

基于 YOLOv4 的杂质玻璃分拣检测研究

杨波¹, 许振明^{1*}, 刘建鑫²

¹上海交通大学中英国际低碳学院, 上海 201306;

²南京电子技术研究所, 江苏 南京 210039

摘要 在废旧玻璃的回收处理流程中, 前端的各种杂质玻璃由人工分拣去除, 自动化程度低。针对此问题, 以上海市某玻璃回收处理企业人工分拣场景为研究背景, 提出了一种基于 YOLOv4 算法的流水线杂质玻璃检测方法。一方面, 实地采集典型样本拍摄图像作为数据集, 采用间接检测两端的方式对细长胶条玻璃类别进行标注; 另一方面, 采用 Kmeans++ 算法重新设定先验框(anchors), 利用 4 倍下采样率的高分辨率特征图进行特征融合以提高对小目标的检测性能。结果表明: 模型检测平均精度均值(mAP)达 97.88%, 其中对胶条玻璃的 AP 为 95.47%; 与其他方法相比, 模型参数量减少 31.11%, 对胶条玻璃的 AP 提高 6.70 个百分点, 检测速度为 42.82 frame/s, 能够满足实时性需求。因此, 所提检测方法适用于自动化分拣杂质玻璃任务的视觉部分。

关键词 目标检测; 背景增强; 垃圾分拣; 杂质玻璃

中图分类号 X705 文献标志码 A

DOI: 10.3788/LOP202259.1815011

Sorting and Detection of Impurity Glass Based on YOLOv4

Yang Bo¹, Xu Zhenming^{1*}, Liu Jianxin²

¹China-UK Low Carbon College, Shanghai Jiao Tong University, Shanghai 201306, China;

²Nanjing Institute of Electronic Technology, Nanjing 210039, Jiangsu, China

Abstract In the recycling process of waste glass, all types of impurities at the front end of the glass are removed manually by sorting. The degree of automation is low. The research background for this study is the manual sorting scene of a glass recycling and processing firm in Shanghai. This paper proposes a method of pipeline impurity glass detection based on the YOLOv4 algorithm. Moreover, typical sample images were collected in the field as the dataset, and the category of thin-tape glass was defined by indirect detection of both ends. Additionally, the Kmeans++ algorithm was used to reset the previous boxes (anchors), and the feature fusion was performed using a high-resolution feature map with a four times down-sampling rate to improve the detection performance for small targets. The results show that the mean average precision (mAP) of the proposed model is 97.88%, and the average precision (AP) for rubber strip glass is 95.47%. Compared to other methods, number of parameters of the proposed model is reduced by 31.11%, the AP for rubber strip glass is increased by 6.70 percentage points, and the detection speed is 42.82 frame/s, which can meet the real-time demand. Therefore, the proposed detection method is suitable for the visual component of the automated task of sorting impurity glass.

Key words object detection; background augmentation; garbage sorting; impurity glass

1 引言

目前, 在废旧玻璃回收处理过程中, 磁选除去磁性金属杂质、风选除去标签和纸片、光电分离不同颜色玻璃颗粒等自动分选设备得到了广泛的应用, 但前端玻璃制品的分类、部分低值和难处理杂质的去除等预处理流程仍是纯手工作业^[1-3]。一方面, 企业面临劳动力成本逐渐提高的压力; 另一方面, 工人面临噪声、粉尘等环境的

健康威胁。考虑到玻璃回收处理行业不断规范带来的技术升级需要, 将以目标检测为引导的自动化系统应用于杂质玻璃分拣过程具有重要意义。近年来, 以深度学习为基础的目标检测技术得到了快速发展和广泛研究^[4-7]。就垃圾识别而言, 刘恩乾^[8]将深度学习目标检测模型应用于生活垃圾分类, 并进行注意力机制、聚类分组归一化等改进。文灿华等^[9]基于 Faster-RCNN 算法对典型生活垃圾进行高精度识别, 对数据进行背景泛

收稿日期: 2021-06-15; 修回日期: 2021-07-26; 录用日期: 2021-08-10

通信作者: *zmxu@sjtu.edu.cn

化,并提出了基于最优检测概率阈值的设置方法。许伟等^[10]提出一种基于YOLOv3,结合MobileNetV3的检测网络,该网络在自制的Garbage数据集上效果显著。岳宝海^[11]利用特征融合的方法对废旧玻璃制品进行智能识别,但研究主要在于特征设计而非深度学习。总的来说,目前垃圾识别方向关注较多的是常见生活垃圾,对废旧玻璃垃圾分类和检测的研究较少。

