



一种基于Faster R-CNN的甲状腺结节超声图像 目标检测改进算法*

郑天雷^{1,2}, 杨娜², 耿诗², 赵先云², 王跃², 程德强¹, 赵蕾^{2△}

1. 中国矿业大学信息与控制工程学院(徐州 221116); 2. 徐州医科大学附属医院 医疗设备管理处 人工智能研究组(徐州 221004)

【摘要】目的 为提高超声图像中甲状腺结节目标检测精度,提出一种基于Faster R-CNN的甲状腺结节目标检测改进算法。**方法** 该算法采用结合了可变形卷积(deformable convolution, DC)的ResNeSt50作为主干网络,提高对形状不规则结节的检测效果。并在主干网络后方引入特征金字塔网络(feature pyramid networks, FPN)和感兴趣区域对齐,前者用于减少甲状腺结节漏检误检现象,后者用于提高小尺寸结节的检测精度。此外,在算法训练的反向传播过程中,使用由锐度感知最小化(sharpness-aware minimization, SAM)改进优化器进行参数更新,提高算法的泛化能力。**结果** 实验采用来自徐州医科大学附属医院及南京市第一医院6261张甲状腺超声图像,对改进算法的有效性进行对比评估。实验表明,该算法具有一定的优化效果,最终在测试集的AP50高达97.4%, AP@50:5:95较原始模型也提升了10.0%。与原始模型和现有模型相比改进算法有着更高的检测精度,能更精准地检测甲状腺结节,特别在较低检测框精度要求下有着较高的召回率。**结论** 本研究提出的改进方法是有效的甲状腺结节目标检测算法,能精准地检测出甲状腺结节。

【关键词】 甲状腺结节 超声图像 Faster R-CNN 目标检测 ResNetSt50

An Improved Object Detection Algorithm for Thyroid Nodule Ultrasound Image Based on Faster R-CNN ZHENG Tianlei^{1,2}, YANG Na², GENG Shi², ZHAO Xianyun², WANG Yue², CHENG Deqiang¹, ZHAO Lei^{2△}. 1. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China; 2. Artificial Intelligence Unit, Department of Medical Equipment Management, Affiliated Hospital of Xuzhou Medical University, Xuzhou 221004, China

△ Corresponding author, E-mail: xyfysbczhaolei@163.com

【Abstract】 Objective To propose an improved algorithm for thyroid nodule object detection based on Faster R-CNN so as to improve the detection precision of thyroid nodules in ultrasound images. **Methods** The algorithm used ResNeSt50 combined with deformable convolution (DC) as the backbone network to improve the detection effect of irregularly shaped nodules. Feature pyramid networks (FPN) and Region of Interest (RoI) Align were introduced in the back of the trunk network. The former was used to reduce missed or mistaken detection of thyroid nodules, and the latter was used to improve the detection precision of small nodules. To improve the generalization ability of the model, parameters were updated during backpropagation with an optimizer improved by Sharpness-Aware Minimization (SAM). **Results** In this experiment, 6261 thyroid ultrasound images from the Affiliated Hospital of Xuzhou Medical University and the First Hospital of Nanjing were used to compare and evaluate the effectiveness of the improved algorithm. According to the findings, the algorithm showed optimization effect to a certain degree, with the AP50 of the final test set being as high as 97.4% and AP@50:5:95 also showing a 10.0% improvement compared with the original model. Compared with both the original model and the existing models, the improved algorithm had higher detection precision and improved capacity to detect thyroid nodules with better accuracy and precision. In particular, the improved algorithm had a higher recall rate under the requirement of lower detection frame precision. **Conclusion** The improved method proposed in the study is an effective object detection algorithm for thyroid nodules and can be used to detect thyroid nodules with accuracy and precision.

【Key words】 Thyroid nodule Ultrasound images Faster R-CNN Object detection ResNetSt50

近年来,甲状腺结节检出率逐年增高^[1-2],对恶性结节患者来讲,早诊早治可以有效降低死亡率^[3]。超声因为无

创伤、无辐射等原因,成为诊断甲状腺结节的首选检查^[4]。在超声图像上,准确判断出结节位置和数量,是进一步诊断的关键。但超声图像存在分辨率低、噪声斑点严重等问题,不同型号设备采集的不同患者的甲状腺结节形态差异较大^[5],这给医生准确诊断甲状腺结节带来极大的挑战。

* 徐州市重点研发计划(No. KC19174)、徐州医科大学江苏省重点实验室开放项目(No. XZSYSKF2021030)和徐州医科大学附属医院院级科研项目(No. 2022ZL26)资助

△ 通信作者, E-mail: xyfysbczhaolei@163.com

随着深度学习在计算机视觉方向的长足进步,目标检测方法在医学领域得到广泛应用^[6-8]。目前,主流的深度学习目标检测算法主要分为单阶段和双阶段检测算法^[9]。单阶段算法采用回归分析思想,结构简单、计算高效,以YOLO^[10]系列和SSD^[11]系列为典型代表。双阶段算法需要生成卷积特征图,获得感兴趣区域后,再进行目标分类^[12]。典型的算法有R-CNN系列^[13]、Mask R-CNN^[14]等。

