

DOI:10.16136/j.joel.2023.12.0861

基于改进 3D UNet 的前列腺 MR 图像分割

桑子江¹, 邵叶秦^{2*}, 许昌炎¹

(1. 南通大学 信息科学技术学院, 江苏 南通 226000; 2. 南通大学 交通与土木工程学院, 江苏 南通 226000)

摘要:针对前列腺磁共振(magnetic resonance, MR)图像边缘模糊、对比度较低,灰度值分布不均而导致分割精度较差的问题,提出了一种结合双路径注意力(dual path attention, DPA)和多尺度特征聚合(multi-scale feature aggregation, MFA)模块的改进 3D UNet 网络模型。首先,对数据集进行重采样和裁剪处理以适应模型输入。然后,在 3D UNet 网络的编码器各层引入 DPA 并添加残差连接,加强特征的编码能力。同时,在网络解码器中加入 MFA 模块,以充分利用空间上下文信息,增强语义信息。最后,在公开数据集 PROMISE12 上进行验证,所提出的模型的 *Dice* 系数为 89.90%, Hausdorff 距离为 9.37 mm。相比较于其他模型,所提出模型的分割结果更优,且参数量和运算量更少。

关键词:前列腺分割; 双路径注意力(DPA); 多尺度特征聚合(MFA); 医学图像分割

中图分类号:TP391 **文献标识码:**A **文章编号:**1005-0086(2023)12-1337-08

Prostate magnetic resonance image segmentation based on improved 3D UNet

SANG Zijiang¹, SHAO Yeqin^{2*}, XU Changyan¹

(1. School of Information Science and Technology, Nantong University, Nantong, Jiangsu 226000, China;

2. School of Transportation and Civil Engineering, Nantong University, Nantong, Jiangsu 226000, China)

Abstract: An improved 3D UNet network model that combines a dual path attention (DPA) module and a multi-scale feature aggregation (MFA) module is proposed to address the problems such as blurred edges, low contrast, and uneven gray value distribution in magnetic resonance (MR) images of the prostate bringing about the poor segmentation accuracy. Firstly, the dataset is resampled and cropped to fit the model input. Then, DPA and residual connection are added to each layer of the 3D UNet network encoder to enhance the feature coding capability. At the same time, an MFA module is added to the network decoder to make full use of spatial context information and enhance semantic information. Finally, the proposed model is validated on the public dataset PROMISE12, with the *Dice* coefficient of 89.90% and the Hausdorff distance of 9.37 mm. Compared with other models, the proposed model has better segmentation results, and the number of parameters and the amount of computation are less.

Key words: prostate segmentation; dual path attention (DPA); multi-scale feature aggregation (MFA); medical image segmentation

0 引 言

前列腺癌是常见的癌症之一,也是导致男性癌症死亡的主要原因之一^[1]。尽早地发现和治療可以降低其死亡率。磁共振(magnetic resonance, MR)成像技术由于其安全性和高清晰度,被广泛应用于前列腺的检查中。在临床应用中,医生需

要对前列腺 MR 图像进行观察,标注并分割出前列腺组织区域,从而对前列腺的病变程度作进一步的诊断和评估。但人工分割需要耗费大量精力,且容易受到个人主观因素的影响。因此,利用计算机辅助医学图像的分割具有重要的临床意义。一些依赖人工设计特征的半自动化分割方法,需要大量的临床经验,限制了其广泛应用。随

