

DOI: 10.19650/j.cnki.cjsi.2006199

基于深度学习的膝关节 MR 图像自动分割方法^{*}

于宁波^{1,2}, 刘嘉男^{1,2}, 高 丽^{1,2}, 孙泽文³, 韩建达^{1,2}

(1. 南开大学人工智能学院 天津 300350; 2. 南开大学 天津市智能机器人技术重点实验室
天津 300350; 3. 北京大学第三医院 运动医学研究所 北京 100083)

摘 要: 膝关节磁共振图像的自动分割具有重要的临床需求, 图像中分割目标的大小不同为精准分割带来了挑战。基于深度学习, 提出一种端到端的 DRD U-Net。以残差模块作为基本模块, 增加了对特征的复用能力。利用并行的扩张卷积模块获取不同的感受野, 克服了 U-Net 模型单一感受野的局限性, 提高了对不同大小目标的分割能力。设计多输出融合的深监督模块, 直接利用不同层次的特征实现了信息互补, 提高了分割区域的连贯性和准确性。在 OAI-ZIB 数据集上测试, 平均分割表面距离为 0.2 mm, 均方根表面距离为 0.43 mm, 豪斯多夫距离为 5.22 mm, 平均戴斯系数 (DSC) 为 93.05%, 重叠误差为 3.86%。相比于基线 U-Net 和其他现有模型, 所提方法在膝关节股骨、胫骨、股骨软骨、胫骨软骨的分割方面都取得了更高的精度。

关键词: 医学图像分割; 膝关节; 磁共振图像; 深度学习; 卷积神经网络

中图分类号: TP29 TH7 **文献标识码:** A **国家标准学科分类代码:** 510.8060

Auto-segmentation method based on deep learning for the knee joint in MR images

Yu Ningbo^{1,2}, Liu Jia'nan^{1,2}, Gao Li^{1,2}, Sun Zewen³, Han Jianda^{1,2}

(1. College of Artificial Intelligence, Nankai University, Tianjin 300350, China; 2. Tianjin Key Laboratory of
Intelligent Robotics, Nankai University, Tianjin 300350, China; 3. Institute of Sports Medicine,
Peking University Third Hospital, Beijing 100083, China)

Abstract: Auto-segmentation of the knee joint in magnetic resonance (MR) images is significant for clinical requirements. However, it is challenging due to that the segmentation targets have dramatically different sizes. In this study, an end-to-end DRD U-Net is proposed, which is based on the deep learning framework. The residual module is used as the basic module in the U-Net model, which increases the ability of reusing feature maps. The parallel dilated convolution modules are used to achieve different receptive fields, which can overcome the limitations of single receptive field in the U-Net model and effectively improve the segmentation capability with targets of different sizes. The multi-output fusion deep supervision module is designed to directly utilize the feature maps of different levels. In this way, the information complementarity is obtained, the consistency and accuracy of the segmented regions are improved. The proposed algorithm is evaluated by using the public OAI-ZIB data set. The average segmented surface distance is 0.2 mm, the root mean square surface distance is 0.43 mm, the Hausdorff distance is 5.22 mm, the average dice similarity coefficient (DSC) is 93.05%, and the volume overlap error is 3.86%. Compared with the conventional U-Net and other currently available models, the proposed DRD U-Net has better segmentation accuracy.

Keywords: medical image segmentation; knee joint; magnetic resonance images; deep learning; convolutional neural network (CNN)

0 引言

骨关节炎是一种慢性退行性关节疾病,导致患者的膝关节疼痛和运动障碍。据统计,我国40岁以上中老年人双膝关节患病率达15.6%^[1]。一方面,膝关节软骨的体积和厚度是早期膝关节炎检测的关键指标之一,也是膝关节炎分级的重要参考^[2]。另一方面,对于晚期膝关节炎患者,疼痛逐渐加重,甚至生活无法自理,需要进行人工膝关节置换术,在机器人辅助人工膝关节置换术中,利用已经分割的患者术前膝关节医学图像,可以进行个性化的假体设计和术前规划,以提高手术的成功率^[3]。因此,膝关节图像的分割对膝关节疾病的检测、治疗与康复都起着关键作用。

磁共振(magnetic resonance, MR)图像可以清楚地显示关节软骨和骨骼等部位,通常由经验丰富的医生逐层进行分割,这可能需要几个小时,而且由于人为误差和主观因素,不同分割者之间和同一分割者不同时期分割结果的可变性相当高^[4],临床上迫切需要计算机辅助分割方法。依照膝关节手术术前规划和个性化假体设计的要求,MR图像中待分割的部位包括股骨(femur bone, FB)、股骨软骨(femur cartilage, FC)、胫骨(tibia bone, TB)和胫骨软骨(tibia cartilage, TC)共4类。

典型的全自动分割方法分别是基于图谱、基于模型和基于像素分类的方法。Tamez-Pena等^[5]提出利用多图谱来分割膝关节软骨,把手动分割的多套MR图像作为参考图像,将待分割的MR图像与所有参考图像进行仿射变换配准,得到多个分割图像,经过边界匹配和离群值检测等后处理,逐像素投票获得最终的分割结果。Seim等^[6]利用统计形状模型(statistical shape modelling, SSM)和广义霍夫变换来分割膝关节,从60套MR图像中训练得到“金标准”的SSM,将待分割图像的胫骨和股骨的SSM与“金标准”进行拟合,然后代入软骨厚度的先验知识,使用多目标和图优化来分割软骨。尽管基于图谱和SSM的方法都能实现自动的膝关节分割,但是由于这两种方法都依赖于膝关节形状的先验知识,当患者个体差异较大时,这两种算法均不能得到满意的结果,并且计算量大,分割时间长。此外,Zhang等^[7]提出使用支持向量机(support vector machine, SVM)和判别随机场(discriminative random fields, DRF)的混合模型来进行逐像素的分类,采用了多种类型的MR图像以便结合多视角信息,在经过配准、灰度值标准化和直方图匹配等预处理后,提取图像中包含外观和形状信息在内的局部结构特征以及全局几何特征,训练提取的特征和与之对应的标签得到分类模型,并在推理过程中使用循环信念传播(loopy belief propagation, LBP)算法,但当只有一种类型