YOLO作为一种快速紧凑的单阶段开源目标检测模型,相比其他模型,同等尺寸下性能更强、检测速度快、容易部署、稳定性高。基于废旧玻璃绿色、智能高效分选技术及成套设备项目,在回收企业实现传送带杂质玻璃自动化分拣的背景下,本文将速度与精度表现优秀的YOLOv4目标检测算法应用于特定种类玻璃的识别与定位过程。目前已有一些垃圾识别的公开数据,但尚没有针对废旧玻璃分类或检测的数据,因此研究内容之一在于制作该项目的数据集。其二,在检测大长宽比的胶条玻璃类别时,单个检测框一般不能提供合理的位置信息,故采用检测两端的方式。胶条玻璃两端的位置信息对于计算其角度更具实际意义。此外,通过更换背景颜色和纹理、引入干扰垃圾、部分遮挡的方式进行数据扩充,增强训练模型的泛化能力。最后,对几种主流的YOLO目标检测网络进行性能对比,以说明改进的合理性。研究内容旨在为整

个玻璃自动分拣项目提供可靠的视觉基础。

2 数据与方法

2.1 数据采集与标注

首先,在上海市某玻璃回收与再利用公司的协助下,分别在传送带、仓库及工人挑拣好的废料收集区内取样。结合实际人工分拣场景,确定搜集水晶玻璃(crystal)、胶边玻璃(glue)、夹胶玻璃(laminated)和胶条玻璃(bar)4种类别,如图1所示,水晶玻璃大体呈方块、三角块及水晶球状;胶边玻璃带有黑色和白色的胶边;夹胶玻璃包括普通夹胶玻璃和钢化夹胶玻璃,处于破碎状态,区别在于钢化型破碎更密集;胶条玻璃为细长条状玻璃夹胶结构,长宽比较大。采集时挑选形状、大小、干净程度及透明度不一的样品以确保样本的代表性。样本总计400余份,拍摄图片时,利用不同背景板、支架和网络摄像头等搭建简单拍摄装置。所选背景为最常见的黑色、绿色传送带背景,光照条件为室内白天和夜间日光灯。采集单个物体不同角度的图像以增加数据多样性,图片原始尺寸为1920 pixel×1080 pixel(网络摄像头)和3648 pixel×2736 pixel(手机),预裁剪为640×640的JPG格式。此外,在拍摄图片时测试集所用的实物样本与训练集验证集不重复,旨在说明结果的可行性。