* E-mail: hnsyk@ntu.edu.cn

收稿日期:2022-12-27 修订日期:2023-04-05

基金项目:国家自然科学基金(61671255)资助项目

着计算机硬件的飞速发展,基于卷积神经网络的分割方法越来越受到研究者的关注。

SHELHAMER 等^[2] 提出一个全卷积网络 (fully convolutional networks, FCN), 在网络最后用卷积层代替传统的全连接层 (full connected layer, FC), 这样不仅能够适应不同尺寸大小的图像输入, 同时还可以进行像素级别的图像分割, 成为了语义分割的旗帜性模型。RONNEBERGER 等^[3] 提出了 UNet 模型, 它采用独特的编码器解码器结构, 编码器不断下采样获得更多的特征信息, 解码器不断上采样便于恢复图像的空间信息, 再通过跳跃连接去补充解码器和编码器的语义歧义, 在细胞图像分割中获得了不错的效果。BAI 等^[4] 改进了 UNet 模型, 通过使用不同大小的空洞卷积来增加网络的感受野, 进而捕获细胞图像中丰富的多尺度特征, 对不规则形状的细胞具有良好的分割精度和泛化能力。由于前列腺 MR 图像是三维图像, 采用传统的 2D 网络结构在训练时需将图像分割成多张切片, 会损失切片之间的空间联系且存在特征关注程度不足的问题^[5]。因此, 许多研究者多采用 3D 网络结构进行 MR 图像的分割。ÇİÇEK 等^[6] 提出的 3D UNet 得到了广泛应用, 它采用 3D 卷积结构来获得更多的特征信息, 加强了图像各层切片的信息交互。MILLETARI 等^[7] 提出了 V-Net, 引入了一种新的 *Dice* 损失函数, 缓解了前景和背景的类别不平衡问题, 可以对输入的前列腺 MR 图像进行端到端的 3D 体素分割。虽然这些方法获得了不错的分割结果, 但因为 MR 图像的对比度较低、边缘模糊等特点, 需要进一步加强网络特征提取能力, 以优化分割效果。ZHU 等^[8] 提出领域迁移的前列腺分割方法, 但需要依赖其他医学数据集, 限制了其适用性。有研究者采用大卷积核^[9] 和增加网络宽度^[10] 的方法来获得更多特征信息, 但会带来参数量和分割时间的大幅增加。SHAO 等^[11] 在网络中使用深度可分离卷积, 虽然理论上可以减少计算量, 但在实际应用中并没有像普通卷积运算那样做专门的优化, 限制了其应用。注意力机制^[12] 在计算机视觉领域大放光彩, 其可以进行重要特征的关注并抑制无关特征, 但大多运用在自然图像中, 在医学图像领域中的应用较少, 且上述分割方法仍旧存在复杂度较高、精度较差等缺陷, 需要进一步研究。

综上, 本文研究了基于改进 3D UNet 的端对端的前列腺 MR 图像的分割模型, 主要研究内容如下:

1) 为了降低 3D UNet 的模型运算量并提高网络的分割能力, 本文优化了模型的结构, 包括调整网络层数和通道数、添加残差连接、改变标准化方式等。

2) 为了进一步提升网络的特征提取能力, 使网络能够关注前列腺边缘等关键区域, 更好地定位和分割前列腺, 本文分析了自然图像领域的注意力机制, 在阐述其局限的基础上, 设计了双路径注意力 (dual path attention, DPA) 模块, 通过通道注意力和全局注意力的结合来准确分割前列腺。

3) 为了加强模型的空间表征能力, 使模型能够适应特征图尺寸的变化, 本文提出了多尺度特征聚合 (multi-scale feature aggregation, MFA) 模块, 通过不同尺寸大小的特征图来获得不同尺度的特征信息, 并对这些多尺度特征图进行了汇集和融合。

1 方 法

1.1 网络结构

本文的网络整体结构如图 1 所示。采用了编码器解码器结构, 左边编码器部分通过下采样获得更多的特征信息, 右边解码器部分通过上采样恢复特征图的原始尺寸大小以便于进行最后的体素分类。为了减少计算量并提升分割性能, 模型基于 3D UNet 网络进行改进。将 DPA 模块添加到网络编码器各层, 加强网络的特征提取。对解码器各层特征图进行 MFA, 更好地进行体素级分类以精细化分割结果。

此外, 本文还对 3D UNet 网络的模块参数和结构进行了优化 (tuned 3D UNet, TUNet)。1) 增加网络深度, 改进的 3D UNet 网络为 4 次下采样和上采样, 网络层数加大, 可以提取更多的特征; 2) 原网络每层 2 个卷积块中的批标准化 (batch normalization, BN) 替换为实例标准化 (instance normalization, IN), ReLU 激活函数替换为 LeakyReLU, 并添加残差连接, 可以加速收敛并减小 Batchsize 超参数的影响^[13]; 3) 调整第 1 层通道数为 16, 跳跃连接用于融合相同分辨率和通道大小的信息, 以缓解计算负担的压力并减小语义歧义。

