

基于级联DDR-UNet++的肝脏肿瘤图像分割方法

扈蕴琨, 王晓艳, 王秀娟

山东第一医科大学(山东省医学科学院)医学信息与人工智能学院, 山东 泰安 271000

【摘要】目的:探讨并解决传统U-Net算法在肝脏肿瘤分割中,肝脏及肿瘤上下文信息缺乏、肿瘤形态差异性大导致的分割精度不足问题。**方法:**提出了一种结合空洞卷积和残差模块的级联肝脏肿瘤分割算法DDR-UNet++。首先,利用LiTS-2017数据集中的CT图像,通过窗位窗宽调整、直方图均衡化和高斯滤波进行预处理,减少噪声并平滑边缘。然后,采用级联肝脏分割模型,增强肝脏区域占比,减轻周围组织干扰,并解决数据不平衡问题。肝脏肿瘤分割模型通过引入可变形空洞卷积和残差网络,扩展感受野,提升特征提取能力。**结果:**DDR-UNet++在LiTS-2017数据集上的Dice相似系数、相对体积差异和Jaccard指数相比于U-Net模型分别提升了4.7%、1.7%和8.5%,有助于克服传统肿瘤分割中的低效性和低准确性,提高早期肿瘤发现率和患者生存率,减轻医生负担。**结论:**该方法通过改进模型结构与分割策略,在一定程度上改善了特征提取能力不强的问题,有效提升了肝脏肿瘤分割的精度和鲁棒性,为临床辅助诊断提供可靠的技术参考。

【关键词】DDR-UNet++; U-Net; 残差模块; 空洞卷积; 肝脏肿瘤分割

【中图分类号】R318; R735.7

【文献标志码】A

【文章编号】1005-202X(2025)07-0901-10

Liver tumor image segmentation method based on cascaded DDR-UNet++

HU Yunkun, WANG Xiaoyan, WANG Xiujuan

College of Medical Information and Artificial Intelligence, Shandong First Medical University & Shandong Academy of Medical Sciences, Tai'an 271000, China

Abstract: Objective To explore and address the issue of insufficient segmentation accuracy in liver tumor segmentation using the traditional U-Net algorithm, which is caused by the lack of contextual information for both the liver and tumor, as well as the large morphological variability of tumors. **Methods** A cascaded liver tumor segmentation algorithm, DDR-UNet++, which integrated dilated convolutions and residual modules was proposed. Firstly, CT images from the LiTS-2017 dataset were preprocessed through window width/level adjustment, histogram equalization and Gaussian filtering to reduce noise and smooth edges. Then, a cascaded liver segmentation model was employed to enhance the liver region proportion, mitigate interference from surrounding tissues and address data imbalance issue. For liver tumor segmentation, deformable dilated convolutions and residual networks were introduced to expand the receptive field and improve feature extraction capability. **Results** DDR-UNet++ outperformed the traditional U-Net on the LiTS-2017 dataset, achieving improvements of 4.7%, 1.7%, and 8.5% in Dice similarity coefficient, relative volume difference, and Jaccard index, respectively. These enhancements contribute to overcoming the inefficiency and low accuracy issues in conventional tumor segmentation, thereby improving early tumor detection rates, enhancing patient survival outcomes, and alleviating the diagnostic burden on clinicians. **Conclusion** The proposed method improves the feature extraction capability to some extent by enhancing the model structure and segmentation strategy, effectively increases the accuracy and robustness of liver tumor segmentation, and provides a reliable technical reference for clinical auxiliary diagnosis.

Keywords: DDR-UNet++; U-Net; residual module; dilated convolution; liver tumor segmentation

【收稿日期】2025-01-25

【基金项目】山东省自然科学基金(ZR2020MF156); 国家级大学生创新创业项目(202410439008); 山东第一医科大学校级教育教学改革研究项目(XM2024045)

【作者简介】扈蕴琨, 研究方向: 医学图像分割、深度学习在医学影像中的应用、医学人工智能, E-mail: 2684489216@qq.com

【通信作者】王晓艳, 博士, 教授, 研究方向: 医学物理学、医学图像处理, E-mail: xywjxc@126.com; 王秀娟, 硕士, 副教授, 研究方向: 医学图像处理, E-mail: 22879867@qq.com

前言

肝脏是人体最大的消化腺^[1],也是新陈代谢的核心器官,参与500余种生化反应。肝脏恶性肿瘤是全球范围内严重威胁人类健康的疾病之一^[2],其在我国的发病率呈逐年上升趋势。在传统的肝脏肿瘤诊断过程中,医生通常依赖手动勾画病灶区域以制定诊

疗方案。然而,该方法存在较强的主观性,并依赖于操作者的经验,导致分割结果的再现性较差。此外,随着临床数据量的增加,手动分割的效率问题愈发突出,进一步增加了误诊和漏诊的风险。因此,如何利用机器学习方式客观且高效地分割肝脏肿瘤并用于临床诊断,已成为医学图像分析领域的重要研究方向。

近年来,随着人工智能技术的不断发展,计算机视觉分割广泛应用于医学图像分割领域。然而,由于多尺度变化以及腹部CT图像边界模糊等问题^[3],语义分割模型在目标定位和分割方面仍然面临获得精确分割的重大挑战。以卷积神经网络(CNN)^[4]为代表的深度学习算法在图像语义分割中取得了显著的性能提升,德国图宾根大学(University of Tübingen)的计算机科学家Olaf Ronneberger于2015年提出U-Net^[5]模型架构并首次用于细胞核的分割。U-Net采用了对称的编码器-解码器结构。编码器部分用于捕捉图像特征,逐步减小空间分辨率;解码器部分则执行上采样操作,逐步将特征图的空间分辨率还原到原始图像的尺寸,这保证了输入输出特征图尺寸的一致性且有助于在解码器中引入来自底层的细节信息,提高了网络的上下文处理能力,成为当前医学数据集分割领域的主流基准模型。在此之后,许多研究人员对U-Net进行不同程度的改进,都取得了比较好的分割效果。Milletari等^[6]在U-Net的基础上引入残差模块并提出ResUNet模型,这使得输入特征可以绕过卷积层而直接与输出特征相聚合,这种跳跃连接能让网络学习到恒等映射,从而能够有效缓解深层网络训练中的梯度消失和退化问题。但仅在U-Net的基础上引入残差模块,该方法未能显著提高对肿瘤复杂边界和小尺寸目标区域的分割能力,特征提取的精度仍有待优化。Yu等^[7]基于U-Net提出一种融合空洞卷积的分割模型AtrousNet,该模型通过设置一定的膨胀率来逐步提高空洞卷积的感受野,从而在不提高参数数量的前提下使模型得到更大范围的上下文信息,有利于捕获长距离依赖关系。但膨胀率较大时,卷积核中的点分布较为稀疏,可能导致梯度传播变得困难,且模型对于细小病灶的分割仍不够精准。Jha等^[8]提出一种在U-Net基础上同时引入空洞卷积和残差网络的ResUNet++模型,空洞卷积使得网络在保持特征图分辨率的同时捕获更大范围的上下文信息,对分割边界和细粒度结构有更好的表现,同时有效缓解了梯度消失问题并使得网络能够更高效地融合低级和高级特征。Xia等^[9]基于DeepLabV3利用生成器和两个判别器来优化分割结果并提出一种将语义分割损失、内容损失和对抗损