的MR图像时,算法精度会明显下降。

近年来,随着卷积神经网络(convolutional neural network, CNN)在医学图像分析领域的发展和應用,基于深度学习的方法已经成为医学图像分割中切实可行的方法^[8-9]。基于深度学习的分割方法是属于像素分类的一种,传统的基于像素分类的方法通常使用手工特征,而基于深度学习的方法能够自主学习图像特征,并且将特征提取和像素分类相结合,在训练过程中相互促进,因此分割精度也较传统算法高。Prasoon等^[10]提出了利用3个二维CNN从MR图像中分割软骨的方法,3个网络分别对应MR图像的3个切面(矢状面、冠状面和横断面),投票产生最终的分割图像,但实验中仅仅分割了胫骨软骨。Lee等^[11]为了提高CNN对小目标(软骨)的分割能力,将膝关节的分割分为两个独立的网络,同时网络的输入均是由数张连续切片组成,以考虑更多的上下文信息。此外,与文献[10]类似,也采用多切面分割的方法提高准确率。Zhou等^[12]将CNN与三维可变形模型和条件随机场(conditional random fields, CRF)相结合,以后处理的方式对膝关节分割结果做精细化处理。Ambellan等^[13]在二维CNN分割的基础上,利用SSM和三维CNN调整股骨和胫骨的分割结果,然后在提取的感兴趣区域(region of interest, ROI)上使用三维CNN分割软骨,但是对于使用非端到端的分割方法,不仅推理阶段时间长而且训练也相对来说更加复杂。

在医生的干预下,通过人机交互可以使用多种算法实现MR图像的半自动分割,半自动方法旨在减少分割过程中用户的交互量,同时允许用户提供一定程度的监督,以调整算法的收敛方向。常见的半自动方法有活动轮廓模型^[14]、区域增长^[15]以及分水岭算法^[16]等。但是由于膝关节MR图像中软骨与周围组织的对比度低,而且对于严重膝关节炎患者来说,软骨不仅薄且不连续,使用半自动方法时需要更多的交互才能完成膝关节图像的分割。

经过上述分析,相比较于传统的半自动/自动分割算法,基于CNN的医学图像分割算法不仅速度快而且精度高。本文在U-Net编-解码架构的基础上,针对膝关节MR图像中有着不同大小的分割目标,设计多尺度上下文特征提取模块以及多输出融合模块,并加强特征的复用,提出端到端的DRD U-Net(dilated residual and deeply-supervised U-Net)用于膝关节图像的分割。

1 DRD U-Net 算法

Ronneberger等^[17]提出了轻量型U-Net,该网络主要用来进行生物医学图像的分割。图1所示为膝关节MR图像。本文在U-Net模型的基础上,结合膝关节图像的

特点,提出了改进的 DRD U-Net 模型以及其核心模块的设计方法。

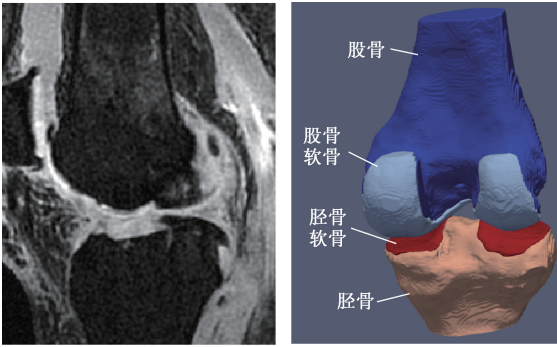


图1 膝关节图谱

Fig. 1 The knee with an atlas for reference

1.1 基本 U-Net 模型

本质上,U-Net 属于编-解码模型,如图2所示。在编码阶段,网络通过级联的卷积模块提取图像的特征信息,并通过对特征图进行最大池化操作以降低分辨率,这样可以增大卷积操作的感受野和获取更多的全局信息。具体来说,编码阶段由4个小模块组成,每一个小模块包含两个级联的 3×3 卷积,其感受野的大小与一个 5×5 的卷积核相同,但能大大减少参数量,随后使用最大池化操作将特征图的分辨率减半,此外在每个小模块中都将通道数加倍,以获取更多的高层语义特征。在解码阶段,采用与编码阶段相对称的设计,只是将最大池化操作替换为反卷积操作以恢复特征图的分辨率,同时,在每个小模块中将通道数减半。

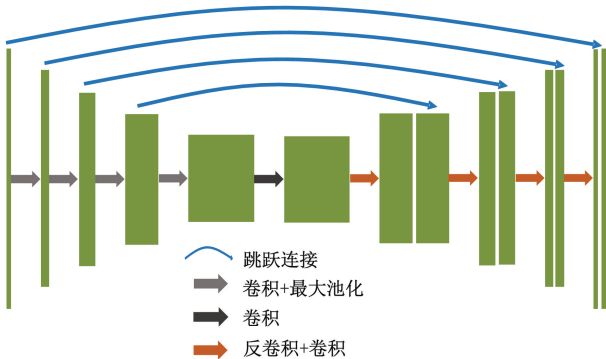


图2 基本 U-Net 模型结构

Fig. 2 Structure of the basic U-Net model

U-Net 网络中的核心操作是在编码层和解码层之间引入跳跃连接,以减少编码阶段进行的池化操作造成的底层特征信息的丢失。但即使在解码阶段对高层语义特征进行反卷积操作,也很难恢复底层特征所包含的空间信息^[18]。高层特征有助于像素分类,底层特征有助于生成精细的边界,跳跃连接通过将编-解码层中对应阶段的

