体操作如图 2 所示,分割结果如图 3 所示。可以发现,空洞卷积作为一种稀疏运算,会带来明显的棋盘伪影问题;

且长距离像素点之间形成了割裂,导致肝脏肿瘤分割的边沿出现伪影,分割精度下降,如图 3(c)所示。

为了解决棋盘伪影现象,依次对正中心像素点进行 $r=1, r=2, r=3$ 的空洞卷积,具体操作如图 4

所示。可以发现,图 4(c)中计算结果的感受野达到了 13×13 ,基本填补了空白像素点,有效去除了棋盘伪影。但需要注意的是,叠加不同空洞卷积的扩张率组合中不能出现相同或大于 1 的公约数。

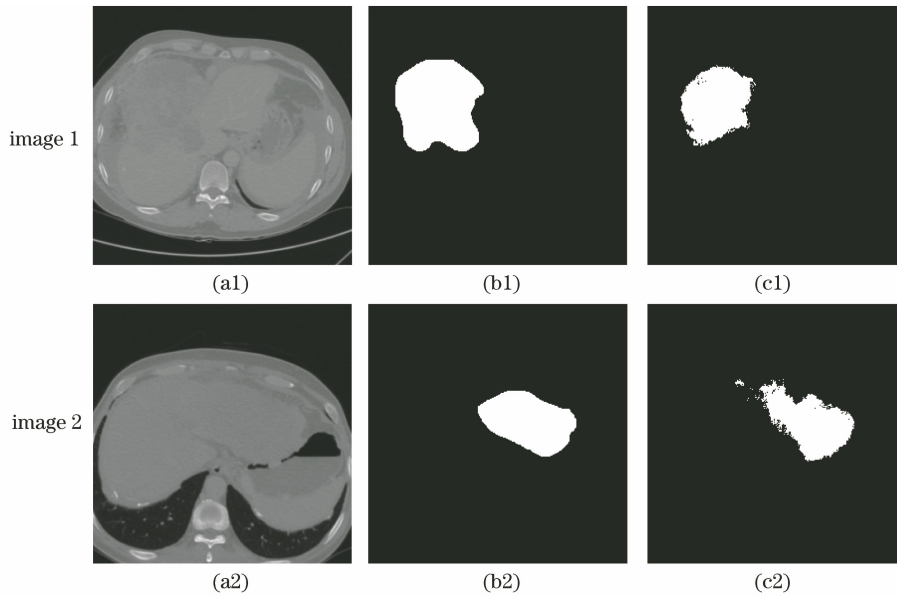


图 3 空洞卷积的分割结果。(a)原始图像;(b)分割真值;(c)棋盘伪影

Fig. 3 Segmentation results of the dilated convolution. (a) Original image; (b) ground truth; (c) grid artifact

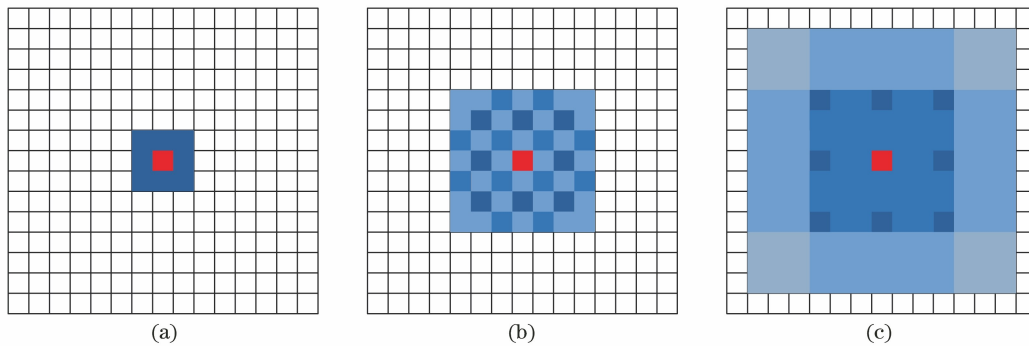


图 4 依次进行不同扩张率的空洞卷积操作。(a) $r=1$; (b) $r=2$; (c) $r=3$

Fig. 4 Dilated convolution operations with different dilated rates in sequence. (a) $r=1$; (b) $r=2$; (c) $r=3$

基于上述分析,提出了一种 T ASD 模块,其结构如图 5 所示。T ASD 模块分别对来自下采样和上采样的特征进行不同扩张率($r=1, r=3, r=6, r=9, r=12$)的空洞卷积,树形结构的计算支路共 5 条,每条支路的输入均为前一条支路的输出,在树形结构的输出端对 5 条支路的结果进行聚合。考虑到过大的扩张率对较小的肿瘤信息丢失比较明显,而肝脏肿瘤中小尺寸的肿瘤占比较大,因此,对 5 条支路的特征图数量进行了合理分配,即为扩张率较小的支路分配较多的特征图数量。假设输入信号的特征图数量为 n ,则 $r=1, r=3, r=6, r=9, r=12$ 的支路对应的特征图数量比例分别为 $n/2, n/4, n/8,$