失相结合的混合加权损失函数。

综上所述,上述模型虽然为医学图像分割领域提供了较为可观的分割效果,然而,这些模型在特定数据下仍面临特征提取困难、计算复杂度高、感受野有限以及损失函数单一等问题。文章受到上述研究方法的启发,提出一种级联肝脏和肝脏肿瘤分割模型(DDR-UNet++),模型基于U-Net模型进行改进,在编码器中引入残差模块并将编码器中的普通卷积替换为可扩充的空洞卷积,提高卷积核的感受野并预防梯度消失问题,同时引入GAN生成对抗网络让分割结果不断逼近真实标签^[10]。本文方法从3个方面对现有模型存在的不足进行改进:(1)为模型的编码器设计残差路径。将U-Net的卷积连接方式改为残差连接以强化特征传递,优化网络训练过程中的梯度流动,避免梯度消失问题。(2)对卷积模块的改进。将U-Net编码器阶段的普通卷积改为具有不同感受野的多尺度卷积,增大感受野而不提高参数量,从而能够捕获各种大小和位置的肿瘤特征,提高分割效率。(3)结合GAN生成对抗网络思想提出一种混合损失函数。使生成器的损失函数最小,不断提高生成器的分割精度;同时使判别器的损失函数最大,使判别器能不断区分出生成特征图与金标准的差异。在生成器损失函数中引入Tversky损失^[11]和二元交叉熵(BCE)损失^[12],形成混合损失函数 G_loss_1 ,引入 G_loss_2 为距离约束函数并通过动态权重调整(α)实现两者的自适应优化。

1 资料与方法

1.1 数据集资料

本文使用2017年LiTS-2017^[13]数据集对所提模型进行训练与评估。该数据集原始包含131个训练图像和70个测试图像,每个训练图像由经验丰富的医生对肝脏和肿瘤区域进行标注,作为分割的金标准。由于部分图像存在质量较差或标签缺失等问题,经数据清洗后实际纳入130例有效训练图像。为保证训练的稳定性与评估的公正性,本文将这130例图像按7:1:2的比例划分为训练集91例、验证集13例和测试集26例。该数据集来自多个临床机构和扫描设备,图像间差异较大,有利于模型的泛化能力提升。

1.2 图像预处理

由于原始图像边界值不明确,肝脏肿瘤分布不明显,直接分割会导致分割精度较低。为提升系统性能,文章对训练图像进行了数据增强,采用RandomHorizontalFlip($p=0.5$)和RandomVerticalFlip($p=0.5$)操作,分别以50%的概率对图像进行水平和

垂直方向的翻转,以增加数据的多样性:采用 RandomRotation(degrees=(-15, 15))在 $[-15^{\circ}, 15^{\circ}]$ 范围内随机旋转图像,以增强模型对不同角度下肿瘤形态的鲁棒性;使用水平和垂直翻转操作进行镜像变换,与随机翻转类似,以提升模型对左右翻转情况下的泛化能力。Perez 等^[14]的研究表明,在图像分类任务中,采用包括裁剪、旋转、翻转在内的传统数据增强方法,可在 ImageNet 子集上将模型的分类准确率提升约 7%,尤其在训练数据量受限的情况下,增强对模型性能提升尤为明显。

此外,为突出肝脏组织用于肿瘤分割,进行了窗宽窗位调整与直方图均衡化预处理,去除无效器官

并增强肝脏对比度。在 $[-100, 400]$ (Hounsfield 单位)范围内进行开窗,随后归一化至 $[-1, 1]$,并减去训练集均值以实现零均值化,从而提高训练稳定性。为进一步优化图像质量,引入高斯滤波器(线性滤波器)进行平滑处理,保留关键特征的同时减少噪声。在图像处理中,每个像素的值会受到由高斯函数决定的加权平均影响。高斯函数通过对滤波窗口中心附近的像素赋予更高权重来实现图像平滑处理。

图 1 为预处理对比结果。可以看出,经过窗位窗宽调整的图像相比于原始 CT 图器官间的对比度更加明显;在此基础上高斯滤波后,肝脏边缘更加光滑,有助于提升后续肝脏分割的稳定性。

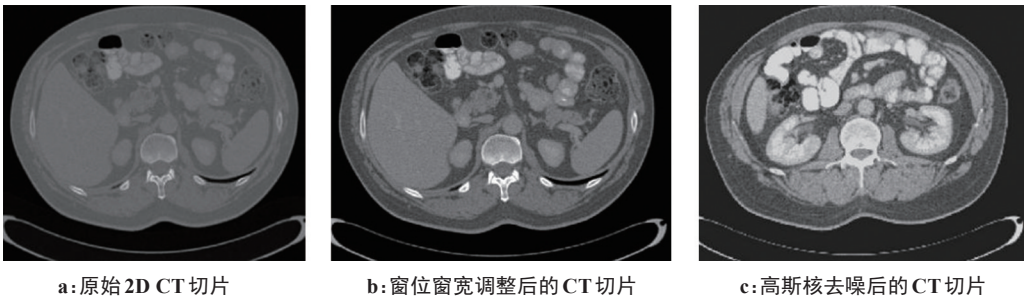


图 1 预处理对比结果

Figure 1 Comparison of preprocessing results

1.3 级联肿瘤分割方法

文章所提出的 DDR-UNet++网络主要由肝脏分

割和肝脏肿瘤分割两部分组成,肝脏肿瘤总体分割流程如图 2 所示。

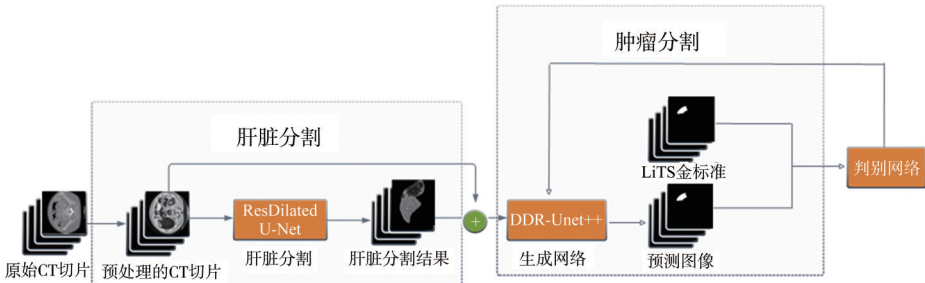


图 2 面向肝脏肿瘤分割的 DDR-UNet++网络架构图

Figure 2 DDR-UNet++ architecture for liver tumor segmentation

1.3.1 肝脏分割方法 在残差结构被提出之前,机器学习中的普遍观点是通过增加卷积层数来提高模型特征提取的精度,从而达到较好的分割效果。如 Simonyan 等^[15]和 Szegedy 等^[16]相继提出 VGG-Net 模型和 GoogLeNet 模型,通过堆叠多个不同尺寸的卷积核和池化操作,同时保持计算效率,从而提高网络对不同尺度特征的学习能力。虽然 VGG-Net、GoogLeNet 模型均在原始模型基础上进行了改良,但其本质依然是增加神经网络的深度,因而会发生特征信息丢失、梯度消失等问题。在 2017 年,Christ

