

# 面向小目标检测的改进 RetinaNet 模型及其应用



罗月童 江佩峰 段昶 周波

合肥工业大学可视化与协同计算研究室 合肥 230601

(ytluo@hfut.edu.cn)

**摘要** 基于深度学习的目标检测算法广泛应用于工业检测,RetinaNet 算法因兼具速度与精度两方面的优势而备受关注,但对于小于  $32 \times 32$  像素的小目标,该算法的检测精度不能满足工业检测的要求。为此,文中以增强小目标的训练为基本思路,针对 RetinaNet 算法进行了如下改进:在采样阶段,将低层特征图 P2 添加到 FPN 中,以确保小目标能被充分采样,同时引入自适应训练样本选择策略,以保证增加特征层之后仍能保持足够快的检测速度;在训练后期采用了损失权重调整策略,用于提高小目标中困难样本的拟合度。针对公共数据集 MS COCO 2017 及实际应用中的 LED 点胶工业数据集,改进后的方法使小于  $32 \times 32$  目标的检测精度分别提高了 4.1% 和 10.7%,这表明改进后的方法能显著提升小目标检测的水平。

**关键词:** 深度学习;小目标检测;RetinaNet;自适应采样

**中图分类号** TP391

## Small Object Detection Oriented Improved-RetinaNet Model and Its Application

LUO Yue-tong,JIANG Pei-feng,DUAN Chang and ZHOU Bo

Visualization & Cooperative Computing, Hefei University of Technology, Hefei 230601, China

**Abstract** Object detection algorithms based on deep learning are widely used in industrial detection. The RetinaNet algorithm has attracted much attention because of its advantages in both speed and accuracy. However, for small objects smaller than  $32 \times 32$  pixels, the detection accuracy of this algorithm cannot meet the requirements of industrial detection. To this end, this article takes the enhancement of small object training as the basic idea, and makes the following improvements to the RetinaNet algorithm: in the sampling phase, the low-level feature map P2 is added to the FPN to ensure that the small object can be fully sampled, and adaptive training sample selection(ATSS) strategy is introduced to ensure that the detection speed is still fast enough after the feature layer is increased; the loss weight adjustment strategy is adopted in the later training stage to improve the fit of difficult samples in small objects. For the public data set MS COCO 2017 and the LED dispensing industrial data set in practical applications, the detection accuracy of this method for objects smaller than  $32 \times 32$  increases by 4.1% and 10.7%, respectively, indicating that this method can significantly improve the detection ability of small objects.

**Keywords** Deep learning, Small object detection, RetinaNet, Adaptive sampling

## 1 引言

目标检测一直是计算机视觉领域的研究热点。得益于深度学习强大的特征表示能力,基于深度学习的目标检测方法取得了传统检测方法难以取得的精确性和鲁棒性。但在实际应用中,往往存在大量的小目标难以被检测出的情况,如在工业检测中的发光二极管(Light Emitting Diode, LED)点胶缺陷检测中,就存在着大量的小缺陷被漏检,其根本原因在于现有的深度学习网络中广泛采用的卷积操作并不能完全提取这些小目标的关键特征<sup>[1]</sup>。因此,开展基于深度学习的小目标检测研究非常具有研究意义和实际应用价值。

当前基于深度学习的目标检测方法主要分为两类,即两阶段和单阶段的目标检测算法。两阶段的检测算法主要有 R-CNN<sup>[2]</sup>, Fast R-CNN<sup>[3]</sup>, Faster R-CNN<sup>[4]</sup>等,这类算法分为两个阶段:前一阶段负责候选区域的生成,通常使用的算法包括 Selective Search<sup>[5]</sup>和 Region Proposal Network<sup>[4]</sup>,其中前者通过分割图像区域和合并相似区域来生成候选区域,后者则利用卷积神经网络生成候选区域;后一阶段对前一阶段生成的候选区域进行精确的预测,包括精确的目标类别预测和位置预测。一般来说,两阶段检测算法的检测精度较高,但检测速度较慢。单阶段检测算法,如 YOLO<sup>[6]</sup>和 SSD<sup>[7]</sup>算法,则不生成候选区域,而是直接在特征图上进行密集的目标类别

到稿日期:2020-09-24 返修日期:2021-02-02 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61602146);国家重点基础研究发展计划(2017YFB1402200);安徽省科技攻关计划(1604d0802009)

This work was supported by the National Natural Science Foundation of China(61602146), National Basic Research Program of China(2017YFB1402200) and Key Science and Technology Program of Anhui Province, China(1604d0802009).