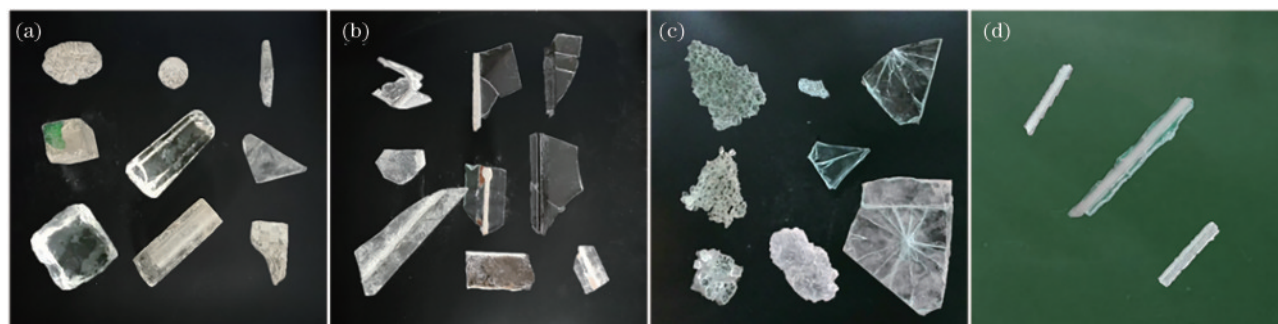


图1 杂质玻璃实物样品。(a)水晶玻璃;(b)胶边玻璃;(c)夹胶玻璃;(d)胶条玻璃

Fig. 1 Impurity glass samples. (a) Crystal; (b) glue; (c) laminated; (d) bar

利用LabelMe标注工具手动标注,格式为Pascal VOC。对于大长宽比的胶条玻璃,考虑三种检测方

式,如图2(a)~(c)所示。图2(a)所示的传统矩形标注方式会导致目标框内绝大部分是背景,且提供的坐标

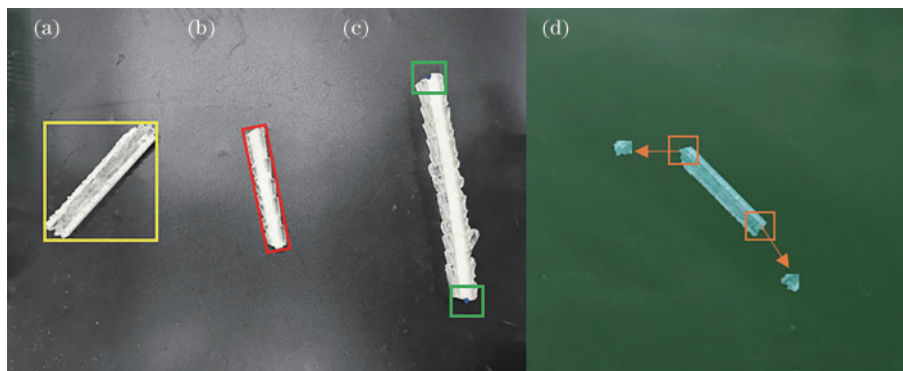


图2 胶条的标注方式与复制。(a)水平矩形;(b)旋转矩形;(c)两端方框;(d)小目标复制

Fig. 2 Three labelling methods for bar and copying. (a) Horizontal box; (b) rotated box; (c) square boxes; (d) small object copying

信息对于分拣执行机构(机械夹爪)来说意义不大。图 2(b)所示的旋转矩形检测方法能够实现更准确的定位^[12],但该方法不适用于其他类别的常规目标,对胶条玻璃类别单独采用该种检测方式会增加模型个数进而增加算法复杂度,不利于实际部署。考虑到实际场景中胶条分布稀疏,相机视野中几乎不会同时出现多个该类目标,故采用标注两端的方式,如图 2(c)所示,利用方框标注两端,边长为胶条宽度,通过检测两端可以间接计算出胶条的中心和角度,为夹取机构提供夹取点和旋转角度等位置信息。鉴于胶条目标尺度较小的特点,将小目标作为掩模(mask)复制 1 倍,以增加胶条正样本个数^[13],如图 2(d)所示。将所标注的 2208 张图片,含 2864 个目标作为初始数据集,各类别基本均衡,训练集、验证集及测试集的图片数量比例为 6:2:2。考虑到原始数据相对较少,训练时对训练集进行实时增强以防止模型过拟合,增强方式为随机缩放 0.5~1、随机水平和垂直翻转、图片色相变换。

2.2 Kmeans++算法

一般地,针对特定应用场景的数据集,重新设定 anchors 能够有效提升模型检测效果。对于数据分布均匀的情况,Kmeans 聚类算法是一种常用的处理方式。相对于传统的 Kmeans 算法,Kmeans++ 算法在初始化聚类中心的方式上做了改进^[14],故本文利用 Kmeans++ 算法聚类设定 anchors。先从数据集中随机选取一个样本作为第一个聚类中心,接着计算每个样本与现有聚类中心的最短距离,表达式为

$$D(x_i) = 1 - \text{IoU}(x_i, x_c), \quad (1)$$

式中: $D(x_i)$ 表示第 i 个样本与当前中心 x_c 的交并比(IoU)距离。接着计算每个样本点被选为下一个聚类中心的概率,表达式为

$$P(x_i) = \frac{D^2(x_i)}{\sum D^2(x)}, \quad (2)$$

最后利用轮盘法选出下一个聚类中心,直至得到 k 个初始聚类中心。用 A_{cc} (accuracy)表示样本与先验框的平均重合程度,表达式为

$$A_{cc} = \frac{\sum \text{IoU}(x_i, x_c)}{N}, \quad (3)$$

式中: N 为总样本数。图 3 表示 A_{cc} 与 k 的关系。根据手肘法则^[15],取 k 值为 6 是合理的。对模型输入尺寸为 416×416 的聚类结果如图 4 所示,anchors 分别为 $(6.0, 6.0)$, $(10.6, 10.6)$, $(17.5, 17.6)$, $(55.4, 44.5)$, $(56.5, 71.8)$, $(92.6, 90.2)$ 。不难发现,尺度较小的胶条玻璃与中等大小的其他三类目标几乎完全隔开。根据数据先验分布,结合 Poly-YOLO 利用单个高分辨率特征层检测特定场合小目标的思路^[16],设置两个检测头,各由 3 个 anchor 负责匹配,即利用一个高分辨率特征层检测小目标,另一个分辨率较低的特征层负责检测其他三类中等大小目标。