等^[17]提出一种基于全卷积神经网络的级联肝脏肿瘤分割系统,使用两级 FCN 分割网络^[18],第一阶段的 FCN 专注于从腹部 CT 图像中分割肝脏,生成肝脏的 ROI 区域,有效富集肝脏区域信息,减少其他器官的影响;第二阶段的 FCN 利用第一阶段生成的肝脏分割结果,进一步分割肝脏区域内的病灶。但该方法级联了两次 U-Net 模型分割肝脏,意味着需要训练和推理两个完整的深度神经网络,这会导致计算资源需求的显著增加和梯度消失问题且级联两次模型会导致高级空间特征的消失。因而文章在肝脏分割模

块中改为使用单次残差网络分割肝脏,通过残差模块^[19]的跳跃短连接使得模型对输入特征的表达更稳定,通过可变形卷积提高网络的特征提取能力而不需要级联两层网络。空洞残差模块示意图如图3所示,该模块由标准卷积、空洞卷积和残差连接组成,以增强特征提取能力并扩大感受野。首先,输入特征图经过归一化(Batch Normalization, BN),并通过ReLU激活函数增强非线性表达能力。随后,特征图进入两个3×3空洞卷积层,其膨胀率(dilation rate)分别设为2和4,以扩大感受野并保留更多上下文信息。此外,为了匹配输入输出通道,残差路径中包含一个1×1卷积,用于通道调整。最终,经过逐元素相加(element-wise addition)融合原始输入与空洞卷积特征,并再次通过ReLU进行非线性变换,确保梯度稳定并提升网络性能。图4展示了残差模块的处理流程。

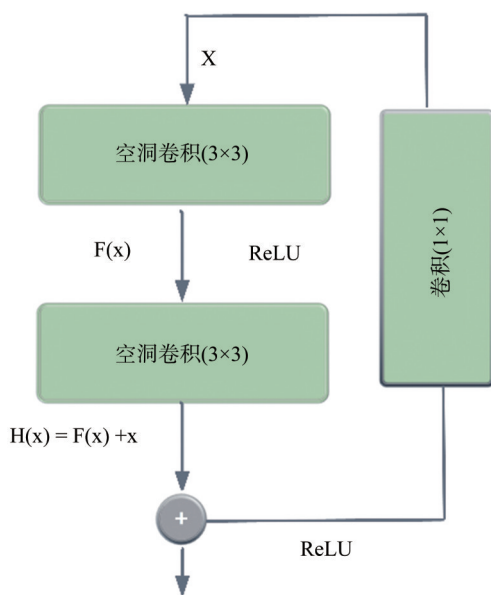


图3 空洞残差模块
Figure 3 Dilated residual module

残差模块的核心思想在于引入跳跃连接,引导网络显式学习输入与输出之间的残差(即差值),并将该残差直接与原始输入相加,从而形成残差映射。其基本形式可表示为:

$$H(x) = x + F(x) \quad (1)$$

底层的映射表示为 $H(x)$,其中, x 是输入数据,而 $F(x)$ 代表网络隐藏层的输出。这种设计的目的是让网络更容易学习到恒等映射(即输入等于输出的映射),从而使底层的信息能够更加顺畅地传递到高层。

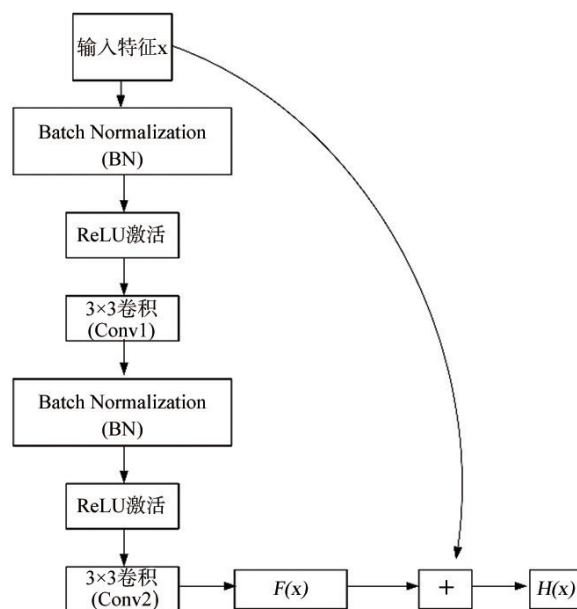
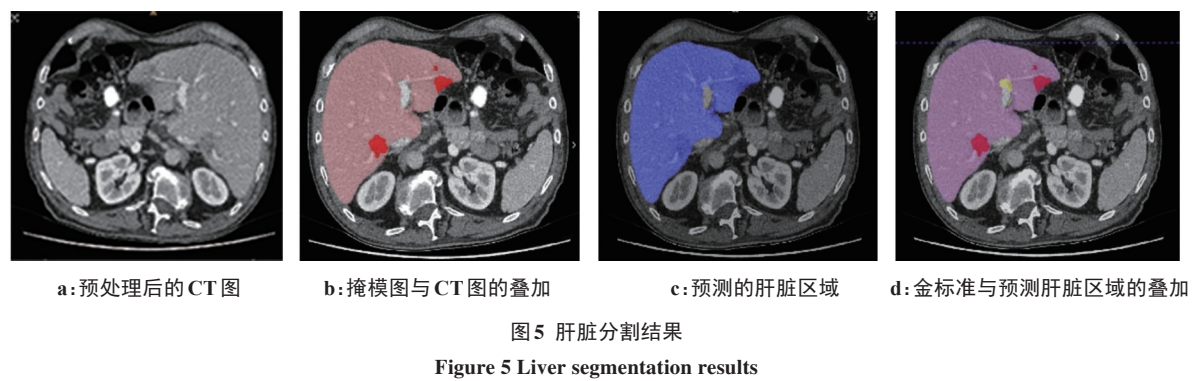


图4 残差模块处理流程
Figure 4 Residual module processing flowchart

为了进一步提高肝脏边缘特征的提取能力,在肝脏分割模块中引入可变形卷积。这种空洞残差模块的设计充分利用了可变形卷积和残差模块的优势,同时扩展了感受野,有效处理了信息丢失和浪费的问题。