近年来,在甲状腺结节超声图像目标检测领域,国内外开展了诸多研究。在单阶段检测算法研究中,WANG等^[15]提出基于YOLOv2的端到端检测网络识别甲状腺结节的位置和类型,ZHANG等^[16]优化改进一个基于YOLOv3的模型,检测超声图像中的甲状腺结节。虽然单阶段检测算法能够较快地实现甲状腺结节的检测,但是其定位精度往往低于双阶段算法。因此越来越多的双阶段算法被应用于甲状腺结节超声图像的目标检测任务。柯威等^[17]采用多尺度方法对Faster R-CNN进行改进,改进模型能够较好地检测甲状腺结节。秦丽娜等^[18]将改进的Cascade Mask R-CNN应用于结节的检测,并基本能满足临床应用需求。尽管这些研究取得了令人鼓舞的结果,但是双阶段算法在甲状腺结节检测中,依然存在较高阈值的情况下检测精度不高的问题。

在医学图像诊断目标检测任务中,检测精度的需求高于检测速度。因此本文选择在双阶段检测中精度表现更为优异的Faster R-CNN作为基础模型,针对甲状腺结节在超声图像中形态多样、尺寸不一等情况,构建一个全新的甲状腺结节目标检测改进算法。采用结合了可变形卷积(deformable convolution, DC)的ResNeSt50作为主干网络,在其后方引入特征金字塔网络(feature pyramid networks, FPN)和感兴趣区域对齐(RoI Align),并在反向传播过程中,使用由锐度感知最小化(sharpness-aware minimization, SAM)改进的优化器进行参数更新。现报道如下。

1 材料和方法

1.1 数据收集及预处理

以甲状腺病理报告为金标准,实验收集5 021例来自徐州医科大学附属医院及南京市第一医院超声科的甲状腺结节患者的76 496张甲状腺超声图像。所有超声图像由飞利浦HD15、GE Voluson S8、飞利浦EPIQ7、西门子ACUSON S3000、飞利浦IU22彩色超声诊断仪采集。经专业超声医师按照以下标准进行图像清洗筛选:①图像中是否含有甲状腺及甲状腺结节;②患者术前采集的,并且有术后病理诊断的超声图像;③患者年龄在18~80岁之间。

最终得到3 907张恶性结节图像及2 354张良性结节图像;然后由两名具备高级职称的超声医师结合病理报告采用Labelme软件对甲状腺结节进行标记;将6 261张图像按照4:1随机划分为训练集和测试集。

标记和审核后的图像在输入网络进行训练和测试前,首先进行图像的标准化处理和数据增强,图像标准化处理主要包括图像尺寸统一为512×512,并将图像像素值归一化至[0, 1]。数据增强主要包括水平镜像、随机缩放、随机亮度增强等。

1.2 模型结构

1.2.1 改进的Faster R-CNN

Faster R-CNN是经典的基于锚的双阶段目标检测网络^[19],它主要由主干网络、区域建议网络(region proposal network, RPN)、RoI Pooling和分类回归网络构成,其中主干网络默认为VGG16^[20],分类网络主要由一个全连接层和Softmax构成,回归网络主要包含一个全连接层和损失函数。

改进的Faster R-CNN在沿用Faster R-CNN双阶段目标检测结构的基础上,在主干网络、RoI特征提取、优化器方面对原始算法进行改进,改进的Faster R-CNN结构如图1所示。改进方法主要包括:①将原始主干网络VGG16修改为ResNeSt50,ResNeSt在继承ResNet残差结构的基础上引入了切分注意力机制(Split-Attention)^[21]。②将ResNeSt50阶段5中所有的普通3×3卷积替换为可变形卷积^[22]。③在模型中加入FPN。④采用Mask R-CNN提出的RoI Align替换RoI Pooling^[14]。⑤在反向传播过程中采用SAM改进优化器^[23]。

1.2.2 ResNeSt

ResNeSt是ResNet的一种改进网络,结合ResNeXt的分组卷积思想以及SENet和SKNet的通道注意力思想^[24]。与VGG16比较,ResNeSt作为Faster R-CNN的主干网络更倾向于提取有效特征,能够提高网络对目标结节的检测能力。

ResNeSt的核心是它的单个区块ResNeSt Block,它主要包括残差结构、分组结构和Split-Attention。ResNeSt Block的总体流程如图2所示:首先按照预设的超参数 K 对输入特征进行分组;然后根据超参数 R 在每个小组中对特征进行切分,生成切分特征注意力向量后对切分特征进行加权求和;最后对每个小组输出的特征进行拼接,通过一个3×3的卷积层后和残差信息相加,输出针对性更强的特征矩阵。

