

多尺度空洞 U-Net 网络的电影 CMR 图像分割

王建军, 李婉晴*, 张 敏

(河北经贸大学 信息技术学院, 河北 石家庄 050061)

摘 要:在电影心脏核磁共振(CMR)图像上准确分割左心室、右心室和心肌是心脏功能评估和诊断的重要步骤。然而,大多数带标注的 CMR 图像数据量较少,无法满足训练需求,同时 CMR 图像中心脏结构复杂,心室及心肌边界不清晰,导致分割效果欠佳。因此,该文提出了一种基于迁移学习和多尺度空洞 U-Net 网络的 CMR 图像分割方法,使用迁移学习,将预训练模型得到的网络参数迁移到目标模型上作为目标模型的初始化参数,提高网络的特征学习能力,解决 CMR 图像数据量不足的问题;在 U-Net 网络中引入多尺度空洞卷积模块,使用空洞卷积代替普通卷积在参数不变的情况下扩大了感受野,并且采用多尺度特征融合提取更加精细的特征,解决 CMR 图像边界曲线欠分割的问题。实验结果表明,该方法能有效实现心脏中左心室、右心室和心肌的准确分割,平均 Dice 系数和 Hausdorff 距离平均值分别为 0.902 和 4.219 mm,对比其他网络分割模型明显提高了分割精度。

关键词:心脏核磁共振图像;U-Net 网络;迁移学习;多尺度空洞卷积;特征融合

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2022)11-0177-06

doi:10.3969/j.issn.1673-629X.2022.11.026

Multi-scale Dilated U-Net Network for Cine CMR Image

WANG Jian-jun, LI Wan-qing*, ZHANG Min

(School of Information Technology, Hebei University of Economics and Business, Shijiazhuang 050061, China)

Abstract: Accurate segmentation of the left ventricle, right ventricle and myocardium on cine cardiac magnetic resonance (CMR) images is an important step in cardiac function assessment and diagnosis. However, most CMR images with annotation have less data to meet the training requirements, while the complex heart structures and unclear boundaries of ventricles and myocardium in CMR images lead to poor segmentation results. Therefore, we propose a CMR image segmentation method based on transfer learning and multi-scale dilated U-Net network. According to transfer learning, the network parameters obtained from the pre-trained model are migrated to the target model as the initialization parameters of the target model to improve the feature learning ability of the network and solve the problem of insufficient data volume of CMR images. The multi-scale dilated convolution module is introduced in the U-Net network. The use of dilated convolution instead of ordinary convolution expands the receptive field with the same parameters, and multi-scale feature fusion is used to extract finer features to solve the problem of under-segmentation of CMR image boundary curves. The experiment shows that the proposed method can effectively achieve accurate segmentation of the left ventricle, right ventricle and myocardium in the heart, and the average Dice correlation coefficient and Hausdorff distance averages are 0.902 and 4.219 mm, respectively, which significantly improve the segmentation accuracy compared with other network segmentation models.

Key words: cardiac magnetic resonance image; U-Net network; transfer learning; multi-scale dilated convolution; feature fusion

0 引言

心血管疾病(cardiovascular disease, CVD)是全球死亡率最高的疾病,心脏影像学是评估和诊断心脏疾病的常规手段,心脏磁共振成像(cardiac magnetic resonance, CMR)被认为是一种无创评估心血管系统功能和结构的方法,对 CMR 图像切片进行病理分析是目前诊断心脏疾病的常用方法^[1]。精确分割左心室

(left ventricle, LV)、右心室(right ventricle, RV)和心肌(myocardium, Myo)是对图像进行病例分析的关键环节,对心血管疾病的诊断和治疗具有重要的意义^[2-4]。但是由于 CMR 图像中心脏与周围组织的对比度差、轮廓不清晰、噪声大及呼吸运动伪影等影响,极大地增加了分割心室和心肌的难度。针对 CMR 图像分割的传统方法有活动轮廓^[5]、水平集^[6]和图割方

