

文章编号: 2095-2163(2021)10-0005-08

中图分类号: TP391.4

文献标志码: A

基于 RB-YOLO 的生活垃圾实时检测算法

董霄霄, 刘振国, 宋滕滕, 何益智, 李 钊

(山东理工大学 计算机科学与技术学院, 山东 淄博 255000)

摘要: 生活垃圾图片背景复杂,同一类垃圾特征差异性大,给垃圾检测带来了极大的挑战。为实现复杂场景下生活垃圾的实时检测,论文提出了一种基于改进 YOLOv3 的生活垃圾检测算法 RB-YOLO。在该垃圾检测模型中,将 YOLOv3 与轻量级卷积神经网络 RepVGG 相结合,解耦模型训练和推理,在保证特征提取能力的同时提高模型推理速度。采用加权双向特征金字塔网络 BiFPN 优化原有的特征金字塔结构,通过双向特征融合充分利用多尺度语义特征。同时引入空洞空间金字塔池化 ASPP 增强特征提取,获取多尺度特征信息。在提出的 DGD 数据集上进行实验分析,与 YOLOv3 模型相比,RB-YOLO 模型在垃圾检测任务中准确率和速度都有所提升,同时在测试集和网络图片上也取得了不错的检测效果,验证了 RB-YOLO 的有效性和鲁棒性。

关键词: YOLOv3 算法; 垃圾检测; RepVGG; BiFPN; ASPP

Real-time detection algorithm of domestic garbage based on RB-YOLO

DONG Xiaoxiao, LIU Zhenguo, SONG Tengting, HE Yizhi, LI Zhao

(School of Computer Science and Technology, Shandong University of Technology, Zibo Shandong 255000, China)

[Abstract] The background of domestic garbage pictures is complicated, and the large differences in the characteristics of the same type of the garbage have brought great challenges to garbage detection. In order to realize the real-time detection of domestic waste in complex scenarios, the paper proposes a domestic waste detection algorithm RB-YOLO based on improved YOLOv3. In this garbage detection model, YOLOv3 is combined with the lightweight convolutional neural network RepVGG to decouple model training and reasoning, and improve the speed of model reasoning while ensuring feature extraction capabilities. The weighted bidirectional feature pyramid network BiFPN is used to optimize the original feature pyramid structure, and multi-scale semantic features are fully utilized through bidirectional feature fusion. At the same time, the hollow space pyramid pooling ASPP is introduced to enhance feature extraction and obtain multi-scale feature information. Based on the above research, experimental analysis is performed on the proposed DGD data set. Compared with the YOLOv3 model, the RB-YOLO model has improved accuracy and speed in garbage detection tasks. Furtherly, it has also achieved good detection results on the test set and network pictures, and verified the effectiveness and robustness of RB-YOLO.

[Key words] YOLOv3 algorithm; garbage detection; RepVGG; BiFPN; ASPP

0 引言

根据全国大、中城市固体废物污染环境防治年报^[1],2019年196个大中城市生活垃圾产生量达到2.3亿吨。而据中国城市环境卫生协会统计,国内每年产生近10亿吨垃圾,并且城镇生活垃圾还在以每年5%~8%左右的速度递增。大量垃圾的堆弃造成国内近三分之二的城市面临垃圾围城的危机,同时导致土地资源匮乏和环境污染加剧等问题。实现可持续发展,建立一个有效的垃圾处理体系至关重要,而垃圾检测则是垃圾处理中的重要一环。根据可回收物、厨余垃圾、有害垃圾、其他垃圾的分类原则对垃圾进行识别和定位,能够在促进资源化管理

的同时提高垃圾处理效率。

在垃圾检测智能化的背景下,研究人员展示了各种深度学习算法在垃圾检测方面的能力。Sousa等人^[2]分别基于形状和材质对食物托盘上的垃圾进行检测和分类,提出了一种基于Faster RCNN^[3](Fast Region-based Convolutional Network method)的分层次的级联网络模型,在自制的数据集上分别达到了86%和80.9%的准确率。文灿华等人^[4]从网络结构出发,分析对比了3种典型的主干网络VGG-16^[5]、ResNet-101^[6]、MobileNet_v1^[7]在Faster RCNN上的性能,实现对多类别垃圾的高精度检测。类似地,Melinte等人^[8]采用SSD-MobileNetV2、SSD-InceptionV2、Faster-RCNN-InceptionV4对城市

