

基于改进 YOLOv5 的茶叶杂质检测算法

黄少华, 梁喜凤*

(中国计量大学机电工程学院, 杭州 310018)

摘 要: 针对现有目标检测算法检测茶叶杂质精度低、速度慢的问题, 该研究提出了一种基于改进 YOLOv5 的茶叶杂质检测算法。采用 K-Means 聚类算法对杂质真实框聚类, 以获取适合茶叶杂质特征的锚框尺寸; 通过在主干特征提取网络 CSPDarkNet 中引入前馈卷积注意力机制 (Convolutional Block Attention Module, CBAM), 将茶叶杂质输入特征图依次经过通道注意力模块和空间注意力模块, 获得特征图通道维度和空间维度的关键特征; 在颈部网络中添加空间金字塔池化 (Spatial Pyramid Pooling, SPP) 模块, 融合并提取不同感受野的关键特征信息; 将普通卷积替换成深度可分离卷积, 增大目标预测特征图的置信度损失权重, 构建了轻量化的改进 YOLOv5 网络结构模型; 分别制作了铁观音茶叶中混合有稻谷、瓜子壳、竹枝和茶梗 4 种杂质的数据集并进行茶叶杂质检测试验。结果表明, 改进的 YOLOv5 比常规 YOLOv5 在茶叶杂质检测中具有更高的置信度分数, 且定位更为准确, 未出现漏检现象。改进 YOLOv5 的多类别平均精度 (Mean Average Precision, mAP) 和每秒传输帧数 (Frame Per Second, FPS) 达到 96.05% 和 62 帧/s, 均优于主流的目标检测算法, 验证了改进算法的高效性和鲁棒性。该研究成果可为提升茶叶制作过程中小目标杂质检测精度与检测速度奠定基础。

关键词: 深度学习; 算法; 目标检测; YOLOv5; 茶叶杂质

doi: 10.11975/j.issn.1002-6819.2022.17.036

中图分类号: TP391

文献标志码: A

文章编号: 1002-6819(2022)-17-0329-08

黄少华, 梁喜凤. 基于改进 YOLOv5 的茶叶杂质检测算法[J]. 农业工程学报, 2022, 38(17): 329-336. doi: 10.11975/j.issn.1002-6819.2022.17.036 <http://www.tcsae.org>
Huang Shaohua, Liang Xifeng. Detecting the impurities in tea using an improved YOLOv5 model[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(17): 329-336. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2022.17.036 <http://www.tcsae.org>

0 引 言

茶叶具有很高的商业价值和药用价值, 在中国具有庞大的消费市场。茶叶分拣是茶叶制作工序中重要环节。茶叶采集过程中, 通常会混入树枝、谷物等杂质, 影响茶叶的口感和品质^[1-4]。为提升茶叶成品质量, 对采摘收集到的新茶炒制成成品茶叶后还需进一步分拣, 剔除掉多余的杂质。传统的茶叶分拣方式采用人工手动分拣剔除杂质, 需要消耗大量人力, 耗时长, 效率低。近年来, 机器视觉逐渐应用于茶叶杂质分拣, 识别精准高, 有助于实现茶叶自动化分拣, 对茶叶分拣技术的产业化具有较高的应用价值^[5]。

基于深度学习的目标检测算法通过强大的卷积神经网络学习茶叶杂质的特征信息, 并使用回归器和分类器对杂质进行定位和分类, 能够实现复杂场景下的茶叶杂质检测。常见的基于深度学习目标检测算法根据是否使用区域建议网络 (Region Proposal Network, RPN)^[6-7] 可分为双阶段目标检测和单阶段目标检测。双阶段检测精度较高但训练复杂, 检测耗时较长^[8]。常见的单阶段检测算法有 SSD 网络^[9-10] 和 YOLO 系列网络^[11]。单阶段检