收稿日期:2021-11-26

修回日期:2022-03-29

基金项目:河北省高等学校科学研究计划项目-科学技术研究青年基金项目(QN2019069)

作者简介:王建军(1965-),男,教授,硕导,研究方向为网络安全、大数据处理算法和计算机应用;通讯作者:李婉晴(1996-),女,硕士研究生,CCF 会员(E1950G),研究方向为医学图像处理。

法^[7]。这些传统的方法过多依赖先验知识,并且存在一定的局限性,很难提高 CMR 分割的准确性。

近年来,基于深度学习的方法在图像分割领域取得了很大的进展。Abdeltawab 等人^[8]使用 FCN (fully convolutional network, FCN)^[9]用于分割,提出了一种用于分割短轴图的密集连接卷积网络。他们采用一种有效的基于 FCN 的方法从心脏 MRI 中提取左心室感兴趣区域(ROI),以缓解类别不平衡的问题,并采用另一种 FCN 基于 ROI 进行左心室腔和心肌分割。此外引入了一个称为径向损失的新损失函数,以最小化 LV 的预测轮廓和真实轮廓之间的距离,取得了很好的分割结果。然而,该网络的下采样过程会导致细节信息丢失,对小物体的分割效果不佳。Cui 等人^[10]使用 U-Net^[11]网络架构,引入图像金字塔和深度监督输出层全自动学习关注各种大小和形状的目标结构,有效提高了心脏分割的准确性。然而,该网络在上采样时会造成图像中心室和心肌边缘细节丢失,存在分割差的问题。

该文提出一个基于迁移学习和多尺度空洞 U-Net 网络 (transfer learning and multi-scale dilated U-Net networks, TLMD-U-Net) 用于左心室、右心室和心肌的分割,利用迁移学习方法将 Xception 模型^[12]在 ImageNet 数据集上训练好的权重信息迁移到目标网络中,解决了 CMR 图像标记数据量少的问题。在 U-Net 网络的下采样中引入多尺度空洞卷积模块 (multi-scale dilated convolution block, MDC-Block) 扩大了感受野,提取了丰富的多尺度 CMR 图像信息;上采样阶段拼接下采样过程中每个池化层前的全部卷积层提取的特征,避免 U-Net 网络仅拼接每个池化层前最后一层特征造成的细节信息丢失问题,提高了 CMR 图像中左心室、右心室和心肌边缘分割的精确度。

1 网络结构和算法原理

1.1 迁移学习

在处理医学分割问题时,数据量不足是限制医学分割研究的一个重要方面^[13]。由于数据需要医学专家手工标记非常耗时耗力,往往获得的数据集仅有几百张或几千张训练图片,难以满足训练需求。然而,大量的数据集能提取更多的信息特征,泛化能力强,较小的数据集容易产生过拟合现象。因此,该文使用迁移学习^[14],将从源数据集上学习到的权值转移到数据量不足的目标网络,对目标数据集进行训练,然后使用一些标记的目标域数据对预训练的目标网络进行微调,使目标域获得更好的学习和表现。

该文采用图 1 所示的迁移学习方法实现 CMR 图像的分割网络。但是由于 ImageNet 数据集中的图像

与 CMR 图像特征差别较大,网络往往不能通过先验知识学习到较多的 CMR 图像特征,所以使用 ImageNet 预先训练好的权值对模型进行初始化后,将预训练目标网络模型的层再使用目标域数据进行逐层微调,更新和优化这些层的参数,加快网络的收敛速度,提高网络的特征学习能力,避免直接对整个网络进行微调产生较大的计算量。迁移学习中的权重初始化和微调过程如图 2 所示。