1.2.3 可变形卷积

甲状腺超声图像存在着结节形状不规则、拍摄角度

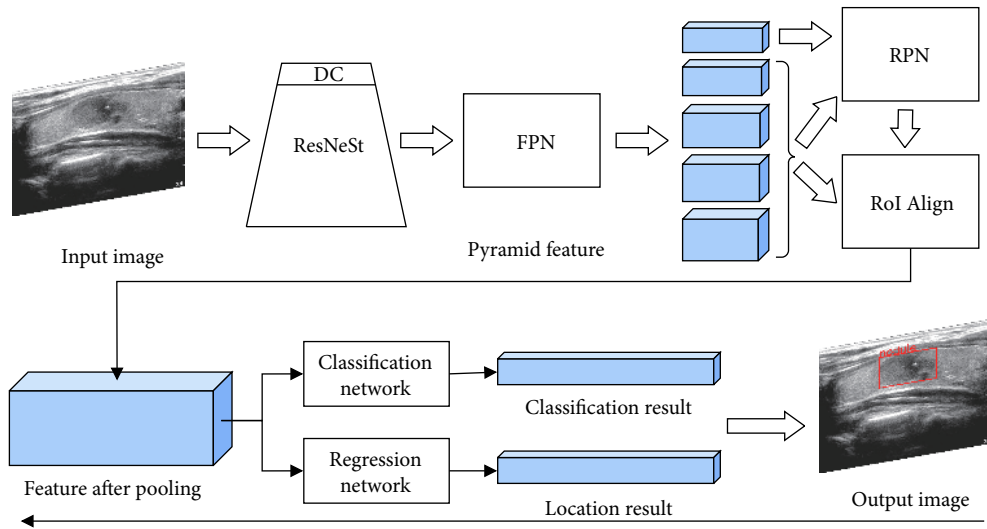


图 1 改进的Faster R-CNN原理图

Fig 1 Schematic diagram of improved Faster R-CNN

DC: deformable convolution; FPN: feature pyramid networks; RPN: region proposal network.

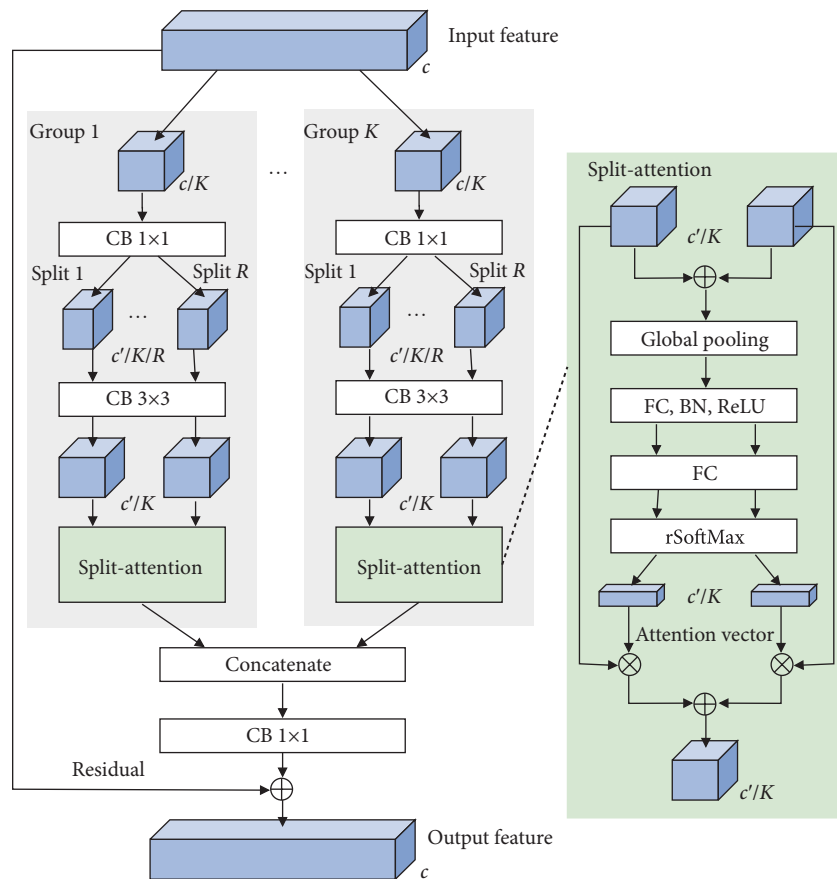


图 2 ResNeSt Block的总体流程

Fig 2 Overall process of ResNeSt Block

FC: fully connected layer; BN: batch normalization; ReLU: rectified linear unit; CB: Conv+BN+ReLU; ⊕: sum of matrix elements; ⊗: multiplication of elements; c: number of channels for output features; c': number of complete channels for process characterization; K: number of groups; R: number of splits.

不一等问题, 这些问题影响传统卷积提取有效特征, 引入 DC可以有效解决上述问题。DC在固定的采样位置基础

上引入了位置偏移, 使卷积能自由地选取所需的有效特征, 在扩大卷积感受野的同时能选取不规则形状区域的

特征、舍弃干扰特征,更有效地获取目标信息。

1.2.4 特征金字塔网络

甲状腺超声图像中还存在大小不一的甲状腺结节,直接采用主干网络生成的尺度单一的特征难以使模型适应不同尺寸目标的检测,引入FPN能够通过结合上下文特征并生成多尺度特征,提高模型对不同尺寸结节的检测能力。

FPN原理如图3所示,FPN不仅给浅层的局部特征带来了全局信息,还生成5组不同尺度的特征,使后方的RPN和RoI Pooling能更好地生成感兴趣特征,因此引入FPN能够降低模型拟合的难度,提高模型在检测不同尺度目标时的鲁棒性。

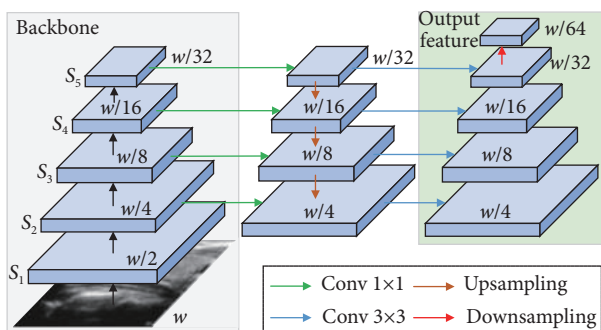


图 3 FPN原理图

Fig 3 Schematic diagram of FPN

w: width of input image; S_n : features of stage n .

