

基于 YOLOv5s 的轻量化可回收饮料瓶颜色识别^{*}

王 振 方海峰 曹 晋 吴群彪
(江苏科技大学机械工程学院 镇江 212003)

摘 要:针对不同颜色的可回收饮料瓶回收价值不同,进而需要解决颜色识别分选问题,提出一种基于 YOLOv5s 的轻量化模型并配合 DELTA 并联机器人分选设备进行智能识别分选。模型减少了原 Backbone 中 C3 数量,并使用 1×1 的卷积核代替了 C3 和 Conv 模块部分 3×3 的卷积核,采用 GhostConv 替代传统 Conv,CIOU 损失增加了检测框尺度、长和宽的损失,提高了矩形框回归效果,选择 CIOW Loss 作为 bounding box 的损失函数,通过对其他传统模型对比实验,验证了模型的有效性。结果表明,参数量和计算量相较原模型分别减少了 33.80% 和 36.84%,对回收饮料瓶颜色的识别时间达到了 0.008 s,识别图片速度 125 张/s,识别精度达到了 97%。较传统模型,改进 YOLOv5s 模型识别准确率更高,识别速度更快。

关键词:YOLOv5s;目标检测;可回收饮料瓶识别

中图分类号:TP391.41 **文献标识码:**A **国家标准学科分类代码:**520.6

Lightweight recyclable bottle color recognition based on YOLOv5s

Wang Zhen Fang Haifeng Cao Jin Wu Qunbiao

(School of Mechanical Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: For different colors of recyclable beverage bottles with different recycling values and the need to solve the color identification sorting problem, a lightweight model based on YOLOv5s and with DELTA parallel robot sorting equipment is proposed for intelligent identification sorting. This model reduces the number of C3 in the original Backbone and uses 1×1 convolution kernel instead of C3 and 3×3 convolution kernel in the Conv module part. GhostConv is used instead of traditional Conv. CIOW Loss increases the Loss of detection box scale, length and width to improve the rectangular box regression effect. CIOW Loss is chosen as the bounding box loss function, and the effectiveness of this model is verified by comparing experiments with other traditional models. The results show that the number of parameters and computation volume are reduced by 33.80% and 36.84%, respectively, compared with the original model, and the recognition time for the color of recycled beverage bottles reaches 0.008 s, which can recognize 125 images per second, and the recognition accuracy reaches 97%. Compared with the traditional model, the improved YOLOv5s model has higher recognition accuracy and faster recognition speed.

Keywords: YOLOv5s; object detection; recyclable bottle recognition

0 引 言

伴随着风电和航空航天领域新技术的快速发展,结构泡沫具有耐高温、成本低和可回收利用等优势逐渐成为主导产品。PET 固废进行深度分拣清洗回收进而生产出用于风电、航空等领域所需求的结构泡沫。PET 固废的主

要来源是消费者丢弃的饮料瓶、打包餐盒等塑料制品^[1]。结构泡沫所需的是纯色塑料颗粒,因此需要将回收的饮料瓶进行颜色分选,部分企业利用气流方式与串联机械臂来实现饮料瓶的分拣,但分类数量较低,一般仅能实现二分类。因此亟需研究可回收饮料瓶的分类方法,以提高分类的效率、增加能源的循环效益。

