

## 综合应用Faster R-CNN和U-net的心脏MRI图像分割

韩俊玲, 李博, 康晓东, 杨靖怡, 刘汉卿, 王笑天

### 引用本文

韩俊玲, 李博, 康晓东, 杨靖怡, 刘汉卿, 王笑天综合应用Faster R-CNN和U-net的心脏MRI图像分割[J]. 计算机科学, 2023, 50(6A): 220600047-9.

HAN Junling, LI Bo, KANG Xiaodong, YANG Jingyi, LIU Hanqing, WANG Xiaotian. [Cardiac MRI Image Segmentation Based on Faster R-CNN and U-net](#) [J]. Computer Science, 2023, 50(6A): 220600047-9.

---

### 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

#### Similar articles recommended (Please use Firefox or IE to view the article)

#### [基于多特征融合的GRU-LSTM大学生就业动态预测](#)

College Students Employment Dynamic Prediction of Multi-feature Fusion Based on GRU-LSTM  
计算机科学, 2023, 50(6A): 220500056-6. <https://doi.org/10.11896/jsjcx.220500056>

#### [基于深度学习的超高频标签识别系统](#)

Tag Identification for UHF RFID Systems Based on Deep Learning

计算机科学, 2023, 50(6A): 220200151-6. <https://doi.org/10.11896/jsjcx.220200151>

#### [CT影像阶段化目标检测方法研究](#)

Study on Phased Target Detection in CT Image

计算机科学, 2023, 50(6A): 220200063-10. <https://doi.org/10.11896/jsjcx.220200063>

#### [基于深度学习的摩托车车道实时检测](#)

Real-time Detection of Motorcycle Lanes Based on Deep Learning

计算机科学, 2023, 50(6A): 220200066-5. <https://doi.org/10.11896/jsjcx.220200066>

#### [基于改进YOLOv5的电动车头盔佩戴检测算法](#)

Electric Bike Helment Wearing Detection Alogrithm Based on Improved YOLOv5

计算机科学, 2023, 50(6A): 220500005-6. <https://doi.org/10.11896/jsjcx.220500005>

# 综合应用 Faster R-CNN 和 U-net 的心脏 MRI 图像分割

韩俊玲<sup>1</sup> 李博<sup>2</sup> 康晓东<sup>1</sup> 杨靖怡<sup>1</sup> 刘汉卿<sup>1</sup> 王笑天<sup>1</sup>

<sup>1</sup> 天津医科大学医学影像学院 天津 300202

<sup>2</sup> 天津医科大学三中心临床学院 天津 300170

(hjl19893381889@163.com)

**摘要** 为解决现有 MRI 神经网络分割中存在因输入端图像信息多样导致分割精度下降的问题,提出了引入 Faster R-CNN 和 U-net 机制的 MRI 图像分割方法。选择公开心脏 MRI 分割挑战赛数据集 ACDC 和 SCD,清洗和修改数据集格式后送入后续神经网络。首先,应用 Faster R-CNN 对目标图像进行检测,以对原始输入图像进行预处理,并去掉冗余的背景信息。其次,对预处理后的图像进行 U-net 分割,同时为检验引入 Faster R-CNN 后,对分割网络的性能和精度是否提高,采用了消融实验和对比实验。消融实验去掉了 U-net 分割网络中的检测裁剪模块,选择 U-net 及其改进网络分别做一组消融实验结果。实验结果表明,新方法的平均交并比和 Dice 系数在 ACDC 数据集上为 0.89 和 0.94,分别提高了 7.3% 和 5%,在 SCD 数据集上为 0.96 和 0.98,分别提高了 5% 和 3%,实现了 MRI 图像的自动预处理和分割。

**关键词:** U-net;Faster R-CNN;MRI;分割算法;深度学习

中图分类号 TP391

## Cardiac MRI Image Segmentation Based on Faster R-CNN and U-net

HAN Junling<sup>1</sup>, LI Bo<sup>2</sup>, KANG Xiaodong<sup>1</sup>, YANG Jingyi<sup>1</sup>, LIU Hanqing<sup>1</sup> and WANG Xiaotian<sup>1</sup>

<sup>1</sup> School of Medical Image, Tianjin Medical University, Tianjin 300202, China

<sup>2</sup> Department of the Third Central Clinical College of Tianjin Medical University, Tianjin 300170, China

**Abstract** In order to solve the problem that the segmentation accuracy of the existing MRI neural network is reduced due to the diversity of input image information. An MRI image segmentation method using Faster R-CNN and U-net mechanism is proposed. Selecting the public cardiac MRI segmentation challenge datasets ACDC and SCD, cleaning and modifying the format of the dataset and sending them to the subsequent neural network. First, Faster R-CNN is applied to target image detection to preprocess the original input image and remove redundant background information. Second, performing U-net segmentation on the pre-processed images. At the same time, in order to test whether the performance and accuracy of the segmentation network are improved after the introduction of Faster R-CNN, ablation experiments and comparison experiments are conducted. In the ablation experiment, the detection and cropping module in the U-net segmentation network is removed, and the U-net and its improved network are selected to do a set of ablation experiments respectively. Experiments show that the average intersection ratio and Dice coefficient of the new method is 0.89 and 0.94 on the ACDC dataset, respectively, which is 7.3% and 5% higher. On the SCD dataset, it is 0.96 and 0.98, which is 5% and 3% higher, respectively. Automatic preprocessing and segmentation of MRI images is achieved.

**Keywords** U-net, Faster R-CNN, MRI, Segmentation algorithm, Deep learning

## 1 引言

临床上,磁共振成像(Magnetic Resonance Imaging, MRI)技术以无电离辐射、多方位成像、软组织敏感的特点弥补了 X 线成像技术的短板,被广泛应用于生殖系统、心脏、肌肉、脑等软组织成像。为满足临床需求,诊断医生要从 MRI 图像中快速识别和准确分割病灶。但医生主观分割的结果不具有可复用性,且由于医疗资源紧张, MRI 图像数据量大,导致诊断医生手动分割任务繁重<sup>[1]</sup>。图像分割技术可以辅助诊断医生完成工作,以更高的效率完成病灶分割和定性、定量分析<sup>[2]</sup>。

传统医学图像分割技术遵循人-机结合的交互式方法,从

技术所利用算法的性质不同,分为基于区域法、边缘分割法和结合区域和边界的分割方法。但无论是哪种算法,分割机制都需要人为设置,且在实际应用中具有依赖性和局限性。

随着计算机硬件设备的日益革新, GPU 的运算性能得到了极大的提高,深度学习技术在近几年内迅猛发展,并促进了以计算机为单一执行者的自动分割技术的长足发展。深度学习是机器学习的一种,相比以往需要专家设计特征的“特征工程”,深度学习具有自主特征学习的特点。模拟人的神经网络,使得机器自主学习,深度学习成为了近年来人工智能领域的研究热点,在医学图像分割领域得到了高度关注和广泛应用<sup>[3]</sup>。卷积神经网络(Convolutional Neural Network, CNN)

基金项目:京津冀协同创新项目(17YEXTZC00020)

This work was supported by the Beijing Tianjin Hebei Collaborative Innovation Project(17YEXTZC00020).