底层细节特征和高层语义特征直接拼接起来,再通过卷积操作进行特征的压缩和融合,最终实现高质量的分割。

1.2 DRD U-Net 模型

虽然 U-Net 在生物医学图像领域取得了很好的效果,但仍有可改进的地方。本文提出的 DRD U-Net 如图3所示,为了实现特征复用和降低优化难度,将原本的卷积模块替换为残差模块^[19]。由于存在大小不一的分割目标,添加多尺度上下文特征提取模块^[20] (atrous spatial pyramid pooling, ASPP),利用具有不同感受野的卷积操作获取丰富的上下文信息。在多输出融合模块中融合不同 S_k 的输出图,其中 S_k 为跳跃连接和随后的特征融合压缩, $k \in [1, 2, 3, 4]$,实现多层次特征的互补,提高分割精度。

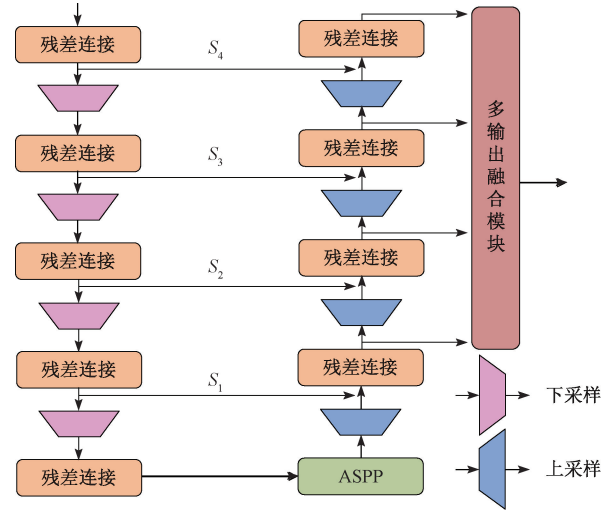


图3 本文提出的 DRD U-Net 模型结构

Fig. 3 Structure of the proposed DRD U-Net model

1.3 残差模块

CNN 的深度至关重要,增加网络深度可以提高网络的表现。但是,一味增加网络深度可能导致梯度消失/爆炸,以至于无法训练。同时也发现,仅仅增加层数的深层 CNN 的效果反而没有相应的浅层 CNN 的效果好,即 CNN 的性能随层数的增加出现了“退化”问题。为此,He 等^[19]提出了深度残差网络。

残差模块降低了神经网络的优化难度,使网络易于训练,而且残差连接有助于特征的复用,使深度网络在不退化的同时也提高了性能。图4(a)所示为 U-Net 中的基本模块,图4(b)所示为结合了 BN (batch normalization) 操作的残差模块,所使用的激活函数为 ReLu。由于在残差模块中发生了维度的变化,即输入的通道数和输出的通道数不匹配,无法直接相加,因此,使用 1×1 卷积完成通道数的变换,同时, 1×1 的卷积不引入过多的参数。

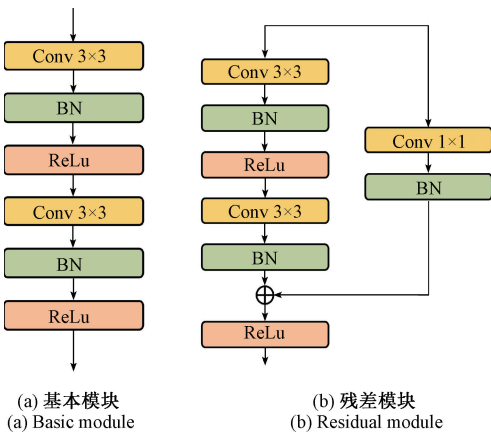


图4 本文采用的残差模块与基本模块的对比

Fig. 4 Comparison of the residual module used in this paper with the basic module

1.4 多尺度上下文特征提取模块

在 U-Net 网络中,使用最大池化对特征图做下采样操作,可以在保持较小卷积核的同时增大卷积操作的感受野,但也在一定程度上损失了空间位置信息。如果不下做下采样操作,而是使用较大的卷积核,会带来成倍的参数增长。因此,本文中引入了扩张卷积,如图5所示。图5(a)所示为卷积核大小为3,扩张率为1,步长为1的二维卷积,等价于常规卷积,此时的感受野为 3×3 区域,参数量为9,图5(b)所示为卷积核大小为3,扩张率为2,步长为1的二维卷积,除了蓝点位置,其余位置的权重均为0,虽然此时的参数量仍为9,但其感受野大小为 7×7 。

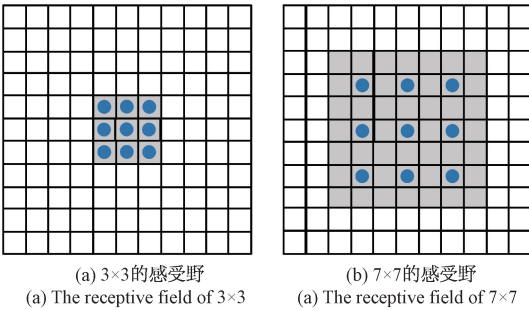


图5 扩张卷积的感受野

Fig. 5 The receptive field of dilated convolution

扩张卷积能够在不增加参数的情况下,扩大感受野,以获取更多的上下文信息。本文使用 ASPP 作为多尺度上下文信息提取模块,如图6所示。设 ASPP 模块的输入为 $x \in \mathbf{R}^{4c \times h \times w}$, ASPP 由4个并行的扩张卷积组成,其扩张率分别为 $[3, 5, 7, 9]$,最终的输出如式(1)所示,其中 $y_i \in \mathbf{R}^{c \times h \times w}$ 。

