

融合递归特征金字塔的病原微生物检测

黄志添¹, 谢怡宁², 赵晶¹, 何勇军³

(1. 哈尔滨理工大学 计算机科学与技术学院, 哈尔滨 150080; 2. 东北林业大学 机电工程学院, 哈尔滨 150006;
3. 哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150006)

摘要:针对病原微生物的检测研究较少且目标尺寸相差较大、背景复杂问题,提出了一种融合递归特征金字塔的病原微生物多尺度检测方法。考虑到有些病原微生物的特征与正常细胞有一定的相似性,在YOLOv5特征提取阶段引入混合注意力模块CBAM,可以增强病原微生物的显著度,降低了病原微生物的误诊率。由于尺寸较小的病原微生物在下采样过程会丢失重要特征,在特征融合阶段使用递归特征金字塔RFP和增加小目标的检测层,可以对小目标的特征进行二次提取,丰富语义信息,提高了小目标的检出率,并在二次特征提取模块中使用深度可分离卷积DSConv替换普通卷积Conv,降低了模型的参数量,提升了性能。实验表明,本文算法在病原微生物数据集上的mAP达到了74.9%,相比基线网络YOLOv5提升了4.9%,小目标滴虫的mAP也提高了6.2%。同时在PascalVOC2007、PascalVOC2012和CDetector公开数据集上比基线网络YOLOv5分别提升了1.9%、3.1%和8%。

关键词:病原微生物;目标检测;递归特征金字塔;小目标;多尺度检测

DOI:10.15938/j.jhust.2023.05.012

中图分类号: TP389.1 **文献标志码:** A **文章编号:** 1007-2683(2023)05-0091-12

Detection of Pathogenic Microorganisms by Fusion of Recursive Feature Pyramid

HUANG Zhitian¹, XIE Yining², ZHAO Jing¹, HE Yongjun³

(1. School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China;
2. College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150006, China;
3. Harbin Institute of Technology, School of Computer Science and Technology, Harbin 150006, China)

Abstract: Aiming at the problems of less research on the detection of pathogenic microorganisms, large differences in target size and complex background, a multi-scale detection method of pathogenic microorganisms fused with recursive feature pyramid was proposed. Considering that the characteristics of some pathogenic microorganisms are similar to normal cells, introducing the mixed attention module CBAM in the feature extraction stage of YOLOv5 can enhance the significance of pathogenic microorganisms and reduce the misdiagnosis rate of pathogenic microorganisms. Since the pathogenic microorganisms with small size will lose important features in the downsampling process, the recursive feature pyramid RFP is used in the feature fusion stage and the detection layer of the small target is added. The detection rate of the target is improved, and the depthwise separable convolution DSConv is used in the secondary feature extraction module to replace the ordinary convolution Conv, which reduces the number of parameters of the model and improves the performance. Experiments show that the mAP of the algorithm in this paper on the pathogenic microorganism dataset has reached 74.9%, which is 4.9% higher than the baseline network YOLOv5, and the mAP of the small target Trichomonas has also

收稿日期: 2022-04-23

基金项目: 国家自然科学基金面上项目(61673142);黑龙江省自然科学基金杰出青年项目(JJ2019JQ0013)。

作者简介: 黄志添(1997—),男,硕士研究生;

谢怡宁(1979—),女,博士,硕士研究生导师。

通信作者: 何勇军(1980—),男,博士,博士研究生导师, E-mail: heyongjun@hit.edu.cn.

increased by 6.2%. At the same time, the public datasets of PascalVOC2007, PascalVOC2012 and CDetector are improved by 1.9%, 3.1% and 8% respectively compared with the baseline network YOLOv5.

Keywords: pathogenic microorganisms; object detection; recursive feature pyramid; small objects; multiscale detection

0 引 言

宫颈感染性疾病^[1]是妇科常见疾病,它是由病原微生物的感染所致。同时,病原微生物的入侵可能会改变细胞核的形态,这给正异常细胞的判别带来一定的影响。检测病原微生物是临床诊断感染性疾病的主要手段,其诊断方法可通过 TCT 检查筛选出病原微生物。传统的病理诊断完全依赖于医生“手动操作、肉眼观察”的人工阅片,这种方式存在误诊和漏诊的情况。如今,随着人工智能的发展,智能诊断系统被广泛应用于宫颈细胞学检验。系统通过人工智能和图像处理等手段自动分析宫颈细胞病理涂片的有形成分,从而对细胞、真菌、滴虫等常见成分进行筛选,帮助病理医生节省 90% 以上的阅片时间,提高诊断效率。医学显微图像的自动化识别也成了热门的研究方向,但目前主要是人体细胞的研究,而对于宫颈的病原微生物检测方面却鲜少有人研究,所以目标检测技术在微生物检测上仍有探索的价值。

经过二十多年的研究与发展,目标检测技术已取得了重大进展。已有的目标检测技术主要分为两种类型:一是基于手工特征的传统方法,例如 Hear 特征 + Adaboost 算法、Hog 特征 + SVM 算法和 DPM 算法;二是基于深度学习技术的方法。传统方法是模仿人类视觉的流程得到特征,Zhang 等^[2]提出了一种结合特征提取、SVM 和卷积神经网络相结合的方法来识别白带显微图像中的白细胞。首先将白细胞进行分割提取,然后提取了待检测目标的 7 个形态学特征,并用 SVM 训练识别。此外,作者还用了一个 4 层的深度学习网络进行额外识别,提升了识别效果。

相对于传统方法,深度学习方法具有较好的灵活性。目标检测任务分为目标定位与分类两个子任务,目标定位是确定输入图像中感兴趣目标的具体位置,目标分类是确定输入图像中感兴趣目标的可能性大小。深度学习方法根据目标检测模型训练方