通信作者:周波(zhoubo810707@163.com)

预测和位置预测。该类算法一般具有较快的检测速度,但是检测精度通常比两阶段算法低。为在保证速度的同时进一步提高检测精度,研究人员针对单阶段检测算法进行了大量研究,RetinaNet<sup>[8]</sup>便是其中的佼佼者,该算法较好地平衡了检测的速度和精度,考虑到了工业检测中对检测速度的要求,因此本文基于 RetinaNet 进行研究,以提升小目标检测的水平。

小目标检测目前已经受到了广泛关注,相关研究可分为两大类:增强小目标信息和增强小目标训练。增强小目标信息的方法主要是通过多尺度图像训练<sup>[9-11]</sup>、多尺度特征层融合<sup>[1,12-14]</sup>和应用超分辨率技术<sup>[15-17]</sup>来强化小目标特征信息;增强小目标训练的方法主要是通过采样过程优化<sup>[18-19]</sup>、训练数据处理<sup>[20-22]</sup>和损失函数改进<sup>[8]</sup>来提高检测算法对小目标特征的学习能力。在前一类研究中,通常是以较大的计算代价换取对小目标检测能力的提升,因此对硬件的要求较高;在后一类研究中,通常无需增加过多的计算代价,对硬件的要求稍低,但目前还没有形成一个公认的解决方案。在工业缺陷检测中,对算法的检测性能要求较高,但硬件条件往往受到一定的成本限制,因此本文考虑从小目标训练增强方面研究小目标检测问题。

综上,为了提高小目标的检测效果、加强目标检测算法在工业检测中的应用,本文拟从 RetinaNet 算法的原理出发,分

析该算法的运行机制,以增强小目标训练为基本思路设计相关改进策略与实验,具体内容如下:1)在特征金字塔(Feature Pyramid Networks,FPN)<sup>[23]</sup>中加入低层特征图 P2,以获取更多的小目标特征;2)采用自适应训练样本选择策略<sup>[14]</sup>,减少每层特征图上的锚框数量,使算法在训练时能够选择合适的训练样本进行训练;3)设计损失权重调整策略,一定程度上解决训练后期的样本极端不均衡问题。

## 2 本文方法

### 2.1 RetinaNet 简介

RetinaNet 算法采用骨干网络 ResNet<sup>[24]</sup> 和特征金字塔<sup>[23]</sup>作为主体框架,两个全卷积(Fully Convolutional Networks,FCN)<sup>[25]</sup>子网络分别负责分类与位置回归,损失函数则采用了 Focal loss<sup>[8]</sup>。

如图 1 所示,预处理后的图像首先经过 ResNet 主干网络进行特征提取,进入 FPN 后得到  $P_3 - P_7$  这 5 层特征图,在每层特征图上生成锚框,使用预先设定的交并比阈值进行锚框与目标的匹配。为使不同大小、不同形状的目标都可以匹配到合适的锚框,需要在所有特征图上设置不同尺度、不同宽高的锚框,随后将特征图输入到目标分类与位置回归的子网络中。