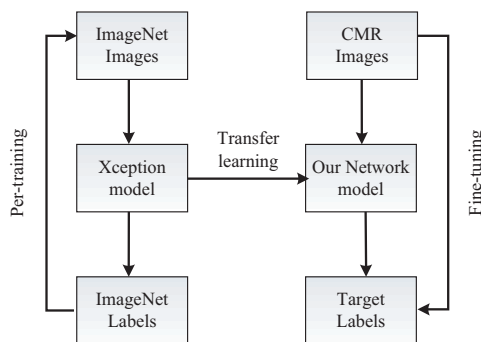


图 1 迁移学习

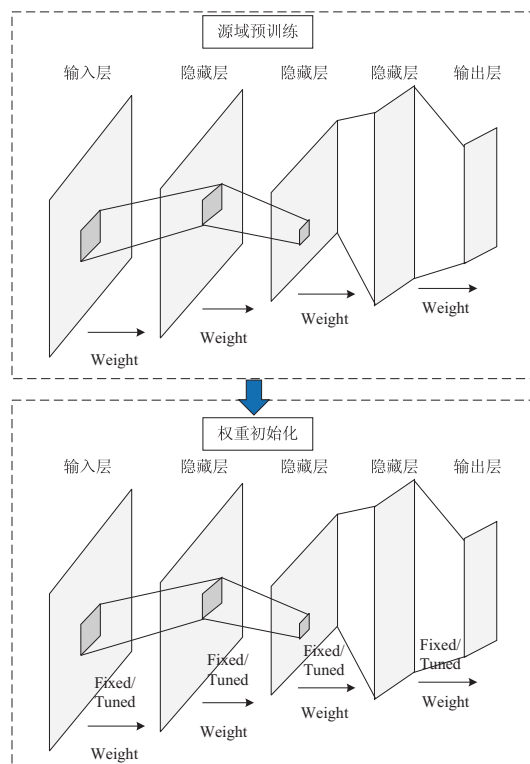


图 2 权重初始化和微调过程

1.2 MD-U-Net 模型

采用 U-Net 网络分割图像时,其编码阶段采用单尺度卷积核提取特征,容易造成细节信息的遗漏;解码阶段跳跃连接只拼接了每个池化层前最后一个卷积层的信息特征,且该层相对其他层分辨率相对较小,保留的信息少,限制了模型性能的提高。因此,该文使用多尺度空洞 U-Net 网络模型 (multi-scale dilated U-Net networks, MD-U-Net),引入多尺度空洞卷积模块,将空洞卷积及多尺度思想引入神经网络,将编码阶段不同

大小膨胀率的空洞卷积提取的信息融合成多尺度特征传播到解码阶段。使用空洞卷积代替普通卷积扩大了感受野,并且将池化层前的所有卷积层的特征进行融合传入解码阶段,有效解决了 CMR 图像切片边界曲线欠分割的问题。

1.2.1 空洞卷积

空洞卷积^[15]在普通卷积中增加权重值为零的空洞,在保持参数不变的同时增加特征信息的感受野,超参数膨胀率 $rate$ 控制着感受野的大小,膨胀率越大感受野越大。如图 3 所示,膨胀率为 $[1,2,3]$ 的空洞卷积,等效的核大小分别是 $[3,5,7]$ 。对于膨胀率为 r 的空洞卷积,实际卷积核计算如公式(1)所示:

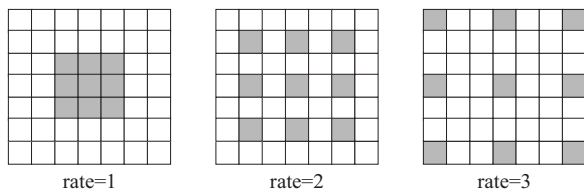


图 3 空洞卷积

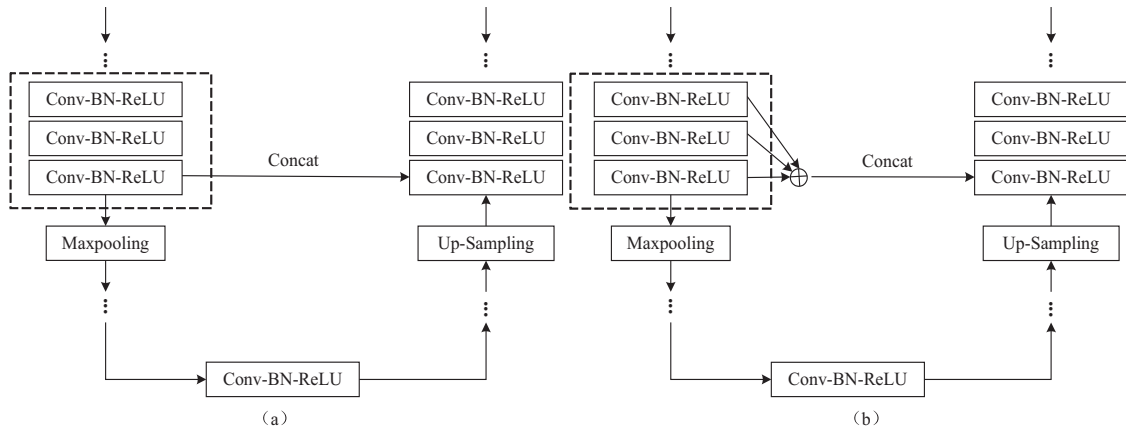


图 4 多尺度特征融合

1.2.3 多尺度空洞卷积模块

多尺度空洞卷积模块 (multi-scale dilated convolution block, MDC-Block) 融合了空洞卷积和多尺度提取特征的优点,首先使用 $r=1$ 、 $r=2$ 和 $r=5$ 三种膨胀率的 3×3 空洞卷积依次进行特征提取,小膨胀率的空洞卷积的感受野大小为 3×3 ,用来获取 CMR 图像的细节特征信息,大膨胀率的空洞卷积感受野大小为 11×11 ,用来获取更具有全局性的特征信息;然后将浅层特征和深层特征进行特征融合,使得编码阶段提取到的特征更加精细准确,解码阶段跳跃连接引入的信息更加完整,提高了 CMR 图像左心室、右心室和心肌的分割精确度;最后将融合的特征采用 1×1 卷积恢复特征图通道维度。多尺度空洞卷积模块结构如图 5 所示。

1.2.4 MD-UNet 网络结构

MD-UNet 网络的整个架构如图 6 所示,使用编码

$$k_r = k + (k - 1) \times (r - 1) \quad (1)$$

其中, k 表示原始卷积核大小, k_r 表示扩张后的卷积核大小。通过设置不同大小膨胀率的空洞卷积能够捕获特征图中的多尺度信息,这有助于对大目标和小目标进行精确分割。

1.2.2 多尺度特征融合

在深度学习神经网络提取特征信息时,随着网络逐渐加深,感受野变大,提取的特征信息更具有全局性。U-Net 网络编码阶段经过 4 次降采样操作,每次降采样池化层前经过多层卷积操作,解码阶段在解池化操作中通过跳跃连接仅拼接了每次池化操作前最深一层的特征图信息(如图 4(a)所示),忽略了包含更大尺寸以及具有更加丰富细节信息的浅层特征,容易造成细节信息的遗漏,影响 CMR 图像分割效果^[16]。为了弥补这一不足,该文将编码阶段每个池化操作前提取的浅层特征和深层特征进行特征融合一起传入到解码阶段(如图 4(b)所示),使得网络提取更加精细的特征,提高 CMR 图像边界分割的精确度。