通信作者:康晓东(kxd2004@126.com)

是深度学习的一种。基于 CNN 的图像分割方法,是目前应用度最高的深度学习分割方法之一<sup>[4]</sup>。CNN 的广泛使用促进了医学图像智能辅助诊断的发展。一种将基本 CNN 的输出视为后续 CNN 附加信息源的级联架构,用于脑肿瘤分割,在 2013 年脑肿瘤分割(Brain Tumor Segmentation, BraTS)数据集上进行训练,相比其他架构速度提高了 30 倍(如 Havaei 等提出的架构)<sup>[5]</sup>。2015 年,Long 等基于 CNN 提出了全卷积神经网络(Fully Convolutional Networks, FCN)<sup>[6]</sup>,用全卷积层替换了 CNN 中的全连接层,实现了随机图片大小输入的语义分割任务。Christ 等<sup>[7]</sup>提出了级联全卷积神经网络(Cascaded Fully Convolutional Neural Networks, CFCNs),用于分割肝脏及肝脏病灶的 CT(Computed Tomography)图像, Dice 系数超过 94%,计算时间低于每体积 100 s。同时,FCN 也具有上采样获取语义信息少、计算性能低的缺点。2015 年 Ronneberger 等提出的 U-net<sup>[8]</sup>解决了 FCN 读取像素位置和上下文信息遗漏造成语义特征丢失的问题<sup>[9]</sup>。U-net 网络采用 U 形镜像结构,编码器和解码器直接跳跃连接,融合了图像细节和全局信息,适用于医学图像分割。一种多阶 U-net 网络模型被提出,用于甲状腺结节和腺体的自动分割,对于边界相对模糊的甲状腺结节图像,取得了较好的分割结果<sup>[10]</sup>。

与自然图像相比,心脏 MRI 图像具有容易受成像设备、成像条件和不同患者等的影响,标准的标注数据少的特点<sup>[11]</sup>。近几年,针对心脏 MRI 成像部位,提出了许多基于 U-net 改进的图像分割方法。Ammar 等<sup>[12]</sup>提出了用于心脏 MRI 图像分割的组合分类器 U-net 变体。分类器集合由 3 个分类器组合而成,即多层感知器、随机森林和支持向量机。数据输入神经网络前进行了 ROI 裁剪。SK-Unet 网络通过在编码器添加挤压和激励残差(Squeeze-And-Excitation Residual, SE-Res)模块,在解码器中添加选择性核来优化原始 U-net 模型。SE-Res 模块应用注意力机制来增强信息特征提取并抑制冗余信息。SK-Unet 网络应用于心脏 MRI 分割,并在 MICCAI 2019 挑战赛中获得比赛第一名<sup>[13]</sup>。强调分割目标特征,弱化冗余的背景信息,在一定程度上增加了分割精度。

除了修改 U-net 网络框架,图像预处理也不失为一种减弱无关背景信息的方法。Liu 等对心脏 MRI 数据集做了手动裁剪预处理,以解决不同设备采集造成图像尺寸不一和背景噪声较大造成网络计算性能降低的问题<sup>[14]</sup>。但手动裁剪处理需要较多的时间和精力。传统的目标区域检测方法,如霍夫变换和 Canny 边缘检测,具有数据集类型选择性,是依赖于人工设计的“特征工程”,特征提取和区域分类效果低,导致检测精度不高<sup>[15]</sup>。Liu 等<sup>[16]</sup>提出了依赖于两阶段 U-net 框架和自适应阈值窗口的心脏自动分割策略。两阶段 U-net 框架包括心脏感兴趣区域(Region of Interest, ROI)检测和心脏边缘的精细分割。自适应阈值窗口方法用于删除数据噪声,同时保留各区域之间的解剖关系。在 MM-WHS 挑战赛(2017)的数据集上进行实验,实验结果表明,心脏分割的 Dice 系数为 79.3%。U-net 网络的分割对象是像素点,重复使用,增加了时间成本。

基于上述内容,本文提出了一种引入 Faster R-CNN 和 U-net 机制的 MRI 医学影像分割方法,以解决医学图像常伴有的目标信息少、背景噪声大的问题。首先,本文采用预先训练过的目标检测模型 Faster R-CNN 对 MRI 数据集中的病灶进行目标检测,将该网络检测出的病灶进行裁剪,裁剪后的图像

作为送入分割网络的原始数据集,将分割结果和未裁剪图像的分割效果进行对比。数据仿真表明,对于心脏 MRI 数据集, Dice 系数(Dice Coefficient)和平均交并比(mean Intersection over Union, mIoU)分别可达 0.94 和 0.89。

## 2 Faster R-CNN 和 U-net 网络模型

本文综合应用的网络模型主要包括 Faster R-CNN 和 U-net 模型。

### 2.1 Faster R-CNN

Ren 等<sup>[17]</sup>于 2016 年提出了 Faster R-CNN 目标检测模型。Faster R-CNN 是一种两阶段检测算法,第一阶段先从原图中生成多个检测框,第二阶段是对每个检测框内的物体分类,同时精细化检测框。如图 1 所示, Faster R-CNN 主要包括提取特征、区域候选网络、目标区池化和目标分类 4 个步骤。

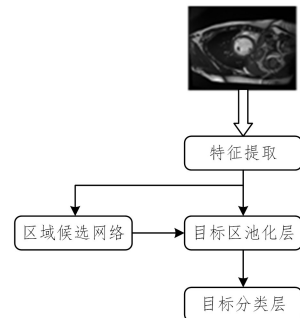


图 1 Faster R-CNN

Fig. 1 Faster R-CNN

#### (1) 提取特征(Convolutional Layer)

提取特征步骤的主体是深度学习网络,基本卷积层、激活函数和池化层组合用于获得图像特征,将图像语义特征输入到后续的区域候选网络和目标区池化。

#### (2) 区域候选网络(Region Proposal Network, RPN)

RPN 层有生成和筛选目标检测框的作用。特征图中每个点生成 9 个矩形框,包含 3 种长宽比。矩形框作为初始检测框,包含检测目标或背景两种可能性分类。通过 Softmax 函数对图像像素点分类,判断其标签类型,产生多个初级目标候选区域框。为了便于 Softmax 函数分类,分类函数前后都设置有重塑层,先变换矩形框维度,便于 Softmax 分类,再恢复原尺度。

为使候选框坐标尽可能接近原始目标检测框,提出了边框回归策略,用于遍历所有候选框,生成精确的候选区域。候选框输入到边框回归函数,采用预测与初始目标框的差距作为监督指标,取得每个矩形框的平移量和变换尺度,最终修正候选框的位置。

#### (3) 目标区池化(ROI Pooling)

目标区池化步骤主要是接收 RPN 层的输出和提取特征层提取的图像原始特征,集合两者信息,得到精确候选区域的特征,并通过映射处理将 RPN 层输出的候选区域尺度统一化,输入目标分类层。

#### (4) 目标分类(Classification)

目标分类步骤的主要作用是再次实现检测框分类和精确定位。通过全连接层和 Softmax 函数识别出候选框的所属类别,与 RPN 层的分类效果不同,目标分类步骤对检测框类别进行分类。对候选框实行二次边框回归处理,生成预测框。