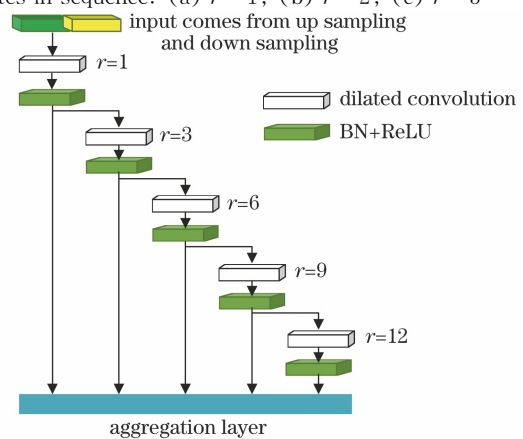


图 5 T ASD 模块的结构

Fig. 5 Structure of the T ASD module

$n/16, n/16$ 。每条支路进行空洞卷积后,还需进行 BN+ReLU 操作。

2.3 网络框架

CT 图像中肝脏肿瘤的特点是尺寸变化大、数据样本少,因此,分割网络的规模不能太大,参数不能太多。本算法基于密集连接网络的思想进行特征重用并控制网络宽度,从而减少模型规模和网络参数,可适用于小样本的医学图像分割领域。为了减少梯度消失的问题,利用残差网络的思想对密集连接网络的 Dense block 进行改进,将 Dense block 中最后输出的级联操作改为加法,提出了一种 RDB。用 RDB 构建的编码器可以提取更多的高层语义特征,在编码器-解码器网络中加入了 TASD 模块,以提取丰富的多尺度信息,网络的总体结构如图 6 所示,其中,Conv 为卷积操作。首先,在图像的某一个尺寸下,用 TASD 模块提取图像的多尺度特征。然后,将多尺度特征融合后送到编码器的 RDB 中,

RDB 学习的对象是含有多尺度特征的图像。最后,对编码器 RDB 的输出进行下采样。传统的池化下采样中没有参数,而本算法使用了步长为 2 的 3×3 卷积进行下采样,其参数通过训练集进行学习。在下采样通道中,下采样率为 16,对于输入分辨率为 512×512 的图像,经过 4 次下采样后,其分辨率为 32×32 ,继续进行下采样会丢失一些较小肝脏肿瘤的信息。为了减少计算量,在解码器网络中不再使用 RDB,而是在上采样之前采用一个 3×3 卷积操作,使高层特征图的数量和低层特征图的数量保持一致;在 3×3 卷积后进行上采样以恢复图像分辨率。将相同数量的低层特征图和高层特征图进行级联融合后输入 TASD 模块,以提取多尺度信息,最终通过一个 1×1 卷积输出分割结果。本算法将 TASD 模块和 RDB 的优势结合在一起,用 RDB 提取高层语义特征,用 TASD 模块提取多尺度信息,有效克服了普通空洞卷积带来的棋盘伪影问题,提升了网络的分割精度。

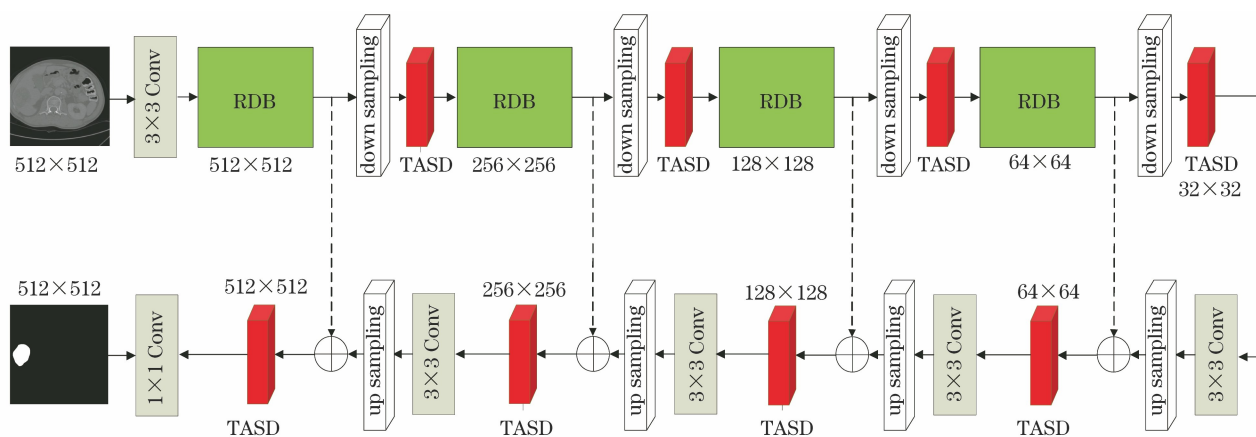


图 6 分割网络的总体结构

Fig. 6 Framework of the segmentation network

3 实验结果与分析

3.1 数据集

实验使用的数据集包括 2012 年至 2017 年在河南省人民医院接受肝癌治疗的 325 例晚期肝细胞肝癌(HCC)患者的 CT 图像,包括 292 名男性和 33 名女性,患者年龄为 18 岁至 81 岁,平均年龄约为 53.23 岁。样本标准:1)病理证实为肝癌且接受手术切除的患者;2)手术前两周在该院接受肝脏多相动态 CT 检查的患者;3)未接受过治疗的患者。

采用多探头行 CT(MDCT)装置进行腹部多相动态对比增强 CT 成像(美国 GE Lightspeed VCT、美国 GE Discovery CT750 HD、荷兰飞利浦

Brilliance16),扫描电压为 120 kV,旋转时间为 0.4, 0.5, 0.75 s;探测器的标准值为 $64 \times 0.625 \text{ mm}$, $128 \times 0.625 \text{ mm}$ 或 $16 \times 0.625 \text{ mm}$,其中,0.625 mm 为每排的宽度;视场为 $(300 \sim 500) \text{ mm} \times (300 \sim 500) \text{ mm}$;图像的分辨率为 512×512 ,切片间距为 5 mm、切片厚度为 5 mm、重建截面厚度为 2 mm。常规非增强 CT 后,静脉注射 1.5 mL/kg 造影剂(德国柏林拜耳, Ultravist 370)经 25~30 s 和 60~70 s 延迟后,以 3.0~3.5 mL/s 的速率用泵式注射器(GE Ulrich CT Plus 150)进行动脉期相位(AP)和门静脉期相位(VP)增强 CT,图像格式保存为 DICOM 格式。