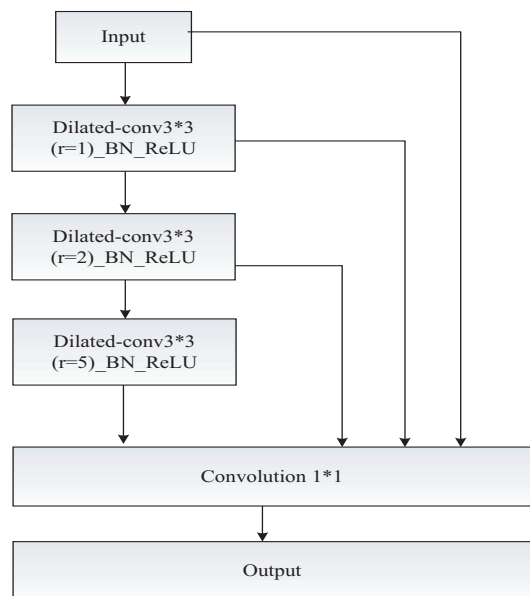


图 5 多尺度空洞卷积模块

器-解码器作为基本网络。编码阶段用于捕获 CMR 图像的上下文信息,解码阶段使用双线性插值方法逐步恢复图像的分辨率大小,且通过跳跃连接与编码阶段同一阶段的特征图进行拼接,减少网络下采样多次卷积池化造成的信息损失,对图像的感兴趣区域精准定位,从而实现左右心室及心肌的精确分割。

网络的编码阶段由多尺度空洞卷积模块和池化层交替构成,每个多尺度空洞卷积模块包括三个不同膨胀率($r=1$ 、 $r=2$ 和 $r=5$)的空洞卷积,并且将模块中卷积层提取的信息进行特征融合,再使用 1×1 的卷积恢复特征图通道维度,能够获得更多 CMR 图像的细节信息,使用批归一化层增加模型的训练速度,提高网络分割精度,并采用 ReLU 激活函数加强网络模型的表达能力^[17]。解码阶段包括卷积块、上采样和 Softmax 输出,通过跳跃连接将每一阶段编码器与解码器进行特征图拼接,使得网络获得更多感兴趣区域(ROI)的特征信息,提高分割准确率,最后使用 Softmax 对像素分类以完成最终的 CMR 图像分割。

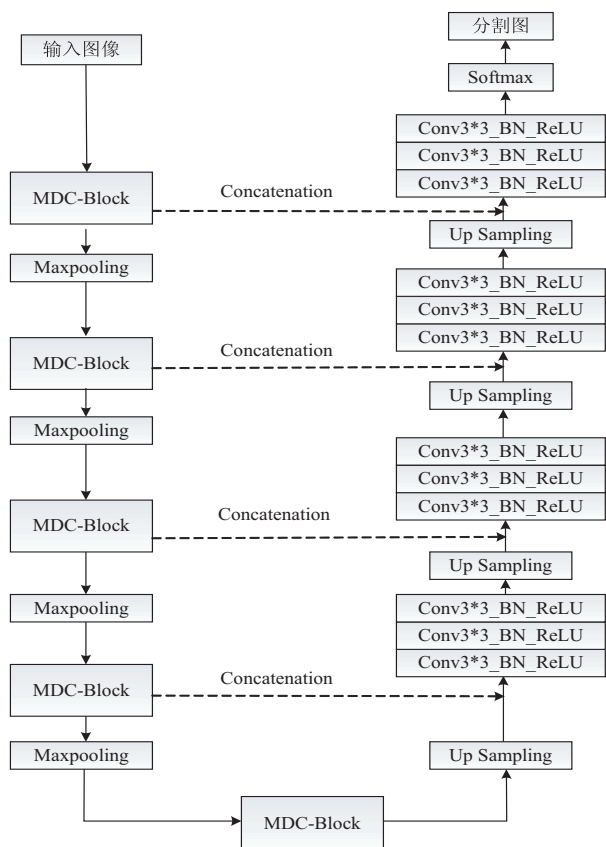


图 6 MD-UNet 网络

2 损失函数

该文采用交叉熵损失函数和 Dice 损失函数的混合损失函数作为分割心室和心肌的损失函数。Dice 损失函数用于衡量两个样本的重叠程度,但是其梯度计算形式比较复杂,反向传播时容易梯度剧烈变化,导

致训练不稳定。交叉熵损失函数梯度计算简单,但不能解决前景类与背景类之间的不平衡问题,并且在分割多类输入图像时,分割性能不佳^[18-19]。将两者损失函数相结合,可以有效提高分割过程中的训练稳定性和分割精度。交叉熵损失函数和 Dice 损失函数如式(2)和式(3)所示:

$$L_{CE} = -\frac{1}{N} \sum_{i \in N} \sum_{c \in M} g_i^c \log(s_i^c) \quad (2)$$

$$L_{Dice} = -\frac{1}{M} \sum_{c \in M} \frac{\sum_{i \in N} g_i^c s_i^c}{\sum_{i \in N} g_i^c + \sum_{i \in N} s_i^c} \quad (3)$$