收稿日期:2022-12-17

^{*} 基金项目:国家自然科学基金青年科学基金(11502098)、江苏省科技成果转化专项资金(BA2017090)、江苏省青蓝工程人才项目资助

现有不少国内外学者对图像分类,视觉系统和定位系统进行研究。吕禾丰等^[2] EIOU 损失函数代替 YOLOv5 算法的 GIOU 损失函数来优化训练模型,提高算法的精度,实现对目标更快速的识别。彭继慎等^[3]以 YOLOv4 为基础在主干网络中加入注意力机制,使用深度可分离卷积代替传统卷积来减少卷积计算量,最终使模型拥有精度高、体积小和检测速度快的优势。曾维等^[4]将不同尺度特征图经过 K-means 聚类算法做锚框处理,利用位置预测实现最终的识别与定位检测结果。李仁鹰等^[5]利用各种轻量级网络的网络结构,对 YOLOv4 主干部分进行改进,使其对图像的局部特征进一步加强,实现模型的压缩和加速。Wang 等^[6]采用凸包原理对相邻的可回收瓶进行分离,使用 ReliefF 算法进行特征提取,最终采用支持向量机(SVM)算法对可回收瓶进行颜色识别。范亭亭^[7]提出了一种基于形态学梯度重构的分水岭图像分割算法,根据可回收饮料瓶的颜色进行分类,降低了重叠饮料瓶对分类的干扰。Fang 等^[8]提出了一种基于支持向量机的线性多层深度结构对可回收饮料瓶进行识别,采用自定义损耗函数控制线性多层支持向量机的深度,并加入了泛化参数,提高了算法的泛化能力。王莉等^[9]为了实现垃圾自动按类处理,采用 YOLOv5s 网络作为垃圾检测与分类模型,在自制垃圾分类数据集上对网络进行训练。然而,这些方法主要针对不同的研究对象,因此对可回收饮料瓶颜色识别不能平衡识别时间和识别精度。

现有算法能够实现可回收饮料瓶的颜色识别,但工业现场在复杂工况下提出更高的生产需求。因此本文提出一种基于 YOLOv5s 轻量化模型对采集到的可回收饮料瓶图像进行识别,通过分选设备进行对应位置抓取,最终实现预期的目标,并通过实验验证了其有效性。

1 可回收饮料瓶识别系统理论

1.1 分选设备搭建

大部分大型产业流程中,采用风选的形式分选轻质固体废物废弃物工作效率高,但针对形状与表面并不规则的可回收饮料瓶,则分选精度迅速下降。DELTA 并联机器人技术逐渐成熟,其运动速度快、抓取更准确,更利于可回收饮料瓶的颜色分类工作,因此本文搭建了 DELTA 机器人分选设备如图 1 所示。

可回收饮料瓶颜色分选设备主要由算法模块、识别模块与分选模块 3 个部分组成,识别模块主要包括图像采集模块与工控机,分选模块则包括传送带、两台 DELTA 并联机器人、PLC 控制柜、吸盘与负压罐。

算法模块实现读取相机图像,对其识别并将对应抓取位置发送至 PLC 控制柜。机器视觉算法识别的精度与速度对可回收饮料瓶的颜色分选影响最大,因此作为本文的主要研究内容。

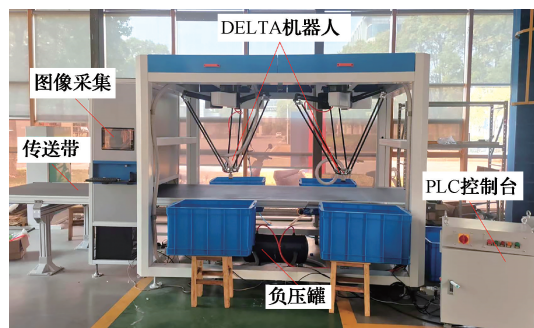


图 1 可回收饮料瓶颜色分选设备

1.2 YOLOv5 理论基础

目前,在 One-Stage 算法中,YOLOv5 是 YOLO 系列最新的目标检测算法,延续了 YOLOv4 的优势,并对其 Backbone、Neck 等方面进行优化^[10]。

在 YOLOv5 算法中共有 4 个网络模型,分别是 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x,它们网络结构基本相似,主要区别为网络的宽度与深度^[11]。