## 2.2 U-net

U-net 是 Ronneberger 等<sup>[8]</sup>于 2015 年提出的以全卷积神经网络为基础的语义分割模型。

如图 2 所示,U-net 网络模型主要由编码器(Encoder)和解码器(Decoder)两部分组成。从编码器上提取出来的局部像素特征在解码器过程中与新的特征图(feature map)进行结合,是 U-net 实现精准定位的主要方法。该方法将图像局部信息和图像全局信息整合,以提高像素点类别预测的准确性。

### (1)编码器

编码器包含 5 个卷积块,单个卷积块有两个卷积层。每个  $3 \times 3$  的卷积层后面采用 ReLU 激活函数对图片进行下采样操作。下采样特征提取模块,提取具有较高水平语义信息的图像特征。前 4 个卷积块中,两个卷积层后有一个  $2 \times 2$  的最大池化层,每次池化操作后都会增加一个特征图通道数。

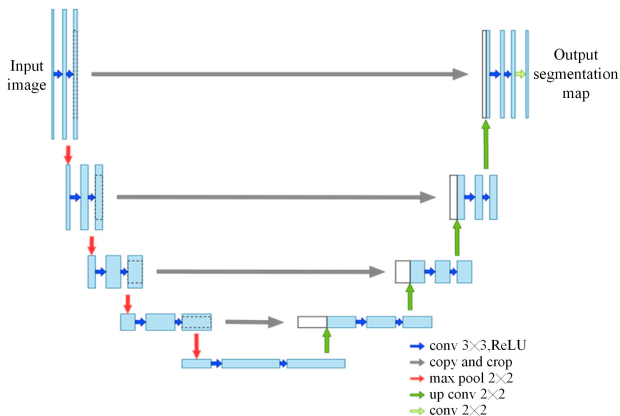


图 2 U-net 网络结构

Fig. 2 Structure of U-net

### (2)解码器

经过编码器对输入图像进行特征提取后,需要将图像尺寸恢复到最初大小,以便于实现进一步的像素分类,称该过程为上采样。上采样中使用反卷积方法恢复图像尺寸,减小图像特征通道数。每一步都有  $2 \times 2$  的反卷积层和两个  $3 \times 3$  的卷积层,激活函数仍然是 ReLU,每一步的上采样都加入了来自相对应收缩路径的特征图。

解码器后面的输出层,是由  $2 \times 2$  大小的卷积层组成,输出通道和分割类别一致,若分割类别有两种,则将特征图映射成二维输出图。

## 3 本文方法

由于心脏 MRI 图像中的心脏目标区域小,深度卷积神经网络提取语义信息较难,影响最终的语义分割效果。因此,本文的心脏 MRI 图像分割综合应用了 Faster R-CNN 和 U-net 机制,自动检测出心脏区域后去除背景信息,再精确分割。具体步骤是:首先,引入 Faster R-CNN 目标检测模型,以对原始输入图像进行预处理,并去掉冗余的背景信息;其次,对预处理后的图像进行 U-net 分割,具体实现包括数据检测裁剪预处理和图像分割两个基本步骤。

本文选用 Faster R-CNN 模型对原始心脏 MRI 图像做裁剪操作。基于卷积神经网络的 U-net 模型是一种端到端的图像分割方法,输入图像的质量可以直接影响网络的学习情况和分割结果。因此往往在输入数据集到 U-net 网络前,需要对其做裁剪等去无关背景信息的操作。Faster R-CNN 模块

能实现自动精确检测裁剪操作。Faster R-CNN 模块中,本文使用 50 层残差网络提取图像语义信息;设计区域候选网络对每个像素点分别生成 9 个矩形框,判别矩形框的标签类别,同时 RPN 中使用边框回归策略精细化初始目标检测框;提出目标区池化层用于提取更多特征图信息,特征图语义信息与初始检测框结合,并统一化检测框尺寸;最后识别检测框类别并实现预测框确定。

通过预训练,得到一个具有对心脏 MRI 图像的心脏区域有检测功能的 Faster R-CNN 网络。Faster R-CNN 网络模块与 U-net 串联,将 Faster R-CNN 模型的输出图像经过裁剪后作为 U-net 分割网络的输入图像。Faster R-CNN 模型批量预测心脏 MRI 图像,获取每个目标检测框的坐标,依据坐标对原始图像做裁剪操作。裁剪后的图像组合成新的数据集,在新数据集的训练集上训练 U-net 分割网络。

为检验本文实验方法的可行性和泛化性,采用对比实验。对比实验中,相同的数据集不引入 Faster R-CNN 对其做预处理工作,直接将数据集输入 U-net 网络,得到分割结果评价指标。另一组对比实验则在不同分割网络中加入 Faster R-CNN 模块,进一步展现经过 Faster R-CNN 图像预处理后的图像分割模型具有更好的效果。分析两组对比实验中的分割结果,最终得到本文方法在分割精度和效率上的优势。

## 4 数据集及其处理

### 4.1 数据及其预处理

本文方法选用心脏自动诊断挑战(Automated Cardiac Diagnosis Challenge, ACDC)数据集、SCD(Sunnybrook Cardiac Data)数据集和来自天津医科大学三中心临床学院的数据集,对数据集中心脏区域做目标检测后送入分割网络模型。SCD 是在 2009 年提出的左心室 MRI 分割挑战赛数据集,包含 45 名不同身体状况患者的心脏 MRI 信息。原始数据集格式为 DICOM 格式,经过图像清洗,本文实验使用了 805 张二维图片。

ACDC 收集了 100 名患者的心脏 MRI 数据作为训练集,每位患者的数据集包含舒张末期(End Diastole, ED)和收缩末期(End Systole, ES)两个时相的 MRI 图像,以及专家标记的左心室、右心室和心肌分割蒙版图。本文将数据集定义为 4 种标签:左心室(Left Ventricle, LV)区域定义为标签 1,右心室(Right Ventricle, RV)区域定义为标签 2,心肌(myocardium)区域定义为标签 3,背景区域则定义为标签 0。ACDC 的分割任务是分割出较为精准的左右心室和心肌区域。

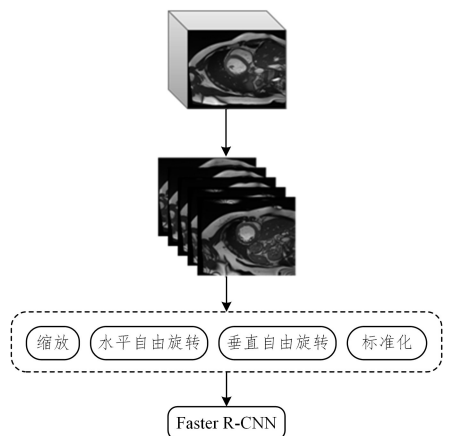


图 3 数据预处理的流程

Fig. 3 Data preprocessing process

原始 ACDC 数据集中单个心脏 MRI 图像和其标签图像均是三维图,保存为 nii 格式。如图 3 所示,本文实验先将三维图像切片为二维图,清洗掉标签图中没有标注出心脏的多余图像。最后得到 1841 张 2D 心脏核磁图像,取其中 920 张图像作为训练 Faster R-CNN 网络的数据集。其余图像经训练好的 Faster R-CNN 检测出目标区域,并按照 7:2:1 的比例划分为训练集、验证集和测试集。