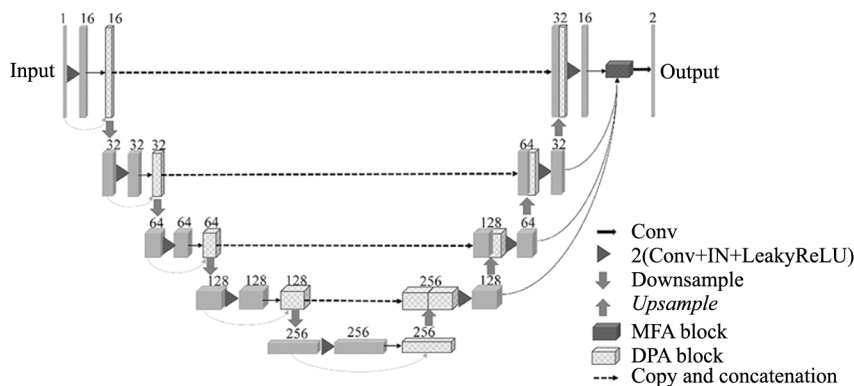


图 1 网络结构

Fig. 1 Network structure

1.2 DPA

SE-Net^[14]作为一种轻量级注意力机制,在计算机视觉领域中得到了较多的使用,它主要由全局平均池化层(global average pooling, GAP)和 FC 组成, GAP 用于压缩图像尺寸,获得各个通道的统计数据, FC 用于通道之间的信息交互。但 SE-Net 的做法可能会损失图像的空间特征信息。受 SE-Net 的启发,本文提出了一种 DPA 模块,如图 2 所示。它由两条路径组成,其中一条路径由 GAP 和 FC 组成,用于获得通道信息;另一条路径由 FC 组成,用于关注图像的全局空间信息。然后通过相加的方式融合来自两条路径的通道信息和全局信息,可以让网络在获得通道注意力的同时兼顾到全局空间特征信息,更好地定位到前列腺区域。最后经过激活函数

获得权重因子,并加权到原始特征图,得到输出结果。

对于输入特征图 $\mathbf{X} \in \mathcal{R}^{C \times D \times H \times W}$, 其中,通道数量为 C , 图像切片数量为 D , 每张切片大小为 $H \times W$, 通道注意力路径(图 2 上半分支)包括两次 FC 和 ReLU 激活函数,其输出结果为 $\mathbf{A}_c$ 。其中, FC 用于线性处理, ReLU 用于非线性处理。计算过程如式(1)和式(2)所示:

$$GAP(\mathbf{X}) = \frac{1}{D \times H \times W} \sum_{k=1}^D \sum_{j=1}^H \sum_{i=1}^W X(i, j, k), \quad (1)$$

$$\mathbf{A}_c = \mathbf{W}_2 \{ \sigma \{ \mathbf{W}_1 [GAP(\mathbf{X})] \} \}, \quad (2)$$

式中, σ 为 ReLU 激活函数; $GAP(\cdot)$ 为全局平均池化层的处理; $\mathbf{W}_1 \in \mathcal{R}^{C \times c}$ 为第一个 FC 的权重矩阵, 本

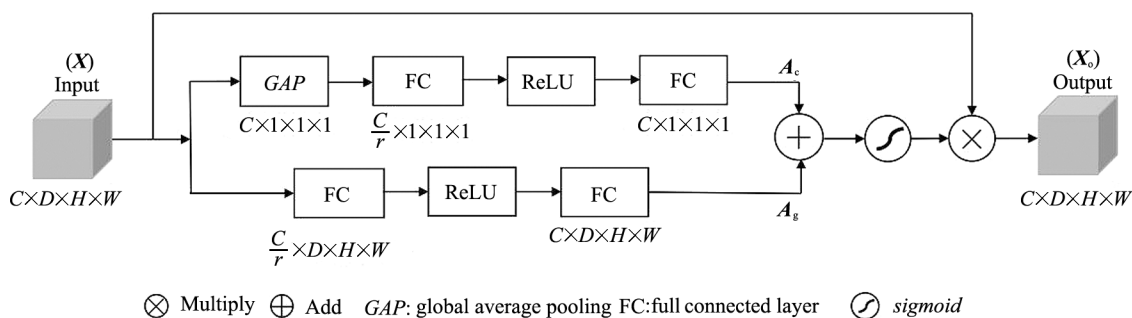


图 2 DPA 模块结构

Fig. 2 Structure of the DPA module

文取压缩因子 r 为 8, 目的是为了压缩通道数量, 减少计算量; $\mathbf{W}_2 \in \mathcal{R}^{C \times c}$ 为第二个 FC 的权重矩阵, 用于进行各个通道的信息交互。

同时, 空间注意力路径(图 2 下半分支)并行处理输入特征图 $\mathbf{X}$, 获得全局特征图 $\mathbf{A}_g$ 。具体实现

如下:

$$\mathbf{A}_g = \mathbf{W}_4 \{ \sigma [\mathbf{W}_3(\mathbf{X})] \}, \quad (3)$$