测省去了候选框生成的工作, 精简了模型参数量, 在牺牲了部分检测精度的前提下, 检测速度显著快于双阶段目标检测。目前大部分检测任务为了追求实时性原则, 均以单阶段检测为主。其中, 以 YOLOv5 为代表的单阶段轻量级网络在综合汇聚了前几代 YOLO 系列网络的基础上, 改良了对小目标的检测速度和精度, 具有较好的检测效果。基于此, 胡根生等^[12]以 YOLOv5 为基线, 利用卷积核组进行了复杂背景下茶尺蠖的检测。尚钰莹等^[13]提出了一种基于 YOLOv5s 深度学习的苹果花朵的检测方法。Bochkovskiy 等^[14]在 YOLOv5 的基础上融合多种策略, 以适应不同场景下的目标检测任务。吕禾丰等^[15]在通过替换 YOLOv5 中损失函数的方法, 提高了交通标志识别精度和准确率。吕军等^[16]提出在 YOLOv5 模型采用 CSPDarknet53 特征提取网络进行茶叶嫩芽检测, 改进了自然光照条件检测的鲁棒性。Zhu 等^[17]构建了 TPH-YOLOv5 网络, 在原始 YOLOv5 基础上增添了 Transformer 模块和 CBAM 卷积注意力模块, 提升了网络模型对小尺度目标的检测性能。徐凤如等^[18]通过采用改进型 YOLOv4-Dense 算法, 研究了茶树茶叶采摘点的定位方法, 具有良好的识别率。以上研究对 YOLOv4 或 YOLOv5 目标检测网络进行了改进和优化。然而在茶叶杂质检测任务中, 茶叶簇杂乱无章, 背景较为复杂, 且杂质目标较小, 混入在茶叶中检测难度大, 树枝等杂质和茶叶颜色相近, 增加了检测难度, 目前的 YOLOv5 网络提取杂质特征能力不足, 检测精度较低, 预测框定位

收稿日期: 2022-06-18 修订日期: 2022-08-27

基金项目: 国家自然科学基金项目 (31971796)

作者简介: 黄少华, 研究方向为农业智能信息处理。

*通信作者: 梁喜凤, 教授, 研究方向为农业机器人、农业智能信息处理等。

Email: lxfcjlu@cjlu.edu.cn

不准确,检测重叠的小目标杂质容易出现漏检等,不能满足小目标茶叶杂质检测需求。

因此,针对以上存在问题,本文提出了一种基于改进 YOLOv5 的茶叶杂质检测算法。采用 K-Means 聚类锚框尺寸,主干网络 CSPDarkNet 引入卷积注意力机制(Convolutional Block Attention Module, CBAM),颈部网络中添加空间金字塔池化(Spatial Pyramid Pooling, SPP)模块,利用深度可分离卷积代替普通卷积,增大小目标预测特征图的置信度损失权重等方面对 YOLOv5 模型进行改进,并对多种茶叶杂质数据集进行试验,验证改进算法的高效性和鲁棒性,以提高茶叶制作中杂质检测的精度与速度。

1 改进的 YOLOv5 网络结构

YOLOv5 模型参数量最少,检测速度快,广泛应用在实际场景中。因此选取 YOLOv5 作为茶叶杂质检测的基础模型。但是,茶叶杂质一般较小,种类复杂,且树枝等杂质颜色与茶叶相近,混入茶叶簇,因此需要在锚框尺寸、关键特性关注度、检测精度等方面进行 YOLOv5 网络改进,从而适应茶叶杂质检测需求。

1.1 先验锚框选取

初始的先验锚框尺寸影响网络训练速度和定位准确性,常规 YOLOv5 锚框针对 COCO (Microsoft Common Objects in Context) 和 VOC (Visual Object Classes) 数据集,而茶叶杂质均为小目标,基于初始锚框的训练不适用于茶叶杂质。如果数据集标注信息针对默认锚框的最佳召回率小于 0.98 时,则需要重新计算符合此数据集的锚框^[19]。YOLOv5 的重新自适应锚框计算采用 K-means^[20-22]重新定义锚框尺寸的聚类算法,实现步骤如下:

- 1) 取出茶叶杂质真实框的宽高值作为坐标点集合,在所有坐标点中随机抽取 K 个初始簇心。
- 2) 计算剩余坐标点与簇心的欧氏距离,将剩余坐标点与最近簇归为同一种类,当所有坐标点归类完成后,根据每个簇中所有坐标点的平均值重新计算每个簇心。
- 3) 反复迭代上述步骤 2,直至簇心不再改变。