YOLOv5s 网络模型主要由 Backbone 与 Head 组成,其模型网络结构如图 2 所示。

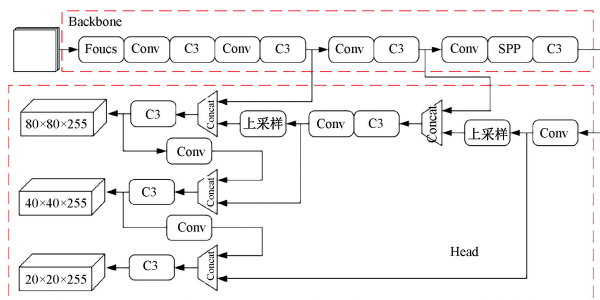


图 2 YOLOv5s 网络结构示意图

主干网络 Backbone 包括了 Focus、Conv、C3、SPP 等结构^[12]。

Focus 模块对图片进行切片操作,将图片数据切分为 4 份,获取没有信息丢失的两倍下采样特征图,再进行 Concat 拼接操作,最终通过卷积输出^[13]。Focus 模块减少了计算量并提高了网络速度,其结构如图 3 所示。

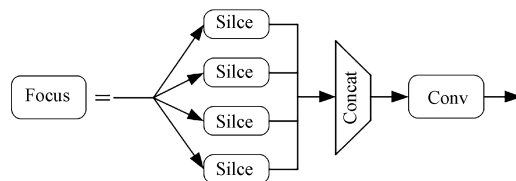


图 3 Focus 模块结构

C3 模块删减了每个瓶颈结构中的一个 Conv,先将基础层的特征映射划分为两部分,再通过跨阶段层次结构将它们合并,从而减少了信息重复的概率,并提高了

YOLOv5s 网络的学习能力^[14],对比如图 4 所示。

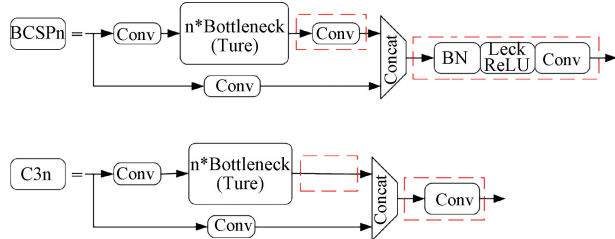


图 4 C3 模块和 BottleCSP 模块对比

SPP 模块只使用一次特征提取避免了反复运算。SPP 模块进行 $5 \times 5, 9 \times 9, 13 \times 13$ 最大池化和一个跳跃链接操作,可以接受不同尺寸的图像,区别于其他网络将图像固定为某一尺寸,使得图像失真、重复特征提取^[15],最终通过 Concat 卷积输出,其结构如图 5 所示。

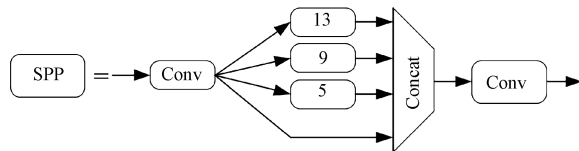


图 5 SPP 网络结构示意图

Head 部分是由 FPN 结合 PAN 组成的,首先 FPN 将高层信息通过上采样的方式进行特征的传递融合,得到预测的特征图^[16]。再通过 PAN 从底部通过下采样的方式向上传递特征,提高了网络的特征融合能力^[17]。

2 改进可回收饮料瓶自动颜色识别模型

2.1 Backbone 优化

为了能够更好的实现实时检测,提高网络的运算速度。在 Backbone 中,本文将各阶段的 C3 模块数量删减至一个。在原网络模型中使用大卷积核能够提高感受野,但会导致计算量增大,不利于模型深度的增加^[18]。本文利用 1×1 的卷积核代替 C3 和 Conv 模块部分 3×3 的卷积核,来缩减网络的训练与识别时间。为了进一步减少计算量,引入 Ghost 模块。