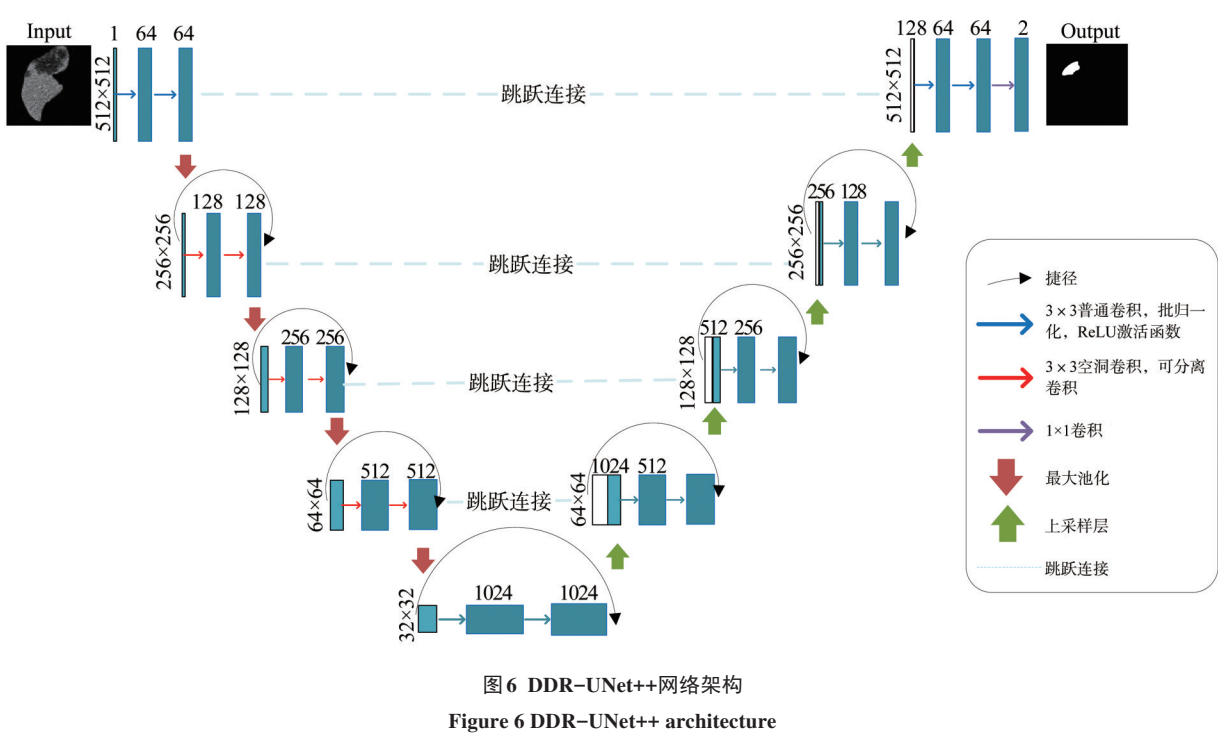
具体操作是在残差网络的编码阶段,将残差模块中的卷积操作替换为可变形卷积;经过反复的实验对比,最终决定在残差网络的编码层使用2个空洞率为3的卷积核进行卷积操作,而在解码层使用两组普通卷积进行卷积操作,通过融合解码层映射和对应的编码层特征,来弥补由于下采样引起的信息损失。在最终输出层,分割网络采用1×1卷积进行像素级分类,实现肝脏分割任务。图5展示了肝脏分割的结果;其中图5a是经过预处理后的CT图;图5b是LiTS-2017数据集提供的掩模图与CT图的叠加,暗红色是肝脏的掩模图,亮红色是肝脏上的肿瘤掩模;图5c是肝脏分割模块预测的肝脏区域;图5d是金标准与预测肝脏区域的叠加,黄色部分是两者的差异部分,可见利用空洞残差模块分割肝脏取得了比较好的分割效果。

1.3.2 肝脏肿瘤分割方法 图6给出了DDR-UNet++肿瘤分割模型的架构和内部有关模块。该网络是一个左右对称的“U”型构造,包括编码器(左)和解码器(右)。其中编码器主要是对图像特征进行提取,解码器则逐级恢复图像分辨率。具体来说,编码器由4组卷积模块组成,其中每组包含两个3×3标准卷积,后接ReLU激活函数和BN处理,以加速训练并稳定梯度分布。其中前3组卷积模块采用空洞卷积



(dilation=3),以扩展感受野,提高对不同尺度肿瘤特征的捕捉能力。同时,每个编码块后均通过2×2最大池化(Max Pooling)进行特征降维,以减少计算复杂度。解码器部分由4组卷积模块组成,与编码器对称,但不再使用空洞卷积,而是采用标准3×3卷积,以在感受野扩展和计算效率之间取得平衡。在上采样过程中,解码模块采用转置卷积(Transposed

Convolution)恢复空间分辨率,并结合1×1卷积进行特征融合,以减少通道维度冗余信息,增强高层语义表达能力。跳跃连接的引入受ResNet架构启发,使得低层特征能够更顺畅地传递至高层,缓解梯度消失问题,并加速模型收敛。此外,残差连接的设计在每个卷积模块内部实现,使得网络在深度增加的同时保持良好的梯度流,确保分割结果更加稳定。



另一方面,为了提高模型的分割精度和泛化能力,本研究在损失函数中结合了Dice损失和交叉熵损失,同时引入GAN判别器损失进行额外监督,以提升模型对小肿瘤区域的识别能力。最终,该模型在LiTS-2017数据集上取得了优于传统方法的分割性能,并表现出较高的临床应用价值。

1.3.3 判别网络方法 文章在肝脏肿瘤分割后引入判别网络模块,用于在训练过程中将分割好的肝脏肿

瘤与CodaLab组织提供的LiTS-2017数据集金标准做对比,使模型能够学习到它们之间的差距并做进一步的调整优化。

(1)判别网络。判别网络由4个特征提取模块和1个特征平滑模块组成,模块结构如图7所示。由DDR-UNet++网络分割出的肝脏肿瘤和LiTS-2017金标准共同构成输入层一起输入卷积神经网络,其中特征提取模块进行一次步长为2的4×4卷积和线性

修正单元(ReLU)操作,并在最后进行一次神经网络的正则化(Dropout),用来修正过拟合问题。模块二由一次步长为2的4×4卷积和ReLU构成,最后经过一个平均池化器将其输出。输出结果是一个[0,1]之