实验选取 AP 图像,由一位超过 12 年读片经验的放射科大夫用开源软件 ITKSNAP3.0 对肿瘤病灶部位进行手动标注,在 AP 模态的 CT 影像上描

绘整个病变的二维感兴趣区域(ROI),结果如图 7 所示。将三维 CT 图像处理为二维切片图像(分辨率

为 512×512),共获取 6146 张图像,取其中 6046 张作为训练集,100 张作为测试集。

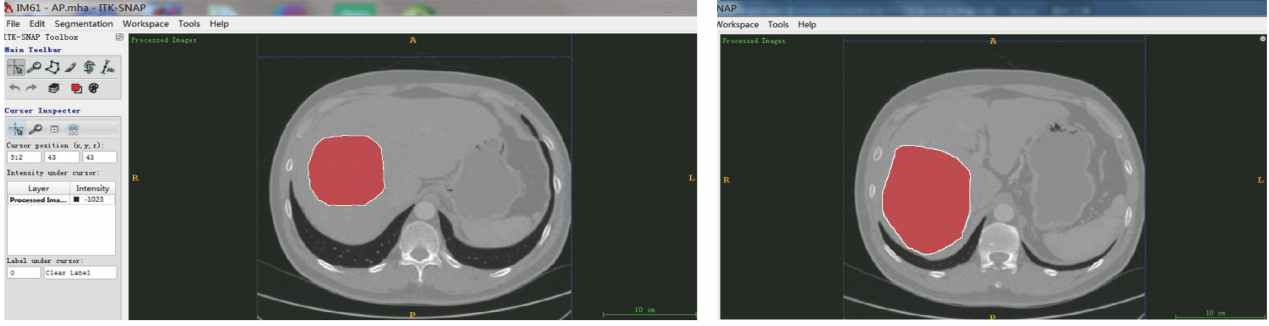


图 7 手动标注的图像数据

Fig. 7 Manually annotated image data

为了加快网络的训练,需对图像进行标准化的预处理,预处理过程中图像的标准化可表示为

$$X_{\text{standardization}} = \frac{X - \mu}{A_{\text{stddev}}}, \quad (7)$$

$$A_{\text{stddev}} = \max\left(\sigma, \frac{1.0}{\sqrt{N}}\right), \quad (8)$$

式中, μ 为图像的均值, σ 为图像的标准方差, N 为图像的像素数。

3.2 分割结果的评价指标

实验用三个经典的评价指标对分割结果进行评价,其中,Dice 相似度系数为分割结果与真值的相似度,像素准确率(PA)为分割正确的像素点占真值的比重,交并比(IOU)为分割结果和真值之间交集和并集的比值,可表示为

$$X_{\text{Dice}} = 2 \times \frac{V_{\text{seg}} \cap V_{\text{gt}}}{V_{\text{seg}} + V_{\text{gt}}}, \quad (9)$$

$$X_{\text{PA}} = \frac{V_{\text{seg}} \cap V_{\text{gt}}}{V_{\text{gt}}}, \quad (10)$$

$$X_{\text{IOU}} = \frac{V_{\text{seg}} \cap V_{\text{gt}}}{V_{\text{seg}} \cup V_{\text{gt}}}, \quad (11)$$

式中, V_{seg} 为分割结果, V_{gt} 为真值,三个评价指标的值越大,表示网络的分割效果就越好。

3.3 网络训练策略

实验环境:机器配置为英特尔的 XEON CPU E5-2630,内存为 64 G,GPU 为英伟达的 GeForce RTX 1080Ti。用 Adam 优化器优化,迭代 300000 次后停止,初始学习率为 0.001,衰减率为 0.995。由于图像的分辨率较高,因此,将批处理数量设置为 1,在编码器中,4 个 RDB 的层数依次设置为 4,6,8,10,网络宽度 $K=12$ 。

为了提高网络的训练效果,防止小样本数据带来的过拟合问题,对原始训练数据集进行了数据增强处理。增强算法包括对原始图像进行任意角度旋转、水平翻转、垂直翻转、缩放、拉伸等,数据增强结果如图 8 所示。其中,图 8(a)为原始图像和手动分割标签,图 8(b)~图 8(h)为对原始图像的增强结果。经过数据增强和扩充后,训练集数据达到了 20×6046 个。

3.4 损失函数

输入的肝脏 CT 图像分辨率为 512×512 ,对肿瘤的尺寸进行统计分析时发现,肿瘤在整个图像中所占的比例比较小,如图 9 所示。其中,占比小于 10%的样本数量超过总样本数量的一半。

网络训练使用的损失函数为焦点损失(FL)^[29],由于训练数据集的图像前景(肿瘤区域)和背景(其他部分)比重不平衡,网络训练结果容易向背景倾斜,导致分割精度下降,因此,将 FL 函数改进为

$$L_{\text{FL}} = \frac{1}{N} \sum_i -$$

$$[\alpha(1-p_i)^\gamma \log(p_i) + (1-\alpha)p_i^\gamma \log(-p_i)], \quad (12)$$