Ghost 模块是 GhostNet 中提出的一个创新模块,可以用较少的参数和计算获得更多的特征图,原理如图 6 所示。

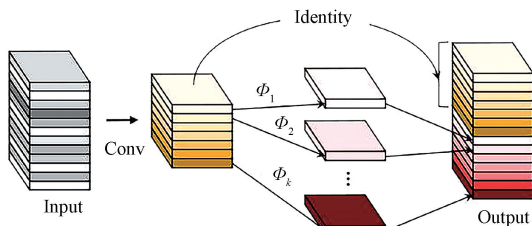


图 6 Ghost 模块原理

Ghost 卷积通过将传统卷积运算与简易线性运算相结合,充分利用了特征提取与特征映射冗余特性之间的关系,有效的减少了模型的参数和浮点数计算的数量,同时保证了检测性能^[19]。改进后的轻量化 YOLOv5s 模型整体结构如表 1 所示。

表 1 轻量化 YOLOv5s 模型整体结构

序号	应用层数	参数	模块	数组对象
0	-1	3 520	Focus	[3,32,3]
1	-1	2 176	GhostConv	[32,64,3,2]
2	-1	18 816	C3	[64,64,1]
3	-1	8 448	GhostConv	[64,128,3,2]
4	-1	74 496	C3	[128,128,1]
5	-1	33 280	GhostConv	[128,256,3,2]
6	-1	296 448	C3	[256,256,1]
7	-1	132 096	GhostConv	[256,512,3,2]
8	-1	656 896	SPP	[512,512,[5,9,13]]
9	-1	1 182 720	C3	[512,512,1,False]
10	-1	131 584	Conv	[512,256,1,1]
11	-1	0	Upsample	[None,2]
12	[-1,6]	0	Concat	[1]
13	-1	361 984	C3	[512,256,1,False]
14	-1	33 024	Conv	[256,128,1,1]
15	-1	0	Upsample	[None,2]
16	[-1,4]	0	Concat	[1]
17	-1	90 880	C3	[256,128,1,False]
18	-1	16 640	Conv	[128,128,1,2]
19	[-1,14]	0	Concat	[1]
20	-1	296 448	C3	[256,256,1,False]
21	-1	66 048	Conv	[256,256,1,2]
22	[-1,10]	0	Concat	[1]
23	-1	1 182 720	C3	[512,512,1,False]

经过计算,轻量化 YOLOv5s 模型层数为 247 层,参数为 4 817 469 个,计算量为 10.8 GFLOPS。原有的 YOLOv5s 模型层数为 283 层,参数为 7 276 605 个,计算量为 17.1 GFLOPS。经比较,轻量化 YOLOv5s 模型参数量减少了 33.80%,计算量减少了 36.84%。

2.2 Loss 优化

YOLOv5 损失函数包括分类损失(classification loss)、定位损失(localization loss)、置信度损失(confidence loss),总损失函数是 3 种损失相加而得到的^[20]。在计算分类损失进行训练时,对每个标签使用二元交叉熵损失,避免了使用 Softmax 函数从而降低了计算复杂度^[21]。

考虑到预测框与真实框的重叠面积、中心点距离、长宽比等因素,本文改进了 YOLOv5s 原有的 GIOU 损失,选择更能体现矩形框回归的 CIOU 损失作为 bounding box 回归的损失函数来提高收敛速度和稳定性,其公式如下:

$$CIOU\ Loss = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (1)$$

式中: $\rho^2(b, b^{gt})$ 分别代表预测框和真实框的中心点的欧氏距离; c 表示预测框和真实框的最小闭包区域的对角线距离; α 表示权重函数; v 表示锚框和目标框之间的宽高比的混合度。

3 YOLOv5s 轻量化可回收饮料瓶识别

3.1 图像数据集的制作

图像集的拍摄效果决定了后期分类误差的大小,为了降低误差,需要模拟工况下的图像拍摄效果。利用可回收饮料瓶颜色分选设备进行数据集的拍摄,每张图片的分辨率大小为 $1\,600 \times 1\,200$ 。数据集包括 5 个种类,分别为淡蓝色、深蓝色、淡绿色、深绿色、透明色。用 make sense 工具对图片中的可回收饮料瓶进行颜色标注,标注例图如图 7 所示,最后输出文件以 txt 作为后缀,作为 labels 文件。