其中, M 是类别数目, N 是每个批次中的像素集, g 是标签图像的 one-hot 值, s 是预测样本为类别 c 的概率值。CMR 图像分割的混合损失函数如式(4)所示:

$$L = L_{CE} + L_{Dice} \quad (4)$$

3 实验部分

3.1 实验数据集和实验环境

实验数据由 55 例病人的短轴 CMR 图像数据组成,其数据存储为 Nifti 格式,短轴切片约 8~10 个,相邻切片之间的距离为 10 mm~20 mm。每个 Nifti 文件包含一个病人心脏不同切面的 25 帧影像数据,每个病例由 200 多张图片组成。33 例 CMR 图像数据为训练集,11 例为验证集,11 例为测试集。图 7 展示了一个病人的一张原始图像和对应的标签图像。

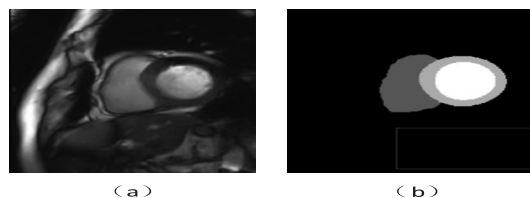


图 7 原始图像(a)和标签图像(b)

实验使用的深度学习框架是 pytorch, torch == 1.2.0, torchvision == 0.4.0 环境,编译器为 Pycharm,并使用 NVIDIA GeForce RTX2080S (8G) 进行训练加速,采用 windows10 操作系统,编程语言 python3.7,每次训练时间约为 12 个小时,采用 Adam 优化器,学习率为 0.01, batch size 为 16, epoch 为 300, 动量参数为 0.9。

3.2 数据预处理

首先提取 ROI 区域,选择一个正方形区域,包括必要的心脏和不可避免的近邻。通过傅里叶分析提取第一谐波图像,在每一个切片相应的心跳频率上捕获最大的活动。假设左心室近似于一个圆,对第一次谐波图像进行 canny 边缘检测。利用圆形霍夫变换方法从边缘映射中计算出左心室的近似半径和中心^[20]。进而裁剪出感兴趣区域,将图像一致调整成 128 *

128。然后将裁剪后的数据归一化,避免像素灰度范围分布较大影响分割效果。另外,使用数据增强对数据进行扩充,防止模型发生过拟合,方法是对输入的 CMR 原图像及其对应的金标准图像进行平移、缩放、水平和垂直翻转。

3.3 评价指标

为了检测模型的分割效果,该文使用 Dice 系数、召回率 (recall)、精确率 (precision) 和 Hausdorff 距离 (HD) 作为评价指标。

Dice 系数表示两个样本重叠程度,取值范围是 $[0,1]$ 。指数越接近于 1,分割结果越好。Dice 系数如公式(5)所示:

$$\text{Dice} = \frac{2TP}{2TP + FP + FN} \quad (5)$$

Hausdorff 距离 (HD) 表示两个样本之间的对称距离,HD 数值越小,方法越好。Hausdorff 距离如公式(6)所示:

$$\text{HD} = \max \left\{ \max_{a \in A} \min_{m \in M} (a, m), \max_{m \in M} \min_{a \in A} (a, m) \right\} \quad (6)$$

精确率 (precision) 表示预测是正的样本被预测正确的概率。精确率如公式(7)所示:

$$\text{precision} = \frac{TP}{TP + FP} \quad (7)$$

召回率 (recall) 表示样本中正样本被预测正确的概率。召回率如公式(8)所示:

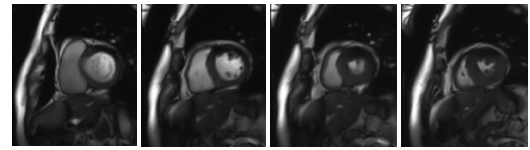
$$\text{recall} = \frac{TP}{TP + FN} \quad (8)$$

其中,TP 表示预测是正例实际上也是正例的样本,FP 表示预测是正例实际上是负例的样本,FN 表示预测是负例实际上是正例的样本, a 、 m 分别是 A 、 M 上的像素点, d 是 a 和 m 之间的距离。

3.4 实验结果及分析

基于改进的 U-Net 网络对 CMR 图像中的左心室、右心室以及心肌的分割结果如图 8 所示,其中(a)表示输入的原始 CMR 图像,里面包含了心脏以及除心脏以外的多个组织器官,(b)和(c)分别表示待分割图像对应的专家标注的标签图像和文中模型的预测分割结果。如图 8 所示,文中算法能准确地从 CMR 图像中分割出心室和心肌,预测结果接近专家手动分割图像。

为了进一步评价文中算法的分割性能,使用 Dice 系数、精确率 (precision)、召回率 (recall) 和 Hausdorff 距离 (HD) 四个指标分别计算左心室 (LV)、右心室 (RV) 与心肌 (Myo) 的分割结果,并且使用 Dice 系数和 Hausdorff 距离两个评价指标将文中算法与其他算法进行了比较。



(a) 原始 CMR 图像



(b) 专家手动分割图像



(c) 文中算法分割图像

图 8 不同 CMR 图像切片分割结果

表 1 为改进的 U-Net 网络模型使用交叉熵和 Dice 组合的混合损失函数对左心室、右心室以及心肌的分割结果。可以看出,左心室的平均 Dice 系数高于心肌和右心室,平均 Hausdorff 距离左心室最低,右心室最高,左心室分割效果好,可能是因为左心室与周围组织之间相对右心室与心肌来说对比比较强烈。右心室有着最低 Dice 系数和最高 Hausdorff 距离,可能和右心室不规则形状以及边界不清晰有关。

表 1 网络分割评估均值

评价指标	Dice	precision	recall	HD/mm
左心室	0.933	0.955	0.928	3.121
心肌	0.889	0.891	0.905	4.452
右心室	0.883	0.883	0.890	5.085
平均	0.902	0.909	0.907	4.219

文中算法与 U-Net、DeepLab 和研究心脏分割的文献[17,21]方法进行了比较,每种算法都使用相同的数据集与损失函数,采用 Dice 系数和 Hausdorff 两种指标评估 CMR 图像中左心室、右心室和心肌的分割精度。结果如表 2 所示,相比于其他实验分割算法,文中算法的分割效果最好,平均 Dice 系数是 0.902,平均 Hausdorff 距离是 4.219 mm。说明文中算法存在一定优势,分割效果比较理想。

表 2 文中算法与其他分割算法对比

model	Dice			HD/mm		
	LV	Myo	RV	LV	Myo	RV
U-Net	0.927	0.871	0.856	4.662	5.235	6.672
DeepLab	0.926	0.868	0.869	4.782	5.251	6.043
Liu ^[17]	0.930	0.873	0.872	3.896	5.070	5.221
Li ^[21]	0.928	0.888	0.865	4.215	5.167	6.390
Ours	0.933	0.889	0.883	3.121	4.452	5.085

4 结束语

针对医学图像标注数据量少难以满足研究需求的问题,使用迁移学习方法,将 Xception 模型在 ImageNet 数据集上训练得到的网络权重参数和结构对目标网络进行迁移,提高目标网络的特征学习能力。针对 U-Net 网络丢失边缘细节信息问题,加入多尺度空洞卷积模块,提高了 CMR 图像切片边界的分割精度。实验结果表明,该网络模型比较完整地保留了 CMR 图像的细节特征信息,能够有效实现左心室、右心室以及心肌的准确分割。

参考文献:

- [1] FU Y B, LEI Y, WANG T H, et al. A review of deep learning based methods for medical image multi-organ segmentation[J]. *Physica Medica*, 2021, 85: 107-122.
- [2] PENSO M, MOCCIA S, SCAFURI S, et al. Automated left and right ventricular chamber segmentation in cardiac magnetic resonance images using dense fully convolutional neural network[J]. *Computer Methods and Programs in Biomedicine*, 2021, 204: 106059.
- [3] KHENED M, ALEX V, KRISHNAMURTHI G, et al. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers[J]. *Medical Image Analysis*, 2018, 51(1): 21-45.
- [4] 王 慧, 王丽嘉. 基于心脏电影磁共振图像的左心肌分割新进展[J]. *中国生物医学工程学报*, 2020, 39(2): 238-246.
- [5] BAI W, SHI W, LEDIG C, et al. Multi-atlas segmentation with augmented features for cardiac MR images[J]. *Medical Image Analysis*, 2015, 19(1): 98-109.
- [6] LIU Y, CAPTUR G, MOON J C, et al. Distance regularized two level sets for segmentation of left and right ventricles from cine-MRI[J]. *Magnetic Resonance Imaging*, 2016, 34(5): 699-706.
- [7] GROGEORGE D, PETITJEAN C, DACHER J N, et al. Graph cut segmentation with a statistical shape model in cardiac MRI[J]. *Computer Vision and Image Understanding*, 2013, 117(9): 1027-1035.
- [8] ABDELTAWAB H, KHALIFA F, FT D, et al. A deep learning-based approach for automatic segmentation and quantification of the left ventricle from cardiac cine MR images[J]. *Computerized Medical Imaging and Graphics*, 2020, 81(1): 101717.
- [9] LONG J, SHELHAMER E, DARRELL T, et al. Fully convolutional networks for semantic segmentation [C]//2015 IEEE conference on computer vision and pattern recognition (CVPR). Berkeley: IEEE, 2015: 640-651.
- [10] CUI H F, CHANG Y W, JIANG L, et al. Multiscale attention guided U-Net architecture for cardiac segmentation in short-axis MRI images[J]. *Computer Methods and Programs in Biomedicine*, 2021, 206(1): 106142.
- [11] 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. Switzerland: Springer International Publishing, 2015: 234-241.
- [12] CHOLLET F. Xception: deep learning with depthwise separable convolutions [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Honolulu: IEEE, 2017: 21-26.
- [13] 黎 英. 迁移学习在医学图像分析中的应用研究综述[J]. *计算机工程与应用*, 2021, 57(20): 42-52.
- [14] ZHU W, BRAUN B, CHIANG L H, et al. Investigation of transfer learning for image classification and impact on training sample size[J]. *Chemometrics and Intelligent Laboratory Systems*, 2021, 211(7639): 104269.
- [15] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [16] 张小娟, 汪西莉. 完全残差连接与多尺度特征融合遥感图像分割[J]. *遥感学报*, 2020, 24(9): 1120-1133.
- [17] LIU D, JIA Z, JIN M, et al. Cardiac magnetic resonance image segmentation based on convolutional neural network[J]. *Computer Methods and Programs in Biomedicine*, 2020, 197(45): 105755.
- [18] TAGHANAKI S A, ZHENG Y, ZHOU S K, et al. Combo loss: handling input and output imbalance in multi-organ segmentation[J]. *arXiv*: 1805.02798v1, 2018.
- [19] 唐 柳, 王晓东, 陈哲彬, 等. 基于 Octave 卷积的超声心动图左心室分割方法[J]. *计算机应用*, 2020, 40(S1): 215-219.
- [20] AMMAR A, BOUATTANE O, YOUSSEFI M, et al. Automatic cardiac cine MRI segmentation and heart disease classification[J]. *Computerized Medical Imaging and Graphics*, 2021, 88: 101864.
- [21] LI Z, LOU Y, YAN Z, et al. Fully automatic segmentation of short-axis cardiac MRI using modified deep layer aggregation [C]//2019 IEEE 16th international symposium on biomedical imaging (ISBI). Venice: IEEE, 2019: 8-11.