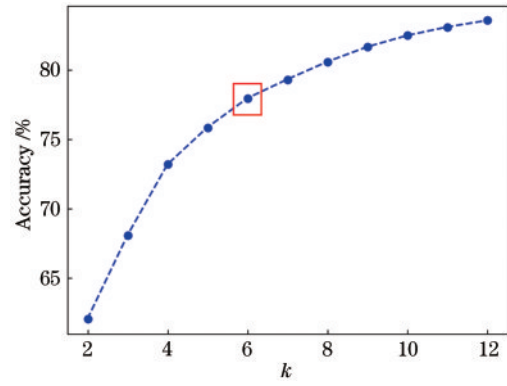


图 3 Kmeans++ 聚类结果与 k 值的关系

Fig. 3 Relationship between clustering effect of Kmeans++ and k

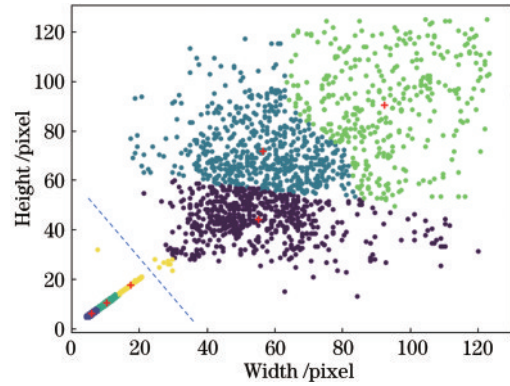


图 4 Kmeans++ 算法的聚类结果

Fig. 4 Clustering result of Kmeans++ algorithm

2.3 YOLOv4 检测算法

YOLOv4 是一个平衡速度与精度的算法^[17]。研究学者在 YOLOv3^[18]的基础上,综合了近年来大量的深度学习技巧,对网络结构、损失函数、后处理、硬件适应性等部分均进行了改进,提出了马赛克(Mosaic)和自对抗训练(SAT)的数据增强方法,提出了修改后的空间注意力机(SAM)、路径聚合网络(PAN)、跨 Batch 的批归一化(BN)等改进措施。如图 5 所示,YOLOv4 利用 P3、P4、P5 三个特征层进行特征融合,预测结果为 y_3 、 y_2 、 y_1 。本文在 YOLOv4 原有 Backbone 的基础上,输出 CSP2 模块后的 4 倍下采样特征图,用更高分辨率的 P2 和 P3 特征层提高小目标检测性能,模型预测输出为 y_4 、 y_3 。

2.4 模型训练与测试

实验平台基于 Ubuntu18.04 系统(Nvidia RTX2080Ti GPU $\times 2$),深度学习框架 PyTorch1.7.0+ CUDA10.2。利用 COCO 数据集上的预训练权重迁移学习加快训练速度,优化方式为 Adam。初始学习率为 0.001, batch size 为 16,学习率在 10^{-3} 、 10^{-4} 、 10^{-5} 、 10^{-6} 之间采用等间隔调整策略(Step-LR)。训练初期,冻结 Backbone 部分参数以加快训练速度,在学习率降低到 10^{-4} 后解冻训练。验证过程中,利用验证集损失(L)、平均精确度(AP)判断收敛程度,测试评