间的数,越接近1则表明经过DDR-UNet++网络模型生成的肝脏肿瘤分割结果与LiTS-2017金标准越接近,反之越接近于0则表明与其越偏离。

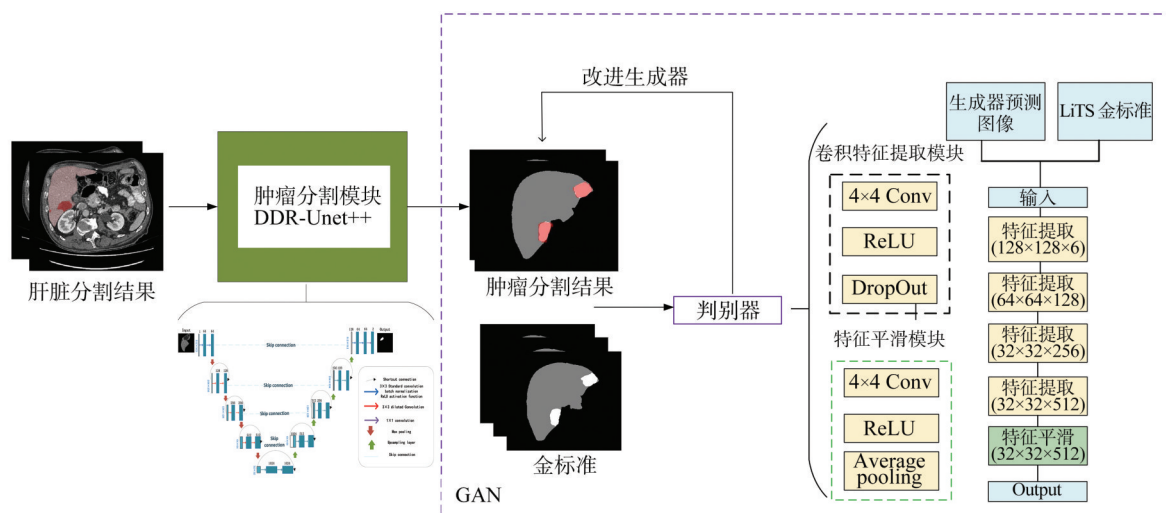


图7 生成对抗网络架构

Figure 7 Generative adversarial network architecture

在判别网络的训练过程中,笔者采用对抗损失函数来优化生成器(DDR-UNet++ 分割模型)和判别器之间的博弈关系。具体来说,判别器的目标是尽可能区分出真实的 LiTS-2017 金标准和生成的分割结果,而生成器则试图欺骗判别器,使其无法正确分辨两者之间的差异。这样的对抗训练方式能够促使生成器不断改进,使得其分割结果更加逼近真实的肿瘤分割金标准,提高模型的鲁棒性和泛化能力。

(2)损失函数。GAN生成对抗网络是一种深度学习模型,于2014年被Goodfellow等^[20]提出。目前,生成对抗网络在图像分割领域的应用越来越广泛;由于采用神经网络技术应用于生物医学影像领域通常需要大量预处理的标注图像样本进行有效训练,这使得该方法具有挑战性且成本高昂。2024年Biswas等^[21]通过引入GAN生成对抗网络对医学图像进行数据增强,有效扩展了肝脏、肺部和脾脏的图像数据集。为提高模型的泛化能力和训练精度,Xia等^[9]提出了GAN生成对抗网络的另一种应用模式即将传统的多分类交叉熵损失函数与生成器输出的损失函数以及判别器输出的对抗损失函数相结合的观点。这种混合损失函数对每个像素进行逐点监督,能够有效捕获类别信息,从而提升分割的精确性;但3个损失函数混合的参数是固定的,无法在训练过程中动态调整参数。文章基于Xia等^[9]提出的方法利用GAN生成对抗网络的思路对生成器的损失函数

G_loss 进行改进。对抗网络中有两个神经网络,生成器(Generator)和判别器(Discriminator)。生成器的目标是生成逼真的图像,以尽可能让判别器认为这些图像来自真实数据分布。而判别器的任务则是尽量识别真实和生成器伪造的图像。因而要使生成器的损失函数最小化从而不断降低生成特征图与金标准的差距,使判别器的损失函数最大化让判别器尽量识别出特征图与金标准的差异。文章的优化目标如式(2)所示:

$$\min_G \max_D L(D,G) = E_{(x,label) \sim p_{data}} [\ln D(label|x)] + E_{x \sim p_{data}} [\ln(1 - D(label|G(x)))] \quad (2)$$

其中, G 是生成网络; D 是判别网络; $E_{(x,label) \sim p_{data}}$ 表示真实数据 x 和真实标签 $label$ 来自于数据分布; $E_{x \sim p_{data}}$ 表示 x 的采样来自真实数据分布。 $G(x)$ 是根据输入数据 x (通常是预处理好的患者CT图像)所生成的图像或伪标签; $D(label|x)$ 表示判别器输入 x 为真实标签($label$)的概率;判别网络的优化目标由两个方面构成:一是对于真实数据 x 和真实标签 $label$, D 的输出 $D(label|x)$ 应尽可能接近1,使判别器能够最大限度的认识什么是真实的数据。二是对于生成器生成的伪标签 $G(x)$, D 的输出 $D(label|G(x))$ 应尽可能接近0,使得判别器能够不断发现预测分割结果和实际分割结果的差别。利用 $\ln$ 函数的特性,如果判别器输出的概率值很低,即认为生成器输出的肿瘤预测与金标

准非常不符,那么 $D(\text{label}|x) \rightarrow 0$, $\ln$ 的结果会无限趋于负无穷,这就会给生成器施加剧烈的惩罚,逼迫使其产生更好的分割结果。生成器的损失函数如式(3)所示:

$$G_{\text{loss}} = a_1 G_{\text{loss}_1} + a_2 G_{\text{loss}_2} \quad (3)$$

其中, G_{loss_1} 是DDR-UNet++进行图像分割时的损失函数。分割损失的目标是提高生成器生成的分割结果与真实标签之间的相似度,确保生成的分割结果尽量准确地反映出真实肿瘤区域。通过对比生成结果与真实标签之间的差异,优化生成器的输出。它通常包含诸如Dice损失、BCE损失等用于度量生成网络输出与真实标签之间差异。但由于Dice损失函数通常不能有效解决肿瘤区域过小、边界模糊、背景复杂、容易假阳性等问题,文章引入能够动态调整假阳性和假阴性权重的Tversky损失代替Dice损失函数,Tversky可以通过设置较大值的 β 来减少假阳性影响,或者设置较大值的 γ 来减少假阴性影响,这在处理不平衡数据时非常有效。如式(4)所示:

$$L_{\text{Tversky}} = 1 - \frac{\sum_i g_i p_i}{\sum_{i: g_i, p_i} + \beta \sum_{i: g_i} (1 - p_i) + \gamma \sum_{i: (1 - g_i), p_i} p_i} \quad (4)$$

除此之外,BCE损失函数作为评估医学图像分割中最常用的损失函数,对图像像素的类别进行概率预测,广泛应用于数据不平衡的数据集中,通过对分割像素和背景像素的加权,有助于损失函数收敛。BCE损失函数的计算公式如(5)所示:

$$L_{\text{BCE}} = -\frac{1}{N} \left(\sum_{i=1}^N g_i \ln(p_i) + (1 - g_i) \ln(1 - p_i) \right) \quad (5)$$

为了更好地结合两者的特性,文章提出了一种混合损失函数,如式(6)所示:

$$G_{\text{loss}_1} = \alpha L_{\text{Tversky}} + (1 - \alpha) L_{\text{BCE}} \quad (6)$$