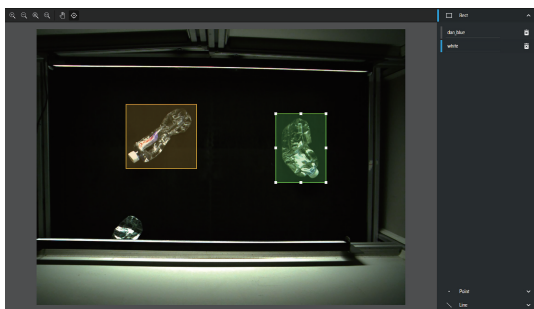


图 7 可回收饮料瓶标注

将所得数据集按照 9:1 的比例划分为训练集和测试集,最终训练集包括 1 350 张图片,测试集包括 150 张图片。

3.2 实验环境及训练

本文实验模型的搭建、训练和测试均在深度学习 PyTorch 框架下完成,采用 Windows10 操作系统,实验具体环境如表 2 所示。

表 2 实验环境配置

参数	配置
操作系统	Windows10
CPU	Intel(R) Core(TM) i5-10400F CPU @ 2.90 GHz
GPU	GeForce GTX 1650 SUPER
Python 版本	3.7
加速环境	CUDA10.2

在轻量化 YOLOv5s 模型训练之前,需要初始化网络参数。YOLOv5s 训练参数设置分别为训练图片大小为 640×640 ,初始学习率为 0.01,权重衰减系数为 0.000 5, Batch 为 12,最大迭代次数为 500。

在大约迭代到 400 次之后,损失函数逐渐保持稳定,稳定数值约为 0.021 9,模型达到最优化,如图 8 所示。与

原网络模型相比,改进后的模型损失收敛的更快,并具有更高的稳定性。

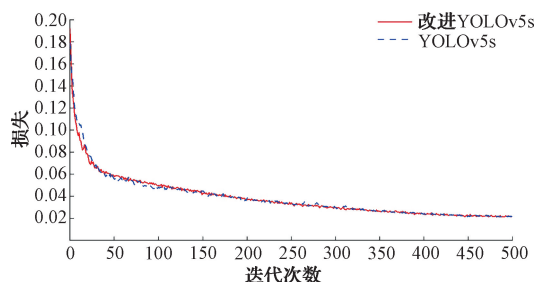


图 8 损失函数变化曲线

3.3 性能指标评价

本文采用 4 种性能指标对实验模型进行评估,分别是精度(Precision)、召回率(Recall)、mAP@0.5 (IoU 设为 0.5 时所有类别的平均 AP)、mAP@0.5:0.95 (在不同 IoU 阈值上的平均 mAP, IoU 取值从 0.5 到 0.95,步长为 0.05)。