为防止网络过拟合,提高泛化性,数据送入分割网络模型前,经过了增广操作。对训练集图像进行缩放、水平自由翻转、垂直自由翻转后再进行标准化。同时,考虑到不同患者、不同心脏收缩时相、不同层面的心脏 MRI 图像和其标签图有序对应,本文对标签图像也做了数据增广处理。

#### 4.2 对图像裁剪分割操作

本文方法引入了 Faster R-CNN 网络对原始数据去冗余

处理,输入 U-net 分割网络的是心脏 MRI 图像的心脏区域裁剪图像。U-net 网络是像素级别的语义分割模型,能实现任意尺寸的图像分割。U-net 网络在上采样过程中,利用反卷积层,将下采样得到的特征图的尺寸还原,最后实现每个像素分类,达到分割的目的。对于目标区域小、成分复杂的医学 MRI 影像,去除背景特征提取的冗余过程,直接将目标区域输入 U-net,卷积层得到的只有心脏的位置特征和特征直接的位置关系。该方法加快了模型分割速度,且提高了心肌和心室区域的分割精度。

本文方法的实验网络模型主要分为裁剪和分割两部分,如图 4 所示。心脏轴位核磁成像要求尽量将心脏位置放在成像视野中央,同时,为防止卷褶伪影的发生,成像区域要适当扩大。最终,获取的心脏 MRI 图像会包含一定的非心脏组织结构。

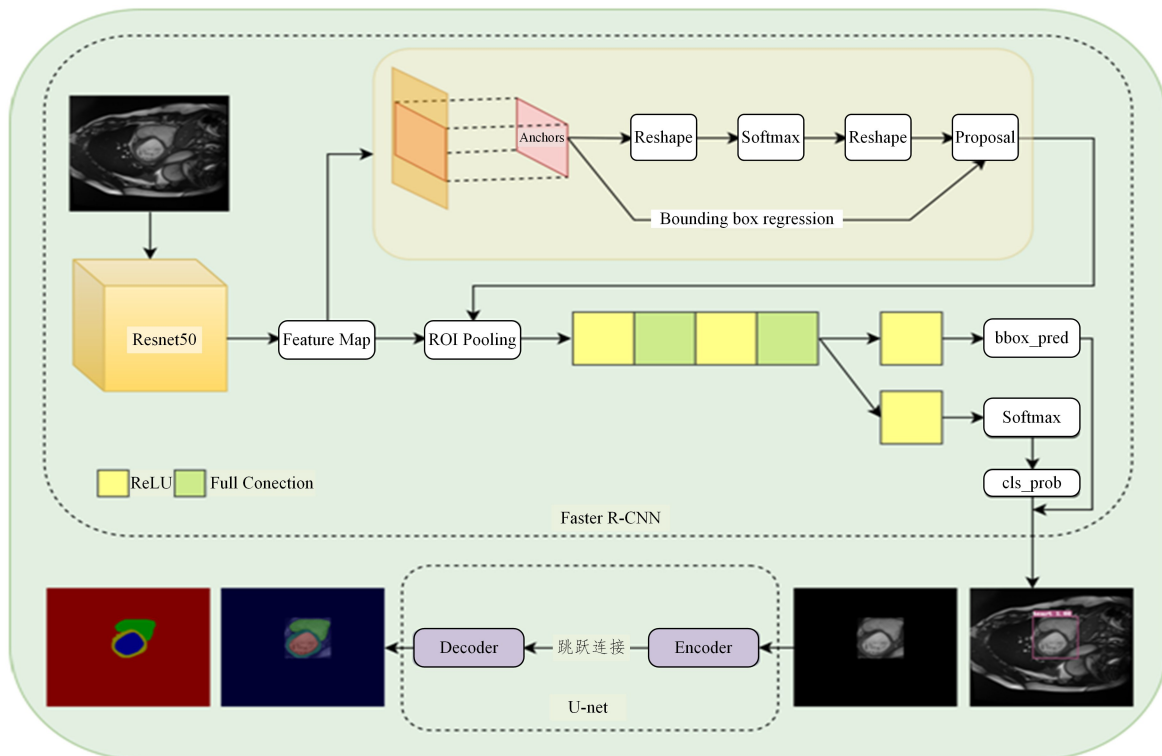


图 4 引入 Faster R-CNN 和 U-net 的语义分割网络

Fig. 4 Semantic segmentation networks with the introduction of Faster R-CNN and U-net

在心脏分割任务中,这些背景信息不仅会消耗算力,也为准确分割带来了困难。图 4 中的裁剪模块,是引入 Faster R-CNN 对包含多种语义信息的原始图像做预处理,不改变图像尺寸,检测出心脏区域,统一背景信息。而在分割模板中,引用经典 U-net 语义分割网络,编码器和解码器在图像通道的升维和降维过程中相互跳跃连接。网络的输入端是数据增强处理后的裁剪图和与其对应的标签图像,语义分割任务完成后输出最终分割结果。

##### (1) 裁剪模块

如图 4 所示,以原始数据集中的标签图为基础,生成心脏检测框,并保存为可扩展标记语言(Extensible Markup Language, XML)格式文件。这里,心脏 MRI 图片和 XML 文件集合制作成符合 Faster R-CNN 目标检测模型的数据集。

残差网络模型由 He 等<sup>[8]</sup>提出,具有捷径连接(Shortcut Connection)的结构特点,解决了 CNN 梯度随着网络加深

消失的问题。本文实验中,为了既增加卷积层数,又提高对心脏区域特征提取的准确度,选择 50 层的残差网络作为 Faster R-CNN 的卷积神经网络。

残差是统计学名词,指预测值和观测值间的差距。残差映射中,一层网络的输出  $y = x_{i+1}$  视为预测值;捷径连接获取上一层网络的输出  $x_i$ , 视为观测值,残差部分即为观测值与预测值的差。式(1)表示残差网的逻辑过程。

$$x_{i+1} = x_i + F(x_i, W_i) \quad (1)$$

其中,  $x_i$  是直接映射,  $F(x_i, W_i)$  是残差部分,反映在图 5 旁支。残差块包含 ReLU 和 BN(Batch Normalization)两种激活函数。ReLU 激活函数模拟整流器,用于激活神经网络,主要优点是加快网络模型的收敛速度<sup>[19]</sup>。BN 结构使卷积层的输入遵循高斯分布,泛化训练中初始参数和学习率的设置<sup>[11]</sup>。

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (2)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (3)$$

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (4)$$

$$y_i = \gamma \hat{x}_i + \beta = \text{BN}_{\gamma, \beta}(x_i) \quad (5)$$

若定义输入数值集合  $B = \{x_1, \dots, m\}$ , 先由式(2)和式(3)计算  $B$  的均值和方差, 式(4)将所得均值和方差变换为 0 或 1。为保留前一层网络学习到的特征分布, 计算可学习参数  $\gamma, \beta$ 。最后输出式(5), 得到  $B$  中每个元素乘以  $\gamma$  后与  $\beta$  的和。BN 函数先将数据归一化, 再使用参数  $\gamma, \beta$  还原数据, 一定程度保留了原来数据的分布特点。BN 激活函数在卷积层上将一整张特征图作为处理对象。残差块中每个 ReLU 函数前都有一个 BN 函数, 共同构成了激活层。