基金项目: 山东省高等学校青年创新团队发展计划资助项目(2019KJN048)。

作者简介: 董霄霄(1996-),女,硕士研究生,主要研究方向:智能信息处理;李钊(1983-),男,博士,副教授,主要研究方向:近似计算、机器学习。

通讯作者: 李钊 Email: lizhao_buaa@126.com

收稿日期: 2021-11-22

垃圾进行检测,研究表明目标检测器的性能取决于数据增强、CNN模型、数据集数量、损失优化、超参数调整、迁移学习。

以上两阶段目标检测算法虽然准确率高,但是训练难度大、检测速度慢,在实际应用中限制了垃圾检测的效率。因此在检测速度方面具有极大优势的YOLO^[9-11](You Only Look Once)引起了广泛研究。许伟等人^[12]为了提高检测速度提出一种改进的MobileNetV3-YOLOv3,在YOLO层加入GRU结构,平均准确率为90.50%,检测速度达到18帧/s,但实验数据集仅包含6种类别约5000张图片。魏铖磊等人^[13]提出ECA_ERFB_s-YOLOv3算法提高检测精度,采用ResNet-50作为特征提取网络,引入多尺度感受野模块和注意力机制,最终达到了89.4%的准确率、17帧/s的检测速度,但是模型较大,不利于边缘设备的部署。同样地,Zhao等人^[14]提出一种用于生活垃圾检测的Skip-YOLO模型,采用密集块提取高维特征图,检测精度提高了22.5%,平均召回率提高了18.6%,但是模型复杂且实时性差。

尽管现有的研究已经将垃圾图像目标检测与深度学习方法相结合,以改进垃圾图像的识别与定位,这一领域依旧存在着一些问题:

(1)基于深度学习的垃圾图像检测的研究脱离实际。目前研究以实验室背景的小数据集为主,与真实场景存在较大差距,难以实际应用工业中。

(2)现有的模型训练与预测运算复杂、时效性差,无法满足实际应用的要求。

(3)生活垃圾数据集具有类内差异性大的特点(例如同属于“果皮果肉”类的香蕉和苹果具有完全不同的颜色、形状等特征),而目前的研究只是从一个广泛的角度进行,未能解决实际生活中的垃圾检测问题。

针对上述问题,本文提出了一种基于改进YOLOv3的复杂场景下生活垃圾检测算法RB-YOLO。以YOLOv3算法为基础,采用轻量级神经网络RepVGG^[15]作为骨干网络,通过结构重参数化解耦模型训练与推理,减少了51%的参数量,提升了43%的速度。采用加权双向特征金字塔(Bi-directional Feature Pyramid Network, BiFPN)^[16]优化特征金字塔(Feature Pyramid Networks, FPN)^[17]结构,通过加权特征融合使得模型能够更好地平衡不同尺度的特征信息,加强低语义和高语义信息共享,提高同类垃圾相似特征的敏感性。引入空洞空间池化金字塔ASPP(Atrous Spatial Pyramid Pooling, ASPP)^[18]多尺度获取特征信息,提高垃圾检测复杂背景下模型特征提取能力。最后结合迁移学习,进一步提高模型准确率。

1 RB-YOLO模型结构

生活垃圾排列无序、形状多变,同类垃圾的特性差别很大,这就产生了一定程度的数据波动,这种数据波动使得网络在学习过程中存在过拟合现象。训练结果缺乏足够的泛化,影响最终平均精度。本文提出了一种用于生活垃圾实时检测的RB-YOLO模型,具体结构如图1所示。采用RepVGG进行特征提取,在训练和推理阶段采用不同的网络结构,保证模型在训练时的准确率和推理时的速度。特征提取网络还结合了ASPP以实现多重感受野,由于 26×26 和 52×52 的检测器拥有足够大的感受野,因此本文仅对 13×13 的检测器添加ASPP模块,在增加少量计算量的同时扩大感受野。完成特征提取之后,采用BiFPN对提取到的特征信息进行加权双向特征融合增强,以充分利用低层次的特征信息和高层次的语义信息。