式中, p_i 为经过 Softmax 函数的输出,即目标被判断为前景(肿瘤区域)的概率, $1-p_i$ 为目标被判断为背景的概率, γ 为超参数,固定大小为 2。 α 为调节系数,可调节背景图像所占比重过大带来的影响,背景图像所占比重越大,调节系数 α 的值就越小。原始 FL 函数中的 α 是固定不变的,改进的 FL 函数中参数 α 为肿瘤所占的比重大小,可表示为

$$\alpha = \frac{C}{512 \times 512}, \quad (13)$$

式中, C 为肿瘤的尺寸。

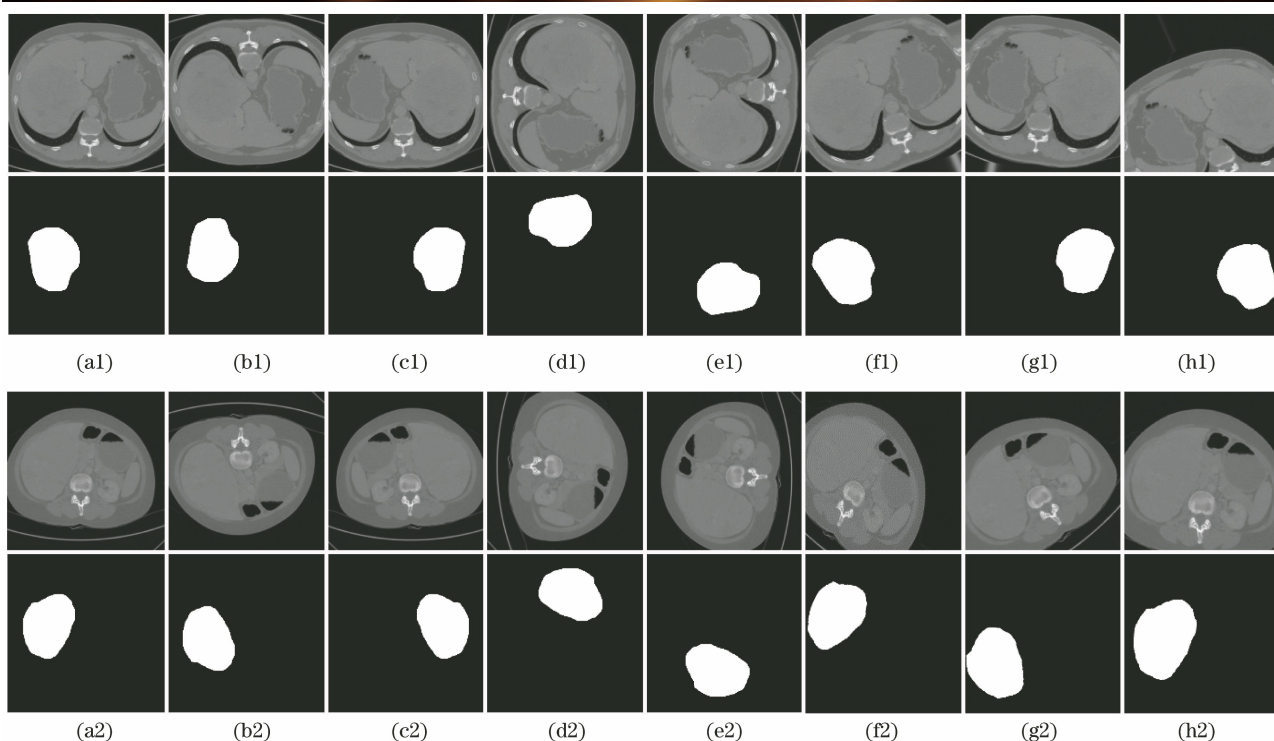


图 8 图像增强的结果。(a)原始图像;(b)上下翻转;(c)左右翻转;(d)顺时针旋转 90°;(e)逆时针旋转 90°;(f)随机缩放旋转 1;(g)随机缩放旋转 2;(h)随机缩放旋转 3

Fig. 8 Result of image enhancement. (a) Original image; (b) flip up and down; (c) flip left and right; (d) clockwise rotate 90°; (e) counterclockwise rotate 90°; (f) random zoom and rotate 1; (g) random zoom and rotate 2; (h) random zoom and rotate 3

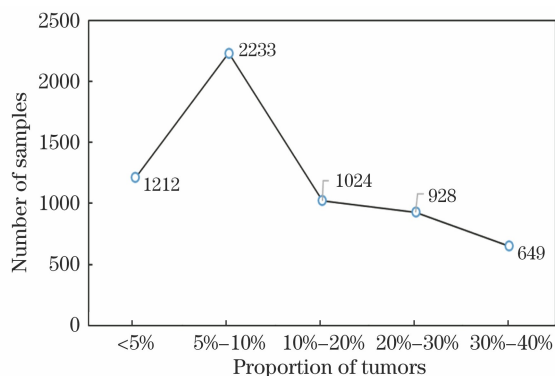


图 9 肿瘤尺寸的统计

Fig. 9 Statistics of the tumor size

3.5 实验结果与分析

3.5.1 本算法与传统分割算法的对比

本算法和传统分割算法的分割结果如表 1 和图 10 所示,可以发现,相比传统分割算法,本算法对不同大小的肿瘤和复杂背景下肝脏病变的分割精度均较高。区域生长算法可根据特定标准合并灰度值相近或类似的像素,为待分割的每个区域定位一个种子像素点或若干个像素点作为生长的起点,然后基于一定的准则对种子像素周围的相似像素点进行识别,并对相似度最高的像素点进行合并,产生分割

表 1 本算法和传统算法的分割结果

Table 1 Segmentation results of our algorithm and traditional algorithms

Algorithm	Dice	IOU	PA
Regional growth	0.5222	0.3603	0.4623
Graph cut	0.5912	0.5034	0.5356
Level set	0.6145	0.5212	0.5629
Ours	0.8026	0.7317	0.7974