式可分为单阶段检测算法和二阶段检测算法两种类型。

基于深度学习的单阶段目标检测算法不需要生成候选框,直接使用预定义的锚框进行分类和回归得到最终的检测结构。主流单阶段目标检测算法包括 YOLO 系列算法^[3-7]以及 SSD 系列算法^[8-10]。两种算法均是对输入图像直接进行特征提取,计算出目标物体的类别概率和位置坐标值。侯剑平等^[11]针对白带显微图像中白细胞、念珠菌、滴虫的自动化识别问题,提出改进的 YOLOv3-tiny 与 Resnet50 相结合的细胞识别算法。用改进的 YOLOv3-tiny 算法进行细胞检测,Resnet50 模型对检测出的细胞进行精确分类,并运用迁移学习的方法进行模型训练,该算法对白细胞、念珠菌、滴虫的识别结果和效率均优于其他算法。王静等^[12]在 YOLOv5 主干特征提取网络的卷积层中添加坐标注意力机制,提升了算法的特征提取能力,同时改变边框回归损失函数,模型检出白细胞的平均精度均值提升了 3.8%。Li 等^[13]在 YOLOv5 特征提取网络对浅层中的跨阶段部分连接(CSP)模块进行了扩展和迭代,进而最大限度地利用了浅层特征,在残差块中引入了改进的注意力模块,以聚焦对象并抑制背景,在 KAIST 数据集上的平均精度提高了 21%。Dong 等^[14]针对 CT 图像中肺结节体积小难以检测的问题,在 YOLOv5 网络的特征提取网络模块中引入了注意力模块,使骨干网络能够专注于感兴趣的区域并提高特征表征能力。此外,在特征融合模块中,对不同层次的特征赋予不同的权重以实现更有效的特征融合。Zhu 等^[15]在 YOLOv5 的基础上增加了一个预测头来检测不同尺度的物体,然后用 Transformer Prediction Heads (TPH) 替换原来的预测头,利用自注意力机制探索预测潜力,还集成了卷积块注意模型(CBAM),以在具有密集对象的场景中找到注意区域。Liang 等^[16]针对 SSD 网中的目标帧冗余问题,利用掩模提取出 ROI 区域,并对原始图像进行预处理,降低了提取图像特征的计算复杂度,然后使用 resnet 网络代替原始的 VGG 网络,修改后的

SSD模型应用于子宫颈上皮细胞识别并实现准确定位。

基于深度学习的二阶段目标检测算法通过生成包含感兴趣区域的高质量候选框,再通过卷积神经网络进一步分类和回归获得最终的检测结果。主流两阶段目标检测算法是基于R-CNN系列算法^[17-19]。该系列算法首先生成候选区域,然后对候选区域进行处理得到最终结果。Zhang等^[20]在Faster R-CNN框架基础上提出了一种用于粘附细胞再检测的循环扫描算法(CSA, cyclic scanning algorithm)。该算法能够在粘附细胞样本有限的条件下有效检测和识别粘附细胞。Hung^[21]使用Faster R-CNN模型实现了显微图像中疟疾感染的血液细胞的检测,并和传统的图像分割、特征提取及随机森林方法实现的目标检测进行了对比,在识别准确率方面得到了很大的提升。尹远来等^[22]基于Faster R-CNN的网络结构改进的模型,引入了可形变卷积网络和特征金字塔网络,实现了对宫颈数字病理图像进行自动识别,改进后的模型能快速收敛。

随着模型层次的不断加深,模型在识别性能上有了提升,但另一面却增加了模型的参数数量和复杂度,而轻量化模型在不损失精度的前提下,减少网络的参数,提高模型的识别速度。Kang等^[23]提出了一种结合轻量级网络和双注意力机制的目标检测算法,使用EfficientNet代替主干网络,以减少参数数量和计算量。Xiong等^[24]提出了一种基于正则卷积构建块的增强搜索空间系列,以提高延迟-准确性权衡目标检测的加速器。Jin等^[25]提出了用于实时和轻量级的移动非对称双流网络,采用IBCMF结构来补偿轻量级骨干网中的信息损失,并使用自适应的空间金字塔模块来加速推理。

单阶段目标检测网络由于检测效率高,准确率不低而得到研究者们的青睐。其中,YOLO系列采用GoogLeNet架构自定义网络DarkNet作为基础网络,它的运行速度相比于选择VGGNet模型作为基础网络的模型快。但是对于病原微生物尺寸相差较大、背景复杂的目标,YOLO系列的检测效果较差,并且在卷积过程中信息丢失较严重。因此,本文提出一种基于递归特征金字塔的YOLOv5病原微生物多尺度检测方法(recursive feature pyramid YOLOv5, RFP-YOLOv5)。首先,在特征提取阶段引入混合注

意力模块(convolutional block attention model, CBAM),加强对目标特征的学习。然后,YOLOv5的第3个检测层上再次上采样得到第4个检测层,从而获得更大尺寸的特征图,充分提取小目标信息。模型还用递归特征金字塔替换YOLOv5的FPN+PAN结构,加强浅层特征与深层特征的融合,实现更精确的病原微生物检测。为了降低模型参数量,对C3模块使用的数量进行了删减,并且在二次特征提取模块用DSConv(depthwise separable convolution)替换普通卷积。

1 病原微生物检测方法

1.1 RFP-YOLOv5 基本网络结构

RFP-YOLOv5主要由骨干网络(Backbone)、递归特征融合网络(RFP)和检测网络(Prediction)组成,其网络结构如图1所示。

首先,模型将裁剪后的WSI图像调整为相同的尺寸输入网络。输入图像被Focus结构切分为4份,在通道维度上进行拼接后,经过卷积操作得到无信息丢失的下采样特征图。然后,特征图在C3(concentrated comprehensive convolution)模块的分支进行卷积操作,得到的残差特征经过CBAM模块,加强对目标特征的表达能力。在特征融合阶段,将FPN融合的特征分别输入到Backbone对应的特征提取层进行特征的二次提取,并用DSConv替换普通卷积,降低模型参数量。最后,通过4个不同尺度的检测层获得预测结果。

1.2 RFP-YOLOv5 特征提取网络

在巴氏涂片中,不同病原微生物的嗜色风格不一样,且不同类之间的尺寸相差较大。因此,在特征提取阶段,本文将Backbone中的Bottleneck CSP替换成C3模块,并引入注意力机制CBAM,形成CBAMC3模块,使模型关注目标的颜色通道和空间位置信息,更加精准地定位和识别感兴趣区域。

如图2所示,CBAMC3分为C3和CBAM两部分。C3模块中的分支卷积操作进行残差特征的学习。为了加强目标特征表达能力,C3模块输出的残差特征图F经过通道注意力分支后得到通道注意特征Mc,将Mc与输入特征F做对应元素相乘操作后得到空间注意力模块的输入特征。输入特征再经

过空间注意力模块分支的池化拼接等操作后得到空间注意特征 M_s 。最后 M_s 与该分支的输入特征做乘法操作得到最终的特征。图像经过 CBAMC3 后进行可视化,并与 YOLOv5 的 Bottleneck CSP 的可视化结果进行比较,结果如图 3 所示。图中(a)表示各微生物的原图,(b)表示各微生物在 CBAMC3 模块

提取后的特征可视化热图,(c)表示各微生物在 Bottleneck CSP 模块提取后的特征可视化热图。在进行两模块的可视化热图对比后可得出,加入 CBAM 注意力模块的模型更侧重于目标特征的提取,忽略无关紧要的背景信息,而且对滴虫小目标信息丢失较轻微。