$$y = \text{concatenate}(y_1, y_2, y_3, y_4) \quad (1)$$

将设计的 ASPP 作为编码和解码阶段的桥梁模块,

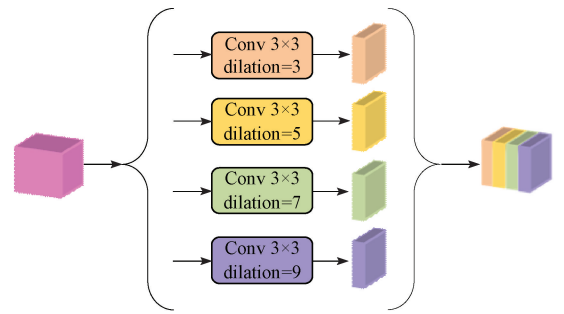


图6 扩张卷积金字塔结构

Fig. 6 The structure of ASPP

并联不同扩张率的扩张卷积以提取更多上下文特征,然后进行特征拼接,充分利用所提取的多尺度特征。

1.5 多输出融合模块

U-Net 网络中为了融合不同阶段的特征,总共使用4次跳跃连接。但在 U-Net 网络中仅在 S_4 阶段输出最后的预测分割图。 S_4 阶段融合了最底层的特征与经过融合压缩后的高层特征,具有良好的空间位置信息和像素类别信息。但在每个阶段,为了进行跳跃连接和控制模型参数,对高层特征进行压缩,丢失了部分特征信息。为此本文引入深监督^[21]以增强网络对不同层次特征的复用能力,设计多输出融合模块,如图7所示。

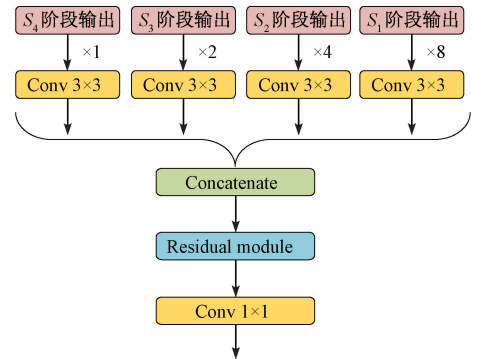


图7 深监督模块

Fig. 7 Deep supervision module

具体而言,对于 S_{1-3} 的输出,使用双线性插值进行上采样,倍数分别为8、4和2, S_4 则不需要上采样操作。随后,分别进行卷积核大小为3的卷积操作(BN层和 ReLu 激活层略去未注)。将卷积的结果拼接起来送入残差模块中,最终使用 1×1 卷积(不含 BN 层,激活函数为 *softmax*)输出预测的分割图,输出通道数为5(4类目标外加背景)。

2 实验与结果

实验中使用的是一台 CPU 为 E5-2678 v3、内存为

128 GB 和显卡为 4 路 2080Ti 的计算机,所有的算法均是在 PyTorch 框架下实现的,学习率为 0.01,学习率衰减系数为 0.5,优化器采用权重衰减为 0.000 1、动量系数为 0.9 的 SGD,同时使用跨 GPU 的同步 BN 操作,批处理大小为 20,预设的训练轮数为 50。

2.1 数据集

本文在公共数据集 OAI-ZIB^[13]上评估了所提出的算法,OAI-ZIB 数据集的详情如表 1 所示。数据集包含 507 套 MR 图像,每套包含 130 张 2D 切片,每张切片的分辨率为 384×384,由专家手动分割出膝关节的不同部位。在实验中将数据集划分为训练集、测试集和验证集,其比例分别约为 60% (300 套)、20% (100 套) 和 20% (107 套)。在验证集上使用早停机制以避免过拟合,在测试集上评估所提出的模型。

表 1 OAI ZIB 数据集详情

Table 1 Details of the OAI ZIB dataset

参数	数值
MR 类型	DESS
采集面	矢状面
分辨率/mm	0.36×0.36×0.7
数量/套	507
性别(男:女)	262:245
年龄/岁	61.87±9.33
骨关节炎分级 (0:1:2:3:4)	60:77:61:151:158

对于 CNN 来说,使用标准化的图像可以提高算法的鲁棒性。对于医学图像而言,其场景基本固定,因此本文统一将膝关节置于图像左侧。此外,由于 MR 图像扫描过程中的静态磁场不均匀等原因,得到的图像在同一组织的不同子区域经常有着不同的灰度值^[22],因此本文对图像进行了磁场偏差场校正,同时,在调整图像灰度至 [0~255] 后,采用保边平滑算法对图像做进一步处理^[23]。如图 8 所示,相比较于原始图像,预处理后的图像的边缘更加明显,软骨部分有更好的区分度。除此之外,本文没有进行其他常见的数据集增强算法,如垂直或水平翻转等。

2.2 损失函数

图像分割中最常见的损失函数是交叉熵 (cross entropy, CE)。然而,当类别极度不平衡时,使用交叉熵作为损失函数容易忽略小目标。在二值分割中,Milletari 等^[24]提出了一种新的基于 Dice 系数的损失函数,以减轻目标像素和背景像素之间的不平衡给网络性能带来的影响。因此,本文使用的损失函数 $\mathcal{L}$ 由两部分组成:



(a) 原始图像 (a) Original image
(b) 预处理后图像 (b) Pre-processed image

图 8 预处理前后对比图像

Fig. 8 Comparison between the original image and the pre-processed image

$$\mathcal{L} = l_{CE} + l_{Dice} \quad (2)$$

$$l_{CE} = -\frac{1}{N} \sum_{n=1}^N y_n \log(p_n) \quad (3)$$

$$l_{Dice} = -\frac{2 \sum_{n=1}^N y_n p_n + \varepsilon}{\sum_{n=1}^N y_n + \sum_{n=1}^N p_n + \varepsilon} \quad (4)$$