结果。从区域增长的过程来看,区域生长算法非常依赖先验知识,即初始像素点的位置。对于一些肝脏 CT 图像,如果图像内部的灰度不均匀或肿瘤边缘的灰度值非常相似,则区域生长算法的分割性能较差。相比区域生长算法,水平集和图割等交互式分割算法的分割效果有所改进,但这类算法对损失函数进行推理的过程中,需要根据先验知识确定肿瘤的位置,以初始化损失函数。水平集理论用于医学图像的分割时依赖图像的低层细节信息,缺乏对图像语义特征的理解。如果图像中出现灰度不均匀或肿瘤边界模糊现象时,分割效果较差;而图割算法必须通过指定肿瘤图像的前景和背景信息进行初始化。这三种算法均为半自动分割算法,需要人工的

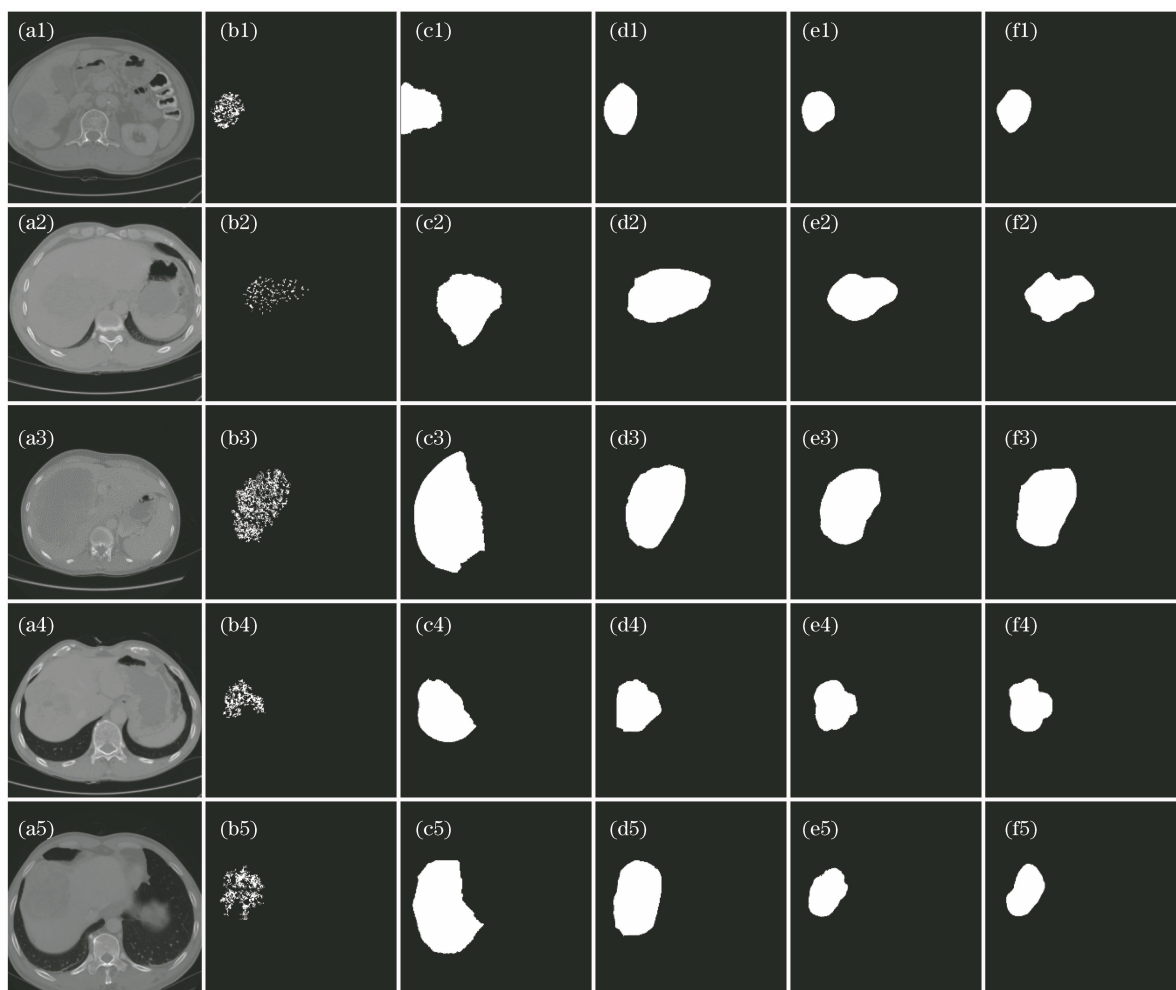


图 10 本算法与传统分割算法的分割结果。(a)原始图像;(b)区域增长;(c)图割分割;(d)水平集分割;
(e)本算法;(f)分割真值

Fig. 10 Segmentation results of our algorithm and traditional segmentation algorithm. (a) Original image;
(b) region growth; (c) graph cut segmentation; (d) level set segmentation; (e) our algorithm; (f) ground truth

交互,而本算法是全自动分割算法。由于 CT 图像中肝脏肿瘤与正常器官的灰度值非常接近,传统阈值分割、分水岭分割、边缘分割等算法的分割效果较差。从表 1 也可以发现,相比传统算法,本算法在 Dice 相似系数、交并比、像素准确率三个指标上均有较大的优势。原因是传统肝脏肿瘤分割算法主要是将灰度、纹理、形状先验、梯度等图像特征应用到各种分割模型算法中,大多需要人工提取特征或先验信息。

3.5.2 本算法与深度学习分割算法的对比

将本算法与一些流行的深度学习分割模型进行对比,包括 UNet^[9]、SegNet^[10]、DeepLabv3^[11]、FC-DenseNet^[12],结果如表 2 和图 11 所示。其中,UNet 是经典的医学图像分割网络,SegNet 是经典的语义分割网络,和 FCN 类似,DeepLabv3 网络中使用了能提取多尺度信息的空洞卷积,提高了分割