由于 YOLOv5 头部网络输出 3 种尺度预测特征图,每个预测特征图均具有大、中、小三种尺度锚框,因此 K 取 9,此时茶叶杂质锚框被分为 9 种不同尺寸的簇。

聚类试验达到稳定后的 9 种锚框尺寸参数如表 1 所示,聚类后的锚框尺寸更符合茶叶杂质的尺寸和宽高比例特征。

表 1 改进的 YOLOv5 锚框尺寸

Table 1 Improved anchor box size for YOLOv5

特征图尺度 Feature map scale	特征图尺寸 Feature map size/ (像素×像素)	锚框尺寸(宽,高) Anchor frame size (width,height)/像素
大 Large	80×80	(113,30), (56,82), (104,60)
中 Medium	40×40	(60,20), (27,46), (49,37)
小 Small	20×20	(17,12), (31,14), (27,22)

1.2 注意力机制

针对茶叶杂质的目标检测任务,除了需要考虑杂质

的语义特征外,同时还需考虑茶叶杂质的空间位置,在 YOLOv5 主干网络中引入注意力机制,能够自适应学习茶叶杂质的关键特征。根据训练获得的注意力权重可增加网络对关键特征的注意力,降低干扰特征的关注度。

CBAM 将通道维度和空间维度相结合^[23],最终权衡得到特征图通道维度和空间维度的关键特征。通道注意力模块对输入茶叶杂质特征图矩阵 F 按照宽高方向进行一次全局平均池化(AvgPool)和全局最大池化(MaxPool),输出 2 个与 F 通道数相等、大小为 1×1 的池化层。然后将 2 个池化层输入多层感知机(Multilayer Perceptron, MLP)分别通过两次共享全连接层,最后将输出的 2 个通道权重层融合后进行非线性激活,得到充分权衡 F 通道维度上的权重层 M_c ,计算式如下:

$$M_c = \text{Sigmoid}(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (1)$$

将 M_c 权重参数与输入的茶叶杂质特征图矩阵 F 在通道上相乘,实现在通道上施加注意力机制,大幅提升特征图通道维度的关键特征利用率。

空间注意力模块对上述通道注意力模块提炼的茶叶杂质特征图矩阵 F' 通道维度上分别进行最大池化和平均池化,获得 2 个通道数为 1、尺寸为 F' 大小的池化层后进行堆叠,利用一次卷积降维至单通道特征图,最后经过非线性激活,增加非线性因素,得到充分权衡 F 空间维度上的权重层 M_s ,计算式如下:

$$M_s = \text{Sigmoid}(f^{7 \times 7}(\text{Concat}(\text{AvgPool}(F'), (\text{MaxPool}(F'))))) \quad (2)$$

式中 $f^{7 \times 7}$ 表示滤波器尺寸为 7×7 的卷积运算。

将 M_s 权重参数与 F' 逐像素相乘,实现在空间维度上施加注意力机制,以提升特征图空间维度的关键特征利用率。因此,将 CBAM 模块引入 YOLOv5 主干网络中能够提升茶叶杂质特征信息的关注度。

1.3 YOLOv5 主干网络改进策略

为验证 YOLOv5 引入注意力机制对茶叶杂质检测性能的影响,在主干网络的 CSP1_1、CSP1_3 与 CSP2_1 3 个跨阶段局部网络(Cross Stage Partial Network, CSPNet)模块后分别添加 CBAM 注意力机制。通过将原图输入不同策略下改进主干网络的最终预测结果图,采用热力图可视化^[24]的方式,直观展示网络更加关注的区域。图 1 对比了常规 YOLOv5 和 3 种改进策略的热力图,其中区域颜色越深代表网络越感兴趣。

从图 1 中可看出,主干网络中引入注意力机制有助于茶叶杂质特征提取效果,且引入 3 个注意力机制 CBAM 的改进策略关注的特征区域最为准确。

1.4 目标检测评价指标

目标检测指标通常有以下 4 种:

1) 单类别精确率:

$$P = T_c / (T_c + F_c) \times 100\% \quad (3)$$

2) 单类别召回率:

$$R = T_c / (T_c + F_n) \times 100\% \quad (4)$$

3) 单类别 P - R 曲线下面积:

$$AP = \int (P, R) \quad (5)$$

4) 多类别平均精度:

$$mAP = \frac{1}{c} \sum_{I=1}^c f(P, R, I) \quad (6)$$

式中 P 表示精确率； R 表示召回率； T_c 表示当前茶叶杂质种类被预测正确的数量； F_c 表示其他茶叶杂质种类但

被预测为当前茶叶杂质种类的数量； F_N 表示当前茶叶杂质种类被预测为其他茶叶杂质种类的数量； AP 表示单类别检测结果， AP 值越大，单类别检测性能越好； AP 反映多类别检测精度； c 表示所有类别数量。

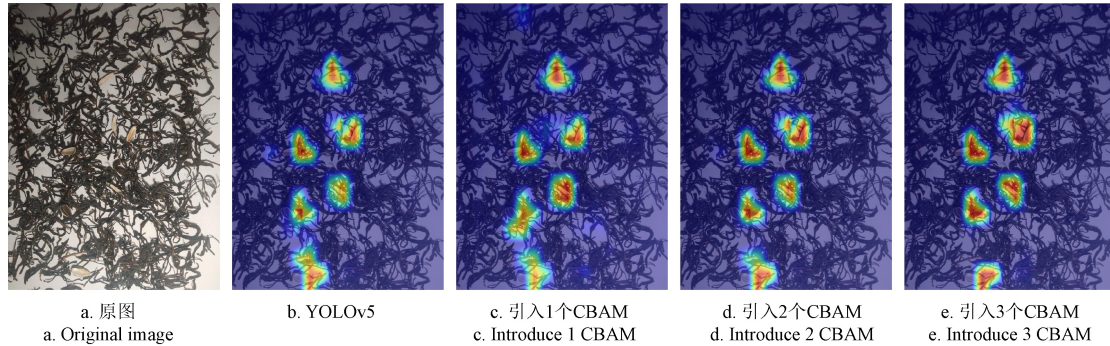


图 1 常规 YOLOv5 与 3 种改进网络热力图

Fig.1 Traditional YOLOv5 and heatmaps of three improved network

1.5 改进的主干网络测试

利用改进的主干网络策略，对茶叶杂质数据集进行试验，保持颈部网络（Feature Pyramid Network, FPN）和头部网络 YoloHead 不变，共训练 500 代数，试验结果如表 2 所示。

表 2 不同策略试验结果

Table 2 Experimental results of different strategies

网络 Network	mAP/%	FPS/(帧·s ⁻¹)
CSPDarkNet	92.14	53
CBAM1	92.68	52
CBAM2	93.04	52
CBAM3	93.62	51

注：mAP 为多类别平均精度，FPS 为每秒传输帧数。下同。

Note: mAP refers to mean average precision of multiple categories, and FPS refers to frame per second. Same below.

从表 2 中可看出，检测精度随着引入 CBAM 模块数的增加而提高，加入 1 个 CBAM、2 个 CBAM 和 3 个 CBAM 的多类别平均精度 mAP 分别提升了 0.54、0.9 和 1.48 个百分点，优于常规 YOLOv5，最终选取加入 3 个 CBAM 策略下的主干网络改进方法。但同时 FPS 略有降低，表明网络检测速度有所降低，需要进一步优化 YOLOv5 网络结构。

2 网络结构优化

2.1 优化方法

在常规 YOLOv5 的颈部网络（FPN）内添加空间金字塔池化（Spatial Pyramid Pool, SPP）模块^[25]，有助于聚合茶叶杂质的不同感受视野特征信息，提取多尺度的关键特征，提升网络的检测精度。因此，在不影响特征图上下采样和通道拼接的前提下，对颈部网络 FPN 中的每一个 CSP 模块后添加 SPP 模块。同时考虑到检测的实时性，对网络参数需要进一步优化，将 FPN 内的普通卷