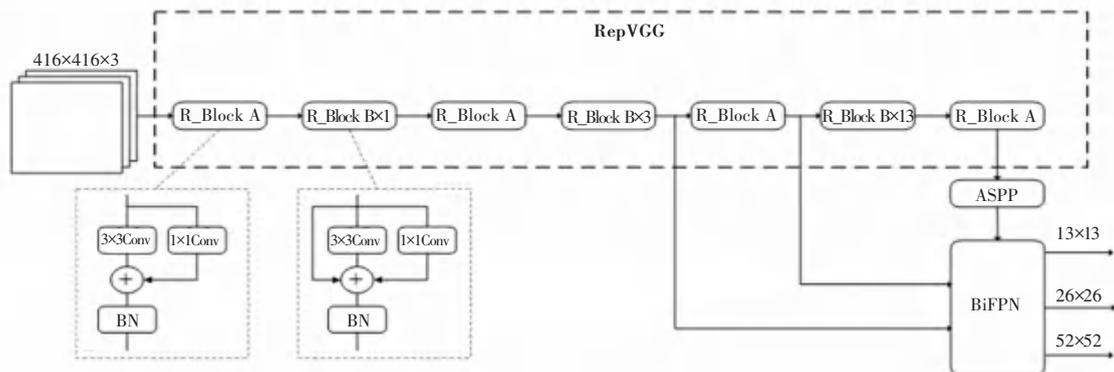


图1 RB-YOLO网络结构图

Fig. 1 Structure of RB-YOLO networks

1.1 改进主干特征提取网络

为了满足检测模型的实时性和准确率, 本文针对 YOLOv3 的主干特征提取网络进行了优化, 采用轻量级卷积神经网络模型 RepVGG 替换原始的 DarkNet-53。相比于 DarkNet-53, RepVGG 具有以下优点: 首先, RepVGG 全部采用 3×3 卷积。与 1×1 卷积和 5×5 卷积相比, 3×3 卷积计算密度 (理论运算量/所用时间) 更大, 计算速度更快; 其次, RepVGG 采用单路架构运算速度快且占用内存少。

RepVGG 网络由 20 多层 3×3 卷积堆叠而成, 分成 5 个 stage, 每个 stage 的第一层是 $stride = 2$ 的降采样, 每个卷积层用 $ReLU$ 作为激活函数。RepVGG 的主要特点是解耦了模型训练和推理, 针对网络训练和推理阶段不同的需求采用不同的网络架构。在训练阶段最重要的是准确率, 因此 RepVGG 网络采用多分支结构, 由 3×3 卷积、 1×1 卷积分支和 identity 的残差分支相结合的 RepVGG Block 组成, 多分支的结构通过应用不同的卷积核获得不同的感受野, 将不同感受野获取到的信息进行融合强化特征提取以提高模型性能。同时, 具有多个分支的残差结构相当于网络具有多条梯度流通过程, 等同于同时训练多个网络并进行融合, 类似于模型集成。在推理阶段更关注速度, 因此 RepVGG 网络采用单路结构, 通过结构重参数化将训练阶段的模型等价转换得到类 VGG 的推理模型, 由 3×3 卷积块构成极大地提高了推理速度, 方便模型部署和加速。在 RepVGG 系列模型中, RepVGG-A0 作为基础模型参数量少且推理速度快, 在获取特征信息方面也有较好的性能。考虑到在实际工业中的应用, 在本文中采用 RepVGG-A0 作为主干网络。

RepVGG 网络中提出的结构重参数化技术, 通过改变结构和参数将网络模型由多分支结构转换为单分支结构, 2 种模型在数学推理上是等价的, 能够提高设备内存利用率, 同时提升模型的推理速度。结构重参数化具体流程如图 2 所示, 主要思路包括卷积层和 BN 层融合 (图 2(a))、卷积层合并 (图 2(b)) 两部分。

1.1.1 卷积层和 BN 层融合

在卷积神经网络中, BN 层 (Batch Normalization) 得到了广泛的应用, 能够提高网络泛化能力, 加快收敛速度, 缓解了梯度弥散问题。在网络的推理阶段, 将 BN 层的运算直接嵌入到卷积层中, 减少运算量, 提升网络的运行速度。