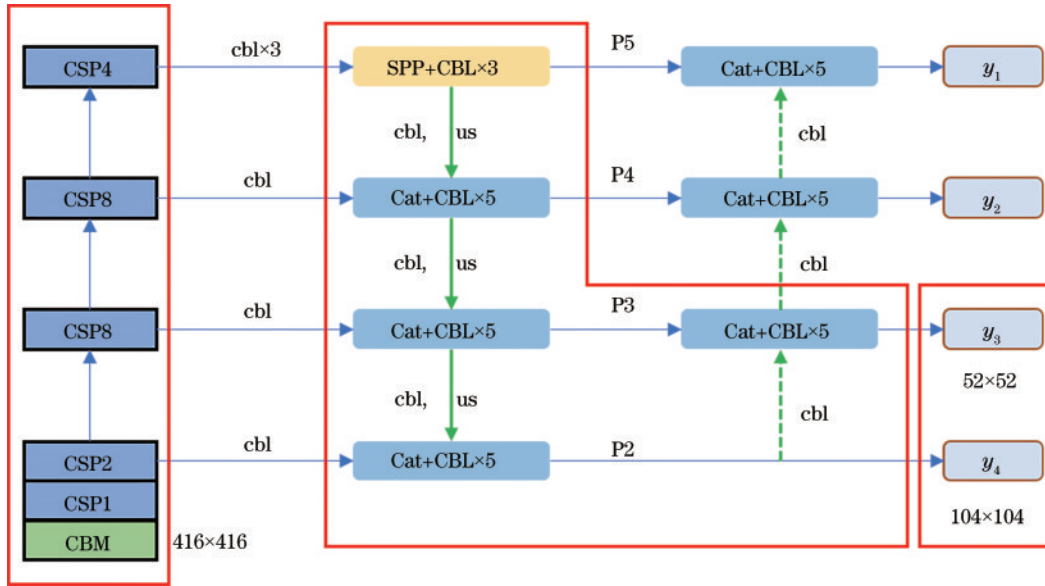


图 5 YOLOv4 网络结构及调整

Fig. 5 YOLOv4 network structure and adjustment

评价指标分别为准确率(P)、召回率(R)、AP。验证集损失分为三个部分^[19], 计算方法为

$$L = L_{\text{conf}} + L_{\text{cls}} + (2 - w_{\text{gt}} \times h_{\text{gt}}) \times L_{\text{IoU}}, \quad (4)$$

$$L_{\text{IoU}} = 1 - \text{IoU}(A, B) + \frac{\rho^2(A_c, B_c)}{c^2} + \alpha \times \nu, \quad (5)$$

式中: L_{conf} 为置信度损失; L_{cls} 为分类损失; L_{IoU} 为坐标回归损失; w_{gt} 和 h_{gt} 为真实框的宽和高; A 表示真实框, B 表示预测框; A_c 和 B_c 代表框的中心; ρ 表示欧氏距离; c 为 A 和 B 最小包围框的对角线长度。 ν 和 α 为长宽比的惩罚因子,表达式分别为

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{w_{\text{gt}}}{h_{\text{gt}}} - \arctan \frac{w}{h} \right)^2, \quad (6)$$

$$\alpha = \frac{\nu}{(1 - \text{IoU}) + \nu}, \quad (7)$$

式中: w 和 h 为预测框的宽和高。准确率和召回率的计算方式为

$$\begin{cases} P = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \\ R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \end{cases}, \quad (8)$$

式中: N_{TP} 表示预测正确的正样本个数; N_{FP} 表示预测错误的正样本,实际为负样本的个数; N_{FN} 表示未被检测出的正样本的个数。AP采用Pascal VOC 2012标准计算^[20]。模型检测速度用帧率(FPS)衡量。

3 分析与讨论

3.1 训练与测试结果

改进后的模型训练过程如图 6 所示,前 40 个 epoch,损失下降明显,当 epoch 达到 60 后,验证集损失与 AP 趋于收敛,选择损失最低的模型作为最终测试

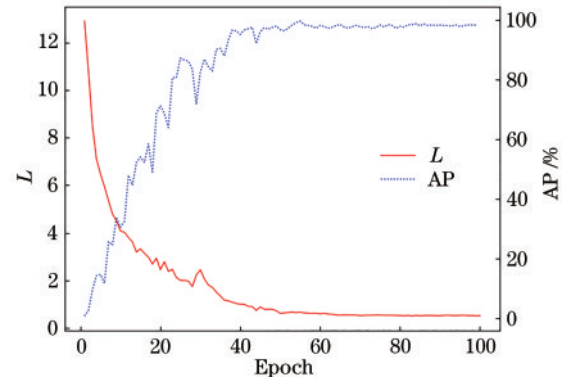


图 6 训练过程中验证集上的损失和 AP

Fig. 6 Loss and AP of validation set during training

模型。

利用测试集评估三种模型,置信度下限为 0.01, 交并比下限为 0.5, 即 AP50, 测试结果如表 1 所示, 其中 YOLOv4 为原始网络, YOLOv4_t 表示改进后的网络, “+”表示结合了 Kmeans++ 算法。结果表明: Kmeans++ 算法对整体检测性能提升明显, 对胶条玻璃类别的 AP 提升了 5 个百分点; 调整后的网络的检测精度相对于原始网络有所提升, 且模型更轻量化。