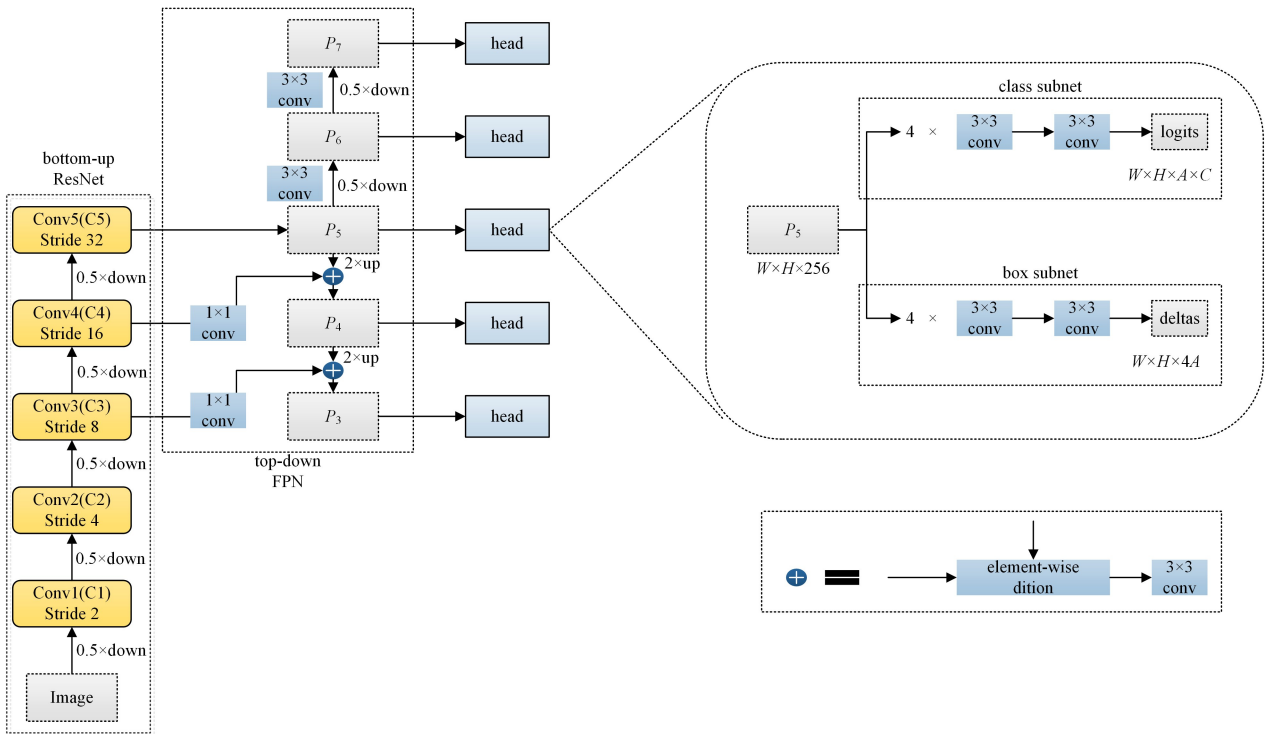


图1 RetinaNet 结构

Fig.1 Architecture of RetinaNet

RetinaNet 算法中的 Focal loss 通过改进交叉熵损失函数,降低训练过程中简单样本的损失权重,使算法能够聚焦于困难样本的学习,这有利于难检测小目标的特征学习。Focal loss 的计算式如式(1)所示:

$$FL(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (1)$$

其中,  $p_i$  指训练样本类别预测的置信度,  $\alpha_i$  和  $\gamma$  为超参数,本文分别采用经验值 0.25 和 2.0。

### 2.2 FPN 调整方法

RetinaNet 算法中,FPN 将低层特征与高层特征融合,在一定程度上增强了低层特征的语义信息,这使得该算法能够检测较小目标且整体精度较高,但依然达不到工业检测的要求,原因在于:在通过 FPN 后生成的中、高层特征图上,小目标的特征已经被高度压缩<sup>[26]</sup>,特征丢失严重,所以小目标难以被检测。另外,由于 RetinaNet 算法的锚框匹配策略,使得

小目标难以得到充分训练,如图 1 所示,通过 FPN 后得到的特征图有  $\{P_3, P_4, P_5, P_6, P_7\}$ , 对应累积步长分别为  $\{8, 16, 32, 64, 128\}$ , 特征图上的每个空间位置上基础锚框的宽高为 4 倍步长, 根据基础锚框尺度、宽高比  $\{1:2, 1:1, 2:1\}$  和缩放比例  $\{2^0, 2^{1/3}, 2^{2/3}\}$  为特征图上的每个空间位置生成 9 个锚框。显然, 对于原图中面积为  $16 \times 16$  的目标, 与之相交锚框的面积最小为  $32 \times 32$  (对应  $P_3$  特征图上的最小锚框), 那么交并比将低于 0.25 (由  $16 \times 16 / 32 \times 32$  得到), 远低于 0.5 这一常规阈值。RetinaNet 算法的匹配策略是首先使用预设的交并比阈值为目标匹配高于阈值的锚框, 然后为没有匹配锚框的目标匹配与其具有最高交并比的锚框, 那么在匹配时, 原图中面积小于  $16 \times 16$  的目标将只能匹配到一个训练样本, 这是不利于小目标训练的。因此, 本文考虑将 C2 层加入 FPN, 并融合生成  $P_2$  特征图 (见图 2),  $P_2$  特征图对应的累积步长为 4。对于原图中面积为  $16 \times 16$  的目标, 与之相交锚框的面积最小为  $16 \times 16$ , 那么可能存在多个交并比高于 0.5 的锚框, 即此类目标可能匹配到多个特征信息丰富的训练样本, 从而可以将更多可辨识小目标的特征纳入训练与检测, 以便利用更多细节特征检测小目标。