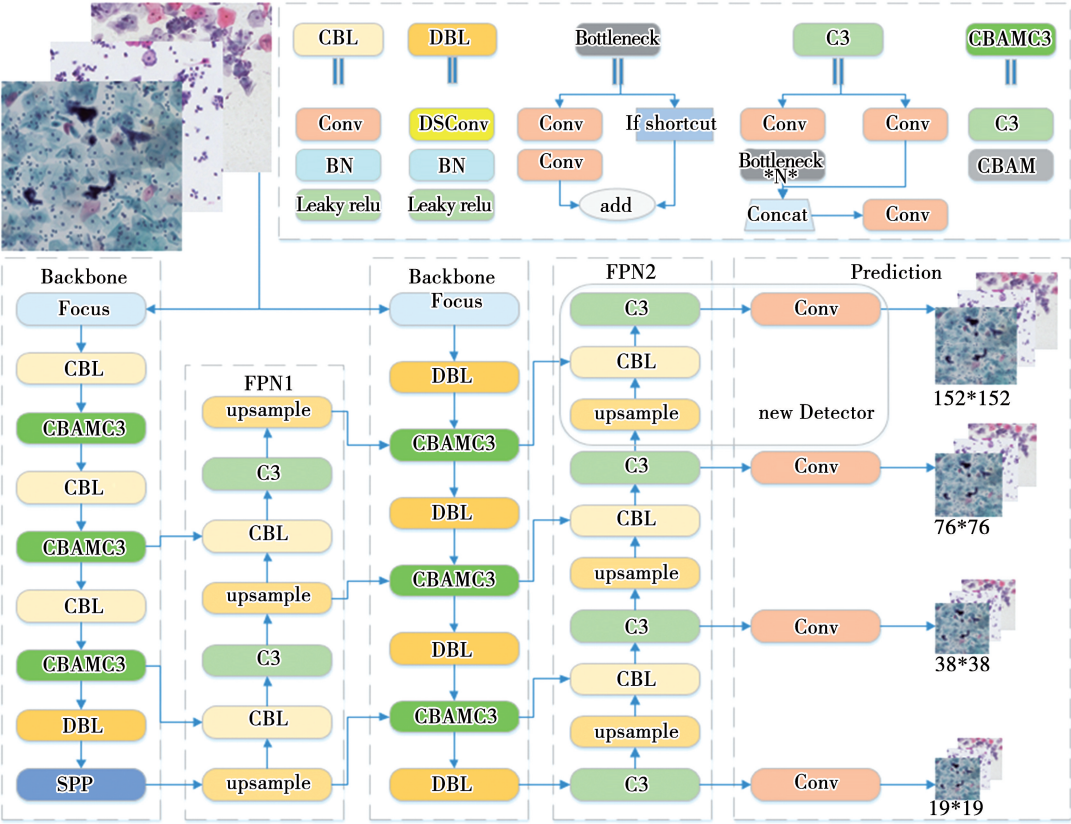


图 1 RFP-YOLOv5 网络结构图
Fig. 1 RFP-YOLOv5 network structure

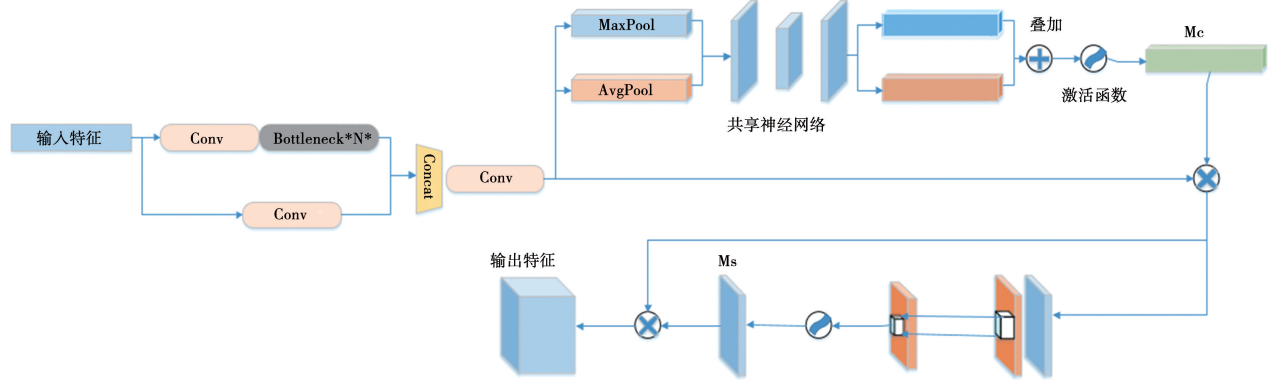


图 2 CBAMC3 结构图
Fig. 2 CBAMC3 structure

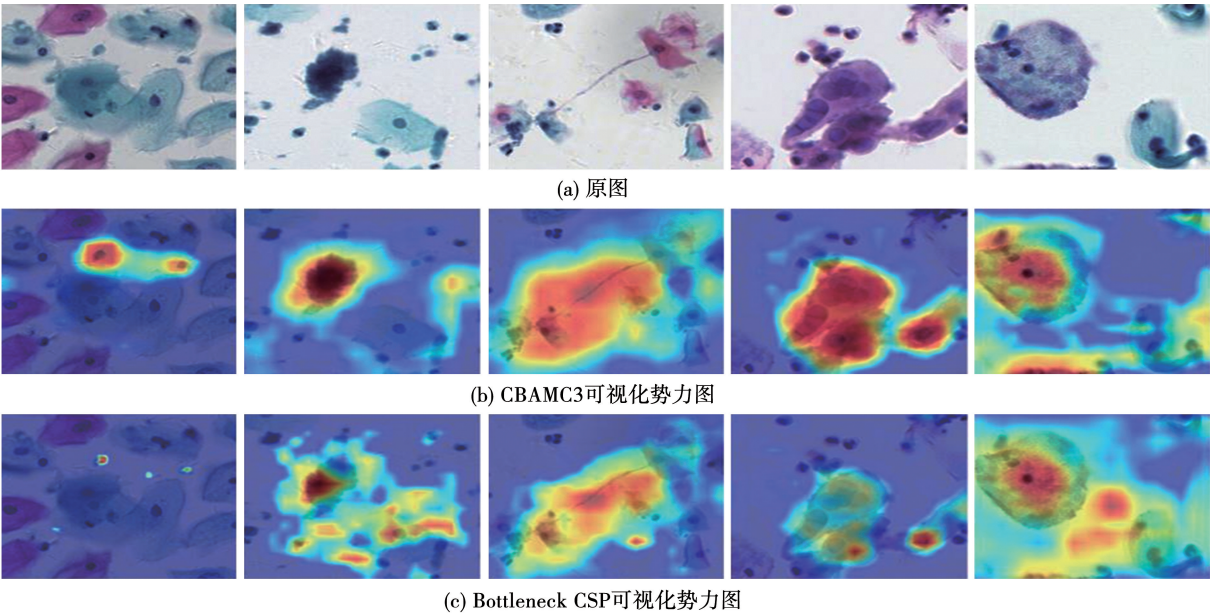


图 3 特征提取可视化热力对比图

Fig.3 Feature extraction visualization thermal comparison map

C3 模块是残差结构,用来学习更多的特征,但结构参数量较大且检测速度慢,在特征提取模块的第九层前已经使用了三层 C3 模块进行学习,且 SPP 模块对特征的全局特征和局部特征进行了融合。为了使模型轻量级化,在模型的特征提取模块去除了第九层的 C3 模块,并在模型二次特征提取模块中使用 DSConv 替换 Conv。DSConv 包括逐通道卷积和逐点卷积,逐通道卷积的每一个卷积核负责一个通道,产生的通道数与输入的通道数一样。逐点卷积大小为 1×1 ,将逐通道卷积的特征图在深度方向上进行加权组合,生成新的特征图,充分利用不同通道在相同空间位置上的特征信息。对于普通卷积,卷积层共 4 个 Filter,每个 Filter 包含 3 个 3×3 的 Kernel,参数量为 $4 \times 3 \times 3 \times 3 = 108$;而 DSConv 的参数量为 $3 \times 3 \times 3 + 1 \times 1 \times 3 \times 4 = 39$,参数量大大下降,如图 4 所示。

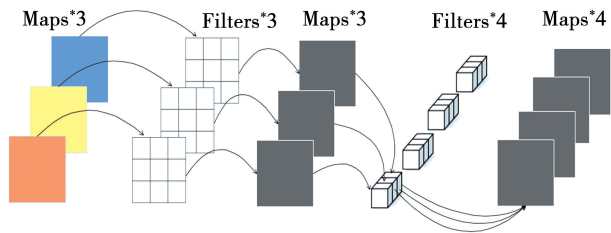


图 4 DSConv 结构图

Fig.4 DSConv structure

1.3 RFP-YOLOv5 特征融合网络

不同病原微生物的大小和形状各异。经测量统

计,滴虫在 $1\,024 \times 1\,024$ 的图像尺寸中标注矩形框面积范围在 $8 \times 12 \sim 20 \times 20$,放线菌矩形框面积范围在 $23 \times 15 \sim 320 \times 280$,疱疹矩形框面积范围在 $18 \times 12 \sim 120 \times 110$,念珠菌矩形框面积范围在 $18 \times 22 \sim 80 \times 103$,细菌矩形框面积范围在 $40 \times 68 \sim 120 \times 90$ 。2019 年 Kisantal 等^[26] 对小、中、大目标进行了区分,如表 1 所示。

表 1 小、中、大目标的区分

Tab.1 Differentiation of small, medium and large targets

	最小矩形框面积	最大矩形框面积
小目标	0×0	32×32
中目标	32×32	96×96
大目标	96×96	$\infty \times \infty$