式中, $\mathbf{W}_3$ 与 $\mathbf{W}_4$ 为权重矩阵。为了获得图像的全局信息, 并未进行 GAP 的处理。

为融合来自通道注意力的特征图和全局空间的

特征图,将两者相加并通过 *sigmoid* 激活函数进行处理获得加权系数,再将此系数加权到原始特征图上,如式(4)所示:

$$\mathbf{X}_o = (\text{sigmoid}(\mathbf{A}_c \oplus \mathbf{A}_g)) \otimes \mathbf{X}, \quad (4)$$

式中,输出特征图为 $\mathbf{X}_o$, $\oplus$ 和 $\otimes$ 分别表示按元素相加和相乘。

1.3 MFA 模块

在经典的 3D UNet 分割网络模型中,解码器通过拼接或相加操作来融合编码器的特征信息并恢复图像尺寸大小。但简单地合并同级分辨率的特征图进行卷积处理可能会导致空间语义信息不足的问题。受 TGA-Net^[15] 的启发,本文提出了 MFA 模块,它通过融合解码器不同层级的特征图来获得不同尺度的特征信息,并进行一系列卷积和非线性处理增强其稳定性。MFA 模块如图 3 所示。

输入的 4 个特征图来自解码器的 4 个不同分辨率级别,用于获得多个尺度的特征信息。用 f_1, f_2, f_3, f_4 分别表示 Input1、Input2、Input3、Input4。MFA 模块输入特征图可表示为:

$$f_i \in \mathbb{R}^{2^{i-1} \times \frac{D}{2^{i-1}} \times \frac{H}{2^{i-1}} \times \frac{W}{2^{i-1}}}, \quad \forall i \in 1, 2, 3, 4, \quad (5)$$

式中, C 表示输入特征图的通道数量, D 为图像的切

片数量, H 和 W 表示切片图的高和宽。

为了能够融合不同尺度特征图的空间上下文信息,需要对不同分辨率的特征图进行上采样处理,分别对 f_2, f_3, f_4 进行 2 倍、4 倍、8 倍的上采样,以统一各个输入的分辨率大小,输出为 f_u , 再通过 $1 \times 1 \times 1$ 卷积调整通道数,然后进行拼接操作融合多个尺度的特征信息。输出 f_{cat} 计算过程如下:

$$\begin{cases} f_{u2} = \text{Upsample}(f_i, n), i = 2, n = 2 \\ f_{u3} = \text{Upsample}(f_i, n), i = 3, n = 4 \\ f_{u4} = \text{Upsample}(f_i, n), i = 4, n = 8 \end{cases}, \quad (6)$$

$$f_{cat} = \text{Concat}(\text{Conv3d}(f_1, f_{u2}, f_{u3}, f_{u4})), \quad (7)$$

式中, $\text{Upsample}(f_i, n)$ 表示对 f_i 进行三线性插值上采样, n 为上采样倍数, Conv3d 表示卷积核大小为 $1 \times 1 \times 1$ 的 3D 卷积, Concat 表示按通道拼接。

为了进一步加强特征信息的提取,增强网络的稳定性,对于拼接后的结果 f_{cat} 还需进行一些卷积块的处理,最后输出 f_o 。如式(8)–(11)所示。为了加速收敛速度并提升分割精度,本文还采用了 DenseNet^[16] 的残差处理,模块内部增加了 3 条残差连接(见图 3 右侧),并将 DenseNet 的拼接操作替换成相加操作以减少计算量。

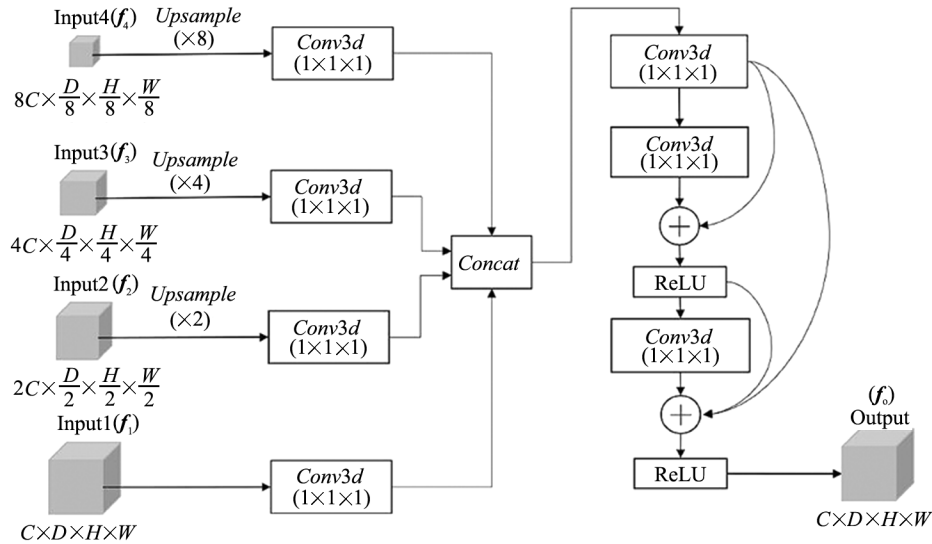


图 3 MFA 模块结构

Fig. 3 Structure of the MFA module

$$\text{res1} = \text{Conv3d}[\text{Conv3d}(f_{cat})] + \text{Conv3d}(f_{cat}), \quad (8)$$