卷积层输出如公式 (1) 所示:

$$y_1 = w * x + b \quad (1)$$

其中, x 为每一层输入; w 为权重; b 为偏置。

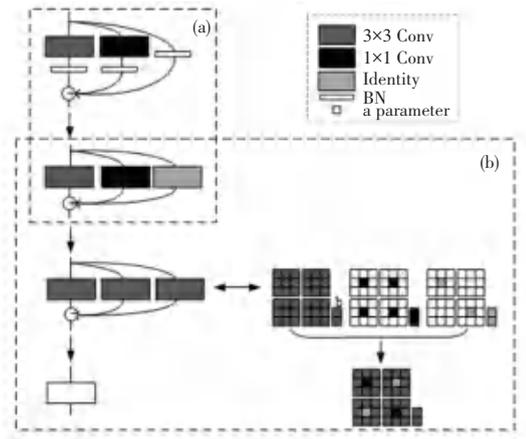


图 2 结构重参数化过程

Fig. 2 Process of structure reparameterization

BN 层主要包括归一化和缩放两部分, BN 层输出的具体计算方式如式 (2) 所示:

$$y_2 = \gamma \frac{y_1 - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (2)$$

其中, γ 为缩放参数; β 为偏移参数; μ 为样本均值; σ 为标准差; ϵ 表示一个极小值 (防止分母为 0)。将卷积层和 BN 层合并, 将式 (1) 代入式 (2) 可得:

$$y_2 = \gamma \frac{w * x + b - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (3)$$

将 BN 层输出 y_2 经过变形可得到如下形式:

$$y_2 = \left(\frac{\gamma}{\sqrt{\sigma^2 + \epsilon}} w \right) * x + \frac{\gamma}{\sqrt{\sigma^2 + \epsilon}} b - \frac{\gamma}{\sqrt{\sigma^2 + \epsilon}} \mu + \beta \quad (4)$$

在推理阶段, BN 层的样本均值和标准差来自于训练样本的数据分布, 缩放参数和偏移参数是和它模型参数一起通过梯度下降方法训练得到, 即 $\mu, \sigma, \gamma, \beta$ 均为固定常数值。则令 $w' = \frac{\gamma w}{\sqrt{\sigma^2 + \epsilon}}, b' = \frac{\gamma}{\sqrt{\sigma^2 + \epsilon}} b$

$- \frac{\gamma}{\sqrt{\sigma^2 + \epsilon}} \mu + \beta$, 得到公式 (5), 因此根据推导可知将

BN 层直接合并到卷积层的计算中, 相当于将卷积核缩放一定倍数, 并对偏置进行一定改变, 省去了整个 BN 层的计算量, 提高了推理速度。数学公式具体如下:

$$y_2 = w' * x + b' \quad (5)$$

1.1.2 卷积层合并

经过卷积层和 BN 层融合之后, 可以得到 3×3 卷积、 1×1 卷积分支和 Identity 分支, 以及 3 个偏置向量。假设输入通道数等于输出通道数, 且每种卷积具有相同的步长。 1×1 卷积和 3×3 卷积层合并过程如图 3 所示。根据图 3 可知, 1×1 卷积可转换为中心值为 1×1 卷积的权重值, 其他位置值为 0 的 3×3 卷积, 1×1 卷积和 3×3 卷积层合并即将 1×1 卷积

核中的数值移动到 3×3 卷积核的中心。而 Identity 分支可以看作是一个核为单位矩阵的 1×1 卷积,进而填充为特殊的 3×3 卷积,将其与输入的特征映射相乘之后,输出数据保持不变。 1×1 卷积和 3×3 卷积层合并过程如图 3 所示。根据卷积的可加性, 1×1 卷积分支 Identity 分支进行等价转换后即可和 3×3 卷积进行合并。最终的偏置向量可由 3 个偏置向量相加得到。