利用改进后的模型进行预测, 置信度阈值取 0.5,

表 1 不同模型在测试集的精度对比

Table 1 Accuracy comparison of different models in testset

unit: %

Method	AP				mAP
	Bar	Crystal	Glue	Laminated	
YOLOv4	88.77	97.31	98.96	98.04	95.76
YOLOv4+	93.96	98.26	99.97	98.04	97.55
YOLOv4_t+	95.47	98.98	99.07	97.99	97.88

统计每个类别正负样本个数,结果如表 2 所示,其中 GT 表示测试集中每个类别的真实目标个数,TP 为正确识别的预测框个数,FP 为识别错误的预测框个数。由表 2 可知,水晶玻璃、胶边玻璃、夹胶玻璃漏检个数分别为 1、1、3,胶条玻璃漏检个数为 7,模型漏检率低于 3%,基本能够全部召回。模型误检个数相对较多,主要来自于部分类间差异不明显的困难样本。

表 2 正负样本分布

Table 2 Positive and negative sample distribution

Type of boxes	Bar	Crystal	Glue	Laminated
GT	262	117	110	102
TP	255	116	109	99
FP	19	9	22	21

3.2 数据泛化

如前文所述,所提模型在现有数据集上误检率较高;图片背景相对单一,对实际应用场景下的复杂背景、干扰、遮挡等鲁棒性不足。采用数据增强的方式进行优化,提高模型的泛化能力。其一:选取不同颜色、纹理的典型传送带背景 24 种,从前文检测结果中分割预测准确的目标(交并比和置信度均大于 0.9)作为前景。将目标随机贴入到新的背景并自动标注,彼此不重叠。考虑到透明度的影响,对水晶玻璃和夹胶玻璃随机调整透明度,得到新背景数据集,记为 A。其二:选取四十余种实际场景下的垃圾(平板玻璃碎块、碎玻璃瓶、陶瓷、石子、瓶盖等),经缩放、旋转、hsv 色域变换等贴入 A 中,得到新的干扰数据集 B。其三:在干扰垃圾贴入过程中对目标进行随机遮挡,遮挡目标占 20%,遮挡程度 O_{oc} 在 0 到 0.3 之间(O_{oc} 计算方法如图 7 所示,A 为检测目标,B 为干扰遮挡),得到遮挡数据集 C。A、B、C 图片数均为 600,目标个数为 3000,分别进行训练与测试,训练集与测试集的图片数量比例为 4:1,结果如图 8 所示。由图 8 可知,纯色、纹理简单背景对精度影响较小,引入干扰后次之,部分遮挡后影响最大,经重新训练后,所提方法的 AP 均有所回升,能够满足实际分拣需求,部分预测实例如图 9 所示。