$$\text{res2} = \text{Conv3d}[\sigma(\text{res1})] + \sigma(\text{res1}), \quad (9)$$

$$\text{res3} = \text{res2} + \text{Conv3d}(f_{cat}), \quad (10)$$

$$f_o = \sigma(\text{res3}), \quad (11)$$

式中, Conv3d 为 $1 \times 1 \times 1$ 大小的 3D 卷积, res1 、

res2 、 res3 为加入残差连接之后的中间输出结果,最后输出 f_o 。尺寸为 $C \times D \times H \times W$ 。

1.4 损失函数

在模型训练过程中,本文采用了混合损失函数^[17],其结合了交叉熵损失和 *Dice* 损失。交叉熵损失从信息论的角度出发,可以度量预测概率分布和

真实分布的差异,它计算的是总体的期望,前景和背景有着相同的加权因子。*Dice* 损失广泛应用于医学图像分割中,是一种基于区域的损失函数,还可以缓解前景和背景的数据分布不平衡的问题。采用混合损失函数可以提升模型训练的稳定性,公式如下所示:

$$loss_{dice} = 1 - \frac{2 \times \sum_{i=1}^N p_i \times g_i + \lambda}{\sum_{i=1}^N p_i + \sum_{i=1}^N g_i + \lambda}, \quad (12)$$

$$loss_{ce} = -\frac{1}{N} \sum_{i=1}^N g_i \ln(p_i), \quad (13)$$

$$Loss = \alpha \times loss_{dice} + \beta \times loss_{ce}, \quad (14)$$

式中, $loss_{dice}$ 为 *Dice* 损失, p_i 和 g_i 分别表示预测图和真实标签图的像素值, λ 为数值很小的平滑因子。 $loss_{ce}$ 为交叉熵损失, N 代表体素个数, $Loss$ 为总体损失函数, α 和 β 为加权系数,本文中均取值为 0.5。

2 实验结果与分析

2.1 数据集及预处理

本文数据集采用 MICCAI2012 挑战赛公开数据集 PROMISE12^[18], 其中含有 50 个包含金标准 (ground truth, *GT*) 的横向 T2 加权前列腺 MR 图像。图像有着不同的体素间隔和尺寸大小,均采集于实际的临床医学环境中。本文对数据集进行了随机划分,训练集与测试集划分比例为 8:2,即训练集为 40 个,测试集为 10 个。

本文对图像的预处理分为 3 个部分,分别为重采样和裁剪、数据增强、归一化。因为数据集有着不同的体素间隔、不同的尺寸大小,所以需要对数据集统一尺寸和体素间隔。本文对数据集进行了重采样处理,通过线性插值的方式统一体素间隔为 (1 mm, 1 mm, 1 mm),更加有利于后续神经网络中 3D 卷积的运算;为了减少前列腺周围无关组织的干扰并降低 GPU 显存的消耗,对图像进行中心裁剪处理,固定图像尺寸为 (96, 96, 96);因数据集样本过少,网络训练很可能会过拟合,所以对数据集进行了翻转、旋转等数据增强处理;为了加速网络收敛和提升模型精度,采用了 Z-score 归一化方法,使前列腺图像体素值服从均值为 0、方差为 1 的正态分布。

2.2 实验环境与评价指标

本文的所有实验均是在相同的硬件和软件环境下进行的。硬件环境采用了一个 RTX 3090 GPU 显卡, 24 G 显存大小,处理器为 Intel(R) Xeon(R)

Platinum 8350C CPU @ 2.60 GHz;软件环境采用了 Ubuntu18.04.5 操作系统,Python 编程语言和 PyTorch 深度学习框架,CUDA 和 CUDNN 版本分别为 11.1 和 8.05,可以更好地加速深度学习算法。在网络训练阶段,本文的实验参数设置为:批处理输入 Batchsize 大小为 4,初始学习率为 0.001,采用 Adam 优化器进行模型训练的更新迭代,总迭代轮数为 500 轮。