其中, α 是取值为0~1的权重因子,用于调节 L_{Tversky} 和 L_{BCE} 所占的比重。

文献[22]提出的基准U-Net结合GAN生成对抗网络的损失函数以及Xia等^[9]提出的混合损失函数中的 a_1 、 a_2 通常是由经验假定的固定参数,这些固定的权重因子不具备自适应调整的能力,从而可能导致模型性能不稳定。文章引入动态权重参数,根据每个epoch的训练效果动态更新权重,增强了模型在训练过程中的自适应程度。

(3)距离约束函数。 G_{loss_2} 为距离约束函数。距离约束损失是为了进一步优化生成器,使其不仅在分割上准确,还能在空间上与真实标签高度一致,减少生成的伪标签和真实标签之间的空间差异。

为进一步提升肝脏肿瘤的分割精度,尤其是在处理小型或形状不规则的病灶时,笔者提出了一种结合对抗约束、像素级距离约束和结构相似性

(SSIM)约束^[23]的新型损失函数:具体而言,该距离约束函数的计算公式如(7)所示:

$$G_{\text{loss}_2} = \lambda_1 \sum \ln D(\text{label}|G(x)) + \lambda_2 \sum \|G(x) - \text{label}\|^2 + \lambda_3 L_{\text{SSIM}}(G(x), \text{label}) \quad (7)$$

其中,参数 λ_1 、 λ_2 、 λ_3 分别表示对抗性约束、像素级距离约束和结构相似性损失的系数,使得每个子项对总损失的贡献可调,能够针对不同任务需求动态调整优化方向。对抗项 $\lambda_1 \sum \ln D(\text{label}|G(x))$ 表示判别网络对生成图像的输出进行的对数概率评估,对于生成器 G 生成的伪标签 $G(x)$,判别器 D 会输出一个概率值 $D(\text{label}|G(x))$ 表示输入图像的分割结果 $G(x)$ 是“真实标签”的可能性。如果判别器输出的概率值很低,那么 $D(\text{label}|G(x)) \rightarrow 0$, $\ln$ 的结果会无限趋于负无穷,这就会给生成器施加剧烈的惩罚,逼迫使其产生更好的分割结果。 $\|G(x) - \text{label}\|^2$ 表示生成图像与真实标签之间的距离的平方。这个距离约束函数的引入旨在通过对生成图像与真实标签之间的距离进行优化,进一步提高分割性能。在整个GAN的损失函数中,这个距离约束函数占据一部分权重,通过调节这个权重,可以平衡分割网络的损失函数和距离约束函数在优化目标中的影响。

2 实验准备与评价指标

2.1 实验设置

为了验证DDR-UNet++网络的有效性,文章利用深度学习框架Pytorch对LiTS-2017数据集实现级联肝脏分割和肿瘤分割。通过上述混合损失函数训练DDR-UNet++级联网络模型,使用Adam优化方法^[24],初始学习率设置为 $1e-4$,批处理数量为20并对全黑或不可使用图像做剔除处理,对于全黑图像的剔除,通过计算每张图像的像素直方图,如果所有像素值均为0(纯黑),则该图像被判定为无效数据并剔除;对于可能因重采样或配准误差导致的异常图像,采用以下方法进行筛选:计算图像的均值和方差,若方差极低,则该图像可能缺乏有效信息,被判定为异常;统计非零像素占比,若低于设定阈值(如5%),则认为该图像可能无实际医学信息,予以剔除。肝脏分割网络训练550代,肝脏肿瘤网络训练700代,同时肝脏肿瘤分割网络使用Early Stopping早停策略防止过拟合,当模型性能不再提高甚至下降时停止训练。由于医学图像数据集数量较少特征单一,为提高CT图的复用程度,增强模型的泛化能力,在模型训练的过程中采用翻转、剪切、缩放、平移等手段增加数据量,提升肿瘤分割的精确度。

2.2 实验环境与参数配置

文章实验的硬件配置:(1)12 vCPU Intel(R) Xeon

(R) Platinum 8352V CPU @ 2.10 GHz; (2) GeForce RTX 3080GPU*2 计算机平台; (3) 内存容量: 48 GB; (4) 集成开发环境: VSCode; (5) 操作系统: Windows10。超参数设置如表1所示。

表1 超参数设置
Table 1 Hyperparameter setting

参数	设置
优化方法	Adam 方法
归一化方法	批量归一化
Dropout	0.5
Batch Size	20
过拟合策略	早停策略

2.3 评价指标

采用Dice相似系数(Dice Similarity Coefficient, DSC)、相对体积差异(Relative Volume Difference, RVD)、Jaccard指数3项指标评估本模型的分割性能。

DSC的公式如式(8)所示:

$$DSC = \frac{2|A \cap B|}{|A| + |B|}$$
 (8)

其中, $|A|$ 是分割结果的区域, $|B|$ 是真实标签的区域, $|A \cap B|$ 表示两者的重叠区域, DSC的范围为0~1, 越接近于1说明分割效果越好。

RVD公式如式(9)所示。RVD用于衡量预测分割结果与真实分割之间的体积差异。 $|A|$ 为预测分割区域的体积, $|B|$ 表示真实分割区域的体积。当 $RVD>0$ 时, 表明预测结果的体积大于真实体积(过分割), 当 $RVD<0$ 时, 表明预测结果的体积小于真实体积。

$$RVD = \frac{|A| - |B|}{|B|}$$
 (9)

Jaccard指数是体积重叠误差, 如式(10)所示。 $|A|$ 表示预测分割区域的集合, $|B|$ 表示真实分割区域的集合, 通过计算重叠区域与总区域的比例, Jaccard指数可以直接反映预测结果与真实结果的吻合程度。

$$Jaccard = \frac{|A \cap B|}{|A \cup B|}$$
 (10)

3 结果与分析

3.1 损失函数权重因子的选择

为测试上述Tversky和BCE混合损失函数中 α 对DDR-UNet++网络分割精度的影响, α 为0~1时, 训练得出对应的Tversky相似系数值, 如图8所示。当 $\alpha=0$ 时, 只使用BCE损失函数, Tversky相似系数值为73.56%, 随着 α 的增大, $\alpha=0.2$ 时Tversky相似系数达到最大值76.15%, 随后逐渐递减, 只使用Tversky损

失函数值($\alpha=1$)时, Tversky相似系数值仅有58.04%。因而文章设置 $\alpha=0.2$ 进行后续实验。

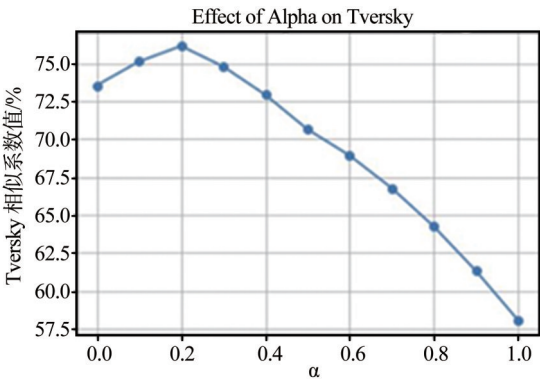


图8 Tversky相似系数值随权重因子 α 变化趋势图
Figure 8 Tversky similarity coefficient changes with the weight factor α