式中: l_{CE} 代表交叉熵损失; l_{Dice} 为扩展的多类别 Dice 损失; N 为标签二值图上的像素总个数; $y_n \in [0,1]$ 代表真值; $p_n \in [0,1]$ 代表网络预测的每个类别的概率值; ε 为常数(1×10^{-5}),避免分母为 0。

2.3 评价指标

在医学图像分割中,由于要分割的部位仅为图像的局部区域,而且医学图像本身也是三维体数据,因此本文采用基于区域和基于表面距离的评价指标^[25]。对于二维切片数据,基于区域的指标包含有戴斯系数 (dice similarity coefficient, DSC)、重叠误差 (volume overlap error, VOE) 和相对误差 (relative volume difference, RVD)。对于三维体数据,采用表面距离作为指标,包括平均表面距离 (average surface distance, ASD)、均方根表面距离 (root mean square surface distance, RMSD) 和豪斯多夫距离 (Hausdorff distance, HD)。在本文中,利用这些指标来评估所提出的模型与现有模型的性能。

$$DSC = \frac{2 \|Y \cap P\|_1}{\|Y\|_1 + \|P\|_1} \times 100 \quad (5)$$

$$VOE = \left(1 - \left\| \frac{P \cap Y}{P \cup Y} \right\|_1\right) \times 100 \quad (6)$$

$$RVD = \left| \frac{\|P\|_1 - \|Y\|_1}{\|Y\|_1} \right| \times 100 \quad (7)$$

式中: Y 代表手工分割的结果; P 代表网络预测的分割结果。

令 $S(Y)$ 和 $S(P)$ 分别为 Y 和 P 表面体素的集合,对

$\forall y \in S(Y)$ 到 $S(P)$ 的距离定义为 y 到 $S(P)$ 上所有点的欧氏距离的最小值,即:

$$d(y, S(P)) = \min_{p \in S(P)} \|y, p\|_2 \quad (8)$$

所以,

$$ASD = \frac{1}{c} (D^1(S(Y), S(P)) + D^1(S(P), S(Y))) \quad (9)$$

$$RMSD = \sqrt{\frac{D^2(S(Y), S(P)) + D^2(S(P), S(Y))}{c}} \quad (10)$$

$$HD = \max \{ \max_{y \in S(Y)} d(y, S(P)), \max_{y \in S(P)} d(y, S(Y)) \} \quad (11)$$

$$D^m(S(Y), S(P)) = \sum_{y \in S(Y)} d^m(y, S(P)) \quad (12)$$

$$c = \text{count}(S(P)) + \text{count}(S(Y)) \quad (13)$$

式中: c 为 P 和 Y 表面体素个数的总和; $\text{count}(\cdot)$ 表示计算集合内元素的个数。

2.4 分割结果定性分析

在相同的实验环境下,分别使用原始 U-Net 和 DRD U-Net 对测试集图像进行了分割,结果如图 9 所示。图 9(a) 和 (c) 显示出所提出的网络对骨骼的分割具有较高的分割精度,这是由于所设计的模块增大了卷积操作的感受野,获取了更多的上下文信息,大大减少了假阳性和假阴性区域,因此,提高了分割区域的连贯性和准确性。在图 9(b) 中, U-Net 预测的分割图出现了明显的股骨软骨断连, DRD U-Net 相对比较完整。而在图 9(d) 中, U-Net 预测的胫骨软骨区域也明显小于对应的 DRD U-Net 和手工分割,说明 DRD U-Net 提高了对小目标的分割能力。

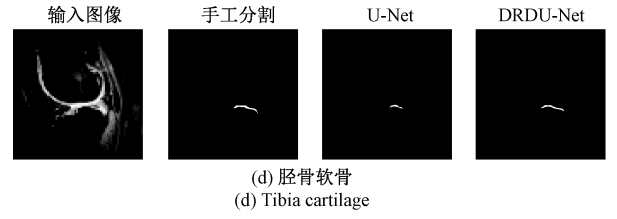


图 9 不同模型下自动分割结果的定性对比

Fig. 9 Qualitative comparison of auto-segmentation results with different models

此外,如图 10 所示,在图 10(a) 中手工分割者将软组织从股骨中标识出来,但在图 10(b) 中分割者只标识出部分软组织,这可能是由于手工标注时出现的人为误差所致。但是,无论是 U-Net 还是所提出 DRD U-Net 均能根据图像特征正确地分割出股骨区域,相比较于手工分割,其一致性较好。

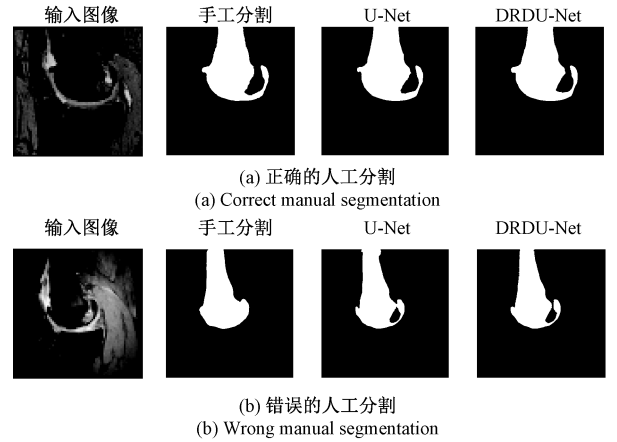


图 10 人工分割存在不一致的情况

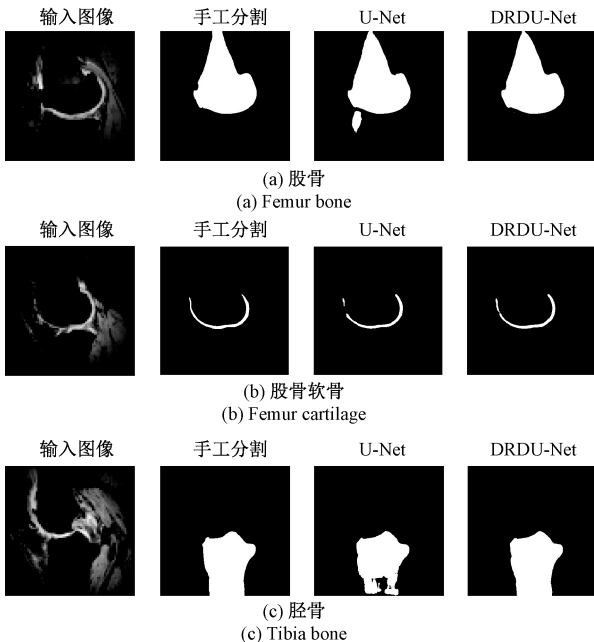
Fig. 10 The inconsistent condition in some manual annotations