精度,FC-DenseNet 则在 DenseNet 的基础上进行改进,在小样本数据集上取得了很好的分割性能。可以发现,在一些肿瘤图像灰度值和周边图像灰度值接近以及肿瘤尺寸较大或较小的情况下,本算法均能取得最好的分割结果,如图 11(a2)和图 11(a5)中的样本。这验证了本算法中 TASD 模块的有效性及其多尺度信息对于分割的重要性。此外,DeepLabv3 算法的分割效果也较好,原因是该模型采用了金字塔池化的空洞卷积,融合了多尺度图像信息。但 DeepLabv3 算法的分割结果中,对于边缘的分割还不够精准。FC-DenseNet 算法的参数较少,对小样本的医学图像分割效果较好。UNet 和 SegNet 算法只是简单通过跳层连接将高低层特征进行融合,没有充分利用上下文信息,也没有提取丰富的多尺度信息,分割效果较差。从模型的规模和运行速度来看,相比 UNet 算法,本算法的网络层数

表 2 本算法和深度学习分割算法的分割结果

Table 2 Segmentation results of our algorithm and deep learning segmentation algorithm

Algorithm	Dice	IOU	PA	Network parameters /M	Running speed /s
UNet ^[9]	0.6562	0.5621	0.7027	7.76	0.5028
SegNet ^[10]	0.6622	0.5528	0.7102	1.425	0.5926
DeepLabv3 ^[11]	0.6929	0.5943	0.7368	115	0.6584
FC-DenseNet103 ^[12]	0.7736	0.7065	0.7682	9.26	0.7969
Ours	0.8026	0.7317	0.7974	2.21	0.5638

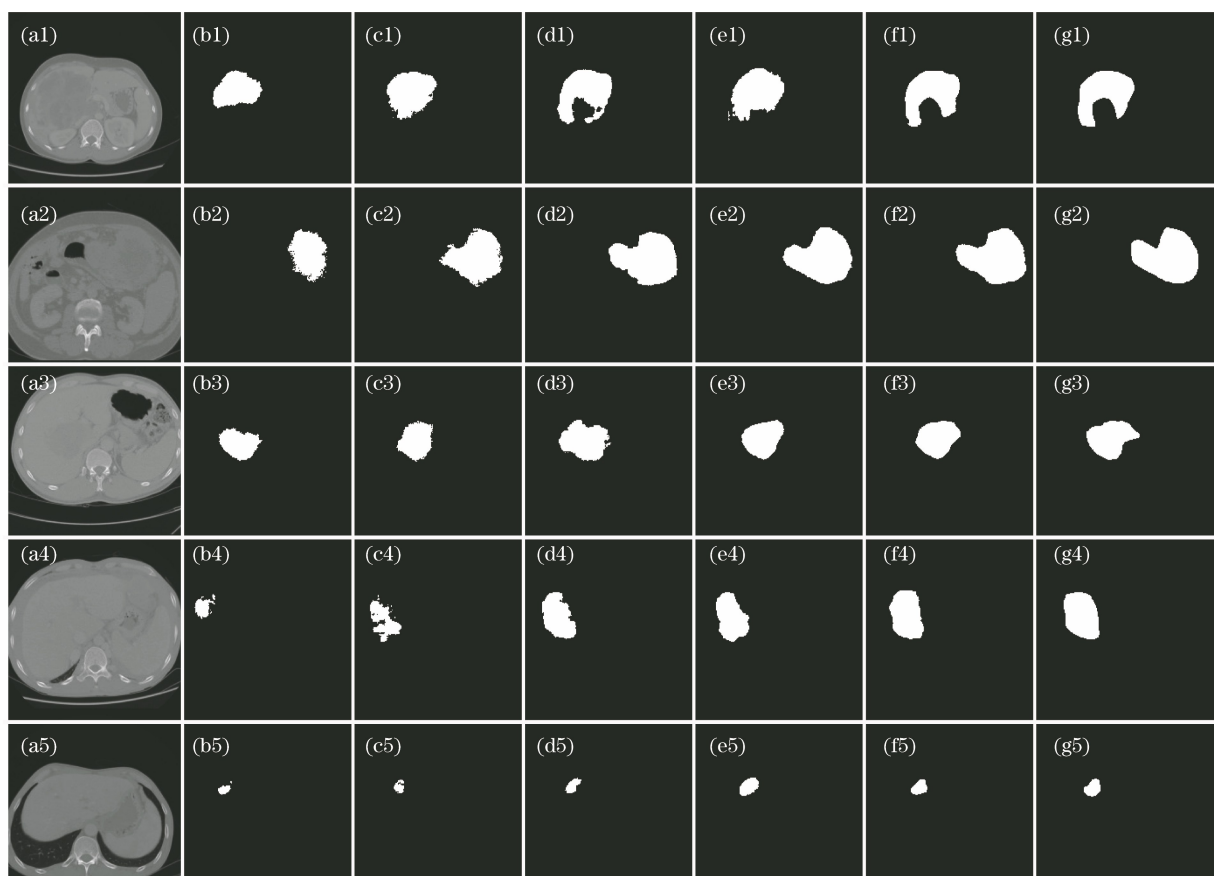


图 11 本算法与深度学习分割算法的分割结果。(a)原始图像;(b) UNet; (c) SegNet; (d) DeepLabv3; (e) FC-DenseNet; (f)本算法;(g)分割真值

Fig. 11 Segmentation results of our algorithm and deep learning segmentation algorithm. (a) Original image; (b) UNet; (c) SegNet; (d) DeepLabv3; (e) FC-DenseNet; (f) our algorithm; (g) ground truth