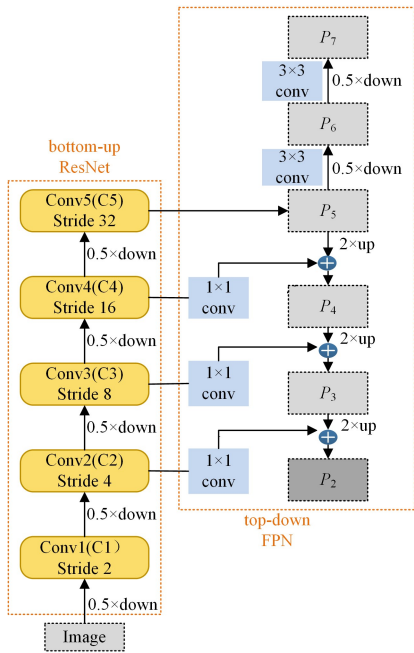


图 2 调整后的特征提取结构

Fig. 2 Adjusted feature extraction structure

### 2.3 基于自适应训练样本选择的锚框策略

将 C2 层加入 FPN 并得到特征图  $P_2$  后, 会带来两个问题: 1) 利用  $P_2$  特征图进行训练和推断都需要更多的卷积操作, 从而导致计算量大幅增加; 2) 与目标重叠的锚框的数量和交并比分布发生较大的变化, 导致原交并比阈值不再适用。为此, 本文采用自适应训练样本选择策略<sup>[19]</sup>来解决上述两个问题。自适应选择策略能够根据锚框与目标框交并比的统计特征为每个目标设定动态交并比阈值, 使目标能匹配到合适的训练样本, 交并比阈值  $t_g$  的计算式如式(2)、式(3)所示:

$$D_g = IoU(C_g, g) \quad (2)$$

$$t_g = Mean(D_g) + Std(D_g) \quad (3)$$

其中,  $C_g$  为根据与目标框的中心距离选择的候选锚框,  $g$  为目标框。

自适应选择策略对于锚框配置具有较强的鲁棒性<sup>[19]</sup>, 其在特征图的每个空间位置设置了单个合适尺度的锚框, 这使得每个预测子网络的输出维度由  $H \times W \times (A \times C)$  和  $H \times W \times 4A$  分别降低为  $H \times W \times C$  和  $H \times W \times 4$  (其中,  $A$  表示特征图的每个空间位置的锚框数量,  $C$  表示目标类别数量), 预测子网络的最后一对卷积操作的计算量降低为原来的  $1/A$ ; 同时, 由于算法会根据数据本身为每个目标计算并设定交并比阈值, 因此我们不再需要根据目标与锚框的交并比分布设计交并比阈值。

### 2.4 损失权重调整策略

RetinaNet 所采用的 Focal loss 降低了大量简单负样本的损失贡献, 从而有效地减小了单阶段算法中正负样本不均衡问题的影响。但是, 随着训练的进行, 模型在拟合越来越多的训练样本之后, 简单样本和困难样本的数据分布也会随之改变, 当进入训练后期, 会存在比训练前期更多的简单样本和更少的困难样本, 而 Focal loss 使用的是静态的超参数, 这些参数并不会随着数据分布的变化而得到调整<sup>[27]</sup>, 进而导致 Focal loss 难以应对训练后期出现的简单样本和困难样本不均衡的问题。

为此, 本文设计了一个损失权重调整策略, 使得在模型的训练后期能进一步提高困难样本的损失权重, 进而更好地应对样本极端不均衡问题, 以提高对难拟合小目标的学习能力。具体过程如下: 当训练至第  $k$  个 epoch 时, 在每次迭代获得预测值后, 首先比较每个训练样本的预测值与目标值, 得到预测标签与目标标签一致且预测框与目标框交并比高于阈值的训练样本集合  $S$ , 接着将训练样本集合  $S$  中没有相匹配样本的目标判定为难拟合目标, 并将与难拟合目标所匹配的训练样本的损失权重放大  $w$  倍, 最后使用调整后的损失权重计算损失。实验中使用网格搜索的方法寻找最优的  $k$  和  $w$  的取值, 置信度低于 0.05 的预测标签取值为背景类别, 交并比阈值取值为 0.5。