根据大尺寸的特征图检测小目标,小尺寸的特征图检测大目标,当输入图像尺寸为 $1\,024 \times 1\,024$ 时,YOLOv5 的最大检测层大小为 76×76 ,每格特征图对应的感受野是 $(1\,024/76) \times (1\,024/76) = 13 \times 13$,再对应到原图中,可以用来检测像素大小在 $(1\,024/1\,024 \times 13) \times (1\,024/1\,024 \times 13) = 13 \times 13$ 以上的目标。为了提高对滴虫的检测,本文算法在 YOLOv5 3 个检测层的基础上增加了一个检测小目标的检测层。新增的检测层是在 76×76 特征层上继续上采样、普通卷积和残差卷积操作,得到 152×152 的特征图,再将卷积操作后的特征与骨干网络中的 152×152 特征图进行 concat 融合,得到第 4 个

检测尺度 152×152 , 对应到原图能检测出像素大小为 6×6 的目标。另外 3 个检测层大小分别为 76×76 、 38×38 和 19×19 。

除此之外,不同医院的巴氏涂片染色风格不一样、不同病人的细胞大小和形态不一样、同种病原微生物在不同人身上呈现的形态不一样,这三种情况给病原微生物的目标检测形成了复杂的背景。念珠菌的典型特征是将表中层细胞串连在一起,因此对于周围细胞特征的提取更重要。而 YOLOv5 在下采样和卷积的过程中局部特征丢失严重,虽然引入了 PAN 自下而上的特征融合来增强信息的提取,但是底层信息与顶层信息还是不能很好的融合。所以 RFP-YOLOv5 用递归特征金字塔 (recursive feature pyramid, RFP) 替换 YOLOv5 的 FPN + PAN 模块,如图 5 所示,将下采样后的特征信息与二次 Backbone 提取的特征融合,丰富特征的语义表达。

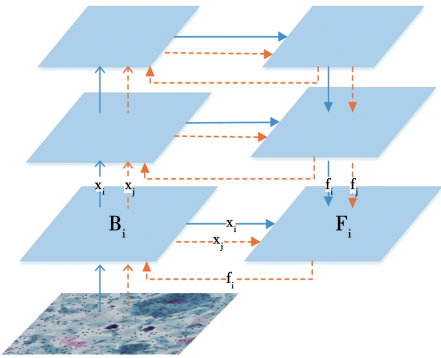


图 5 RFP 结构图
Fig. 5 RFP structure

RFP-YOLOv5 的 RFP 包括两个骨干特征提取模块和两个特征融合模块。RFP 是建立在特征金字塔之上的。RFP 将第一次特征融合模块输出的特征额外反馈合并到第二次的特征提取模块中,以实现查看图像两次或多次。第一次模块特征提取融合可由表达式(1)、(2)所示、第二次模块特征提取融合可由表达式(3)~(5)所示,第二次 Backbone 特征提取不仅包含了第一次特征融合后的信息,还包括了从原始图像提取的该层特征。

$$x_i = B_i(x_{i-1}) \quad (1)$$

$$f_i = F_i(f_{i+1}, x_i) \quad (2)$$

$$x_j = B_i(x_{i-1}), j < 3 \quad (3)$$

$$x_j = B_i(x_{j-1}, f_i), j > = 3 \quad (4)$$

$$f_j = F_i(f_{j+1}, x_j) \quad (5)$$

其中: x_i, x_j 分别表示第一次特征提取的第 i 层、第二次特征提取的第 j 层; B_i 表示第 i 层 Backbone 网络结构; f_i, f_j 分别表示第一次特征融合中第 i 层网络特征经第 i 层 FPN 融合后的融合特征、第二次特征融合中第 j 层网络特征经第 j 层 FPN 融合后的融合特征; F_i 表示第 i 层 FPN 融合结构。