1.2.5 感兴趣区域对齐

Faster R-CNN使用RoI Pooling框选甲状腺结节特征时存在两次量化误差,容易导致特征丢失,而RoI Align不对坐标进行取整,能够规避量化误差带来的影响。因此引入RoI Align可以抑制由特征丢失导致的甲状腺结节误检漏检现象。

RoI Align通过候选框位置信息和池化区域数量计算每个裁剪特征块位置,避免只裁剪整数区域。在计算采样点特征值时,使用线性插值法对采样点周围特征值进行采样,有效使用输入特征中的所有特征,在提高目标边缘检测精度的同时,尤其提高模型对小目标的检测能力。

1.2.6 锐度感知最小化

甲状腺超声图像总体形态较为单调且样本数量有限,模型容易发生过拟合现象,引入SAM可以缓解模型过拟合程度,提高模型泛化能力。SAM在最小化模型损失值的同时最小化损失锐度,使损失能降到一个较为平坦的低损失区域,提高模型对整体数据分布的拟合效果。虽然SAM进行了两次传播,训练时间翻倍,但它使损失区间更为平坦稳定,提高模型对测试集结节的检测能力。

1.3 实验环境

实验硬件环境: CPU: Intel Xeon Gold 6230 2.1 GHz 20/40×2, GPU: NVIDIA Quadro GV100 32 G×2, 内存: 384 G。

实验软件环境: 操作系统: Ubuntu 20.04 (Linux 5.11.0-41-generic), 编程语言: Python 3.6.15, 深度学习框架: Pytorch 1.8.2+cu111。

1.4 评价指标

实验使用的评价指标为精确率-召回率(precision-recall, PR)曲线、平均精度(average precision, AP), 其中AP为PR曲线下面积。PR曲线以召回率为横坐标、精确率为纵坐标绘制而成。其中召回率和精确率公式分别为:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

式中, TP为真阳性, 表示将正样本正确识别为正样本的个数, FN为假阴性, 表示将正样本错误识别为负样本的个数, FP为假阳性, 表示将负样本错误识别为正样本的个数。

为更全面地评测不同改进方法的有效性, 实验采用不同交并比阈值下的PR曲线及精度综合评测, 其中计算精度的交并比包括单一阈值以及连续阈值。将交并比设置为0.5和0.75可以得到AP50和AP75, 计算从0.5开始以0.05为步长递增至0.95交并比阈值下所有精度的均值可以得到AP@50:5:95。

1.5 多模型对比分析

为验证模型改进的有效性, 实验对比不同现有模型和本文改进模型在本数据集中的精度差异。实验的现有模型包括: 常用于超声图像目标检测的经典模型YOLO v3, 具有更好前景背景平衡能力的RetinaNet, COCO2017等自然图像数据集中表现优异的YOLOX, 具有更好的分类和定位协调能力的TOOD, 以及具有较强特征提取倾向性的Swin Transformer。其中YOLOX实验对比了不同大小的YOLOX-S、YOLOX-M、YOLOX-L、YOLOX-X, 本研究选择在本数据集中性能最佳的YOLOX-X; Swin Transformer采用的基础检测框架为优化后的Faster RCNN。

2 结果

2.1 模型参数设计

实验统一对所有网络进行20轮的迭代训练, 训练基础学习率为0.0001, 每一轮学习率衰减10%, 基础优化器默认采用Adam, 批数量统一设为4, 目标框分类损失函数

统一为交叉熵损失, 回归损失函数统一为L1损失, 模型训练整体的损失函数为:

$$L = \frac{1}{N_{cls}} L_{cls} + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg} \quad (3)$$

式中, i 为小批量中锚的索引, p_i^* 作为锚的标签, 当锚为阳性时 p_i^* 为 1, 当锚为阴性时 p_i^* 为 0, L_{cls} 表示分类损失, L_{reg} 表示回归损失, λ 为平衡参数, N_{cls} 和 N_{reg} 都用于归一化。使用的ResNeSt的超参数 K 为 1, R 为 2。RPN的非极大值抑制 (non-maximum suppression, NMS) 阈值统一设为 0.7, 模型的NMS阈值统一设为 0.5。

2.2 实验结果分析

2.2.1 SAM有效性验证实验

图4给出了以Faster R-CNN为基础模型加入SAM前后每迭代1轮的训练集及测试集损失值曲线。从图中可以看出, 在第10轮左右, 测试集损失值降低幅度明显衰减, 在第14轮左右, 测试集损失值已无明显下降趋势, 为避免测试集过拟合现象, 实验统一以20轮为迭代次数。相比不包含SAM的模型, 图中包含SAM的模型的训练集曲线与测试集曲线更为接近, 差值更小, 可以看出加入SAM后模型过拟合程度有所降低; 在第12轮之后, SAM在测试集上的损失总体都明显低于基础模型, 且损失最低值也更低, 可看出加入SAM后模型抗过拟合能力有所提升。