为了客观评价前列腺图像的分割结果,本文选用了 4 种医学图像分割中常用的评价指标,分别为 *Dice* 系数 (*dice coefficient*, *Dice*)、豪斯多夫距离 (*hausdorff distance*, *HD*)、精确度 (*Precision*) 和敏感性 (*Sensitivity*)。公式如下所示:

$$Dice = \frac{2(GT \cap Pre)}{GT \cup Pre} = \frac{2TP}{2TP + FN + FP}, \quad (15)$$

$$Sensitivity = \frac{TP}{TP + FN}, \quad (16)$$

$$Precision = \frac{TP}{TP + FP}, \quad (17)$$

$$HD(Y, \hat{Y}) = \max_{y \in Y} \{ \min_{\hat{y} \in \hat{Y}} d(y, \hat{y}), \max_{\hat{y} \in \hat{Y}} \{ \min_{y \in Y} d(\hat{y}, y) \} \}, \quad (18)$$

式中, *GT* 表示分割金标准, *Pre* 表示预测分割结果, *TP* 和 *FN* 分别表示真阳性和假阴性,表示实验中被正确分类为前列腺区域的体素和被错误分类为无关背景组织区域的体素, *FP* 表示假阳性,即错误分类为前列腺区域的体素。*Dice* 系数用于评估分割结果与真实标签的区域重叠程度,取值范围为 [0, 1],值越高,表明与金标准重合程度越高。敏感性和精确度用于评估分割的完整性和准确性。*HD* 则是用来计算分割边界和真实边界的距离,值越小,说明分割边界与真实边界的差异越小。其中, *Y* 和 $\hat{Y}$ 分别为金标准分割轮廓的集合和预测图轮廓的集合, $d(y, \hat{y})$ 表示预测分割图边界像素点 y 与金标准分割图边界像素点 $\hat{y}$ 的欧式距离。

2.3 模型消融实验

为了验证所提方法中各模块的有效性,对改进的 3D UNet 模型在 PROMISE12 数据集上设计了消融实验,实验结果见表 1。网络以第 1 章方法中提到的优化的 3D UNet 网络 (tuned 3D UNet, TUNet) 模型作为基准网络,在此基础上加入了 DPA 模块、MFA 模块,验证各个模块对模型性能的影响。

从表 1 中第 2 行可以看出,相比较于标准的 3DUNet 网络,经过优化过的 3DUNet 在多个指标

性能上都得到了提升,其中 *Dice* 系数提升了 2.19%, *HD* 降低了 1.26 mm, 敏感性提高了 1.35%, 精确度提升了 2.60%, 说明经过优化的 3D UNet 可以在一定程度上提升模型的分割性能。

表 1 各模块消融实验对比

Tab. 1 Comparison of ablation experiments of each module

Method	<i>Dice</i> /%	<i>HD</i> /mm	<i>Sensitivity</i> /%	<i>Precision</i> /%
3D UNet	86.13	10.24	89.01	84.20
TUNet	88.32	8.98	90.36	86.80
TUNet+DPA	89.52	8.34	88.90	90.38
TUNet+MFA	89.63	10.52	89.74	89.86
TUNet+DPA+MFA	89.90	9.37	89.37	90.64

表中第 3 行加入了 DPA 模块,进一步提升了模型的分割性能,其中, *Dice* 系数提升了 1.20%, *HD* 降低了 0.64 mm, 特别注意的是在精确度指标上有了很大的提升,提升了 3.58%。这说明模型获得了比较准确的分割结果,误分割率较低,表明了 DPA 模块的有效性。

表中第 4 行加入了 MFA 模块,相比于第 2 行中优化的 3D UNet 模型, *Dice* 系数和精确度分别提高了 1.31% 和 3.06%。MFA 模块融合了来自解码器的多个尺度的特征信息,增强了空间上下文的建模能力,因而带来了 *Dice* 系数和精确度指标上的提升。但因其只关注到了多尺度的空间特征信息,在细节特征方面的关注尚有欠缺,所以在 *HD* 指标上不如添加 DPA 模块的结果。

表中最后一行为本文提出的改进 3D UNet 模型的分割指标结果,可以看到 *Dice* 系数和精确度有着最好的结果,分别达到 89.90% 和 90.64%, *HD* 也与只加 MFA 模块相比降低了 1.15 mm。这说明结合 DPA 和 MFA 可以有效提升 *Dice* 和 *HD* 指标的性能。虽然敏感性指标提升不明显,但对于前列腺分割任务而言,分割的准确性,如 *Dice* 系数和精确度是最为重要的指标。通过消融实验表明,本文提出