2 实验与分析

2.1 病原微生物数据集构建方法

宫颈中常见的病原微生物主要包括:滴虫、细菌、放线菌、念珠菌、疱疹,5 种病原微生物的主要形态如图 6 所示。滴虫形态上呈梨形或者椭圆形,嗜蓝色,细胞核淡染且偏位;细菌主要是检测球杆菌,嗜蓝色,犹如细小的沙粒覆盖在细胞表面;放线菌是细丝状病原体,有成锐角的分支缠绕成团,形状类似“棉花团样”;念珠菌菌丝的特征是竹节状,横跨许多中表层细胞,巴氏染色呈嗜伊红色到灰棕色;疱疹病毒感染后细胞核呈毛玻璃样外观,像病理科固态的石蜡,略带折光性。不同病原微生物会引起不同类型的感染。

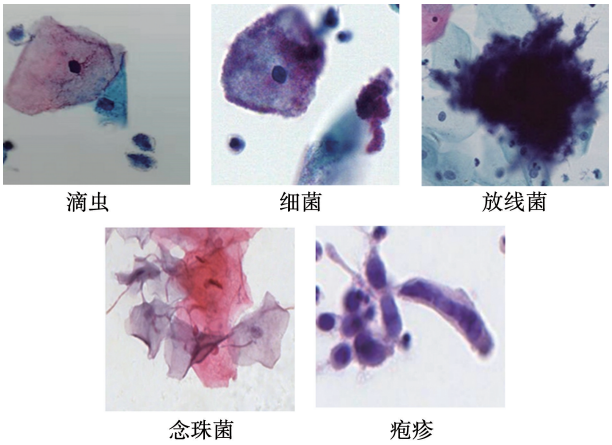


图 6 5 种病原微生物典型形态

Fig. 6 Five typical forms of pathogenic microorganisms

本文宫颈病原微生物的数据集来源主要有:哈尔滨医科大学附属第二医院、黑龙江精准元医学检验所、哈尔滨口岸医院。医院使用 TCT(thinprep cytologic test,液基薄层细胞检查)技术、巴氏染色和封片制作宫颈细胞病理涂片,然后经过 WISLEAP-10 扫描仪得到宫颈细胞病理图像 WSI。由于 WSI 的尺寸过大,所以通过动态裁剪方法后,共获得 12 233

张1 024 × 1 024分辨率的图片:滴虫3 720张、放线菌 2 641张、念珠菌1 518 张、细菌4 214 张和疱疹 140 张。对裁剪的图像用 labelImg 进行人工标注生成 xml 的标注格式,并通过专业的病理医生核对标注的类别正确性。最后将 5 类病原微生物图片分别按 8:2 的比例分成训练集和测试集,如表 2 所示。

2.2 模型训练

RFP-YOLOv5 的深度学习框架是 Pytorch,模型训练的初始参数:epoch 为 300 轮,lr 为 1×10^{-5} , batch_size 为 8, img_size 为 1 024 × 1 024。RFP-YOLOv5 使用 K-means 对标注的候选框进行聚类,得到初始锚框如表 3 所示。

表 3 锚框表		
Tab.3 Anchor box table		
特征图	感受野	锚框
152 × 152	最小	(8,15) (13,26) (24,17) (25,27)
76 × 76	小	(29,37) (53,31) (70,68) (78,116)
38 × 38	中	(117,85) (126,162) (178,116) (188,171)
19 × 19	大	(165,231) (278,197) (398,403) (1000,990)

RFP-YOLOv5 模型采用精确度 (Precision, P)、召回率 (Recall, R)、平均精度均值 (mean average precision, mAP) 来评价模型的优劣,计算公式分别如下:

$$P = \frac{TP}{TP + FN}$$
(6)

$$R = \frac{TP}{TP + FP}$$
(7)

$$AP = \int_0^1 P(R)$$
(8)

$$mAP = \frac{1}{c} \sum_{j=1}^c AP_j$$
(9)

其中:TP 为真阳性;FN 为假阴性;FP 为假阳性;c 为类别数。

2.3 实验结果与分析

训练完模型后,分别对 RFP-YOLOv5 模型和 YOLOv5 模型的 mAP 进行比较,具体如图 7 所示。修改后的 RFP-YOLOv5 模型在各类别的识别能力都比 YOLOv5 模型好,其中小目标滴虫提高了 6.2%,模型的 mAP 总提高了 4.9%,由于疱疹的数量较少,所以 mAP 较低。

表 2 数据集分布情况		
Tab.2 Distribution of datasets		
	训练集/张	测试集/张
滴虫	2976	744
放线菌	2 113	528
念珠菌	1 214	304
细菌	3 371	843
疱疹	112	28
总数	9 786	2 447

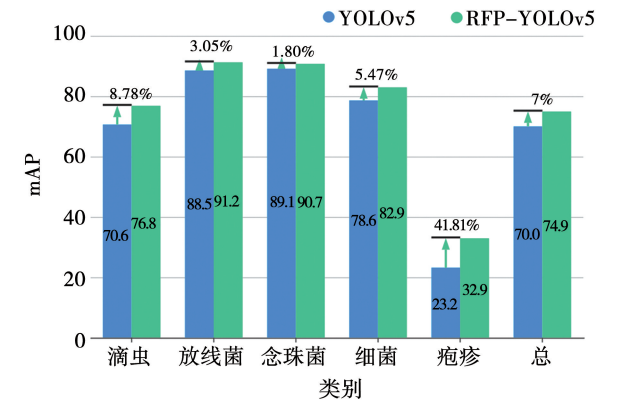


图 7 各类别 mAP 比较图
Fig.7 Map of each category

2.4 消融实验

为了验证本文算法的有效性,文中设置了 8 组消融实验,其中包含了通过减少模块数量和引入 DSCConv 减少参数量来实现模型轻量化的实验,实验采用平均精度均值 mAP、各类别 mAP、精确率 P、召回率 R、参数量 parameters 和推理时间来评价模型的优劣性,具体如表 4 所示。