较多,但在 RDB 和 TASD 模块中控制了网络的宽度,RDB 中的参数 $K=12$,从而减小了网络的参数量,但运行速度略慢于 UNet 算法。相比其他三种算法,本算法在参数数量和运行速度方面的表现均较好。从表 2 可以发现,相比 UNet、SegNet 和 DeepLabv3 算法,本算法在 Dice 相似度系数、交并比和像素准确率三个指标上均有较大优势,验证了本算法的有效性。

3.5.3 网络分模块对分割性能的影响

为了测试本算法中 FL 函数、TASD 模块及

RDB 对分割性能的影响,用交叉熵损失(CEL)函数取代 FL 函数,CEL 函数可表示为

$$L_{CEL} = \frac{1}{N} \sum_i L_{CEL_i} = \frac{1}{N} \sum_i -[y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)] \quad (14)$$

在分割网络中分别去掉 TASD 模块和 RDB,得到的实验结果如图 12 和表 3 所示。可以发现,未添加 TASD 模块时,网络的分割边缘出现了明显的模糊伪影;添加 TASD 模块后,分割性能在 Dice 相似度系数、交并比和像素准确率三个指标上均有一定的

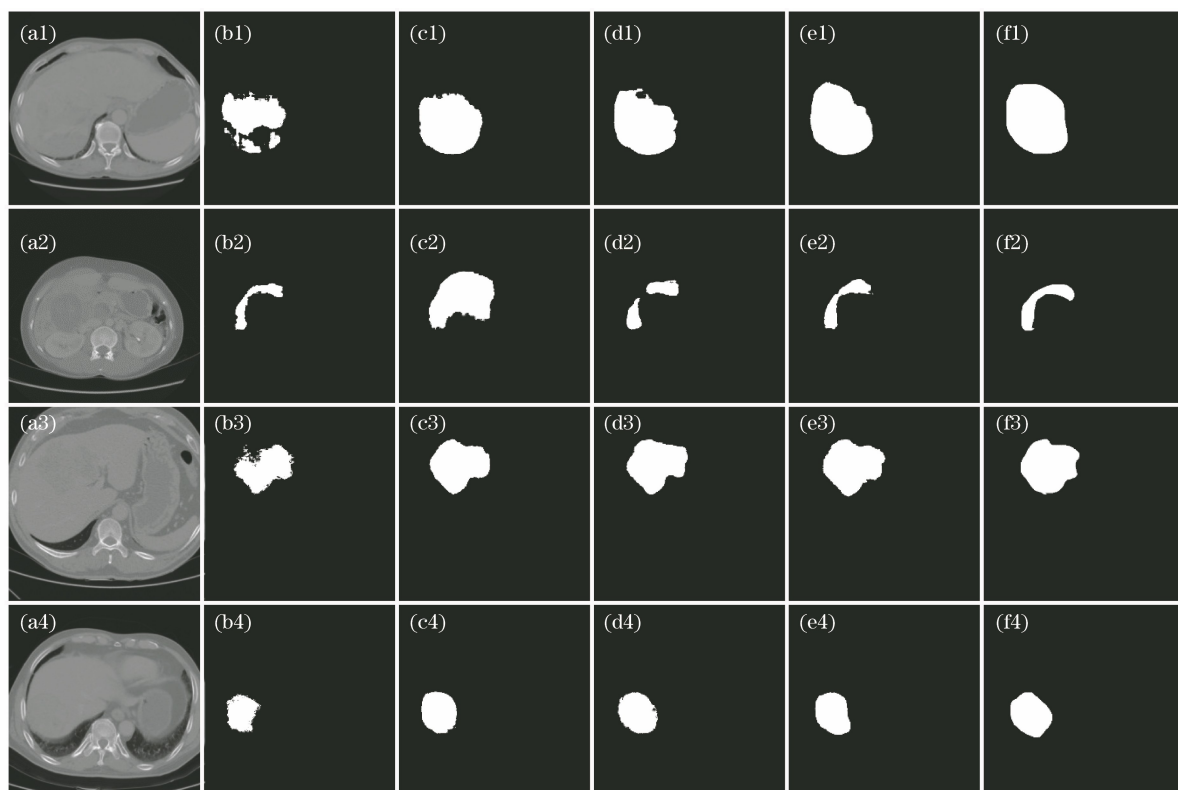


图 12 不同模块对分割性能的影响。(a)原始图像;(b)CEL 函数;(c)去除 RDB;(d)去除 T ASD 模块;(e)本算法;(f)分割真值

Fig. 12 Effect of different modules on segmentation performance. (a) Original image; (b) CEL function; (c) remove RDB; (d) remove T ASD module; (e) our algorithm; (f) ground truth

表 3 不同模块对分割性能的影响

Table 3 Effect of different modules on segmentation performance

Module				Dice	IOU	PA
CEL	FL	RDB	T ASD			
+		+	+	0.7125	0.6613	0.7234
	+	+		0.7422	0.6927	0.7549
	+		+	0.7386	0.6822	0.7458
	+	+	+	0.8026	0.7317	0.7974

提升。未添加 RDB 时,网络对肿瘤灰度和周围组织接近的情况,分割结果不够精准;添加 RDB 后,分割性能在 Dice 相似度系数、交并比和像素准确率三个指标上也有一定的提升。用 CEL 函数代替 FL 函数后,网络的分割性能有明显下降。这表明本算法中动态调整前景和背景样本不均衡的调节系数 α 对 FL 函数的改进是有效的。