残差网络的主要结构如图 5 所示, 每个残差块由卷积层路径和捷径连接组成。本文实验将多个残差块堆叠, 构建了 50 层的残差网络。其中 50 层残差网络是在 34 层网络的基础上用三层的瓶颈块替换了两层块, 即残差块中有卷积核大小分别为  $1 \times 1 \times 64, 3 \times 3 \times 64, 1 \times 1 \times 256$  的 3 个卷积层。

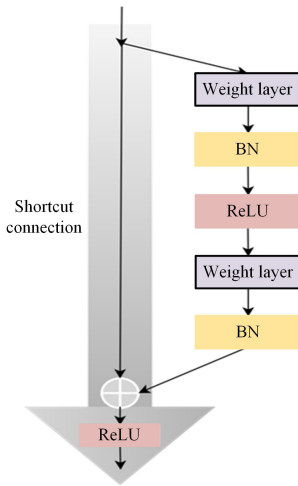


图 5 残差块结构

Fig. 5 Residual block structure

结合 50 层残差网络提取的图像语义信息, 对图像中每个点都生成多个矩形 anchors 作为原始检测框, 接下来的任务是对这些检测框进行一一修正。每个点的多个 anchors 分别输入 RPN 和感兴趣区域池化层。RPN 通过 Softmax 函数获得 anchors 是检测目标或背景的分类, 同时学习 anchors 与真实监督框的偏移量, 得到用于预测边框回归监督边框的变换量, 实现所有 anchors 边框回归。Proposal layer 整合边框信息, 过滤得到包含检测目标和回归最优的所有 anchors。但这些候选框受具体检测目标形状的影响而大小各异。

为给后续卷积操作输入固定大小的图像结构, 本文方法使用了候选框池化层, 以实现候选区域的固定形状输出。之后用目标分类层继续准确化目标检测边框。最后, Faster R-CNN 用连接层和 Softmax 分类候选框, 输出检测框是心脏还是背景的概率向量, 同时再次利用边框回归获取更高精度的检测框。

实验中对所有数据训练 40 轮, 训练和评估同时进行, 每 500 步评估且保存一次训练参数, 训练网络收敛后获取最优参数。使用赋给最优参数的 Faster R-CNN 对将要输入分割

网络的数据集的目标区域做预测和裁剪, 得到如图 6(b) 中背景为全黑的图像。

## (2) 心脏分割模块

将裁剪处理后的图像做数据增广, 输入语义分割网络模型 U-net, 如图 4 所示。U-net 网络延续了 Hinton 等<sup>[20]</sup>提出的编码和解码的思想, 整个网络呈 U 型对称结构。网络一共 23 层, 输入图像可以是任意大小, 主要包括下采样和上采样路径。

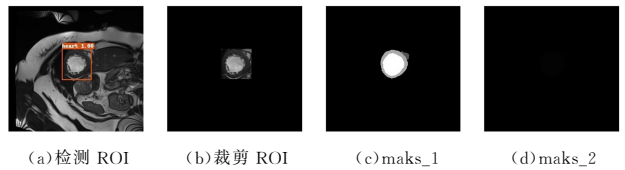
网络越深, 提取的图像特征分辨率越低, 语义信息越好; 相反, 网络越浅, 提取图像语义水平越低, 但分辨率越高。为了获取较高的图像语义信息, 本文方法是在下采样网络中加入池化层或采取较多的卷积块, 来降低图像分辨率。U-net 编码器可以输出每个卷积块获取的图像特征, 保留不同深度获取的图像语义信息和分辨率, 并将其作为解码器的输入。

为了处理更多的细节信息, U-net 分割网络采用了跳跃连接, 跳跃连接是编码器在增加图像特征通道数的同时提供不同深度的特征给解码器。解码器的每一步都有反卷积层, 减半特征通道数量, 加倍特征图大小。

## 5 实验结果与分析

### 5.1 实验数据及其评价指标

本文实验使用的 GPU 版本是 Tesla V100, 配置 16GB 的 Video Mem, 16GB 的 RAM 和 100GB 的 Disk。使用 nibabel 医学图像处理库, 选择随机梯度下降方法 (Stochastic Gradient Descent, SGD) 为优化器。Faster R-CNN 网络选择分段衰减和线性学习率热身两种学习率优化策略。



(a)检测 ROI (b)裁剪 ROI (c)maks\_1 (d)maks\_2

图 6 检测、裁剪后的 ROI, 以及两种像素设定的标签图像

Fig. 6 Detected, cropped ROI, label image with two pixel settings

预训练模型 Faster R-CNN 识别出心脏, 如图 6(a) 所示, 同时记录下检测框的 4 个点坐标, 将其他区域像素值设定为 0, 结果如图 6(b) 所示。保留非心脏背景区域, 凸显心脏在整张图像中的位置信息。相反, 第一组对比实验中, 不对原始数据输入 Faster R-CNN 网络且不对其做检测裁剪处理, 直接将其制作成满足分割网络模型要求的原图像和标签图像, 输入 U-net 分割网络模型。

本文实验中的分割网络 U-net 要求输入标签图像每个点的灰度值作为标签, 且灰度值要连续。图 6(c) 是切片原始三维标注图得到的二维标签图像, 每个灰度值代表一种标签, 图像中有 4 种像素值, 分别为 0, 85, 170 和 255, 分别代表背景、右心室、心肌和左心室。图 6(d) 是将前 4 种灰度值分别对应更改为 0, 1, 2 和 3 后的图像, 标签表示意义对应不变。

评价医学图像分割时一般采用差异实验法, 将分割结果图和标准分割图做比较。与传统分割方法相比, 本文实验选用平均交并比和 Dice(dice coefficient) 系数为评价指标。本质还是对预测的标签图和标准标签图进行差异性评估。Kappa 系数, 用于一致性检验, 可以用来衡量分类的效果。

上述 3 种评估方法都是基于混淆矩阵的。对于二分类

问题,混淆矩阵如表 1 所列。

表 1 混淆矩阵  
Table 1 Confusion matrix

混淆矩阵		预测值	
		Positive	Negative
真实值	Positive	TP	FN
	Negative	FP	TN

表 1 中一共设定了 4 个基本指标:真阳性(True Positive, TP),预测值为阳性,真实值也为阳性;假阳性(False Positive, FP),预测值为阳性,真实值为阴性;真阴性(Ture Negative, TN),预测值为阴性,真实值为阴性;假阴性(False Negative, FN),预测值为阴性,真实值为阳性。

本文实验中,ACDC 数据集涉及 4 种元素分类,用标签 0,1,2,3 分别对应背景、右心室、心肌和左心。以计算左心室的混淆矩阵为例,将真实值和预测值的阳性都定义为左心室,阴性都定义为不是左心室,将其转化成二分类问题,再进行进一步解决。