2.5 分割结果定量分析

将本文所提出的算法与原始的 U-Net、U-Net++^[26] 和 Attention U-Net^[27] (简称 Att U-Net) 算法进行比较,后者都是最新的针对 U-Net 改进的端到端网络。网络参数遵循原论文,实验中所有模型采用统一的训练策略及超参数设置,其中初始滤波器的个数均为 32。结果如图 11 所示,具体数值如表 2 所示,加粗表示的是最优指标。可以看到,本文提出的 DRD U-Net 模型在总共 24 个指标中,有 23 个处于领先地位。在关键指标 DSC 和 HD 上都处于最优,相比较于原始网络 U-Net 平均提高了 0.3% 和减少了 3.69 mm 误差。

2.6 消融实验

实验中通过删除部分模块来探究删除的模块对网络性能的影响,即进行了消融实验。



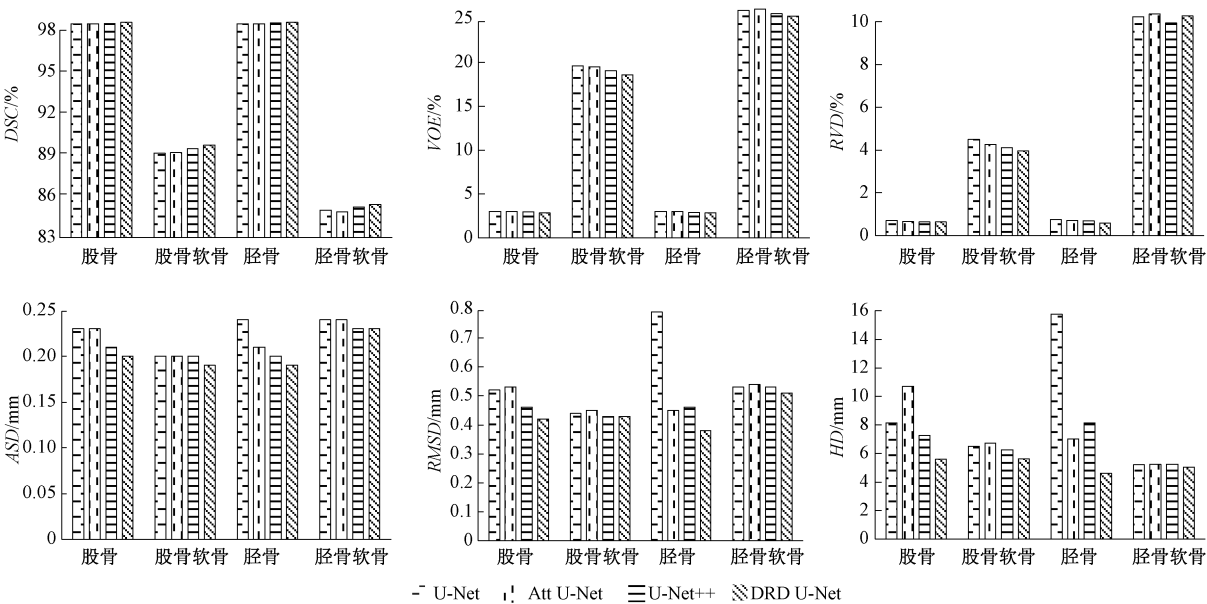


图 11 不同模型下自动分割结果的量化对比

Fig. 11 Quantitative comparison of auto-segmentation results with different models

表 2 不同模型的定量对比 (均值±方差)

Table 2 Quantitative comparison using different models (mean ± variance)

部位	模型	$DSC/\%$ (↑)	$VOE/\%$ (↓)	$RVD/\%$ (↓)	ASD/mm (↓)	$RMSD/\text{mm}$ (↓)	HD/mm (↓)
fb	U-Net	98.46±0.29	3.03±0.57	0.72±0.54	0.23±0.07	0.52±0.3	8.13±7.03
	U-Net++	98.51±0.31	2.93±0.59	0.64±0.51	0.21±0.06	0.46±0.16	7.28±5.29
	Att U-Net	98.46±0.31	3.03±0.6	0.67±0.53	0.23±0.06	0.53±0.24	10.71±9.34
	DRD U-Net	98.56±0.3	2.85±0.58	0.64±0.51	0.2±0.05	0.42±0.11	5.59±2.52
fc	U-Net	89.1±2.51	19.57±4.02	4.49±3.47	0.2±0.06	0.44±0.2	6.49±4.06
	U-Net++	89.42±2.58	19.05±4.14	4.11±3.1	0.2±0.06	0.43±0.19	6.27±3.77
	Att U-Net	89.15±2.63	19.48±4.21	4.27±3.3	0.2±0.06	0.45±0.21	6.72±5.63
	DRD U-Net	89.68±2.54	18.62±4.12	3.96±3.39	0.19±0.05	0.42±0.16	5.63±2.35
tb	U-Net	98.46±0.34	3.04±0.66	0.75±0.65	0.24±0.13	0.79±1.04	15.78±20.28
	U-Net++	98.53±0.34	2.89±0.66	0.68±0.58	0.2±0.06	0.46±0.28	8.13±9.93
	Att U-Net	98.47±0.35	3.01±0.68	0.71±0.61	0.21±0.07	0.45±0.21	7.04±7.78
	DRD U-Net	98.57±0.34	2.81±0.65	0.58±0.56	0.19±0.06	0.38±0.11	4.62±2.44
tc	U-Net	84.96±4.13	25.93±6.28	10.22±6.81	0.24±0.08	0.53±0.23	5.22±2.17
	U-Net++	85.19±4.08	25.58±6.22	9.94±6.57	0.23±0.08	0.53±0.21	5.26±2.24
	Att U-Net	84.86±4.19	26.08±6.35	10.37±6.96	0.24±0.08	0.54±0.23	5.23±2.08
	DRD U-Net	85.38±4.13	25.28±6.27	10.27±7.19	0.23±0.07	0.51±0.2	5.04±2.14