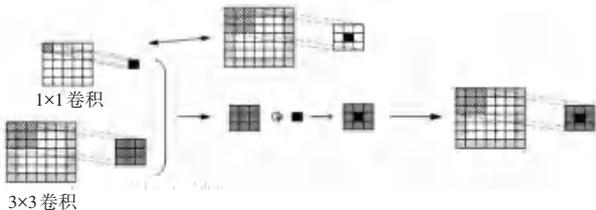


图 3 1×1 卷积和 3×3 卷积层合并

Fig. 3 Combination of 1×1 convolution and 3×3 convolution layer

1.2 引入空洞空间卷积金字塔池化

在进行了主干特征提取之后,本文引入了 ASPP 模块。ASPP 在 SPP (Spatial Pyramid Pooling)^[19] 模块基础上引入了空洞卷积^[20],可以增强模型对不同尺度分割目标的感知能力。空洞卷积是在标准卷积中添加空洞来增加感受野,可以有效缓解下采样导致的空间分辨率降低的问题。对于给定的输入,ASPP 采用多个具有不同膨胀率的并行扩张卷积层,能够在不降低特征图分辨率的情况下扩张感受野,使得计算成本不显著增加。ASPP 通过不同膨胀率空洞来获取长跨度的上下文语义信息,这不仅提高了准确率,而且速度更快。

针对垃圾检测背景复杂的问题,ASPP 能够通过不同的感受野将局部信息和全局信息进行融合,丰富了特征图的表达能力,提高复杂场景下垃圾检测能力。在该模型中,使用全局平均池化、一个普通卷积和 3 个空洞卷积 $\{Rate = 3, 5, 7\}$ ($Rate$ 为空洞率)。图 4 展示了包含 3 个空洞卷积的 ASPP 结构,将图像级特征的全局平均池化与空洞卷积的结果进行融合得到最终结果。

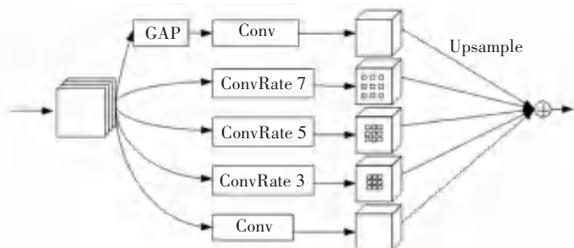


图 4 ASPP 结构图

Fig. 4 ASPP structure chart

1.3 加强多尺度特征融合

YOLOv3 采用 FPN 来检测大小不同的物体,不同的尺度输出特征的在融合时贡献相等,对有效信息没有特殊关注,同时 FPN 获取到的特征信息受到单向信息流的限制。BiFPN 主要包括双向多尺度连接和加权特征融合两部分。首先,以类似残差连接的方式在输入和输出节点之间增加了一条额外的边,在增加少量计算的代价下融合更多特征;同时借鉴 PANet^[21] 提出的自上而下和自下而上的双向路径作为基础层,多次重复以实现更有效的特征融合。其次,考虑到不同尺度特征图重要性不同,BiFPN 使用加权融合的方法来融合不同分辨率的特征层,通过可训练的权重,调节特征在特征融合中的贡献,可以用于解决 FPN 中的平等贡献问题,针对多尺度的目标细粒度检测任务有很好的效果。

针对生活数据集同类别差异性大的问题,本文采用 BiFPN 对提取到的特征进行加强,将主干的 C3 到 C5 层的特征传递到 BiFPN 作为选定的多尺度特征层。自上而下的特征融合通过上采样来完成,高层特征通过上采样调整分辨率,进而与低层特征信息进行融合,对融合后的特征进行 3×3 卷积以消除混叠效应。自下而上的特征融合通过下采样来完成,将自上而下融合得到的特征通过池化操作调整分辨率与高层特征进行再融合,最终得到 3 个尺度不同的输出特征 P3、P4、P5,具体结构如图 5 所示。

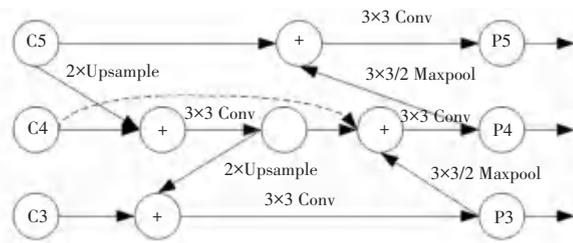


图 5 RB-YOLO 中的 BiFPN 结构

Fig. 5 BiFPN structure in RB-YOLO

2 实验和结果

2.1 实验数据集

本文所采用的数据集共 34 319 张图片,主要有 3 个来源。一是 2020 年深圳举办“华为云杯·生活垃圾图片分类”大赛公布的 VOC 格式的生活垃圾图像数据集(以下简称为“华为云垃圾数据集”),该数据集有 4 个大类和 47 个小类,总计 19 655 张图像。二是由 Mittal G 等人发布的 COCO 格式的 TACO 数据集,共由 1 500 张图片和 4 784 个标注。三是通过网络爬虫、百度飞桨平台搜集到 13 164 张生活垃圾