的融合 DPA 和 MFA 模块的改进 3D UNet 模型有着更好的分割性能。

2.4 与主流方法的对比

为了更进一步地验证本文所提方法的有效性,本文与主流的分割方法进行了比较,主要对比方法为 FCN、3D UNet、V-Net,并用参数量 (parameters, *Params*)、运算量 (floating-point operations, *FLOPs*)、*Dice* 系数、*HD* 等指标进行了性能的分析。测试集的平均分割结果如表 2 所示。

由表 2 可以看出,本文模型的分割方法在多个指标上均优于其他方法,其中 *Dice* 为 89.90%, *HD* 为 9.37 mm, 敏感性为 89.37%, 精准度为 90.64%。而且相比较于 3D UNet 和 V-Net, *Dice* 分别提高了 3.77% 和 3.54%。运算量分别减少了 87.29% 和 83.39%, 参数量分别减少了 56.12% 和 89.12%, 提升了分割精度并极大地减少了计算量。虽然 FCN-8s 的运算量和参数量均最低,但其对底层信息的利用程度不足,分割指标较差, *Dice* 只有 80.09%。为了更直观地表明本文方法的有效性,选取了测试集的几个前列腺切片进行了可视化的展示,如图 4 所示。

图 4 中,第 1 列为原始图像,第 2 列为金标准。其余 4 列分别为主流分割方法与本文分割方法的分割结果图。从整体来看,FCN 表现最差,分割图中出现了锐利的边缘以及误分割,分割结果比较粗糙。3D UNet 和 V-Net 分割结果定位不够精准且边缘部分分割不够好,相比于金标准来说,还有不少的差距。本文方法在外观上更贴近于真实金标准。

Case09 的前列腺图像亮度较低,对比度较低,肉眼都很难观察到前列腺组织,3D UNet 和 V-Net 对边缘的分割不够精细,本文提出模型的分割边缘更好。

对于 Case13 而言,前列腺组织较小,考验模型对小目标的分割能力,可以看到 3D UNet 和 V-Net 对前列腺判定不够准确,没有很好地定位到前列腺的组织区域,分割结果较差,本文方法有着更优的分割

表 2 与主流方法的测试结果对比

Tab. 2 Comparison of test results with state-of-the-art methods

Method	<i>FLOPs</i> /G	<i>Params</i> /M	<i>Dice</i> /%	<i>HD</i> /mm	<i>Sensitivity</i> /%	<i>Precision</i> /%
FCN-8s	24.77	0.99	80.09	16.03	81.37	79.23
3D UNet	3 200.47	16.32	86.13	10.24	89.01	84.20
V-Net	2 448.28	65.82	86.36	12.43	86.65	86.64
Our method	406.74	7.16	89.90	9.37	89.37	90.64

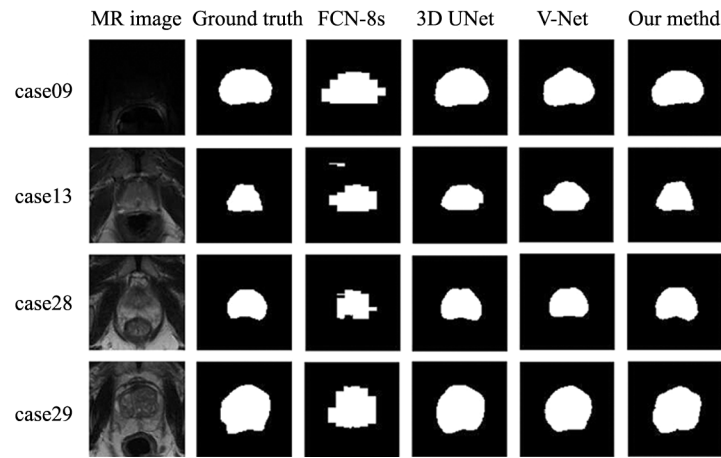


图 4 不同方法分割结果对比图

Fig. 4 Comparison of segmentation results by different methods

结果。

Case28 边缘模糊,没有清晰的边界,分割较困难,可以观察到,3D UNet 和 V-Net 的分割边缘存在误分割和欠分割等情况,本文所提方法有着更优的分割结果。

Case29 前列腺组织较大,属于大目标分割,且前列腺组织内部的灰度值分布不均匀,甚至出现了亮度明暗的两级分化的情况,3D UNet 和 V-Net 边缘分割较差,本文方法更贴近于专家手工分割的金标准。