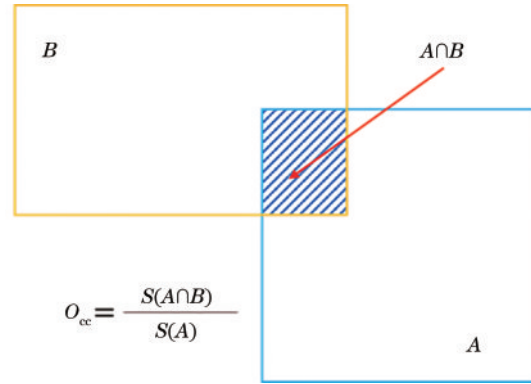


图 7 遮挡比例计算方式

Fig. 7 Calculation method of occlusion proportion

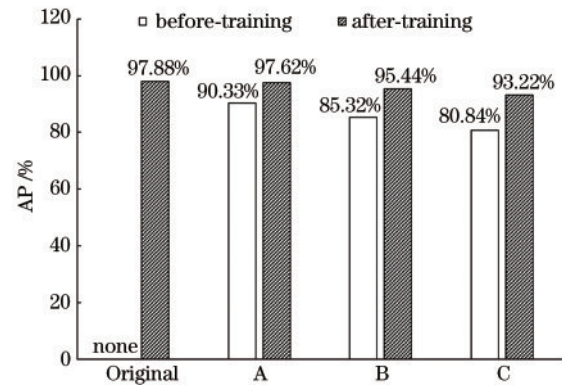


图 8 不同增强类型训练前后的 AP

Fig. 8 AP under different augmenting types before & after training

3.3 不同模型检测对比

在前文研究基础上,对几种常用的 YOLO 模型进行训练与测试,对比模型参数量、计算量、测试集 mAP 和模型推理速度,统计结果如表 3 所示。其中数据为 2208 张图片,模型是基于 PyTorch 构建的,Params(训练参数量)、Multi-Add(乘加运算量)是基于 Torchinfo 模块计算的。表 3 的模型均利用 Kmeans++ 聚类 anchors。实验结果表明:相对于 YOLOv3、YOLOv4 算法,调整后的算法参数量更少、速度相当、精度更高;相对于 tiny 版本的 YOLOv3、YOLOv4 精度,调整后的算法更能满足实际需求,速度上显得更慢,但检测帧率为 42.82 frame/s,能够达到实时性要求。

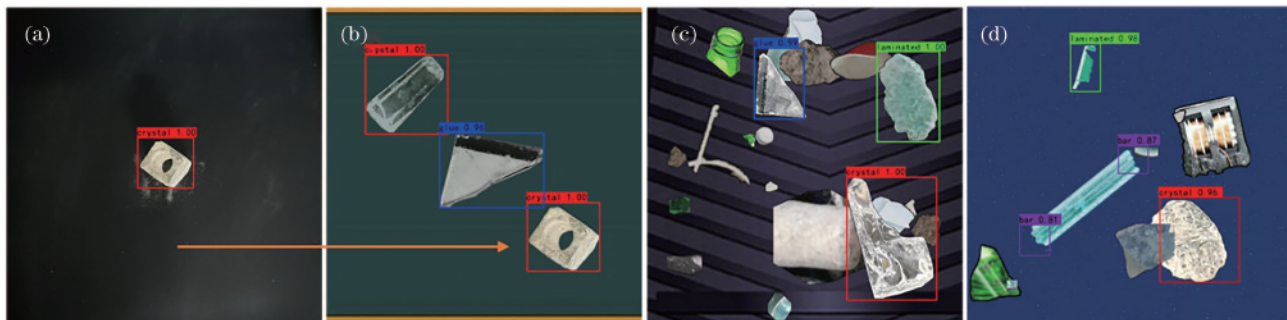


图 9 目标贴入与预测实例。(a)(b)目标贴入标注;(c)干扰;(d)遮挡

Fig. 9 Target pasting and detecting examples. (a)(b) Target attached with label; (c) interference; (d) occlusion

表 3 典型网络参数大小及检测性能

Table 3 Parameter size and detection performance of typical networks

Parameter	YOLOv4	YOLOv3	YOLOv4-tiny	YOLOv3-tiny	YOLOv4_t
mAP	0.96	0.92	0.88	0.86	0.98
Speed/(frame·s ⁻¹)	40.54	38.51	118.67	158.25	42.82
Multi-Add/10 ⁹	29.78	32.66	3.39		28.96
Params/10 ⁶	255.82	246.16	23.52	17.90	176.22

4 结 论

在自制废旧杂质玻璃数据集的基础上,利用 Kmeans++ 算法进行先验框聚类并对现有的 YOLOv4 网络进行适应性改进,结果表明所提算法的检测平均精度为 97.88%,检测速度为 42.82 frame/s。所提算法对小目标胶条玻璃的 AP 为 95.47%,相对原始网络提高了 6.70 个百分点。在对几种典型的检测算法进行对比测试后发现,所提算法具有精度更高、对小目标更友好的优势。此外,在原始图片的基础上进行不同程度的数据泛化,复杂背景、干扰及部分遮挡的情况下,所提算法的 mAP 在 90% 以上,说明速度与精度均能满足实际需求。后期的工作在于实地采集更多复杂场景的图片以完善数据集以及轻量化模型。