为了验证图像预处理对网络性能的影响,分别将原始图像和预处理过的图像作为实验数据,在基线模型 U-Net 上进行评估,前者称为 U-Net,后者称为 U-Net w/pre。原始图像由窗位为 200 和窗宽为 400 的窗函数获得,预处理图像获取流程如 2.1 节所示。为了探究网

络的参数量对网络性能的影响,将 U-Net 的初始滤波器的个数增加至 64,称为 U-Net w/64,采用预处理后的图像作为训练数据。各个部位分割指标的均值和各网络的参数量如表 3 所示。

从表 3 中可以得到,经过图像预处理以后,各个指标

表 3 基线模型上的消融实验

参数	U-Net	U-Net w/pre	U-Net w/64	DRD U-Net
DSC/(↑)	92. 75	92. 80	92. 86	93. 05
VOE/(↓)	12. 89	12. 80	12. 70	12. 40
RVD/(↓)	4. 05	3. 95	3. 90	3. 86
ASD/mm(↓)	0. 23	0. 22	0. 22	0. 20
RMSD/mm(↓)	0. 57	0. 50	0. 50	0. 43
HD/mm(↓)	8. 91	7. 59	7. 69	5. 22
参数量/M	9. 80	9. 80	28. 90	14. 30

均有所改善。将 DRD U-Net 与 U-Net w/64 比较可以看到,前者的参数量只有后者的 1/2 左右,但每个评估指标均优于后者,表明了单纯增加滤波器个数并不能有效提高网络性能。

此外,本文提出的 DRD U-Net 也进行了消融实验。在基本 U-Net 模型的基础上,将增加了残差模块的模型称为 Res U-Net,在残差模块基础上增加了 ASPP 模块的模型称为 DR U-Net,本文所提出的新模型是 DRD U-Net。实验结果如表 4 所示,所使用的模块均能逐步提高分割精度。

表 4 所提出模型上的消融实验

参数	Res U-Net	DR U-Net	DRD U-Net
DSC/(↑)	92. 87	92. 97	93. 05
VOE/(↓)	12. 68	12. 53	12. 40
RVD/(↓)	3. 92	3. 80	3. 86
ASD/mm(↓)	0. 22	0. 21	0. 20
RMSD/mm(↓)	0. 49	0. 44	0. 43
HD/mm(↓)	6. 87	5. 29	5. 22

3 结 论

膝关节 MR 图像中存在着不同大小的分割目标,为精准分割带来了挑战。本文在 U-Net 深度卷积神经网络的基础上,提出了端到端的 DRD U-Net。通过使用残差模块增强了特征的复用能力并降低了模型的优化难度,利用扩张卷积模块 ASPP 获取不同尺度的感受野以适应不同大小的分割目标,设计的多输出融合的深监督模块能够直接利用不同层次的特征实现信息互补。消融实验表明,所使用的残差模块、ASPP 模块和深监督模块均能逐步提高膝关节 MR 图像的分割精度。在公开数据集上评估了本文算法和其他的最新分割算法。实验结果的对

比分析表明,本文算法在膝关节股骨、胫骨、股骨软骨、胫骨软骨的分割方面都取得了更好的分割精度。

参考文献

[1] 薛庆云,王坤正,裴福兴,等. 中国 40 岁以上人群原发性骨关节炎患病状况调查[J]. 中华骨科杂志, 2015, 35(12): 1206-1212.

XUE Q Y, WANG K ZH, PEI F X, et al. The survey of the prevalence of primary osteoarthritis in the population aged 40 years and over in China[J]. Chinese Journal of Orthopaedics, 2015, 35(12): 1206-1212.

[2] LI X, MAJUMDAR S. Quantitative MRI of articular cartilage and its clinical applications[J]. Journal of Magnetic Resonance Imaging, 2013, 38(5): 991-1008.

[3] 雷静桃,唐明瑶,王君臣,等. 机器人辅助膝关节置换术的术前规划研究综述[J]. 机械工程学报, 2017, 53(17): 78-91.

LEI J T, TANG M Y, WANG J CH, et al. Review of the preoperative planning of robot assisted knee arthroplasty[J]. Journal of Mechanical Engineering, 2017, 53(17): 78-91.

[4] HEIMANN T, MORRISON B J, STYNER M A, et al. Segmentation of knee images: A grand challenge[C]. Proceeding MICCAI Workshop on Medical Image Analysis for the Clinic, 2010: 207-214.

[5] TAMEZ-PENA J G, FARBER J, GONZALEA P C, et al. Unsupervised segmentation and quantification of anatomical knee features: data from the osteoarthritis initiative [J]. IEEE Transactions on Biomedical Engineering, 2012, 59(4): 1177-1186.

[6] SEIM H, KAINMUELLER D, LAMECKER H, et al. Model-based auto-segmentation of knee bones and cartilage in MRI data[J]. Proceeding Medical Image Analysis for the Clinic: A Grand Challenge, in conjunction with MICCAI, 2010: 215-223.