Precision 表示在可回收饮料瓶颜色数据集中,识别出有各个可回收饮料瓶的颜色数量,反应识别结果的正确性比率,如图 9 所示。

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

式中: TP 表示预测为可回收饮料瓶,识别是可回收饮料瓶的数量; FP 表示为预测不是可回收饮料瓶,识别不是可回收饮料瓶的数量。

图 9 中原 YOLOv5s 模型波动的频率较大,本文模型在稳定的条件下,先达到了峰值,并与原模型精度相当。

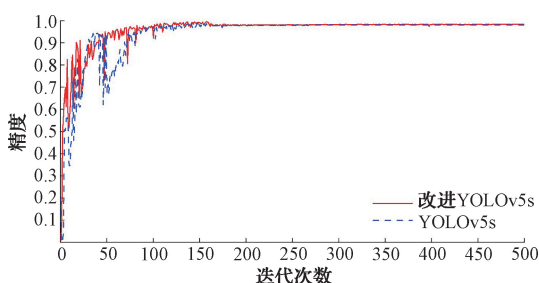


图 9 精准度曲线

Recall 表示在全部可回收饮料瓶颜色数据集中,正确识别出可回收饮料瓶的数量比值,如图 10 所示。

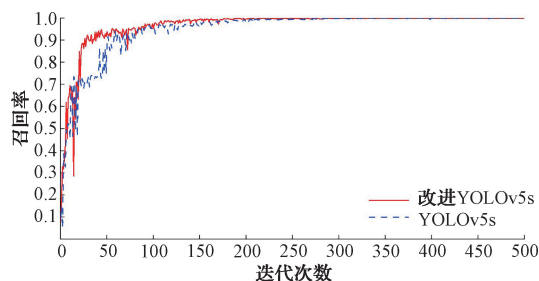


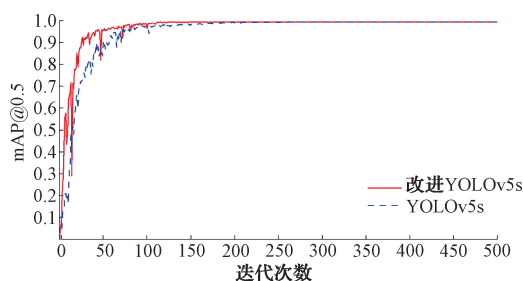
图 10 召回率曲线

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

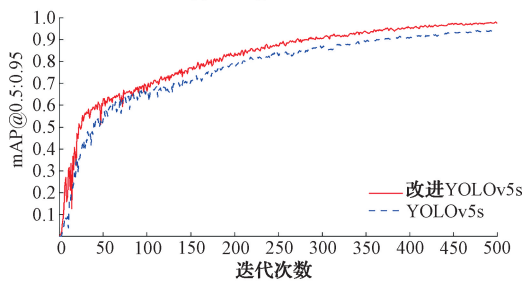
式中: FN 表示预测为可回收饮料瓶, 识别不是可回收饮料瓶的数量。

图 10 中本文模型在前期明显优于原模型, 振荡幅度相对较低, 后期达到的稳定值也略高于原模型。

平均精度均值曲线如图 11 所示, 与原模型相比, 本文模型 mAP 值在任何阶段始终高于原模型, 且达到的稳定值更快。



(a) $mAP@0.5$ 变化曲线



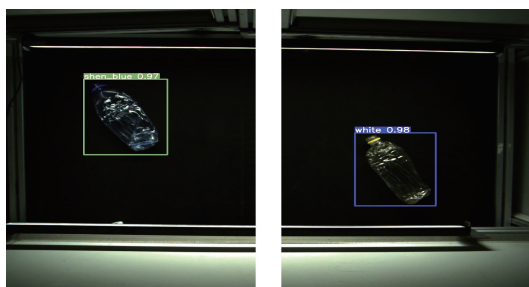
(b) $mAP@0.5:0.95$ 变化曲线

图 11 平均精度均值曲线

从各个性能指标评估结果可以得出, 轻量化 YOLOv5s 网络模型对于可回收饮料瓶颜色识别在训练阶段表现出比较理想的结果。

4 实验结果分析与比较

将网络模型训练完成之后, 选取最优权重, 把测试数据集导入最优模型中识别。在单一工况下识别, 如图 12 所示, 深蓝色饮料瓶的置信度达到了 97%, 透明色饮料瓶置信度达到了 98%。

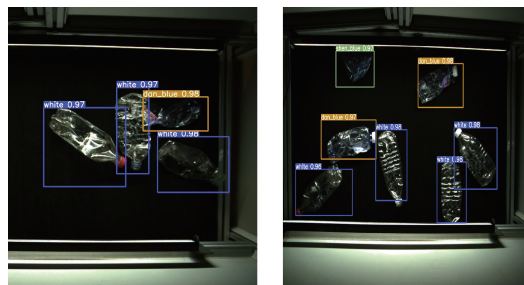


(a) 单一深蓝色

(b) 单一透明色

图 12 单一工况识别

在实际情况下可回收饮料瓶数量众多并存在堆叠现象, 本文模拟复杂工况下对可回收饮料瓶进行识别, 如图 13 所示。在可回收饮料瓶堆叠和数量众多的情况下, 也能够很好的识别出目标, 置信度达到了 96%~98%。