## 3 实验验证与分析

本文针对提出的构想进行了实验验证, 数据集选择了 MS COCO 公开数据集, 该数据集存在一定比例小于  $32 \times 32$  的小目标, 可用于测试小目标检测能力; 另外, 本文还基于实际应用项目中的 LED 点胶缺陷数据集测试本文算法。实验使用基于 pytorch 深度学习框架的开源工具包 MMDetection<sup>[28]</sup>, 工具包版本为 2.5.0, ResNet50 作为 RetinaNet 的骨干网络, 使用 ImageNet 预训练模型。硬件条件如下: 17-7700K CPU, Titan X GPU, 32GB 内存。各实验超参数如下: batch size 为 8, epoch 为 6, initial learning rate 为 0.002 (当训练至第 5 个和第 6 个 epoch 时, 学习率分别除以 10), linear warmup 为初始训练的前 500 步。图像短边缩放到 400 个像素 (长边则按短边缩放比例进行缩放), 图像按 50% 的概率随机水平翻转, 特征图上每个空间位置铺设单个锚框的长宽为 4 倍步长, 在本文实验的模型性能评估中, 置信度阈值取值为 0.05, 交并比阈值分别取  $[0.5, 0.55, 0.6, \dots, 0.95]$ 。

由于本文的研究内容主要是针对小目标检测, 本文实验结果仅统计尺度小于  $32 \times 32$  的目标计算评估指标。评估指标使用平均精度 (Average Precision, AP) 和平均召回率 (Average Recall, AR)。平均精度是综合查准率和查全率的

性能指标,其定义如式(4)所示:

$$AP = \sum_n (R_n - R_{n-1}) P_n \quad (4)$$

其中,  $R_n, P_n$  分别表示第  $n$  个检测阈值下的查全率和查准率。平均召回率是综合查全率和定位精度的性能指标,其定义如式(5)所示:

$$AR = 2 \int_{0.5}^1 recall(o) do \quad (5)$$

其中,  $o$  表示交并比阈值,  $recall(o)$  表示交并比阈值  $o$  下的查准率。

### 3.1 MS COCO 数据集上的实验

在 MS COCO 2017 数据集中,根据 MS COCO 测试标准中的尺度划分,尺度小于  $32 \times 32$  的 small 类型目标比例为 41.43%、尺度大于  $32 \times 32$  且小于  $96 \times 96$  的 medium 类型目标比例为 34.32%、尺度大于  $96 \times 96$  的 large 类型目标比例为 24.24%。本次实验使用 MS COCO 2017 train(118287 张图片)作为训练集,使用 MS COCO 2017 val(5000 张图片)作为测试集,用于模型性能的评估。在 MS COCO 数据集上的实验中,通过网格搜索(起始 epoch 取值集合为 {4, 5, 6}, 权重放大倍数取值集合为 {2.0, 3.0, 4.0, 5.0, 6.0})发现,当损失权重调整起始于第 6 个 epoch,且困难样本的损失权重放大倍数为 2.0 时,模型的性能达到最优。实验的模型性能评估结果如表 1 所列,其中  $AP_s$  和  $AR_s$  分别指模型在尺度小于  $32 \times 32$  目标上的平均精度和平均召回率,改进 1 指将低层特征图  $P_2$  加入 FPN,改进 2 指自适应训练样本选择策略,改进 3 则对应损失权重调整策略。根据该组实验测试结果,分别绘制交并比阈值为 0.5 和 0.7 时的 PR(Precision-Recall) 曲线,如图 3 所示。

表 1 MS COCO 数据集的实验结果

Table 1 Experimental results of MS COCO data set

情况	改进 1	改进 2	改进 3	$AP_s/\%$	$AR_s/\%$	$fps$
情况 1	×	×	×	11.6	20.1	35
情况 2	✓	×	×	13.2	25.2	21
情况 3	✓	✓	×	15.3	29.4	25
情况 4	✓	✓	✓	15.7	29.4	25