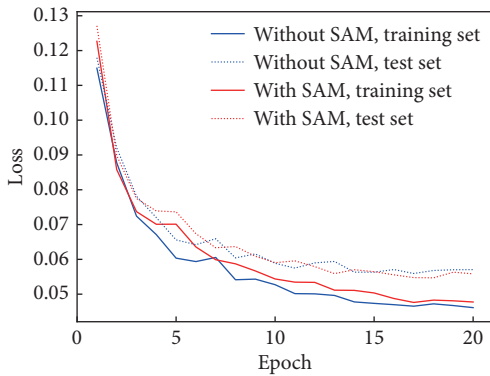


图 4 加入SAM前后的损失值曲线图
Fig 4 Curve of loss value before and after adding SAM

2.2.2 模型改进方法有效性验证实验

图5显示了基础模型及不同改进模型的PR曲线。在对检测框精度要求较低的交并比阈值0.5的标准下, 6种模型都能较好地检测出结节; FPN能更有效地适应不同尺寸结节的识别, 加入FPN后模型召回明显提高, 保证绝大部分目标都能被检测; 有着更好的特征提取倾向性的ResNeSt和可变形卷积能提高模型检测结果的可信度, 在相同的召回下有着更高的精确率。从对检测框精度要求较高的AP75的PR曲线可以看出, RoI Align由于没有丢失

目标边缘特征, 在高精度检测方面明显优于RoI Pooling; ResNeSt和可变形卷积的特征提取倾向性以及SAM的泛化性也都提高了模型的检测精度; 可能由于适应检测尺度的范围较广, 模型的针对性有所发散, 加入FPN后高精度检测性能没有明显提升。

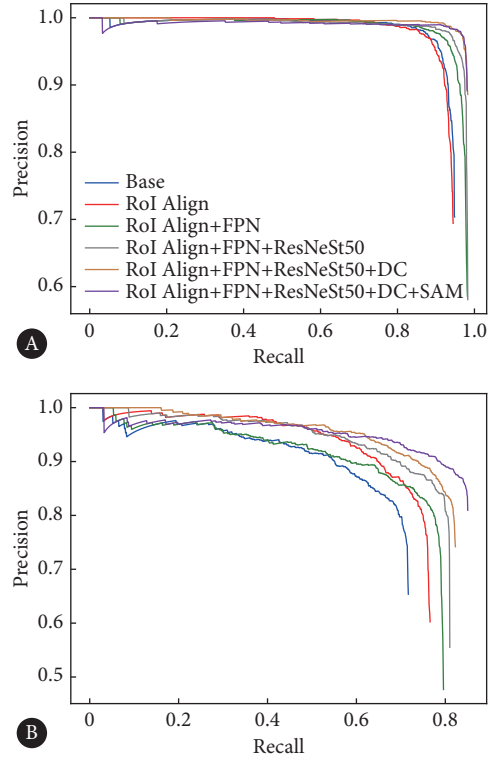


图 5 基础模型及不同改进模型的PR曲线图
Fig 5 Precision-recall curves of the basic model and different improved models

A, Precision-recall (PR) curve (AP50); B, PR curve (AP75). The legends for graph B are the same as those in graph A.

表1展示了基础模型及不同改进模型的平均精度。从表中可看出, 改进方法总体而言均具有一定提升。其中第一次改进, 即加入RoI Align的AP@50:5:95提升最为明显, 相比Faster R-CNN基础模型提升4.4%, 同时采用了RoI Align、FPN、ResNeSt50、DC、SAM的模型有着最佳

表 1 基础模型及不同改进模型的平均精度对比

Table 1 Comparison of the average precision between the basic model and different improved models

Model	AP50/%	AP75/%	AP@50:5:95/%
Base	94.0	66.8	58.7
RoI Align	93.6	73.0	63.1
RoI Align+FPN	97.1	73.9	64.3
RoI Align+FPN+ResNeSt50	97.4	77.3	66.4
RoI Align+FPN+ResNeSt50+DC	97.9	79.0	67.3
RoI Align+FPN+ResNeSt50+DC+SAM	97.4	81.3	68.7

的综合性能,相比Faster R-CNN基础模型提升10.0%。综合而言,FPN能够保证目标结节的基本检测能力,RoI Align能明显提高目标结节的检测精度,ResNeSt、可变形卷积、SAM能提高结节的综合检测性能。

多模型对比分析结果如表2,对比现有模型,本文改进模型的AP50为97.4%,在平均精度指标上优于以上现有模型,对甲状腺结节具有良好的检测能力。

表 2 不同现有模型和改进模型的平均精度对比

Table 2 Comparison of the average precision between different existing models and the improved model

Model	AP50/%	AP75/%	AP@50:5:95/%
YOLOv3	96.2	64.8	58.3
YOLOX	97.0	76.2	66.4
RetinaNet	95.6	69.4	61.4
TOOD	97.1	64.5	64.5
Swin Transformer	97.4	74.7	64.1
Improved model	97.4	81.3	68.7