图9是DDR-UNet++网络权重因子为0.2时损失值(Loss)随训练轮数(Epochs)的变化, 蓝色折线表示训练过程中Loss的变化趋势。从图中可以看出, 模型在训练初期(前50个Epochs内)Loss值从接近1.0快速下降至0.6左右, 显示出良好的初始学习能力。随着训练轮数的增加, Loss值继续稳步下降, 在前100个Epochs后基本稳定在0.1~0.3, 并伴有小幅波动。整个训练过程中Loss呈现明显的下降趋势, 表明模型正在有效学习并逐步收敛。在训练后期, Loss值趋于平稳, 波动幅度较小, 说明模型已基本收敛。结合Early Stopping早停策略的使用, 当模型在验证集上的性能不再提升时会自动停止训练, 有效防止了过拟合现象的发生。

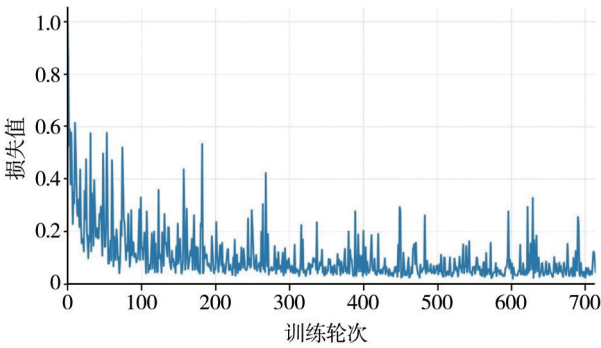


图9 损失值随训练轮数变化图
Figure 9 Loss value changes with the epoch

3.2 消融实验

为验证各模块对肝脏肿瘤分割性能的影响, 本

文以传统 U-Net 模型作为基准, 设计了系列消融实验。实验分别引入不同结构模块进行对比, 包括: 在 UNet 基础上引入残差模块的 ResUNet、在 UNet 基础上引入空洞卷积的 AtrousNet、同时引入残差模块和空洞卷积的 ResUNet++, 以及在 ResUNet++ 基础上进一步加入对抗损失机制的最终模型 DDR-UNet++。

从表 2 可知, 基础 U-Net 模型在 DSC、RVD 和 Jaccard 3 个指标上分别为 0.921、0.025 和 0.853, 性能相对较低。在 U-Net 模型基础上引入残差结构后的 ResUNet 在 DSC 指标上提升至 0.935, RVD 降低至 0.016, 表明残差连接增强了特征流动和训练稳定性。在 U-Net 模型的基础上引入空洞卷积后的 AtrousNet 在 DSC 上升至 0.938, RVD 降至 0.014, 有效增强了模型对上下文信息的感知能力。通过对比可以得知, 引入残差模块和空洞卷积均可在不同程度上优化 U-Net 模型, 且引入空洞卷积的效果相对更显著。

表 2 消融实验对比结果
Table 2 Ablation experiment results

模型	DSC	RVD	Jaccard
U-Net	0.921	0.025	0.853
ResUNet	0.935	0.016	0.878
AtrousNet	0.938	0.014	0.883
ResUNet++	0.955	0.010	0.914
DDR-UNet++	0.968	0.008	0.938

ResUNet++同时引入了残差网络和空洞卷积, 模

型性能显著提升, DSC 达到 0.955, Jaccard 提升至 0.914, 说明相比于单一引入残差网络或空洞卷积, 同时引入的效果更好, 体现出特征提取和空间信息捕捉能力的协同增强。本文提出的 DDR-UNet++ 在 ResUNet++ 模型的基础上引入对抗损失函数, 进一步优化边界细节与结构表达能力, 实现了 DSC 为 0.968、RVD 为 0.008、Jaccard 为 0.938 的最优表现, 证明了本文所提方法在精度、鲁棒性及边界细节刻画上都展现出最优的肝脏肿瘤分割效果。

3.3 实验结果比较

不同网络模型的可视化分割结果如图 10 所示, 从实验图像中可以看出, 传统 U-Net 分割出的肝脏肿瘤边界模糊, 没有分割出肿瘤的边缘细节, 出现了将背景区域预测为肿瘤区域的问题。相比于 U-Net, 由于 AtrousNet 和 ResNet 分别引入了空洞卷积模块和残差网络, 更丰富地提取了肝脏肿瘤特征, 因此肿瘤边缘细节较好, 分割效果相较于 U-Net 有明显提升, 但依然出现了微小肿瘤分割效果差, 肿瘤轮廓不清晰的问题。DA-UNet 综合引入了空洞卷积和残差模块, 且比单个引入空洞卷积和残差模块的模型效果更好。DDR-UNet++ 相较于 DA-UNet 引入了 GAN 生成对抗网络以学习预测模型与金标准之间的差异, 实验数据表明 DDR-UNet++ 既能够识别出面积较大的肿瘤, 也能相对完整地分割出微小肿瘤, 且肿瘤边缘细节清晰, 分割出的肿瘤结果最接近 LiTS-2017 数据集提供的金标准。