图片,大部分从百度的垃圾数据集中挑选而来,小部分来自网络爬取并进行手工标注。在数据处理过程中,本文发现华为云垃圾数据集和 TACO 数据集小部分图像存在标注错误、漏标的问题,而百度数据集漏标更为严重。针对此,本文将问题数据通过 Labellmg 重新标注,并且对数据集的类别进行了统一,最终得到了本文数据集 Domestic Garbage Dataset (DGD),将原始数据集按照 4:1 的比例分为训练集和测试集。

2.2 实验环境

本文实验在 Window 系统上进行评估。实验使用了深度学习框架 Pytorch 1.6 和编译器 Python 3.7,硬件环境为 Intel (R) Core (TM) i5 - 10400F CPU、GTX2070s 显卡、8 GB 显存以及 32 GB 内存的电脑。

在网络参数设置方面,训练时输入图像设置为 416×416 ,采用 Adam 作为优化函数对模型进行训练。基于迁移学习的思想采用冻结训练,对预训练网络模型权重参数进行层迁移,冻结预训练网络模型中主干网络的权重参数,初始学习率为 $1e-3$,epoch 为 50, Batch_Size 为 32;解冻后初始学习率设为 $1e-4$,epoch 为 150(在冻结训练模型的基础上再训练 100 epoch), Batch_Size 为 16。为了增强模型的鲁棒性和泛化能力,对训练数据集采用了简单的在线数据增强,包括图像缩放、图像翻转等。

2.3 评价指标

本文实验采用了 mAP (Mean average precision)、FLOPs (Floating-point Operations)、FPS (Frames Per Second)、Params (parameters) 作为模型性能评价指标。其中, mAP 指所有目标平均准确率的均值,综合表示模型性能; FLOPs 是浮点运算数,用来衡量模型的复杂度; FPS 即每秒传输帧数,可以验证模型的实时性; Params 为参数量,代表模型所需要的算力。

2.4 实验结果与分析

针对以上改进,本文设计了 2 组实验进行验证。一是对 RB-YOLO 进行消融实验以验证每个改进点的有效性,二是与其他算法综合对比验证所提出算法的可行性。

2.4.1 消融实验结果

为了验证所提出方法的可行性,在提出的 DGD 数据集上进行了一系列对比实验,见表 1。表 1 中,“√”表示添加该模块,否则为不添加。实验考虑了 RepVGG、BiFPN、ASPP 三个因素对实验结果的影

响,在对比实验中逐个添加改进模块以验证单个变量对实验数据的影响。测试集图片大小为 416×416 ,训练周期为 150 个 epoch。

本文实验将 YOLOv3 作为基础模型,消融实验结果见表 1。消融实验检测的对比仿真结果如图 6 所示。图 6 中,第 1、2、3 列原图片来自数据集测试集,第 4 列原图片来自网络搜索。对比第 1 组和第 2 组实验数据,结果表明相比 YOLOv3, RepVGG-YOLO 模型准确率提升了 0.88%,速度提升了 27 FPS, FLOPs 减少了 19.9 G, RepVGG-A0 作为特征提取网络在保证准确率的同时将检测速度提升了约 1.7 倍。改进的特征融合 BiFPN 也起到了不错的效果,从第 2 和第 3 组实验数据可以看出对比原模型精度提升了 0.85%,且参数量进一步减少了 3.64 M,有利于工业端的部署,但是增加了部分运算量。从图 6(c)可以看出,改进特征融合对于发生形变的易拉罐具有更好的检测效果,在一定程度上能够使得检测结果更加优化。第 2 组和第 4 组实验验证了 ASPP 的可行性,对比 RepVGG-YOLO,精度提升了 1.31%,速度减少了 8 FPS。从图 6(d)可以看出,加入 ASPP 后能够检测出背景复杂的陶瓷器皿和被遮挡的污损用纸,而其它模型(a)、(b)、(c)均有不同程度的漏检。1、5 组实验可以看出,本文提出算法 RB-YOLO 准确率比 YOLOv3 提升了 4.31%,速度提升了 16 FPS,计算量减少了 17.52 G,参数量减少了约 51%。由图 6(e)的检测结果同样可以看出,本文算法的检测结果更加细致和全面,且对于网络搜索的图片具有良好的检测效果,证明了本文所提出算法的可行性和鲁棒性。