#### (1)平均交并比

平均交并比是语义分割的标准度量,计算出预测区域和实际区域的交集除以预测区域和实际区域的并集,取所有类别的交集和并集之比的平均值。具体验算过程为:

$$mIoU = \frac{1}{n+1} \sum_{i=0}^n \frac{TP}{FN+FP+TP} \quad (6)$$

其中, $n$  表示分割的类别, $(n+1)$ 则是类别数加背景。实验中 ACDC 数据集分割对象包含 4 类,分别是左心室、右心室、心肌和背景。平均交并比则是计算 4 种分割类别的交并比求和后的平均值。mIoU 一般在评估分割模型时计算,数值会对模型学习不断增加到稳定,记录下来的数值变化会是一条非线性曲线。

#### (2)Dice loss

Dice 系数是一种表示集合相似度的度量函数,其取值范围在 0 到 1 之间,数值越接近于 1,表示结果越好,表达式如下:

$$Dice = \frac{2|X \cap Y|}{|X| + |Y|} \quad (7)$$

其中, $X$  表示图像中的真实区域, $Y$  表示分割网络模型的预测区域, $|X \cap Y|$ 是真实区域和预测区域的交集, $|X| + |Y|$ 则表示两区域的集合。

$$d = 1 - Dice \quad (8)$$

将 1 减去  $Dice$  定义为 Dice 损失,如式(8)所示,Dice loss 与 Dice 系数的取值范围相同,但其值越接近于 0,表示结果越好。

#### (3)Kappa 系数

Kappa 系数是衡量分类精度的指标,计算结果通常落在 0~1 之间,值为 0.61~0.80 表示高度一致性,0.81~1 表示几乎完全一致。具体计算式如式(9)和式(10)所示:

$$k = \frac{p_o - p_e}{1 - p_e} \quad (9)$$

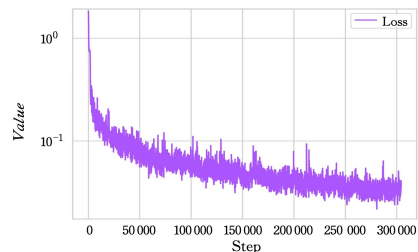
$$p_e = \frac{\sum_{i=1}^n (a_i \times b_i)}{n^2} \quad (10)$$

其中, $i$  表示分类任务个数; $a_i$  是每一类真实样本的数量; $b_i$  是每一类预测样本的数量,总样本数为  $n$ ;  $p_o$  表示总体分类精度,即用所有正确分类样本数量和总样本数相除。

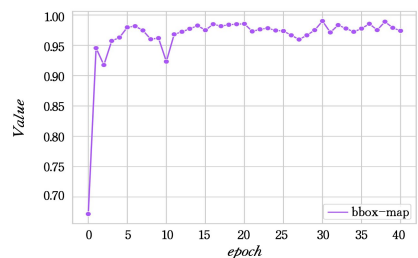
## 5.2 数据处理

### (1)心脏区域检测实验

本文实验优选 Faster R-CNN 检测网络模型。Faster R-CNN 进行了 40 轮训练后收敛,网络训练的同时在验证集上进行验证,得到的结果如图 7 所示。图 7(a)给出了训练过程中 Faster R-CNN 的损失函数,它记录了网络的收敛和拟合情况。图 7(b)给出了对所有数据集学习 40 次过程中计算验证集 map 的变化曲线,训练到 35 轮时 bbox\_map 值变化逐渐趋于平稳态,验证集中 map 可以达到 99.0%。



(a)损失函数曲线



(b)bbox\_map 曲线

图 7 Faster R-CNN 损失函数和 map 曲线

Fig. 7 Faster R-CNN loss function and map curve

### (2)对比分割实验

实验时,将经过 Faster R-CNN 网络预处理后的数据集送入 U-net 网络,同时设定两组对比实验。对比实验 1,不添加 Faster R-CNN 模块做图像预处理;对比实验 2,选择不同的分割网络添加 Faster R-CNN 模块。

对比实验 1 中图像增强方式、网络的学习率、优化器和损失函数均保持不变,则可得两种方法在验证集上的 mIoU 结果,并如图 8、图 9 所示。图 8 给出了 U-net 网络对 ACDC 数据集训练所得的曲线。图 8(a)是分别训练两个 U-net 网络时的损失函数记录,其中红色曲线对应经过裁剪处理的图像输入 U-net 学习所得的损失函数,红色曲线下落速度和收敛速度都比蓝色曲线快;图 8(b)给出了两个 U-net 网络训练过程中,每 500 个 batch\_size 评估一次验证集后记录的 mIoU,红色曲线对应裁剪处理图像输入 U-net 学习后对验证集的 mIoU,最高可达 0.89,蓝色曲线则对应未经处理图像输入 U-net 学习后对验证集的 mIoU,可达 0.817。图 9 给出了 U-net 网络对 SCD 数据集训练所得的曲线。与图 8 同理,红色曲线代表加入 Faster R-CNN 检测模块后的分割结果曲线,蓝色则代表对比实验 1 所得实验结果。

图 8(a)和图 9(a)中红色损失函数曲线下落速度快,最小值也明显小于蓝色曲线。图 8(b)和图 9(b)中红色平均交并比曲线上速度更快,最高 mIoU 也优于蓝色曲线。分析曲线后可以基本明确,去除了背景信息的图像输入分割网络模型后,网络的收敛速度加快,模型分割性能提升,分割精度亦有巨大的提高。

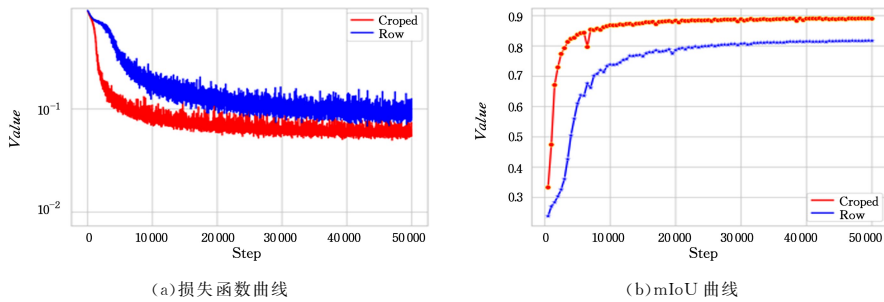


图 8 基于 ACDC 数据集的 U-net 损失函数和 mIoU 曲线(电子版为彩图)

Fig. 8 U-net loss function and mIoU curve based on ACDC dataset

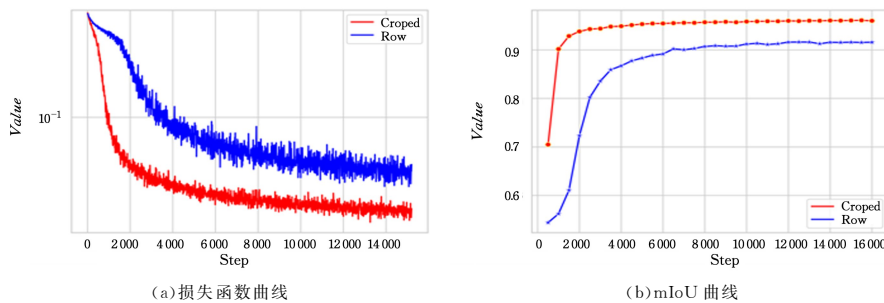


图 9 基于 SCD 数据集的 U-net 损失函数和 mIoU 曲线(电子版为彩图)