2.2.3 检测结果可视化分析实验

图6给出了不同优化方法的可视化效果,与表1的顺序相同,依次增加了改进方法。从第1组图像(图6A~6C)可以看出,RoI Pooling的量化误差容易导致特征丢失,导致检测结果不完整,RoI Align则可以有效抑制特征丢失,生成边界完整的检测结果。从第2组图像(图6D~6F)可看出,FPN的多尺度特性保证了模型基本的检测性能,减少目标漏检现象。从第3组图像(图6G~6I)可看出,ResNeSt的通道注意力机制过滤了部分无关信息,舍去甲状腺腺体外部与结节相似的目标,提高模型检测的精确度。从第4组图像(图6J~6L)可看出,可变形卷积能更好地适应形状不规则的结节的检测。从第5组图像(图6M~6O)可看出,SAM更好的泛化性使模型对较为少见的困难样本也能有着较好的检测效果。

3 讨论

本研究以Faster R-CNN为基础,针对甲状腺超声图像中结节尺寸不一、形状不规则等特点及Faster R-CNN固有缺陷进行改进,设计了一种更为适合甲状腺超声图像结节检测的模型。首先模型采用ResNeSt替换原始的VGG16,引入残差结构及通道注意力机制,在缓解梯度消失的同时使特征矩阵包含更多感兴趣信息,提高了模型提取有效特征的能力。其次为了能适应形状不规则结节的检测,模型还采用可变形卷积代替ResNeSt第5阶段的普通卷积,在扩大卷积感受野的同时引入了空间注

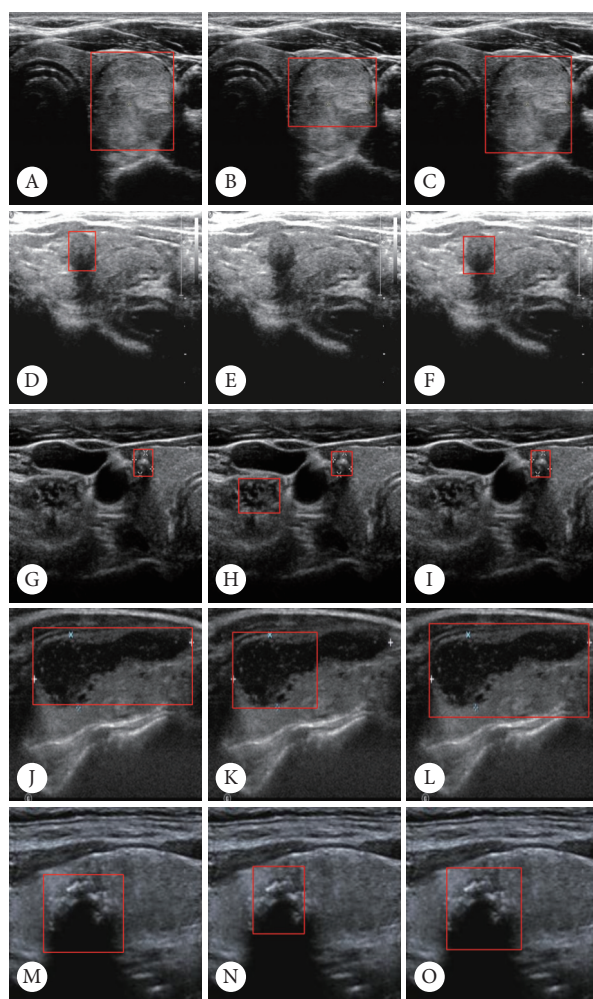


图 6 不同优化方法的可视化效果图

Fig 6 Visual renderings of different optimization methods

A, D, G, J and M, Labeled images; B, RoI Pooling; C, RoI Align; E, FPN not included; F, FPN included; H, VGG16; I, ResNeSt50; K, ordinary convolution; L, deformable convolution; N, SAM not used; O, SAM used. Red boxes indicate the location of nodule.

意力机制,使得模型能舍弃干扰特征,自由地提取所需的不规则形状区域特征。然后FPN加入模型之后,模型对不同尺寸结节的检测适应性更强,使生成感兴趣特征时依赖的特征从单一尺度的特征变为多尺度地包含上下文信息的特征。同时为了缓解小尺寸结节的漏检误检,还采用RoI Align代替RoI Pooling,避免了RoI Pooling的量化误差带来的特征丢失现象。最后采用SAM改进优化器,在最小化模型损失值的同时最小化损失锐度,使损失能降到一个较为平坦的低损失区域,抑制模型过拟合,提高模型泛化能力。

本文还与近几年甲状腺结节超声图像相关目标检测文章的检测结果进行了对比。利用在目标检测领域常用的AP或mAP评价指标对算法性能进行了评估,mAP为所有类别AP的均值,当检测目标只有一个类别时,mAP值

等同于AP值^[25]。秦丽娜^[18]开发并验证了一种改进的Cascade Mask R-CNN模型,用于检测与识别甲状腺超声图像良恶性结节,mAP50值为87.1%。柯威等^[17]提出了一种改进的Faster R-CNN模型,能够识别甲状腺乳头状癌超声图像特征,识别精度mAP50为73.8%。郑英豪^[26]以IoU-Net网络为基础,设计了一种用于完成高质量的甲状腺结节自动检测与识别的Trident R-CNN框架,其AP50为89.8%。张静漪等^[27]利用Cascade Mask R-CNN模型对正常甲状腺、甲状腺局限性病变和甲状腺弥漫性病变等4种甲状腺超声图像进行分类,平均检测精度mAP50为84.5%。WU等^[28]提出了一种能利用视频帧之间的上下文关系Cache-Track模型,该模型在检测和计数甲状腺结节的基础上,还可以跟踪和监测周围组织,mAP50值为88.2%。本研究所提改进模型的AP50值最高,达到了97.4%,与性能最好的文献^[28]相比较提高了9.2%。由此表明本研究所提改进模型的目标检测效果最佳。