表 1 消融实验

Tab. 1 Ablation experiments

序号	RepVGG	BiFPN	ASPP	mAP/ %	FPS	FLOPs/ G	Params/ M
1				69.52	37	32.93	61.78
2	√			70.40	64	12.92	28.24
3	√	√		71.25	57	23.65	24.60
4	√		√	71.71	56	15.41	40.94
5	√	√	√	73.83	53	25.59	36.50

2.4.2 与其他模型对比实验结果

为了验证算法的优越性,将 RB-YOLO 与常见的目标检测算法 YOLOv3、SSD、Faster RCNN 在 DGD 数据集上进行实验对比,所采用的框架均为 Pytorch,默认输入图像大小为 416×416 ,训练周期为 150 个 epoch。表 2 展示了 RB-YOLO 算法与其

他算法的对比,图7展示了RB-YOLO与其他算法的检测效果(从左到右依次为Faster RCNN、SSD、

RB-YOLO检测效果图;第1、2行原图片来自数据集测试集,第3行原图片来自网络搜索)。



图6 消融实验检测结果对比

Fig. 6 Comparison of detection results of ablation experiments

从表2和图7可以看出,和其他几种算法相比,本文提出的RB-YOLO不论在速度、还是准确率上都具有巨大优势,更好地实现了速度和精度的均衡。在检测速度方面,RB-YOLO检测速度最快,达到了53 FPS,比SSD算法高出27 FPS,快了大约1倍;比Faster RCNN算法高出32 FPS,快了大约1.5倍;比YOLOv3高出16 FPS,快了大约0.4倍。在准确率上,RB-YOLO同样优于其他算法,达到73.83%,同时对图7进行综合分析也能够看出,相比Faster RCNN和SSD,RB-YOLO在3组图片中能够较为细致、全面地检测出图片中的目标,并且能够正确分类。Faster RCNN虽然检测出较多的目标,但是没有正确分类,SSD则没有检测出任何目标。3组不同来源的数据证明RB-YOLO不仅在DGD数据集上取得了良好的效果,同时也具有良好的鲁棒性,能够

被广泛应用。在参数量方面,RB-YOLO的参数量明显小于YOLOv3和Faster RCNN,但是比SSD略大6.61 M,因为RB-YOLO中ASPP模块的引入增加了大量参数。综合分析,基于RB-YOLO的垃圾检测算法在快速检测的同时保证了垃圾检测的准确率,能够较好地完成垃圾检测任务。

表2 不同目标检测算法结果比较

Tab. 2 Comparison of results of different target detection algorithms

Model	Input size	mAP/ %	FPS	Params/ M
SSD	512 * 512	64.14	26	29.89
Faster RCNN	416 * 416	66.74	21	137.63
YOLOv3	416 * 416	69.52	37	61.78
RB-YOLO (proposed)	416 * 416	73.83	53	36.50