积 Conv 替换成深度可分离卷积 Dwise^[26]，能够保证在精度不降低的前提下，大幅减少网络运算量，提升检测速度。

在改进主干网络加入 3 个 CBAM 策略基础上，改进并优化的网络整体结构如图 2。

2.2 平衡不同尺度损失

YOLOv5 的头部网络 YoloHead 共输出 3 个不同尺度大小的预测特征图，其中 YoloHead1、YoloHead2 和 YoloHead3 分别为 80×80、40×40 和 20×20（像素）的特征图，对输入图像的感受野逐渐增大，分别用于预测小目标物体、中等目标物体和大目标物体。由于茶叶杂质均为小目标，增大 YoloHead1 的损失权重，能够提高 YOLOv5 检测小目标的精度。因此，经过训练试验，对 3 个预测特征图（YoloHead1、YoloHead2 和 YoloHead3）的置信度损失设置不同的权重，获得总置信度损失如下：

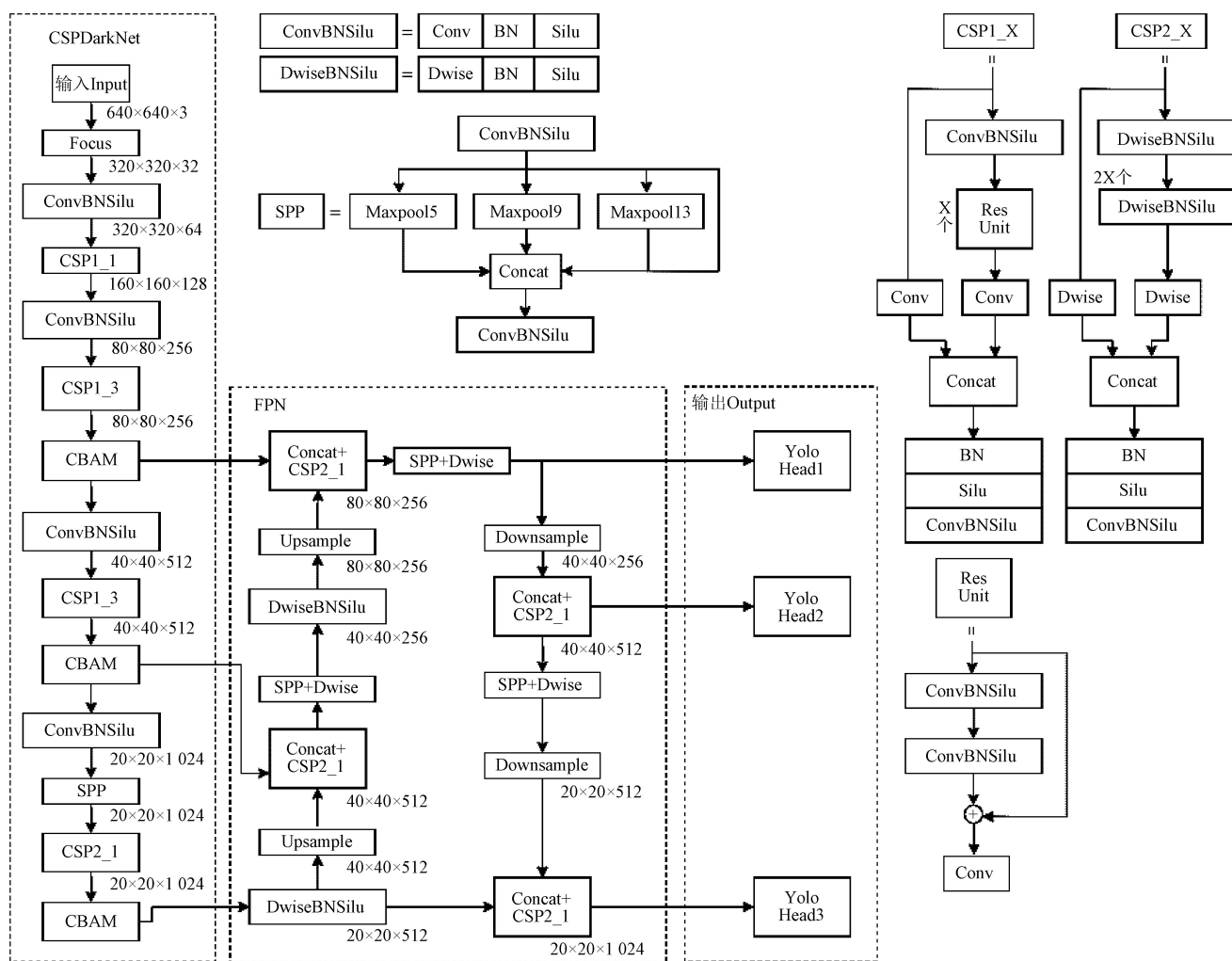
$$L_{obj} = 5.0L_{obj}^{YoloHead1} + 2.5L_{obj}^{YoloHead2} + 0.5L_{obj}^{YoloHead3} \quad (7)$$