为了验证此算法的有效性,在实验中使用Faster R-CNN算法和改进Faster R-CNN算法,得到的AP50分别是94.0%和97.4%,提高了3.4%,并且AP@50:5:95也提升了10.0%。综合评价每一步改进措施,也可以看出模型的性能在逐步提高,这些结果都可以表明本文的改进方法有助于提高算法的检测能力。同时在与当前目标检测领域先进算法以及近年来具有代表性的甲状腺结节超声图像目标检测文献中的结果进行对比后,表明本文所提出的改进模型具有最高的目标检测能力。

但该算法在取得优异性能的同时也存在着一定不足。一是甲状腺结节超声图像的诊断主要取决于医生的专业知识和经验,该研究只邀请两名医生对结节区域进行标记,为了减少主观因素带来的样本偏差,可能需要增加参与标记的医生人数。二是市面上的超声设备种类众多,不同品牌和型号的检查设备所采集的超声图像特点不一,为了提高实验结果的鲁棒性,需要在图像采集过程中丰富数据来源、收集更多的超声图像。三是由于研究的回顾性和收集图像的跨度较长,导致无法统计出所有结节的具体大小和中位值信息,对此,计划实施一项前瞻性研究,以目前研究作为基础,进一步改进并进行分层验证。四是本回顾性研究只统计大概1/5、比较有代表性的结节的尺寸,缺乏大部分结节的实际尺寸信息,对于直径大于4 cm或小于5 mm的结节检出情况,在后续研究中将重点关注。

针对甲状腺结节超声图像目标检测任务,由甲状腺结节大小形状各异导致的结节检测精度偏低问题,本研究以甲状腺结节超声图像为研究对象,对经典目标检测

算法Faster R-CNN的主干网络、感兴趣区域特征提取方式、优化器进行改进,以实现甲状腺结节高精度检出。实验证明改进算法具有一定的优化效果。在甲状腺结节超声图像的数据集上原始的Faster R-CNN算法得到的AP50为94.0%,改进算法的AP50是97.4%,提高了3.4%,并且AP@50:5:95也提升了10.0%。同时,与现有模型和既往研究相比,改进算法的检测精度更高。故可认为本研究提出的改进方法是有效的甲状腺结节能目标检测算法,能精准地在超声图像上检测出甲状腺结节。

* * *

作者贡献声明 郑天雷负责论文构思和研究方法,杨娜负责数据审编,耿诗负责数据分析,郑天雷和赵蕾负责经费获取,杨娜和王跃负责调查研究,程德强和赵蕾负责研究项目管理和监督指导,耿诗和赵先云负责软件,王跃负责验证,杨娜和耿诗负责可视化,郑天雷、杨娜、耿诗、赵先云、王跃和赵蕾负责初稿写作,郑天雷、程德强和赵蕾负责审读与编辑写作。所有作者已经同意将文章提交给本刊,且对将要发表的版本进行最终定稿,并同意对工作的所有方面负责。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] VACCARELLA S, FRANCESCHI S, BRAY F, *et al.* Worldwide thyroid-cancer epidemic? The increasing impact of overdiagnosis. *New Engl J Med*, 2016, 375(7): 614-617. doi: 10.1056/NEJMp1604412.
- [2] HAUGEN B R, ALEXANDER E K, BIBLE K C, *et al.* 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid*, 2016, 26(1): 1-133. doi: 10.1089/thy.2015.0020.
- [3] 高明. 甲状腺结节和分化型甲状腺癌诊治指南. *中国肿瘤临床*, 2012, 39(17): 1249-1272. doi: 10.3969/j.issn.1000-8179.2012.17.001.
- [4] ZHENG Z, SU T, WANG Y, *et al.* A novel ultrasound image diagnostic method for thyroid nodules. *Sci Rep*, 2023, 13(1): 1654. doi: 10.1038/s41598-023-28932-2.
- [5] LIN X, ZHOU X, TONG T, *et al.* A super-resolution guided network for improving automated thyroid nodule segmentation. *Comput Meth Prog Bio*, 2022, 227: 107186. doi: 10.1016/j.cmpb.2022.107186.
- [6] XU X, WANG C, GUO J, *et al.* DeepLN: a framework for automatic lung nodule detection using multi-resolution CT screening images. *Knowl-Based Syst*, 2020, 189: 105128. doi: 10.1016/j.knsys.2019.105128.
- [7] KUWANA R, ARIJI Y, FUKUDA M, *et al.* Performance of deep learning object detection technology in the detection and diagnosis of maxillary sinus lesions on panoramic radiographs. *Dentomaxillofac Rad*, 2021, 50(1): 20200171. doi: 10.1259/dmfr.20200171.
- [8] 王嘉良, 罗健旭, 刘斌, 等. 基于 R-FCN 算法的糖尿病眼底病变自动诊断. *计算机工程与应用*, 2020, 56(4): 109-114. doi: 10.3778/j.issn.1002-8331.1811-0035.
- [9] 许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述. *计算机工程与应用*, 2021, 57(8): 10-25. doi: 10.3778/j.issn.1002-8331.2012-0449.