[7] ZHANG K, LU W, MARZILANO P. Automatic knee cartilage segmentation from multi-contrast MR images using support vector machine classification with spatial dependencies[J]. Magnetic Resonance Imaging, 2013, 31(10): 1731-1743.

[8] COURTIOL P, MAUSSION C, MOARII M, et al. Deep learning-based classification of mesothelioma improves prediction of patient outcome [J]. Nature Medicine, 2019, 25(10): 1519-1525.

[9] 戴相昆,王小深,杜乐辉,等. 基于三维 U-NET 深度卷积神经网络的头颈部危及器官的自动勾画[J]. 生

- 物医学工程学杂志, 2020, 37(1): 136-141.
- DAI X K, WANG X SH, DU L H, et al. Automatic segmentation of head and neck organs at risk based on three-dimensional U-NET deep convolutional neural network[J]. Journal of Biomedical Engineering, 2020, 37(1): 136-141.
- [10] PRASOON A, PETERSEN K, IGEL C, et al. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2013: 246-253.
- [11] LEE H, HONG H, KIM J. BCD-NET: A novel method for cartilage segmentation of knee MRI via deep segmentation networks with bone-cartilage-complex modeling[C]. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018: 1538-1541.
- [12] ZHOU Z, ZHAO G, KIJOWSKI R, et al. Deep convolutional neural network for segmentation of knee joint anatomy[J]. Magnetic Resonance in Medicine, 2018, 80(6): 2759-2770.
- [13] AMBELLAN F, TACK A, EHLKE M, et al. Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the osteoarthritis initiative[J]. Medical Image Analysis, 2019, 52: 109-118.
- [14] 张永德, 彭景春, 刘罡, 等. 基于水平集的前列腺磁共振图像分割方法研究[J]. 仪器仪表学报, 2017, 38(2): 416-424.
- ZHANG Y D, PENG J CH, LIU G, et al. Research on the segmentation method of prostate magnetic resonance image based on level set[J]. Chinese Journal of Scientific Instrument, 2017, 38(2): 416-424.
- [15] 胡正平, 谭营. 基于SVDD的交互式区域增长图像分割算法[J]. 仪器仪表学报, 2006, 27(S3): 2114-2115.
- HU ZH P, TAN Y. Interactive region growing for image segmentation based on SVDD[J]. Chinese Journal of Scientific Instrument, 2006, 27(S3): 2114-2115.
- [16] GRAU V, MEWES A U J, ALCANIZ M, et al. Improved watershed transform for medical image segmentation using prior information[J]. IEEE Transactions on Medical Imaging, 2004, 23(4): 447-458.
- [17] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015: 234-241.
- [18] DROZDZAL M, VORONTSOV E, CHARTRAND G, et al. The importance of skip connections in biomedical image segmentation[M]. Deep Learning and Data Labeling for Medical Applications, Berlin: Springer, 2016: 179-187.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [20] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 801-818.
- [21] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [22] LIKAR B, VIERGEVER M A, PERCUS F. Retrospective correction of MR intensity inhomogeneity by information minimization[J]. IEEE Transactions on Medical Imaging, 2001, 20(12): 1398-1410.
- [23] BONARETTI S, GOLD G E, BEAUPRE G S. PyKNEEr: An image analysis workflow for open and reproducible research on femoral knee cartilage[J]. Plos one, 2020, 15(1): e0226501. DOI:10.1101/556423.
- [24] 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), IEEE, 2016: 565-571.
- [25] TAHA A A, HANBURY A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool[J]. BMC medical imaging, 2015, 15(1): 29.
- [26] ZHOU Z, SIDDIQUEE M M R, TAJBAKSH N, et al. Unet++: A nested u-net architecture for medical image segmentation[M]. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Berlin: Springer, 2018: 3-11.
- [27] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention u-net: Learning where to look for the pancreas[J]. arXiv preprint arXiv, 1804.03999, 2018: 1-10.

作者简介



于宁波 (通信作者), 2003 年于天津大学获得学士学位, 2005 年于香港科技大学获得硕士学位, 2011 年于苏黎世联邦理工学院获得博士学位, 现为南开大学人工智能学院副教授, 主要研究方向为医疗康复机器人。

E-mail: nyu@nankai.edu.cn

Yu Ningbo (Corresponding author) received his B. Sc. degree from Tianjin University in 2003, M. Sc. degree from the Hong Kong University of Science and Technology in 2005, and Ph. D. degree from ETH Zurich in 2011. He is currently an associate professor in the College of Artificial Intelligence at Nankai University. His main research interests include medical and rehabilitation robotics.



刘嘉男, 2018 年于重庆邮电大学获学士学位, 现为南开大学人工智能学院硕士研究生, 主要研究方向为医学图像配准与分割。

E-mail: ljnnzb@mail.nankai.edu.cn

Liu Jianan received his B. Sc. degree from

Chongqing University of Posts and Telecommunications in 2018. He is currently a M. Sc. candidate in the College of Artificial Intelligence at Nankai University. His main research interests include medical image registration and segmentation.



韩建达, 1990 年于西安交通大学获得学士学位, 1994 年于中国科学院沈阳自动化研究所硕士研究生毕业 (直博), 1998 年于哈尔滨工业大学获得博士学位, 现为南开大学杰出教授, 主要研究方向为机器人自主行为与人机协作/共融方法, 医疗康复机器人、地面移动及飞行器机器人技术与系统。

E-mail: hanjianda@nankai.edu.cn

Han Jianda received his B. Sc. degree from Xi'an Jiaotong University in 1990, M. Sc. degree from Institute of Automation, Chinese Academy of Sciences in 1992, and Ph. D. degree from Harbin Institute of Technology in 1998. He is currently an outstanding professor at Nankai University. His main research interests include robot autonomy and human-robot coordination, medical and rehabilitation robotics, mobile and flying robotics.