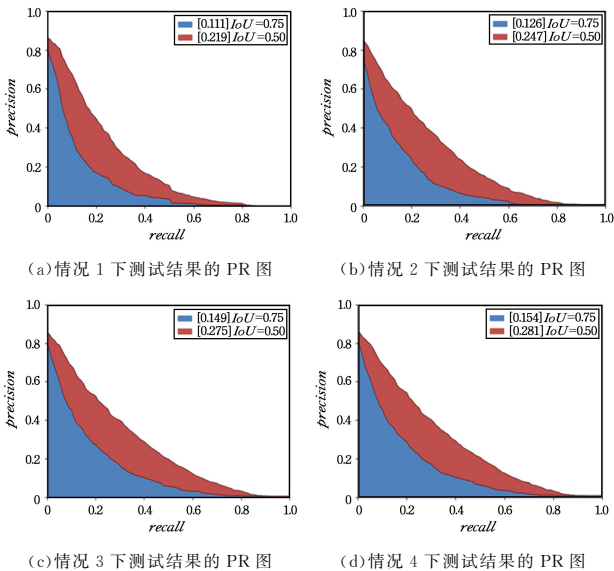


图 3 MS COCO 实验中 4 种情况下测试结果的 PR 图

Fig. 3 PR chart of the test results under 4 conditions in MS COCO experiments

### 3.2 LED 点胶缺陷数据集上的实验

为了进一步验证改进方法在工业检测中的实际效果,使用以小目标为主的 LED 点胶缺陷数据集进行了实验。LED 经过点胶工艺处理后,产品表面可能会出现气泡、异物等缺陷,而这些缺陷的存在会影响产品的光源亮度、使用寿命和安全性,因此存在小目标检测需求。本文检测的 LED 点胶缺陷数据包含共计 400 样本数据,缺陷类型包括黑色异物、白色异物和气泡,图片分辨率为  $400 \times 400$  像素。在这 400 个样本数据中共有 2417 个缺陷,存在 97% 的缺陷尺度小于  $32 \times 32$  像素。在 LED 点胶缺陷数据集上的实验中,通过网格搜索(起始 epoch 取值集合为 {4, 5, 6}, 权重放大倍数取值集合为 {2.0, 3.0, 4.0, 5.0, 6.0})发现,当损失权重调整起始于第 3 个 epoch,且困难样本的损失权重放大倍数为 4.0 时,模型性能达到最优。本文实验将 400 个样本数据随机划分为 4 等份数据,分别进行 4 组实验,其中第  $k$  组实验中取第  $k$  份数据作为测试集,用于模型性能评估,剩余 3 份数据合并作为训练集。实验的模型性能评估结果如表 2 所列,改进 1 指将低层特征图  $P_2$  加入 FPN,改进 2 指自适应训练样本选择策略,改进 3 则对应损失权重调整策略。另外,取 4 组实验中的第一组实验测试结果,并分别绘制交并比阈值为 0.5 和 0.7 时的 PR(Precision-Recall) 曲线,如图 4 所示。

表 2 LED 数据集的实验结果

Table 2 Experimental results of LED data set

情况	改进 1	改进 2	改进 3	$AP_s/\%$		$AR_s/\%$		$fps$
				mean	standard deviation	mean	standard deviation	
情况 1	×	×	×	33.4	2.0	46.1	4.9	42
情况 2	✓	×	×	40.1	1.7	50.1	1.7	31
情况 3	✓	✓	×	43.7	1.7	54.1	1.5	34
情况 4	✓	✓	✓	44.1	1.4	54.7	1.4	34

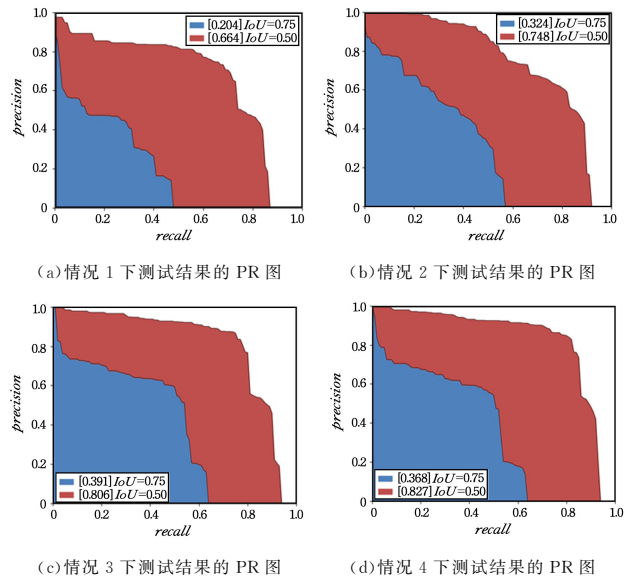


图 4 LED 第一组实验中 4 种情况下测试结果的 PR 图

Fig. 4 PR chart of the test results under 4 conditions in the first set of LED experiments