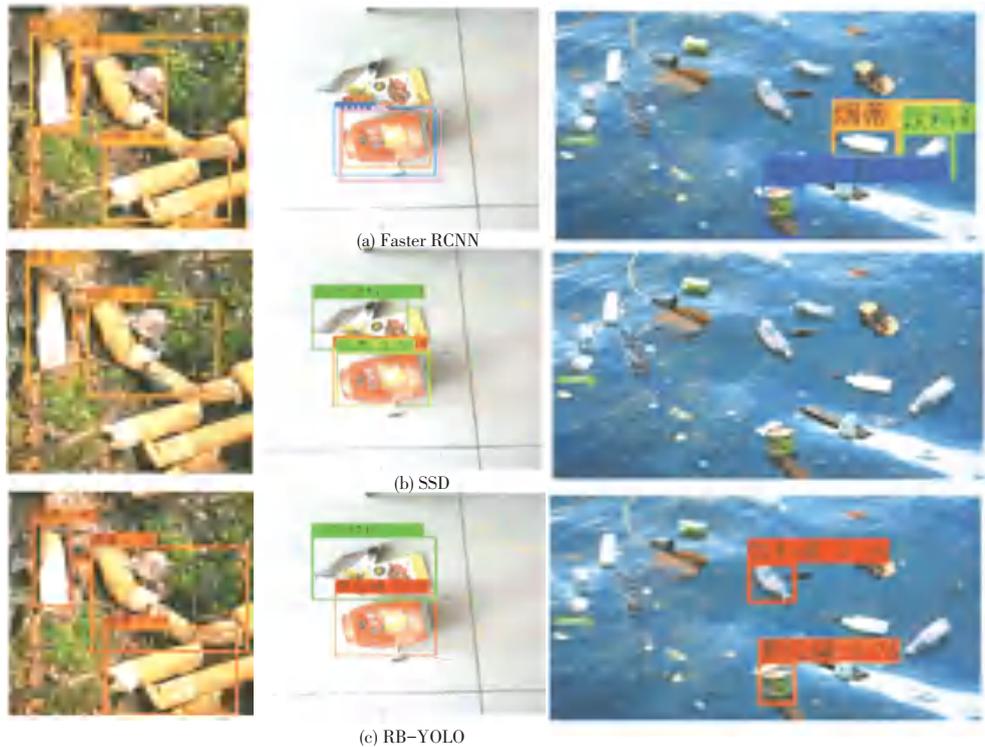


图 7 不同目标检测算法检测结果对比

Fig. 7 Comparison of detection results of different target detection algorithms

3 结束语

针对生活垃圾数据集存在的背景复杂、类内差异大的问题,进一步实现生活垃圾的快速检测,本文提出了一种基于改进 YOLOv3 的轻量级垃圾检测算法。该算法将特征提取网络 Darknet-53 替换为 RepVGG,在保持精度不变的同时大幅减少计算量和提升速度;采用 BiFPN 加强特征融合,进一步减小网络模型的参数量;引入 ASPP 实现局部信息和全局信息的有效融合,更好地对复杂背景下的目标进行检测。在 DGD 数据集上, RB-YOLO 算法的准确率和速度均优于其他算法,但是根据实际检测效果来看,依旧存在一定的漏检情况,因此,下一步的研究方向是在保持检测速度的同时提高模型的准确率。

参考文献

- [1] 中华人民共和国生态环境部. 2020 年全国大、中城市固体废物污染环境防治年报 [EB/OL]. [2020-10-28]. <https://www.mee.gov.cn/hjzl/sthjzk/gtfwzrfz/>.
- [2] SOUSA J, REBELO A, CARDOSO J S. Automation of waste sorting with deep learning [C]//2019 XV Workshop deVisão Computacional (WVC). São Bernardo do Campo, Brazil: IEEE, 2019: 43-48.
- [3] 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 and Machine Intelligence, 2016,

39(6): 1137-1149.

- [4] 文灿华, 李佳, 董雪. 基于 Faster RCNN 的生活垃圾智能识别 [J]. 激光与光电子学进展, 2020, 57(20): 139-145.
- [5] SIMONYAN K, ZISSERMAN A. Very deepconvolutional networks for large-scale image recognition [C]//3rd International Conference on Learning Representations. San Diego: International Conference on Learning Representations, 2015: 1-4.
- [6] 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. Las Vegas: IEEE Computer Society, 2016: 770-778.
- [7] HOWARD A G, ZHU M, CHEN B, et al. MobileNet: efficient convolutional neural networks for mobile vision applications [EB/OL]. [2021-10-28]. <https://arxiv.org/abs/1704.04861>.
- [8] MELINTE D O, TRAVEDIU A M, DUMITRIU D N. DeepConvolutional Neural Networks object detector for real-time waste identification [J]. Applied Sciences, 2020, 10(20): 1-18.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE Computer Society, 2016: 779-788.
- [10] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//Proceedings of 30th IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, United States: Institute of Electrical and Electronics Engineers Inc., 2017: 7263-7271.
- [11] REDMON J, FARHADI A. YOLOV3: An incremental improvement [C]// Proceedings of 2018 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Computer Society, 2018: 1-6.

(下转第 19 页)