Fig. 9 U-net loss function and mIoU curve based on SCD dataset

具体实验结果如表 2 所列。在 ACDC 数据集裁剪处理后的分割结果中,Dice 系数提高了 5%,Kappa 系数提高了 6%;在 SCD 数据集裁剪处理后的分割结果中,Dice 系数提高了 3%,Kappa 系数提高了 6%。

表 2 分割结果

Table 2 Segmentation results

		Dice 系数	Kappa 系数	mIoU
ACDC	对数据集进行裁剪处理	0.94	0.94	0.890
	不对数据集进行裁剪处理	0.89	0.88	0.817
SCD	对数据集进行裁剪处理	0.98	0.96	0.960
	不对数据集进行裁剪处理	0.95	0.90	0.910

对比实验 2 的所有分割网络都添加了 Faster R-CNN 检测裁剪模块,具体分割网络有 U-net, attention-Unet 和 U-net3plus。

分割网络模型,抑制输入图像的不相关区域,突出特定待分割区域的局部显著特征。软注意力机制的使用,弥补了传统 U-net 网络提取冗余信息较多的缺点,有效抑制了无关区域的激活。U-net3plus 则是全尺度跳跃连接和深度监督特点的改进 U-net 网络,能够从多尺度图像中提取足够的语义信息。全尺度跳跃连接把多尺度特征图的多级语义信息结合;同时深度监督从特征图中学习层次表示,优化器官边界分割。U-net3plus 对医学图像有较高的分割精度,优化的 U-net 分割网络分别添加本文实验提出的检测裁剪模块,在 ACDC 数据集上进行训练,网络收敛后对其分割效果进行评估,不同评估指标的评估结果如图 10 所示。增加 Faster R-CNN 裁剪模块的 3 种分割网络的 Dice 系数、Kappa 系数和平均交并比这 3 个评估指标都优于不添加 Faster R-CNN 模块的指标。

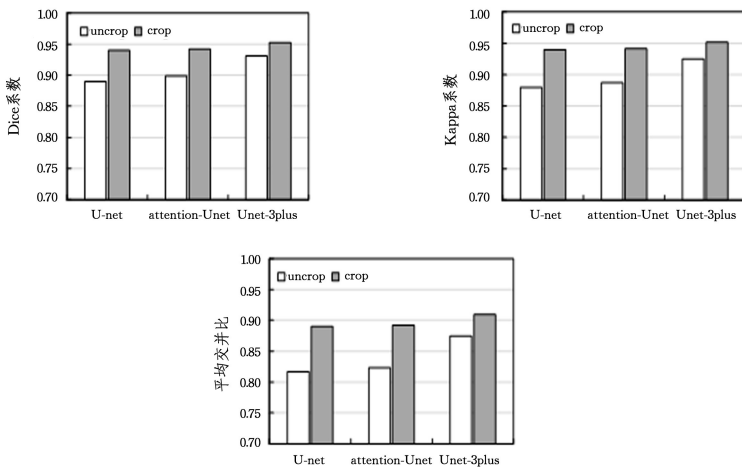


图 10 U-net 及其改进网络的评估

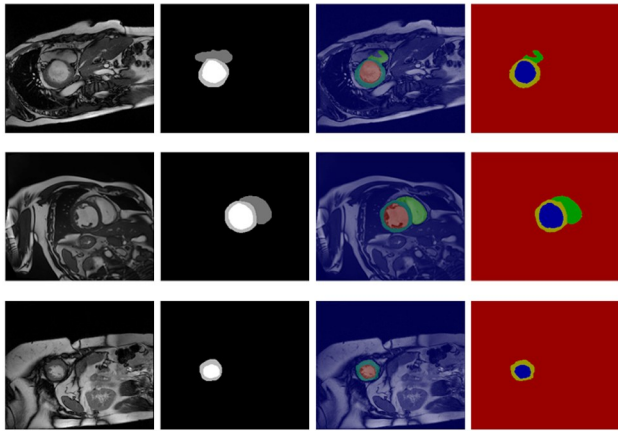
Fig. 10 Evaluation of U-net and its improved network

### 5.3 实验结果分析

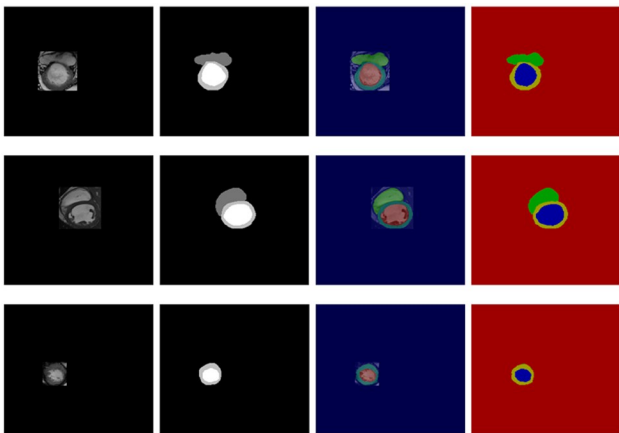
分别对引入 Faster R-CNN 做裁剪预处理和未裁剪处理

的两种数据集训练好的网络模型做验证,验证结果如图 11 所示。选择 3 张验证图像,分别输入两个有不同参数的 U-net

网络模型。图 11(a)给出了未裁剪图像的原图、初始标签图、原始图像训练 U-net 网络后得到最优模型对原图的预测图,以及预测标签图。图 11(b)给出了裁剪后图像、初始标签图、裁剪图像训练 U-net 网络后得到最优模型对裁剪图的预测图,以及验证标签图。



(a)引入 Faster R-CNN 前



(b)引入 Faster R-CNN 后

图 11 两种方法的预测结果

Fig. 11 Prediction results of both methods

### (1)网络性能对比

网络模型的学习速度一直是衡量模型性能的一大重要指标,也是该模型是否可以应用于临床诊断的选择标准之一。真实临床场景中,由于病人病情需要和医疗资源间的矛盾,要求诊断医生或智能诊断系统在尽量短的时间内做出医学影像诊断。本文方法为提高网络性能提出了一种新思路。如图 9(b)所示,对比实验在验证集计算 mIoU 达到 0.4 时,在相同的时间内本文方法在验证集的 mIoU 达到 0.8 以上。

### (2)分割表现对比

结合 Faster R-CNN 和 U-net 网络对心脏 MRI 图像做分割,对比只用 U-net 网络的分割方法,不仅 mIoU, Dice 系数和 Kappa 系数分别提高了 5.0%, 6.0% 和 7.3%, 而且从图 11 的测试结果可视化来看,对比实验中,右心室区域的分割有识别不全或者边缘处理粗糙的表现,同时左心室心肌在预测图中的分割厚度与在标签图中的实际分割厚度有一定差距,没有很好地反映心肌厚度;相反,应用本文实验方法,预测图中所显示出的右心室分割区域更为精准,左心室心肌厚度与原来标签图中的心肌厚度基本保持一致,且没有颗粒性的粗糙边缘。