### 3.3 实验分析与结论

根据表 1、表 2 中的结果可得出如下结论:1) 对比情况 1

和情况 2 发现,改进 1 在提升模型在小目标上的平均召回率和平均精度的同时,降低了模型的检测速度,说明 FPN 中加入低层特征图 P2 虽然有利于小目标的检测,但同时也带来了较大的计算量;2)对比情况 2 和情况 3 发现,改进 2 不仅提升了模型的检测速度,还提升了模型在小目标上的检测精度,其中 MS COCO 实验中平均精度和平均召回率分别提升了 2.1%和 4.2%,LED 实验中平均精度和平均召回率分别提升了 3.6%和 4.0%,说明自适应训练样本选择策略在降低计算量的同时,使小目标匹配到了合适的训练样本;3)对比情况 3 和情况 4 发现,改进 3 使模型在小目标上的检测精度得到了进一步的提升,说明损失权重调整策略使部分难拟合小目标得到了相对充分的特征学习;4)对比情况 1 和情况 4,即原方法与改进方法相比,改进方法能够捕获更多的小目标(具体细节见图 5、图 6)。另外,从图 3 和图 4 的 PR 图可以直观看出,组合使用文中的改进策略能够有效提高检测算法对小目标的检测性能。

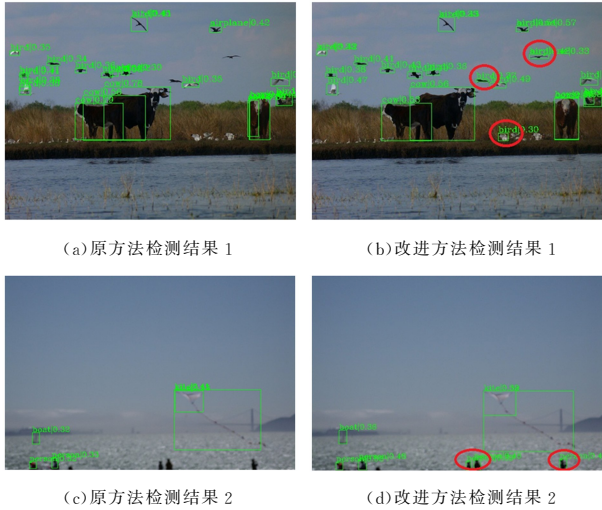


图 5 MS COCO 检测结果对比

Fig. 5 Comparison of MS COCO test results

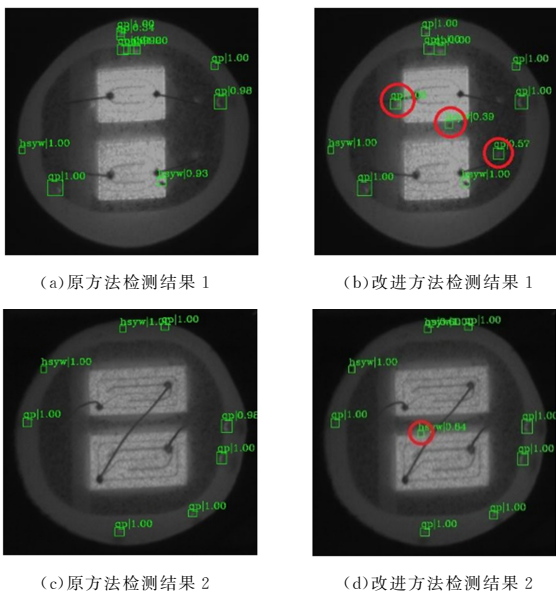


图 6 LED 检测结果对比

Fig. 6 Comparison of LED test results

由此可知,改进方法使用 P2 特征图和自适应训练样本选择策略使小目标得到了更好的采样,采用基于 Focal loss 的损失权重调整策略减小了训练后期样本极端不均衡的影响,使部分难拟合的小目标得到相对充分的特征学习,在通用目标检测和特定目标检测两种任务上,改进方法都使模型更适用于检测小目标。