4 结 论

为了实现对 CT 图像中肝脏肿瘤的自动分割,提出了一种基于 RDB 和 T ASD 模块的分割算法。首先,采用编码-解码结构,在编码器中设计了 RDB,在

输出端用加法取代了级联操作,保证特征重用的同时解决了梯度消失问题。然后,在 RDB 之间用步长为 2 的卷积操作取代了池化下采样,保留了图像更多的细节信息;在解码器模块中,为了能精准恢复出图像的细节,通过跳层连接融合高、低层特征。最后,在整个编码器-解码器网络上,用 T ASD 模块提取丰富的多尺度信息,有效消除了普通空洞卷积带来的棋盘伪影现象,显著提升了网络的分割精度。为了减少计算量,在解码器网络中用 3×3 卷积操作代替 RDB。考虑到肿瘤尺寸在整个图像中所占比例过低的问题,对 FL 函数进行了改进,用肿瘤面积占整个图像面积的比重作为调节系数,解决了前景和背景图像不均衡造成的分割精度下降问题。

参 考 文 献

- [1] 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.
- [2] Lu X Q, Wu J S, Ren X Y, et al. The study and application of the improved region growing algorithm

- for liver segmentation[J]. *Optik*, 2014, 125(9): 2142-2147.
- [3] 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*, 2012, 12(3): 464-473.
- [4] Li B N, Chui C K, Chang S, et al. A new unified level set method for semi-automatic liver tumor segmentation on contrast-enhanced CT images [J]. *Expert Systems with Applications*, 2012, 39(10): 9661-9668.
- [5] Zareei A, Karimi A. Liver segmentation with new supervised method to create initial curve for active contour [J]. *Computers in Biology and Medicine*, 2016, 75: 139-150.
- [6] Shimizu A, Narihira T, Furukawa D, et al. Ensemble segmentation using AdaBoost with application to liver lesion extraction from a CT volume[EB/OL]. [2021-02-25]. https://www.researchgate.net/publication/28359593_Ensemble_segmentation_using_AdaBoost_with_application_to_liver_lesion_extraction_from_a_CT_volume.
- [7] Conze P H, Noblet V, Rousseau F, et al. Random forests on hierarchical multi-scale supervoxels for liver tumor segmentation in dynamic contrast-enhanced CT scans[C]//2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), April 13-16, 2016, Prague, Czech Republic. New York: IEEE Press, 2016: 416-419.
- [8] Vorontsov E, Abi-Jaoudeh N, Kadoury S. Metastatic liver tumor segmentation using texture-based omnidirectional deformable surface models[M]//Yoshida H, Näppi J J, Saini S. *Abdominal imaging. Computational and clinical applications. Lecture notes in computer science*. Cham: Springer, 2014, 8676: 74-83.
- [9] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[EB/OL]. (2015-05-18)[2021-02-25]. <https://arxiv.org/abs/1505.04597>.
- [10] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [11] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-06-17)[2021-02-25]. <https://arxiv.org/abs/1706.05587>.
- [12] Jégou S, Drozdal M, Vazquez D, et al. The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1175-1183.
- [13] Christ P F, Elshaer M E A, Ettliger F, et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields[M]//Ourselin S, Joskowicz L, Sabuncu M R, et al. *Medical image computing and computer-assisted intervention-MICCAI 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9901: 415-423.
- [14] Chlebus G, Meine H, Moltz J H, et al. Neureal network-based automatic liver tumor segmentation with random forest-based candidate filtering [EB/OL]. (2017-06-02)[2021-02-25]. <https://arxiv.org/abs/1706.00842>.
- [15] Han X. Automatic liver lesion segmentation using a deep convolutional neural network method[EB/OL]. (2017-05-24)[2021-02-25]. <https://arxiv.org/abs/1704.07239>.
- [16] 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.
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.
- [17] 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.
- [18] Li X M, Chen H, Qi X J, 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.
- [19] Gu Z, Cheng J, Fu H, et al. CE-Net: context encoder network for 2D medical image segmentation[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(10): 2281-2292.
- [20] Ai L M, Li T D, Liao F Y, et al. Magnetic resonance brain tumor image segmentation based on attention U-Net[J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141030.
艾玲梅, 李天东, 廖福元, 等. 基于注意力 U-Net 的脑肿瘤磁共振图像分割[J]. *激光与光电子学进展*, 2020, 57(14): 141030.

- [21] Chen Y, Wang K, Liao X, et al. Channel-UNet: a spatial channel-wise convolutional neural network for liver and tumors segmentation [J]. *Frontiers in Genetics*, 2019, 10: 1110.
- [22] Zhang W X, Zhu Z C, Zhang Y H, et al. Cell image segmentation method based on residual block and attention mechanism [J]. *Acta Optica Sinica*, 2020, 40(17): 1710001.
张文秀, 朱振才, 张永合, 等. 基于残差块和注意力机制的细胞图像分割方法 [J]. *光学学报*, 2020, 40(17): 1710001.
- [23] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 833-851.
- [24] Zheng T Y, Tang C, Lei Z K. Multi-scale retinal vessel segmentation based on fully convolutional neural network [J]. *Acta Optica Sinica*, 2019, 39(2): 0211002.
郑婷月, 唐晨, 雷振坤. 基于全卷积神经网络的多尺度视网膜血管分割 [J]. *光学学报*, 2019, 39(2): 0211002.
- [25] Wu C Y, Yi B S, Zhang Y G, et al. Retinal vessel image segmentation based on improved convolutional neural network [J]. *Acta Optica Sinica*, 2018, 38(11): 1111004.
吴晨玥, 易本顺, 章云港, 等. 基于改进卷积神经网络的视网膜血管图像分割 [J]. *光学学报*, 2018, 38(11): 1111004.
- [26] Feng B W, Lü X Q, Gu Y, et al. Three-dimensional parallel convolution neural network brain tumor segmentation based on dilated convolution [J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141009.
冯博文, 吕晓琪, 谷宇, 等. 基于空洞卷积的三维并行卷积神经网络脑肿瘤分割 [J]. *激光与光电子学进展*, 2020, 57(14): 141009.
- [27] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [28] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [29] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318-327.