(a) 堆叠情况

(b) 大量情况

图 13 复杂工况识别

为了验证先进性, 与 k 最邻近(KNN)、SVM、卷积神经网络(CNN)和视觉几何组(VGG)模型比较的结果如表 3 所示。

表 3 各算法分类效果对比

实验模型	训练精度/%	预测时长/s	预测精度/%
KNN	93.4	0.109	92.3
SVM	96.4	0.016	96.5
CNN	95.8	0.230	91.8
VGG	97.7	0.053	96.8
YOLOv5s	98.23	0.032	97
本文	98.46	0.008	97

根据实验结果的对比发现, 由于 KNN 算法受特异点影响较大, 导致不同的训练集训练出的结果相差较大, 所以分类精度较低。SVM 算法需要对图像进行前期预处理工作。CNN 算法可以直接对图像集进行处理, 但是预测时长最久, Softmax 分类器的分类效率较低。VGG 算法在复杂工况下对可回收饮料瓶的颜色识别效果略低于 YOLOv5s 模型。轻量化 YOLOv5s 网络模型对所有颜色的可回收饮料瓶相比于原模型训练精度与预测时长分别提高了 0.23% 与 0.024 s, 识别图片速度 125 张/s, 达到了 0.008 s。

网络的结构变化会对网络的性能进行影响, 通过消融实验进一步证明网络的有效性, 如表 4 所示。

表 4 消融实验性能对比

网络模型	精度/%	$mAP@0.5$ /%
YOLOv5s-SE	95.67	94.56
YOLOv5s-CBAM	96.84	96.28
YOLOv5s-MobileNetV2	94.93	95.31
本文	98.46	98.35

消融实验结果表明, 本文的网络模型比使用注意力机

制优化 Backbone 骨干结构的网络模型精度和 mAP 值较高。相较于轻量化模型 YOLOv5s-MobileNetV2, 该模型精度和 mAP 值分别提升了 3.53% 和 3.04%。在分选设备实际操作中对所有颜色饮料瓶识别率达到了 98%。

5 结 论

本文轻量化 YOLOv5s 模型应用于可回收饮料瓶的颜色识别分类, 更利于 PET 固废的纯色颗粒回收, 简化后工序, 生产结构泡沫, 并且实现了颜色识别的高效性与快速性, 结果表明, 减少了原网络模型 C3 的数量, 并将 1×1 的卷积核代替 C3 和 Conv 模块部分 3×3 的卷积核, 引入 Ghost 模块, 使得参数量减少了 33.80%、计算时间减少了 36.84%, 加快了迭代速度。采用效果更有的 CIOU 损失替换了原模型的 GIOU 损失作为 bounding box 回归的损失函数, 使得轻量化 YOLOv5s 模型损失函数收敛的更快, 并具有更高的稳定性, 稳定值约为 0.0219。将轻量化 YOLOv5s 模型和 KNN、SVM、CNN、VGG 模型进行分类效果对比。实验结果表明, 与原模型相比, 对可回收饮料瓶颜色训练精度提高了 0.23%, 识别时间提高了 0.024 s。轻量化 YOLOv5s 模型识别准确率更高, 识别速度更快。