表 4 消融实验
Tab.4 Ablation experiment

模型	mAP/%						P/%	R/%	parameters	推理时间/ (ms/张)
	放线菌	细菌	念珠菌	滴虫	疱疹	总				
YOLOv5	88.5	78.6	89.1	70.6	23.2	70.0	83.4	84.7	7 114 785	10.6
DCYOLOv5	87.9	77.7	88.9	70.4	24.1	69.8	84.6	81.4	5 886 216	8.9
DCYOLOv5 + CBAM	88.8	80.1	89.3	72.0	25.3	71.1	85.4	84.6	5 923 392	9.1
DCYOLOv5 + RFP	89.4	79.7	89.2	72.7	30.5	72.3	88.4	85.9	8 990 362	11.7
DCYOLOv5 + 多尺度	88.7	79.1	88.8	73.2	24.7	70.9	81.6	86.0	6 264 416	9.6
DCYOLOv5 + CBAM + RFP	89.6	80.2	89.5	73.5	30.6	72.7	87.5	85.2	9 045 320	12.2
DCYOLOv5 + CBAM + 多尺度	89.7	80.6	89.5	73.9	28.8	72.5	84.5	88.1	6 275 462	9.6
DCYOLOv5 + CBAM + RFP + DSConv	90.8	82.1	90.0	75.8	32.3	74.2	89.8	87.2	6 330 056	9.8
DCYOLOv5 + CBAM + RFP + 多尺度	90.2	81.8	89.9	76.0	31.6	73.9	89.3	86.9	9 088 142	12.6
RFP-YOLOv5	91.2	82.9	90.7	76.8	32.9	74.9	90.9	88.4	6 372 878	10.1

由表 4 数据可知,DCYOLOv5 表示从基准 YOLOv5 模型特征提取模块中去除了第 9 层的 C3 模块,降低了模型参数量, mAP 只下降了 0.2%。引入混合注意力机制 CBAM 后,模型的 mAP 提升了 1.1%,说明 CBAM 能有效的抑制无用的背景信息。给 YOLOv5 增加一个小目标的检测层后 mAP 提升了 0.9%。引入 RFP 网络结构后 mAP 提升了 2.3%,说明 RFP 能增强特征的语义表达。随着模块的增加,模型的 mAP 值都有一定的提升。实验结果表明,本文提出的模型在病原微生物检测任务上性能更好,RFP-YOLOv5 与基线模型 YOLOv5 进行对比, mAP 提高了 4.9%, P 、 R 分别提升了 7.5%、3.7%,说明预测的类别和定位更接近真实值,通过对模型的简化和使用 DSConv 卷积,使得模型的参数量降低,推理时间也快了 0.5 ms/张。

2.5 对比实验结果与分析

为了验证 RFP-YOLOv5 在病原微生物检测上的有效性,本文还与主流的目标检测算法进行比较,主要有 YOLOv5、faster R-CNN^[19]、SSD^[8]、YOLOv3^[5]、YOLOX^[7]、Fcos^[27]、Cascade R-CNN^[28]、PP-YOLO^[29]。实验以 mAP 和推理时间作为评价指标,结果如表 5 所示。由表可知,虽然 RFP-YOLOv5 的推理时间较长,但模型的 mAP 比其他目标检测模型都

有较大的提升,说明 RFP-YOLOv5 的模型性能更好。为了更直观的看到各模型的检测效果,实验中用 RFP-YOLOv5、YOLOv5、SSD、YOLOv3、YOLOX、Fcos、Cascade R-CNN 和 PP-YOLO 去检测不同类别病原微生物,检测效果如图 8 所示。

表 5 与现有模型对比
Tab.5 Comparison with existing models

模型	mAP /%	推理时间/ (ms/张)
faster R-CNN	62.5	25.35
SSD	64.5	20.05
YOLOv3	62.2	12.79
YOLOX	72.6	21.90
fcos	48.7	8.79
Cascade R-CNN	59.9	11.88
Yolov5	70.0	10.6
PP-YOLO	72.8	14.23
RFP-YOLOv5	74.9	10.1

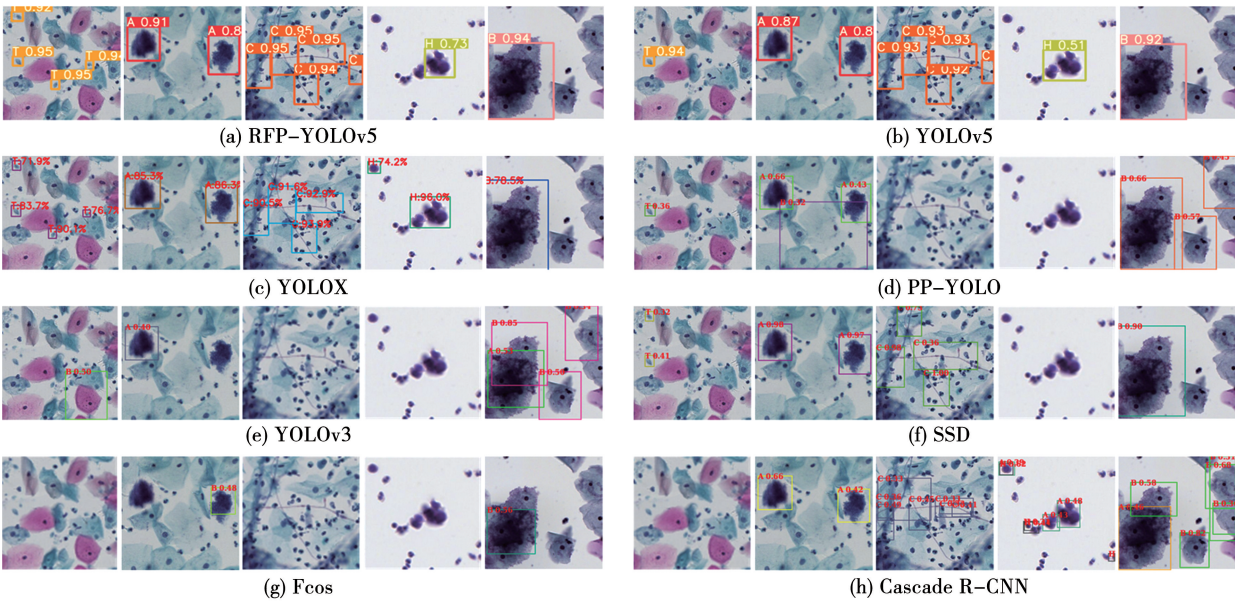


图 8 模型检测对比图

Fig.8 Comparison diagram of model detection

图 8 中标注框的标签 A 表示放线菌、B 表示细菌、C 表示念珠菌、H 表示疱疹、T 表示滴虫。可得出 YOLOv5 对于滴虫这类的小目标存在漏检的情况。YOLOx 能检测出各类别但准确率低于 RFP-YOLOv5。PP-YOLO 和 YOLO3 都存在对小目标漏检和误检。SSD 对放线菌和细菌的准确率较高,但对滴虫和疱疹识别较低。Fcos 性能较差,对于一些目标较大的放细菌都识别错误。Cascade R-CNN 存在重复识别的现象且识别准确率较低。相比较下,RFP-YOLOv5 在各类别的识别准确率上都比其他模型较高,RFP-YOLOv5 能检出 YOLOv5 漏检的滴虫小目标。