**结束语** 本文在分析目标检测算法 RetinaNet 的基础上,以增强小目标训练为出发点,将低层特征图 C2 输入 FPN 得到检测特征图 P2,使用自适应训练样本选择策略来降低计算量,提出并采用损失权重调整策略,使小目标得到充分训练。根据在 MS COCO 数据集和 LED 点胶缺陷数据集上的实验结论可知,在 FPN 中加入低层特征图 P2 后有利于更好地检测小目标,而使用自适应训练样本选择不仅可以减少计算量,还能为小目标选择合适的训练样本。另外损失权重调整策略能使一些难拟合的小目标更好地被拟合,以上策略都有效提高了小目标的检测效率。但同时我们也认识到,虽然增强小目标训练能有效提升小目标检测的性能,但小目标有限的分辨率和信息<sup>[26]</sup>才是影响小目标检测的关键因素,因此未来的工作将探索如何有效增强小目标特征或减少小目标特征的丢失,以便进一步提高对小目标的检测精度。

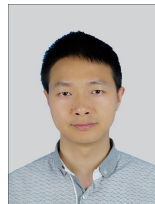
### 参考文献

- [1] HU G X, YANG Z, HU L, et al. Small object detection with multiscale features[J]. International Journal of Digital Multimedia Broadcasting, 2018(2): 1-10.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [3] GIRSHICK R. Fast r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [4] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[C]// Advances in Neural Information Processing Systems, 2015: 91-99.
- [5] UIJLINGS J R R, VAN DE SANDE K E A, GEVERS T, et al. Selective search for object recognition[J]. International Journal of Computer Vision, 2013, 104(2): 154-171.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]// European Conference on Computer Vision. Springer, Cham, 2016: 21-37.
- [8] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [9] SINGH B, DAVIS L S. An analysis of scale invariance in object detection snip[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 3578-3587.
- [10] SINGH B, NAJIBI M, DAVIS L S. SNIPER: Efficient multi-

- scale training[C]// Advances in Neural Information Processing Systems. 2018;9310-9320.
- [11] HU P, RAMANAN D. Finding tiny faces[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017;951-959.
- [12] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018;8759-8768.
- [13] QIN Z, LI Z, ZHANG Z, et al. ThunderNet: Towards Real-Time Generic Object Detection on Mobile Devices[C]// Proceedings of the IEEE International Conference on Computer Vision. 2019; 6718-6727.
- [14] PANG J, CHEN K, SHI J, et al. Libra r-cnn: Towards balanced learning for object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019;821-830.
- [15] BAI Y, ZHANG Y, DING M, et al. Sod-mtgan: Small object detection via multi-task generative adversarial network[C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018;206-221.
- [16] LI J, LIANG X, WEI Y, et al. Perceptual generative adversarial networks for small object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017;1222-1230.
- [17] NOH J, BAE W, LEE W, et al. Better to Follow, Follow to Be Better: Towards Precise Supervision of Feature Super-Resolution for Small Object Detection[C]// Proceedings of the IEEE International Conference on Computer Vision. 2019;9725-9734.
- [18] ZHU C, TAO R, LUU K, et al. Seeing Small Faces from Robust Anchor's Perspective[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018;5127-5136.
- [19] ZHANG S, CHI C, YAO Y, et al. Bridging the Gap Between Anchor-based and Anchor-free Detection via Adaptive Training Sample Selection[J]. arXiv:1912.02424, 2019.
- [20] CHEN Y, ZHANG P, LI Z, et al. Stitcher: Feedback-driven Data Provider for Object Detection[J]. arXiv:2004.12432, 2020.
- [21] YU X, GONG Y, JIANG N, et al. Scale Match for Tiny Person Detection[C]// The IEEE Winter Conference on Applications of Computer Vision. 2020;1257-1265.
- [22] KISANTAL M, WOJNA Z, MURAWSKI J, et al. Augmentation for small object detection[J]. arXiv:1902.07296, 2019.
- [23] 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.
- [24] 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.
- [25] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015;3431-3440.
- [26] DUAN K, DU D, QI H, et al. Detecting small objects using a channel-aware deconvolutional network[C]// IEEE Transactions on Circuits and Systems for Video Technology. 2019.
- [27] LI B, LIU Y, WANG X. Gradient harmonized single-stage detector[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2019;33;8577-8584.
- [28] CHEN K, WANG J, PANG J, et al. MMDetection: Open mmlab detection toolbox and benchmark[J]. arXiv:1906.07155, 2019.



**LUO Yue-tong**, born in 1978, Ph.D, professor, is a member of Chinese Computer Society, Computer Aided Design and Graphics Committee. His main research interests include image processing and scientific visualization.



**ZHOU Bo**, born in 1981, Ph.D, associate professor. His main research interests include digital terrain analysis and object detection.