参 考 文 献

- [1] SCHMIDT S, LANER D, VAN EYGEN E, et al. Material efficiency to measure the environmental performance of waste management systems: A case study on PET bottle recycling in Austria, Germany and Serbia[J]. Waste Management, 2020, 110: 74-86.
- [2] 吕禾丰, 陆华才. 基于 YOLOv5 算法的交通标志识别技术研究[J]. 电子测量与仪器报, 2021, 35(10): 137-144.
- [3] 彭继慎, 孙礼鑫, 王凯, 等. 基于模型压缩的 ED-YOLO 电力巡检无人机避障目标检测算法[J]. 仪器仪表学报, 2021, 42(10): 161-170.
- [4] 曾维, 尹生阳, 张凤. 基于计算机视觉的垃圾塑料瓶识别与定位算法研究[J]. 电子测量技术, 2021, 44(23): 12-17.
- [5] 李仁鹰, 钱慧芳, 郭佳豪, 等. 基于 M-YOLOv4 模型的轻量化目标检测算法[J]. 国外电子测量技术, 2022, 41(4): 15-21.
- [6] WANG Z K, PENG B B, HUANG Y J, et al. Classification for plastic bottles recycling based on image recognition [J]. Waste Management, 2019, 88: 170-181.
- [7] 范亭亭. 基于颜色特征的回收塑料瓶分类算法研究[D]. 邯郸: 河北工程大学, 2018.
- [8] FANG H F, CAO J, CAI L H, et al. The recognition of plastic bottle using linear multi hierarchical SVM classifier[J]. Journal of Intelligent and Fuzzy Systems, 2021(10): 1-14.
- [9] 王莉, 何牧天, 徐硕, 等. 基于 YOLOv5 网络的垃圾分类和检测[J]. 包装工程, 2021, 42(8): 50-56.
- [10] 张宏群, 班勇苗, 郭玲玲, 等. 基于 YOLOv5 的遥感图像舰船的检测方法[J]. 电子测量技术, 2021, 44(8): 87-92.
- [11] WANG H, ZHANG S, ZHAO S H, et al. Real-time detection and tracking of fish abnormal behavior based on improved YOLOv5 and SiamRPN++ [J]. Computers and Electronics in Agriculture, 2022, 192: 106512.
- [12] LI S, GU X, XU X, et al. Detection of concealed cracks from ground penetrating radar images based on deep learning algorithm[J]. Construction and Building Materials, 2021, 273: 121949.
- [13] 杨晓玲, 江伟欣, 袁浩然. 基于 yolov5 的交通标志识别检测[J]. 信息技术与信息化, 2021(4): 28-30.
- [14] 彭成, 张乔虹, 唐朝晖, 等. 基于 YOLOv5 增强模型的口罩佩戴检测方法研究[J]. 计算机工程, 2022, 48(4): 39-49.
- [15] SRI J S, ESTHER R P. Little YOLO-SPP: A delicate real-time vehicle detection algorithm[J]. Optik, 2021, 225: 165818.
- [16] 谈世磊, 别雄波, 卢功林, 等. 基于 YOLOv5 网络模型的人员口罩佩戴实时检测[J]. 激光杂志, 2021, 42(2): 147-150.
- [17] 尹靖涵, 瞿绍军, 姚泽楷, 等. 基于 YOLOv5 的雾霾天气下交通标志识别模型[J]. 计算机应用, 2022, 42(9): 2876-2884.
- [18] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [19] WU L G, ZHANG L, SHI J H, et al. Damage detection of grotto murals based on lightweight neural network [J]. Computers and Electrical Engineering, 2022, 208: 108237.
- [20] 赵永强, 饶元, 董世鹏, 等. 深度学习目标检测方法综述[J]. 中国图象图形学报, 2020, 25(4): 629-654.
- [21] BOSMAN A S, ENGELBRECHT A, HELBIG M. Visualising basins of attraction for the cross-entropy and the squared error neural network loss functions[J]. Neurocomputing, 2020, DOI: 10.1016/j.neucom. 2020. 02. 113.

作 者 简 介

王振, 硕士研究生, 主要研究方向为机器学习、人工智能识别与检测等。
E-mail: 851682878@qq.com

方海峰,副教授,硕士生导师,主要研究方向为机器自动化、图像识别及机器视觉等。

E-mail:fanghale@163.com

曹晋,博士研究生,主要研究方向为图像处理及人工智能,深度学习模糊化处理等。

E-mail:1534357324@qq.com

吴群彪,博士,副教授,主要研究方向为深度学习识别、机器自动化处理等。

E-mail:just_wqb@just.edu.cn