通过分割结果的可视化进一步表明了本文分割方法的有效性。通过模型方法的对比分析和可视化展示,证明了本文方法无论是从定性还是定量性能指标上都有着更优的结果。

3 结 论

本文针对前列腺图像分割精度较差和参数量较多等问题,提出了端到端的改进的 3D UNet 网络,首先优化了 3D UNet 网络并加入了残差连接,加速了网络的收敛。然后在此基础上增加了本文提出的 DPA 模块和 MFA 模块,DPA 模块用于加强网络的编码能力,提取更多的关键特征信息,MFA 模块用于提取空间的上下文信息,获得更好的分割结果。在 PROMISE12 数据集的实验结果表明,与其他主流分割方法相比,本文所提出方法能够更好地分割前列腺图像,对于对比度较低、边缘模糊的图像也有着不错的分割结果,可以辅助医生进行临床诊断及治疗。但本文没有对前列腺分割结果进行后处理操作,接下来的工作将进一步验证模型在其他数据集上的泛化能力。

参考文献:

[1] CULP M B, SOERJOMATARAM I, EFSTATHIOU J A, et al. Recent global patterns in prostate cancer incidence and mortality rates[J]. European Urology, 2020, 77(1): 38-52.

[2] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.

[3] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015, October 5-9, 2015, Munich, Germany. Berlin: Springer, 2015: 234-241.

[4] BAI H, LU C, MA M, et al. An improved U-Net for cell confluence estimation[J]. Optoelectronics Letters, 2022, 18(6): 378-384.

[5] ZHOU W, TAO X, WEI Z, et al. Automatic segmentation of 3D prostate MR images with iterative localization refinement[J]. Digital Signal Processing, 2020, 98: 102649.

[6] ÇIÇEK Ö, ABDULKADIR A, LIENKAMP S S, et al. 3D U-Net: Learning dense volumetric segmentation from sparse annotation[C]// Medical Image Computing and Computer-Assisted Intervention, October 17-21, 2016, Athens, Greece. Berlin: Springer, 2016: 424-432.

[7] MILLETARI F, NAVAB N, AHMADI S A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation[C]// 2016 Fourth International Conference on 3D Vision (3DV), October 25-28, 2016, Stanford, CA, USA. New York: IEEE, 2016: 565-571.

[8] ZHU Q, DU B, YAN P. Boundary-weighted domain adaptive neural network for prostate MR image segmentation

- [J]. IEEE Transactions on Medical Imaging, 2020, 39(3): 753-763.
- [9] JIA H, CAI W, HUANG H, et al. Learning multi-scale synergic discriminative features for prostate image segmentation[J]. Pattern Recognition, 2022, 126: 108556.
- [10] JIA H, SONG Y, HUANG H, et al. HD-Net: hybrid discriminative network for prostate segmentation in MR images[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2019, October 13-17, 2019, Shenzhen, China. Berlin: Springer, 2019: 110-118.
- [11] SHAO D G, HUANG J H, XU H. Segmentation of prostate image based on U-Net of multi-scale dilated separable convolution[J]. Journal of Optoelectronics • Laser, 2022, 33(5): 554-560.
邵党国, 黄俊辉, 徐慧. 基于多尺度空洞分离卷积的 U-Net 分割前列腺图像[J]. 光电子 • 激光, 2022, 33(5): 554-560.
- [12] DAI Y, GIESEKE F, OEHMCKE S, et al. Attentional feature fusion[C]//IEEE/CVF Winter Conference on Applications of Computer Vision, January 3-8, 2021, Waikoloa, HI, USA. New York: IEEE, 2021: 3560-3569.
- [13] ISENSEE F, JAEGER P F, KOHL S A A, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nature Methods, 2021, 18(2): 203-211.
- [14] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7132-7141.
- [15] TOMAR N K, JHA D, BAGCI U, et al. TGANet: Text-guided attention for improved polyp segmentation[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2022, September 18-22, 2022, Singapore. Cham: Springer, 2022: 151-160.
- [16] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 4700-4708.
- [17] TAGHANAKI S A, ZHENG Y F, ZHOU S K, et al. Combo loss: Handling input and output imbalance in multi-organ segmentation[J]. Computerized Medical Imaging and Graphics, 2019, 75: 24-33.
- [18] LITJENS G, TOTH R, VAN DE VEN W, et al. Evaluation of prostate segmentation algorithms for MRI: The PROM-ISE12 challenge[J]. Medical Image Analysis, 2014, 18(2): 359-373.

作者简介:

邵叶秦 (1978—), 男, 博士, 副教授, 硕士生导师, 主要从事数字图像处理 and 人工智能方面的研究。