### (3)分割方法的泛化性

ACDC 数据集的分割任务中包含 4 个分割对象;SCD 数据集的分割任务只包含左心室和背景两个分割对象,该数据集采集对象的身体状况不定,患有心脏病或健康。两种数据集的选用,增加了分割网络学习心脏 MRI 图像的多样性。在多样性实验数据的基础上,本文提出的分割方法仍然表现出了良好的分割精度和分割效率。

对以 U-net 为基础的改进分割网络 attention-Unet 和 U-net3plus 分别增加 Faster R-CNN 检测裁剪模块,在 ACDC 数据集上训练和评估网络性能,与没有 Faster R-CNN 模块的 attention-Unet 和 U-net3plus 在 ACDC 数据集上的分割结果做对比,与 Faster R-CNN 串联的分割网络表现出了良好的分割精度。因此,本文提出的实验方法不仅有数据集泛化性,也有分割网络适用泛化性。

**结束语** 本文结合 Faster R-CNN 和 U-net 网络,采用联合网络的形式,实现对心脏 MRI 原始图像的分割。临床中获取的心脏 MRI 图像不仅包含心脏的语义信息,还包括心脏周围纵膈等其他组织信息,且心脏 MRI 可以多方位成像,针对多信息的心脏 MRI 图像,如果直接送入深度学习网络,冗余的背景信息会占用一定的计算空间和产生背景噪声,导致计算时间加长,分割精度降低。针对这个问题,本文提出先采用预训练模型 Faster R-CNN 对原始图像进行心脏区域检测,然后将检测所得的图像做裁剪,再选用适用于医学影像分割的 U-net 分割网络模型,对裁剪图像做分割处理。在 ACDC 数据集和 SCD 数据集上对本文方法进行实践,设计了对比实验,验证了本文方法的可行性。

实验说明,本文方法不仅提高了分割精度,而且压缩了 U-net 分割网络的计算数据量,从而缩短了分割时间,提高了分割效率。此外,引入 Faster R-CNN 网络模型后,不需要设计额外的图像裁剪处理,就能获得较为准确的目标区域框,减少在心脏分割过程中的人为操作,提高了对 U-net 分割不同心脏方位的泛化度。

本文方法对心脏 MRI 图像分割效率和精度都有一定的提升。但因选用数据集的局限性,在一定程度上限定了 Faster R-CNN 的目标检测精度,从而影响了 U-net 的分割结果。同时,在预处理阶段,本文对三维图片进行了二维化处理,进而清洗掉标签图中没有标注出心脏的多余图像。此人工方法无形中增大了工作量,导致模型无法很好地被推广到临床实践中。因此,提出三维心脏 MRI 图像分割网络具有更高的临床价值<sup>[21]</sup>。对三维心脏 MRI 数据集的目标区域先进行检测后分割,以提高分割精度,是课题组今后的重要研究内容。

此外,医学心脏 MRI 领域的分割任务不仅包括能分割出心脏不同的区域,还需要对疾病的诊断有帮助,本文方法还将面向多种疾病影像的分割优化。

## 参考文献

- [1] KANG X D. Medical Image Processing[M]. People's Medical Publishing House, 2009:200.
- [2] SUTTON R T, PINCOCK D, BAUMGART D C, et al. An overview of clinical decision support systems: benefits, risks, and strategies for success[J]. Npj Digital Medicine, 2020, 17(2020):3.
- [3] HESAMIAN M H, JIA W, HE X, et al. Deep Learning Techniques for Medical Image Segmentation: Achievements and Chal-

- lenges[J]. *Journal of Digital Imaging*, 2019, 32(4):582-96.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks[J]. *Communications of the ACM*, 2012, 60(6):84-90.
- [5] HAVAEI M, DAVY A, WARDE-FARLEY D, et al. Brain tumor segmentation with Deep Neural Networks[J]. *Medical Image Analysis*, 2017, 35:18-31.
- [6] SHELHAMER E, LONG J, DARRELL T. Fully Convolutional Networks for Semantic Segmentation[J]. *Ieee Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4):640-51.
- [7] CHRIST P F, ELSHAER M, ETTLINGER F, et al. Automatic Liver and Lesion Segmentation in CT Using Cascaded Fully Convolutional Neural Networks and 3D Conditional Random Fields[C] // *proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention-MICCAI*, 2016:415-423.
- [8] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015 Lecture Notes in Computer Science*, 2015, 9351:234-241.
- [9] TIAN X, WANG L, DING Q. Overview of Image Semantic Segmentation Methods Based on Deep Learning [J]. *Journal of Software Science*, 2019, 30(2):440-468.
- [10] WANG B, YUAN F Q, CHEN Z R, et al. Automatic segmentation method for multi-level U-Net thyroid ultrasound images [J]. *Computer Engineering and Applications*, 2023, 59(5):205-212.
- [11] IOFFE S, SZEGEDY C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[J]. *JMLRorg*, 2015, 37:448-56.
- [12] AMMAR A, BOUATTANE O, YOUSSEFI M. Automatic cardiac cine MRI segmentation and heart disease classification[J]. *Comput Med Imaging Graph*, 2021, 88:101864.
- [13] WANG X Y, YANG S, FANG Y Q, et al. SK-Unet: An Improved U-Net Model With Selective Kernel for the Segmentation of LGE Cardiac MR Images[J]. *IEEE Sensors Journal*, 2021, 21(10):11643-53.
- [14] LIU C, LIN N, CAO Y J, et al. Seg-CapNet: Neural Network Model for Cardiac MRI Image Segmentation [J]. *Chinese Journal of Imaging and Graphics*, 2021, 26(2):452-463.
- [15] ZOU Z, SHI Z, GUO Y, et al. Object Detection in 20 Years: A Survey[J]. *arXiv*:1905.05055, 2019.
- [16] LIU T, TIAN Y, ZHAO S F, et al. Automatic Whole Heart Segmentation Using a Two-Stage U-Net Framework and an Adaptive Threshold Window[J]. *IEEE Access*, 2019, 7:83628-83636.
- [17] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. *Ieee Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6):1137-1149.
- [18] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[J]. *IEEE*, 2016:1063-6919.
- [19] YANG H J, WANG E S, SUI Y X, et al. Simplified residual structure and fast deep residual network [J]. *Journal of Jilin University (Engineering Edition)*, 2022, 52(6):1413-1421.
- [20] HINTON G E, KRIZHEVSKY A, WANG S D. Transforming Auto-Encoders[C] // *Proceedings of the Artificial Neural Networks and Machine Learning-ICANN 2011-21st International Conference on Artificial Neural Networks*. Berlin: Springer, 2011, 6791:44-51.
- [21] ZHANG X Y, WANG B, AN W C, et al. 3D U-Net++ glioma segmentation network based on fusion loss function [J]. *Computer Science*, 2021, 48(9):187-193.



**HAN Junling**, born in 1999, undergraduate. Her main research interest is medical image processing.



**KANG Xiaodong**, born in 1964, Ph.D., professor, is a member of China Computer Federation. His main research interests include medical image processing and medical information system integration.