2.6 PascalVOC 数据集实验结果与分析

本文除了在宫颈病原微生物数据集上验证改进模型 RFP-YOLOv5 的优越性,还在目前常用的目标

检测公开数据集 PascalVOC2007 和 PascalVOC2012 上进行了实验。两数据集分别按照 8:2 划分为训练集和测试集,具体如表 6 所示。

表 6 数据集划分情况

Tab.6 Distribution of datasets

	训练集/张	测试集/张	总数/张
PascalVOC2007	8823	1140	9963
PascalVOC2012	9821	1719	11540

然后,用 RFP-YOLOv5 与多个目标检测模型训练两个数据集,并通过各类别的平均准确率均值进行比较,结果分别如表 7、表 8 所示。同时对目前较新的目标检测方法进行了模型总 *mAP* 的比较,结果如表 9 所示。

表 7 PascalVOC2007 目标检测 *mAP* 比较

Tab.7 PascalVOC2007 target detection *mAP* comparison

模型	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	<i>mAP</i> / %
	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv	
faster R-CNN	76.5	79.0	70.9	65.5	52.1	83.1	84.7	86.4	52.0	81.9	73.2
	65.7	84.8	84.6	77.5	76.7	38.8	73.6	73.9	83.0	72.6	
SSD	76.3	84.6	79.3	64.6	47.2	85.4	84.0	88.8	60.1	82.6	77.1
	76.9	86.7	87.2	85.4	79.1	50.8	77.2	82.6	87.3	76.6	
YOLOv5	91.5	85.6	80.6	68.3	73.0	87.8	88.8	90.5	73.3	77.9	80.5
	69.8	84.2	89.5	85.6	85.8	55.8	83.7	66.8	90.4	80.5	
RFP-YOLOv5	93.8	85.8	84.1	68.5	75.7	85.8	89.3	93.6	72.6	83.3	82.4
	68.2	85.9	93.1	87.5	88.4	58.8	85.4	74.3	91.6	82.3	

表 8 PascalVOC2012 目标检测 *mAP* 比较

Tab. 8 PascalVOC2012 target detection *mAP* comparison

模型	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	<i>mAP</i> / %
	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv	
faster R-CNN	86.5	81.6	77.2	58.0	51.0	78.6	76.6	93.2	48.6	80.4	73.8
	59.0	92.1	85.3	84.8	80.7	48.1	77.3	66.5	84.7	65.6	
SSD	87.9	82.8	73.8	61.5	45.3	81.4	75.6	92.6	57.4	78.3	75.4
	65.0	90.8	86.8	84.2	81.5	50.3	78.1	75.3	85.2	72.5	
YOLOv5	84.1	79.2	71.8	53.6	57.1	89.8	84.7	89.4	62.4	77.4	74.6
	59.2	79.1	80.0	80.2	86.1	55.7	82.6	55.9	79.4	82.4	
RFP-YOLOv5	88.2	83.3	76.0	59.1	57.9	92.5	86.1	91.2	63.9	81.3	77.7
	68.5	84.6	79.3	85.6	86.5	57.2	82.8	66.9	82.3	80.8	

由表 7、表 8 分析可知,改进后的 RFP-YOLOv5 比对比的目标检测模型都有较好的性能。在 PascalVOC2007 数据集 20 个类别中,RFP-YOLOv5 检测出的 16 个类别 *mAP* 都高于其他模型,而且模型 *mAP* 比基线 YOLOv5 高出了 1.9%。在 PascalVOC2012 数据集中,由于数据集的背景复杂度较高,所以识别出来的类别 *mAP* 基本都低于 PascalVOC2007。与比较的目标检测模型相比,RFP-YOLOv5 的 *mAP* 还是最高,比基线模型 YOLOv5 高出了 3.1%。综合分析可得,修改后的 RFP-YOLOv5 模型更能增强特征的语义表达,提高目标的检测能力。

表 9 目标检测在公共数据集的效果对比

Tab. 9 Effect comparison of target detection in public dataset

算法	PascalVOC	PascalVOC
	2007	2012
StairNet ^[30]	78.8	76.4
CollaborativeR-CNN ^[31]	77.5	75.7
MFCN ^[32]	80.1	77.4
Tiny RetinaNet ^[33]	71.4	73.8
RFP-YOLOv5	82.4	77.7

由表 9 分析可知,本文提出的算法与最近的目标检测算法相比较,在 PascalVOC2007 与 PascalVOC2012 两个数据集的表现上都是较优的。

2.7 CDetector 数据集实验结果与分析

本文算法除了在 PascalVOC 数据集上验证改进

模型 RFP-YOLOv5 的性能,还在病理公开数据集 CDetector^[34]做了对比。CDetector 数据集有 7410 张宫颈显微图像,包含了 ASC-US、ASC-H、Lsíl、Hsíl、SCC、AGC、滴虫、念珠菌、细菌、疱疹、放线菌共 11 个类别。将数据集分为训练集 6 666 张和验证集 744 张进行实验,实验结果如表 10 所示,可得出 RFP-YOLOv5 模型比对比模型的性能都较好。

表 10 目标检测在 CDetector 数据集的效果对比

Tab. 10 Effect comparison of target detection in CDetector dataset

模型	<i>mAP</i> /%
YOLOv5	51.9
YOLOx	49.4
PP-YOLO	45.4
YOLOv3	26.1
Comparison detector ^[34]	48.8
RFP-YOLOv5	59.9

3 结 论

宫颈病原微生物的存在会引起各种宫颈感染性疾病,所以对病原微生物的检测具有研究价值。目标检测技术已广泛被应用到宫颈有形成分的智能检测,但是对于 TCT 片的病原微生物智能检测这一任务研究几乎为零。同时,病原微生物的尺寸大小不一和背景复杂等原因,现有的目标检测模型对小目标的检测还有待提升。所以本文提出了基于递归特

征金字塔的病原微生物多尺度检测方法。为了解决滴虫目标过小,在YOLOv5的基础上引入小目标检测层来增强滴虫的检测。为了解决降采样过程中会把滴虫小目标的信息给丢失,在YOLOv5中引入RFP网络结构,对图像的信息进行二次提取,丰富特征的语义表达。考虑到疱疹跟正常细胞有一定的相似性,所以在Backbone模块引入CBAM注意力模块来增强对目标特征的提取,并引入DSConv卷积替换普通卷积,极大的减少了参数量,提高了性能。经实验验证,本文方法在病原微生物数据集和Pascal-VOC2007数据集和PascalVOC2012数据集上都有提高,其中,与基线YOLOv5相比,本文方法在病原微生物数据集、PascalVOC2007、PascalVOC2007和CDetector 4个数据集上的*mAP*分别提升了4.9%、1.9%、3.1%和8%。

参考文献:

- [1] 陈锐,廖秦平. 重视中国女性生殖道感染及阴道微生物诊治[J]. 中华检验医学杂志,2018,41(4):251.
CHEN Rui, LIAO Qiping. Pay Attention to the Diagnosis and Treatment of Female Reproductive Tract Infection and Vaginal Microecology in China[J]. Chinese Journal of Laboratory Medicine,2018,41(4):251.
- [2] ZHANG J, ZHONG Y, WANG X, et al. Computerized Detecion of Leukocytes in Microscopic Leukorrhea images [J]. Medical Physics,2017,44(9):4620.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]// Computer Vision & Pattern Recognition. IEEE, 2016, 91:779.
- [4] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger [C]//IEEE Conference on Computer Vision & Pattern Recognition. IEEE,2017:7263.
- [5] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement [J]. arXivpreprints:1804.02767,2018.
- [6] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal Speed and Accuracy of Object Detection [J]. arXiv Preprint arXiv:2004.10934,2020.
- [7] GE Z, LIU S, WANG F, et al. YoloX: Exceeding Yolo Series in 2021 [J]. arXiv Preprint arXiv: 2107.08430, 2021.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector [C]//European Conference on Computer Visio,2016:21.
- [9] FU CY, LIU W, RANGA A, et al. Dssd:Deconvolutional Single Shot Detector [J]. arXiv Preprint arXiv: 1701.06659,2017.
- [10] LI Z, ZHOU F. FSSD: Feature Fusion Single Shot Multi-box Detector [J]. arXiv preprint arXiv: 1712.00960,2017.
- [11] 侯剑平,王超,赵万里,等. 基于深度学习的白带显微图像细胞识别 [J]. 计算机应用与软件, 2021, 38(9):232.
HOU Jianping, WANG Chao, ZHAO Wanli, et al. Cell Recognition in Leucorrhea Microscopic Image Based on Deep Learning [J]. Computer Applications and Software, 2021,38(9):232.
- [12] 王静,孙紫云,郭苹,等. 改进YOLOv5的白细胞检测算法 [J]. 计算机工程与应用,2022,58(4):134.
WANG Jing, SUN Ziyun, GUO Ping, et al. Improved Leukocyte Detection Algorithm of YOLOv5 [J]. Computer Engineering and Applications,2022,58(4):134.
- [13] LI S, LI Y, LI Y, et al. YOLO-FIRI: Improved YOLOv5 for Infrared Image Object Detection [J]. IEEE Access, 2021,99:1.
- [14] DONG X, XU N, ZHANG L, et al. An Improved YOLOv5 Network for Lung Nodule Detection [C]//2021 International Conference on Electronic Information Engineering and Computer Science (EIECS), IEEE, 2021:733.
- [15] ZHU X, LYU S, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision,2021:2778.
- [16] LIANG G, ZHONG J, LONG M. Cervical Epithelial Cell Recognition Based on Improved Single Shot MultiBox Detector [C]//2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), IEEE,2020:1953.
- [17] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2015,38(1):142.
- [18] GIRSHICK R. Fast R-CNN [C]//Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015:1440.
- [19] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2017,39(6):91.
- [20] ZHANG J, HU H, CHEN S, et al. Cancer Cells Detection in Phasecontrast Microscopy Images Based on Faster R-CNN [C]//2016 9th International Symposium on Com-

- putational Intelligence and Design, Hangzhou, Zhejiang, China; Institute of Electrical and Electronics Engineers Inc, 2016:363.
- [21] HUNG J, GOODMAN A, LOPES S. Applying Faster R-CNN for Object Detection on Malaria Images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017:56.
- [22] 尹远来, 赵磊. 基于改进的 Faster R-CNN 模型的异常鳞状上皮细胞检测[J]. 智能计算机与应用, 2021, 11(2):7.
- YIN Yuanlai, ZHAO Lei. Detection of Abnormal Squamous Epithelial Cells Based on Improved Faster R-CNN Model[J]. Intelligent Computers and Applications, 2021, 11(2):7.
- [23] KANG S, HWANG J, CHUNG K. DLNet: Domain-specific Lightweight Network for On-Device Object Detection[C]//2022 International Conference on Information Networking (ICOIN), 2022:335.
- [24] XIONG Y Y, LIU H X, GUPTA S, et al. Mobile Dets: Searching for Object Detection Architectures for Mobile Accelerators[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021:3824.
- [25] JIN X, YI K, XU J. MoADNet: Mobile Asymmetric Dual-Stream Networks for Real-Time and Lightweight RGB-D Salient Object Detection[J]. in IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(11):7632.
- [26] KISANTAL M, WOJNA Z, MURAWSKI J, et al. Augmentation for Small Object Detection[J]. arXiv Preprint arXiv: 1902.07296, 2019.
- [27] TIAN Z, SHEN C, CHEN H, et al. Fcos: A Simple and Strong Anchor-free Object Detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4):1922.
- [28] CAI Z, VASCONCELOS N. Cascade R-CNN: Delving into High Quality Object Detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 6154.
- [29] LONG X, DENG K, WANG G, et al. PP-YOLO: An Effective and Efficient Implementation of Object Detector[J]. arXiv Preprint arXiv:2007.12099, 2020.
- [30] WOO S, HWANG S, KWEON I S. Stairnet: Topdown Semantic Aggregation for Accurate One Shot Detection[C]//IEEE Winter Conference on Applications of Computer Vision (WACV). Nevada, USA, 2018:1093.
- [31] 李志欣, 陈圣嘉, 周韬, 等. 协同级联网络 and 对抗网络的目标检测[J]. 计算机科学与探索, 2022, 16(1):217.
- LI Zhixin, CHEN Shengjia, ZHOU Tao, et al. Object Detection Based on Cooperative Cascade Network and Adversarial Network[J]. Computer Science and Exploration, 2022, 16(1):217.
- [32] 杨爱萍, 鲁立宇, 冀中. 多层特征图堆叠网络及其目标检测方法[J]. 天津大学学报(自然科学与工程技术版), 2020, 53(6):647.
- YANG Aiping, LU Liyu, JI Zhong. Multilayer Feature Graph Stacked Network and Target Detection Method[J]. Journal of Tianjin University (natural science & engineering), 2020, 53(6):647.
- [33] CHENG M, BAI J, LI L, et al. Tiny-RetinaNet: A one-stage Detector for Realtime Object Detection[C]//Eleventh International Conference on Graphics and Image Processing (ICGIP 2019), 2020:113730R.
- [34] LIANG Y, TANG Z, YAN M. Comparison Detector for Cervical Cell/Clumps Detection in the Limited Data Scenario[J]. Neurocomputing, 2021, 437(3):195.

(编辑:温泽宇)