参 考 文 献

- [1] 严陈玲. 德国柏林市生活垃圾分类经验及启示[J]. 中国环保产业, 2020(4): 35-39.
Yan C L. Experience and enlightenment of domestic waste classification in Berlin, Germany[J]. China Environmental Protection Industry, 2020(4): 35-39.
- [2] 刘志海. 我国废玻璃回收利用综述[J]. 玻璃, 2018, 45(10): 1-8.
Liu Z H. Review of waste glass recycling in China[J]. Glass, 2018, 45(10): 1-8.
- [3] 陈桂宏, 彭竞仪, 廖雨森. 有色玻璃回收新技术[J]. 节能, 2020, 39(7): 104-106.
Chen G H, Peng J Y, Liao Y S. New technology for recycling colored glass[J]. Energy Conservation, 2020, 39(7): 104-106.
- [4] 王康如, 谭锦钢, 杜量, 等. 基于迭代式自主学习的三维目标检测[J]. 光学学报, 2020, 40(9): 0915005.
Wang K R, Tan J G, Du L, et al. 3D object detection based on iterative self-training[J]. Acta Optica Sinica, 2020, 40(9): 0915005.
- [5] 戴媛, 易本顺, 肖进胜, 等. 基于改进旋转区域生成网络的遥感图像目标检测[J]. 光学学报, 2020, 40(1): 0111020.
Dai Y, Yi B S, Xiao J S, et al. Object detection of remote sensing image based on improved rotation region proposal network[J]. Acta Optica Sinica, 2020, 40(1): 0111020.
- [6] 来文豪, 周孟然, 胡锋, 等. 基于多光谱成像和改进 YOLO v4 的煤矸石检测[J]. 光学学报, 2020, 40(24): 2411001.

- Lai W H, Zhou M R, Hu F, et al. Coal gangue detection based on multi-spectral imaging and improved YOLO v4 [J]. Acta Optica Sinica, 2020, 40(24): 2411001.
- [7] 王一同, 周宏强, 闫景道, 等. 基于深度学习算法的计算光学研究进展[J]. 中国激光, 2021, 48(19): 1918004.
Wang Y T, Zhou H Q, Yan J X, et al. Advances in computational optics based on deep learning[J]. Chinese Journal of Lasers, 2021, 48(19): 1918004.
- [8] 刘恩乾. 基于深度学习的生活垃圾分类和检测[D]. 太原: 山西大学, 2020.
Liu E Q. Classification and detection of household garbage based on deep learning[D]. Taiyuan: Shanxi University, 2020.
- [9] 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): 201014.
Wen C H, Li J, Dong X. Intelligent domestic garbage recognition based on faster RCNN[J]. Laser & Optoelectronics Progress, 2020, 57(20): 201014.
- [10] 许伟, 熊卫华, 姚杰, 等. 基于改进 YOLOv3 算法在垃圾检测上的应用[J]. 光电子·激光, 2020, 31(9): 928-938.
Xu W, Xiong W H, Yao J, et al. Application of garbage detection based on improved YOLOv3 algorithm[J]. Journal of Optoelectronics·Laser, 2020, 31(9): 928-938.
- [11] 岳宝海. 基于特征融合的玻璃制品分拣方法研究[D]. 吉林: 东北电力大学, 2020.
Yue B H. Research on glass products sorting method based on feature fusion[D]. Jilin: Northeast Dianli University, 2020.
- [12] Yang X, Yan J C. Arbitrary-oriented object detection with circular smooth label[M]//Vedaldi A, Bischof H, Brox T, et al. Computer vision-ECCV 2020. Lecture notes in computer science. Cham: Springer, 2020, 12353: 677-694.
- [13] Kisantal M, Wojna Z, Murawski J, et al. Augmentation for small object detection[EB/OL]. (2019-02-19)[2021-06-05]. <https://arxiv.org/abs/1902.07296>.
- [14] Arthur D, Vassilvitskii S. k-means++: the advantages of careful seeding[C]//Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 7-9, 2007, New Orleans, Louisiana, USA. Philadelphia: SIAM, 2007: 1027-1035.
- [15] Elbow Method for optimal value of k in Kmeans[EB/OL]. [2021-05-08]. <https://www.geeksforgeeks.org/elbow-method-for-optimal-value-of-k-in-kmeans>.
- [16] Hurtik P, Molek V, Hula J, et al. Poly-YOLO: higher speed, more precise detection and instance segmentation

- for YOLOv3[J]. Neural Computing and Applications, 2022: 1-16.
- [17] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2021-05-04]. <https://arxiv.org/abs/2004.10934>.
- [18] Redmon J, Farhadi A. Yolov3: an incremental improvement[EB/OL]. (2018-04-08)[2021-05-04]. <https://arxiv.org/abs/1804.02767>.
- [19] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [20] Everingham M, Winn J. The pascal visual object classes challenge 2012 (voc2012) development kit[EB/OL]. [2021-06-14]. http://host.robots.ox.ac.uk/pascal/VOC/voc2012/devkit_doc.pdf.