- [10] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 779–788. doi: [10.1109/CVPR.2016.91](https://doi.org/10.1109/CVPR.2016.91).
- [11] LIU W, ANGUELOV D, ERHAN D, *et al.* Ssd: Single shot multibox detector//Computer Vision–ECCV 2016: 14th European Conference, Amsterdam: Springer, 2016: 21–37. doi: [10.1007/978-3-319-46448-0_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- [12] XU J, REN H, CAI S, *et al.* An improved faster R-CNN algorithm for assisted detection of lung nodules. *Comput Biol Med*, 2023, 153: 106470. doi: [10.1016/j.compbiomed.2022.106470](https://doi.org/10.1016/j.compbiomed.2022.106470).
- [13] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation//Proceedings of the IEEE conference on computer vision and pattern recognition. Columbus: IEEE, 2014: 580–587. doi: [10.1109/CVPR.2014.81](https://doi.org/10.1109/CVPR.2014.81).
- [14] HE K, GKIOXARI G, DOLLAR P, *et al.* Mask R-CNN//Proceedings of the IEEE international conference on computer vision. Venice: IEEE, 2017: 2961–2969. doi: [10.1109/ICCV.2017.322](https://doi.org/10.1109/ICCV.2017.322).
- [15] WANG L, YANG S, YANG S, *et al.* Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network. *World J Surg Oncol*, 2019, 17(1): 12. doi: [10.1186/s12957-019-1558-z](https://doi.org/10.1186/s12957-019-1558-z).
- [16] ZHANG L, ZHUANG Y, HUA Z, *et al.* Automated location of thyroid nodules in ultrasound images with improved YOLOv3 network. *J Xray Sci Technol*, 2021, 29(1): 75–90. doi: [10.3233/XST-200775](https://doi.org/10.3233/XST-200775).
- [17] 柯威. 基于深度卷积神经网络的甲状腺乳头状癌超声图像识别的研究. 广州: 广东工业大学, 2018. doi: [CNKI:CDMD:2.1018.865877](https://doi.org/CNKI:CDMD:2.1018.865877).
- [18] 秦丽娜. Cascade Mask R-CNN模型的改进及其在甲状腺良恶性结节识别中的应用研究. 南昌: 南昌大学, 2020.
- [19] REN S, HE K, GIRSHICK R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE T Pattern Anal*, 2017, 39(6): 1137–1149. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
- [20] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition//Proceedings of the 3rd International Conference on Learning Representations. San Diego: DBIP, 2014. doi: [10.48550/arXiv.1409.1556](https://doi.org/10.48550/arXiv.1409.1556).
- [21] HE K, ZHANG X, REN S, *et al.* Deep residual learning for image recognition//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 770–778. doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [22] ZHU X, HU H, LIN S, *et al.* Deformable ConvNets V2: more deformable, better results//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 9308–9316. doi: [10.1109/CVPR.2019.00953](https://doi.org/10.1109/CVPR.2019.00953).
- [23] KWON J, KIM J, PARK H, *et al.* ASAM: adaptive sharpness-aware minimization for scale-invariant learning of deep neural networks//International Conference on Machine Learning. PMLR, 2021: 5905–5914. doi: [10.48550/arXiv.2102.11600](https://doi.org/10.48550/arXiv.2102.11600).
- [24] LI X, WANG W, HU X, *et al.* Selective kernel networks//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Long Beach: IEEE, 2019: 510–519. doi: [10.1109/CVPR.2019.00060](https://doi.org/10.1109/CVPR.2019.00060).
- [25] 张泽苗, 霍欢, 赵逢禹. 深层卷积神经网络的目标检测算法综述. *小型微型计算机系统*, 2019, 40(9): 1825–1831. doi: [10.3969/j.issn.1000-1220.2019.09.004](https://doi.org/10.3969/j.issn.1000-1220.2019.09.004).
- [26] 郑英豪. Iou-Net模型的改进及其在甲状腺良恶性结节检测与识别中的研究. 南昌: 南昌大学, 2021.
- [27] 张静漪, 罗燕, 刘加林, 等. 卷积神经网络识别正常及异常甲状腺超声图像的价值. *四川医学*, 2021, 42(3): 305–309. doi: [10.16252/j.cnki.issn1004-0501-2021.03.021](https://doi.org/10.16252/j.cnki.issn1004-0501-2021.03.021).
- [28] WU X, TAN G, ZHU N, *et al.* CacheTrack-YOLO: real-time detection and tracking for thyroid nodules and surrounding tissues in ultrasound videos. *IEEE J Biomed Health*, 2021, 25(10): 3812–3823. doi: [10.1109/JBHI.2021.3084962](https://doi.org/10.1109/JBHI.2021.3084962).

(2022 – 12 – 30收稿, 2023 – 08 – 11修回)

编辑 余琳



开放获取 本文遵循知识共享署名—非商业性使用4.0国际许可协议(CC BY-NC 4.0), 允许第三方对本刊发表的论文自由共享(即在任何媒介以任何形式复制、发行原文)、演绎(即修改、转换或以原文为基础进行创作), 必须给出适当的署名, 提供指向本文许可协议的链接, 同时标明是否对原文作了修改; 不得将本文用于商业目的。CC BY-NC 4.0许可协议访问<https://creativecommons.org/licenses/by-nc/4.0/>。

© 2023 《四川大学学报(医学版)》编辑部 版权所有