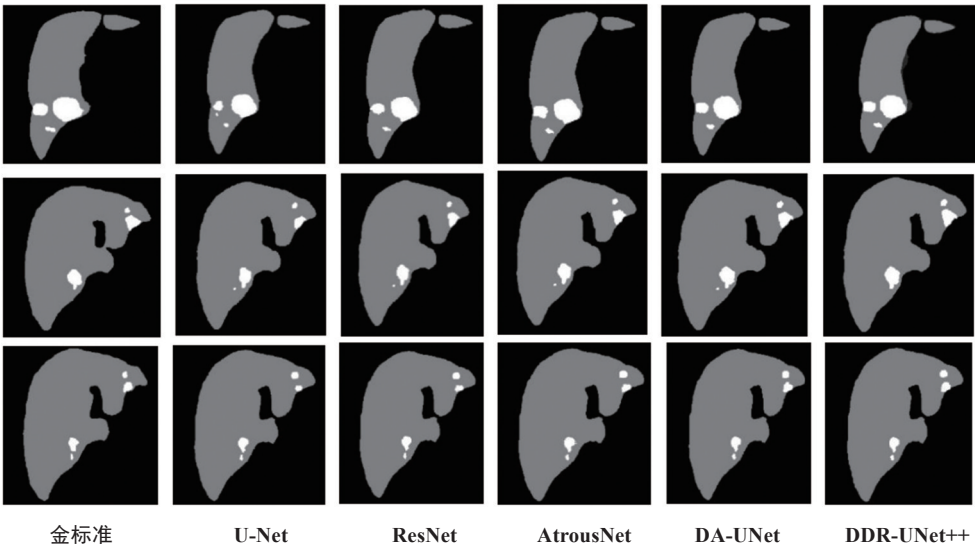


图 10 不同算法分割结果

Figure 10 Segmentation results obtained by different methods

第一到第三行图像分别代表 3 位不同患者的肝脏和对应肿瘤掩膜的叠加

4 结 论

文章针对现有的语义分割算法进行肝脏肿瘤分割时存在肿瘤提取特征不充分、肿瘤边界模糊、分割精度低和欠分割、过分割等问题提出DDR-UNet++肝脏肿瘤分割。该算法以U-Net网络为基准引入残差模块避免因模型参数量过大而出现梯度消失问题,并将卷积方式改为可变形的空洞卷积,在不增大参数量的同时扩展了感受野最大程度地富集肿瘤信息。为了提高模型的泛化能力和训练精度,文章将传统的多分类交叉熵损失函数与生成器输出的内容损失函数以及判别器输出的对抗损失函数相结合,并提出了动态调参机制,构建了一个加权损失函数。定性和定量实验表明,文章提出的深度网络架构可以稳定地提高语义分割模型的性能,帮助模型学习分割效果与金标准之间的差距从而不断向其靠拢。实验结果表明,本文方法能够有效分割出边界清晰的肿瘤图像,有效解决肝脏肿瘤分割中存在的欠分割和过分割问题,展现出较好的分割性能。文章采用二维处理方式对肝脏肿瘤进行分割,虽然能够提供一定的分割效果,但由于忽略了三维结构的完整性和上下切片之间的联系,可能导致特征利用不充分,特别是对于肿瘤在不同切片之间的空间分布关系。因此文章后期的工作是在空间上同时充分考虑深度、高度和宽度上的信息并在三维空间内进行卷积操作以捕捉肝脏和肿瘤在不同切片之间的上下文信息,从而进一步提高分割精度。

【参考文献】

- [1] 赵天业, 吴燕华, 潘禹辰, 等. SCIE收录的胃肠肝病学期刊发表的中国内地作者的肝脏疾病研究文献计量分析[J]. 临床肝胆病杂志, 2022, 38(4): 872-877.
Zhao TY, Wu YH, Pan YC, et al. A bibliometric analysis of liver disease research articles published by Chinese mainland authors in Gastroenterology & Hepatology journals indexed in Science Citation Index Expanded[J]. Journal of Clinical Hepatology, 2022, 38(4): 872-877.
- [2] 姜宇朗, 王铮, 孙明瑜. 基于生物信息学分析和实验验证筛选影响肝癌患者预后的铁死亡调控基因[J]. 海南医学院学报, 2024, 30(9): 674-682.
Jiang YL, Wang Z, Sun MY. Screening of ferroptosis-related genes affecting the prognosis of hepatocellular carcinoma patients based on bioinformatics analysis and experimental validation[J]. Journal of Hainan Medical University, 2024, 30(9): 674-682.
- [3] Goldman LW. Principles of CT and CT technology[J]. J Nucl Med Technol, 2007, 35(3): 115-128.
- [4] Alzubaidi L, Zhang JL, Humaidi AJ, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions[J]. J Big Data, 2021, 8(1): 53.
- [5] Ronneberger O, Fischer P, Brox T, et al. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [6] Milletari F, Navab N, Ahmadi SA. V-Net: fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV). Piscataway, NJ, USA: IEEE, 2016: 565-571.
- [7] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[EB/OL]. (2015-11-23). <https://arxiv.org/abs/1511.07122v1>.
- [8] Jha D, Smedsrud PH, Riegler MA, et al. ResUNet++: an advanced architecture for medical image segmentation[C]//2019 IEEE International Symposium on Multimedia (ISM). Piscataway, NJ, USA: IEEE, 2019: 225-2255.
- [9] Xia KJ, Yin HS, Qian PJ, et al. Liver semantic segmentation algorithm based on improved deep adversarial networks in combination of weighted loss function on abdominal CT images[J]. IEEE Access, 2019, 7: 96349-96358.
- [10] Hitaj B, Ateniese G, Perez-Cruz F, et al. Deep models under the GAN: information leakage from collaborative deep learning[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. New York, NY, USA: ACM, 2017: 603-618.
- [11] Salehi SSM, Erdogmus D, Gholipour A. Tversky loss function for image segmentation using 3D fully convolutional deep networks [C]//Machine Learning in Medical Imaging. Cham: Springer International Publishing, 2017: 379-387.
- [12] Li QF, Jia X, Zhou JC, et al. Rediscovering BCE loss for uniform classification[EB/OL]. (2024-03-12). <https://arxiv.org/abs/2403.07289>.
- [13] Bilic P, Christ P, Li HB, et al. The liver tumor segmentation benchmark (LiTS)[J]. Med Image Anal, 2023, 84: 102680.
- [14] Perez L, Wang J. The effectiveness of data augmentation in image classification using deep learning[EB/OL]. (2017-12-13). <https://arxiv.org/abs/1712.04621>.
- [15] Simonyan K, Vedaldi A, Zisserman A. Deep inside convolutional networks: visualising image classification models and saliency maps[EB/OL]. (2013-12-20). <https://arxiv.org/abs/1312.6034v1>.
- [16] Szegedy C, Liu W, Jia YQ, et al. Going deeper with convolutions [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2015: 1-9.
- [17] Christ PF, Ettlinger F, Grün F, et al. Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks[EB/OL]. (2017-02-23). <https://arxiv.org/abs/1702.05970>.
- [18] Dai JF, Li Y, He KM, et al. R-FCN: object detection via region-based fully convolutional networks[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2016: 379-387.
- [19] Xiao X, Lian S, Luo ZM, et al. Weighted Res-UNet for high-quality retina vessel segmentation[C]//2018 9th International Conference on Information Technology in Medicine and Education (ITME). Piscataway, NJ, USA: IEEE, 2018: 327-331.
- [20] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Commun ACM, 2020, 63(11): 139-144.
- [21] Biswas A, Maity SP, Banik R, et al. GAN-driven liver tumor segmentation: enhancing accuracy in biomedical imaging[J]. SN Comput Sci, 2024, 5(5): 652.
- [22] 邓鸿, 邓雅心, 丁廷波, 等. 基于生成对抗网络的肝脏CT图像分割[J]. 北京生物医学工程, 2021, 40(4): 367-376.
Deng H, Deng YX, Ding TB, et al. Liver CT image segmentation based on generative adversarial network[J]. Beijing Biomedical Engineering, 2021, 40(4): 367-376.
- [23] Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM[C]//2010 20th International Conference on Pattern Recognition. Piscataway, NJ, USA: IEEE, 2010: 2366-2369.
- [24] Zhang ZJ. Improved Adam optimizer for deep neural networks[C]//2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). Piscataway, NJ, USA: IEEE, 2018: 1-2.

(编辑: 薛泽玲)