式中 L_{obj} 为茶叶杂质检测网络的总置信度损失，其上标表示不同预测特征图。

3 茶叶杂质检测试验

3.1 茶叶杂质数据集制作

茶叶在采摘和炒制过程中通常会混有一些杂质，例如谷物、茶梗、树枝等，因此本次试验以市场购买的铁观音茶叶为研究对象，在茶叶中掺杂现实场景中容易混入稻谷、瓜子壳、竹枝和茶梗 4 种杂质，制作了具有杂质的茶叶杂质检测数据集。采用分辨率为 150 dpi 的超低畸变相机，拍摄了 4×500 共 2 000 张包含 4 种茶叶杂质的图像。通过水平镜像、上下翻转、亮度变换和随机对比度等数据增强方式将茶叶杂质数据集扩充至 4 000 张图像，按照 8:1:1 的比例随机划分训练集、验证集和测试集。利用 Labelimg 软件对其进行标注，生成杂质真实框坐标和种类信息的标签文件。



注：FPN 为特征图金字塔网络结构；CBAM 为前馈卷积注意力机制；SPP 为空间金字塔池化结构；CSP 为跨阶段局部网络。

Notes: FPN is Feature Pyramid Network; CBAM is convolutional block attention module; SPP is a spatial pyramid pool structure; CSP is cross stage partial network.

图 2 改进并优化的 YOLOv5 网络结构
Fig.2 Improved and optimized YOLOv5 network structure

3.2 试验环境

试验环境配置：深度学习框架 Pytorch-GPU V1.7.1，显卡 Nvidia GeForce GTX 1660，显存 6 GB，处理器 Inter Core i7-9700k 3.60 GHz 八核，内存 16 GB，神经网络加速库 CUDNN7.6.4 和 CUDA10.1，所用到的相关库 Pillow8.0.1、Opencv4.4.0、numpy1.19.4 和 matplotlib3.3.3；操作系统为 Windows10，编程语言为 Python3.7；光照环境为 1 000 lx 光源；茶叶随机铺设一层。

3.3 网络训练

将改进后的 YOLOv5 网络在图形处理器 (Graphics Processing Unit, GPU) 上进行训练，采用随机梯度下降优化器 (Stochastic Gradient Descent, SGD) 更新参数，初始学习率设置成 0.001，学习率衰减系数为 0.000 1，批尺寸 (Batch_size) 为 8，共训练 500 个世代，损失值变化如图 3 所示。从图 3 中可以看出，随着训练代数的增加，改进 YOLOv5 的损失值逐渐降低，整体趋于稳定，最终损失值收敛于 0.3 附近，证明改进的网络达到了较优的训练结果。

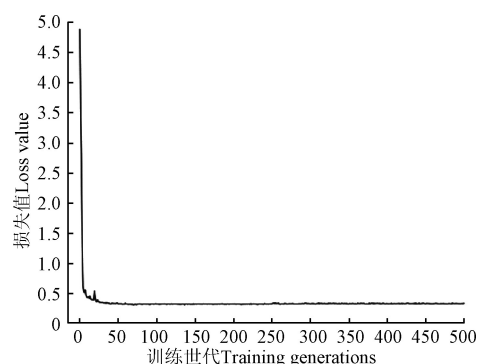
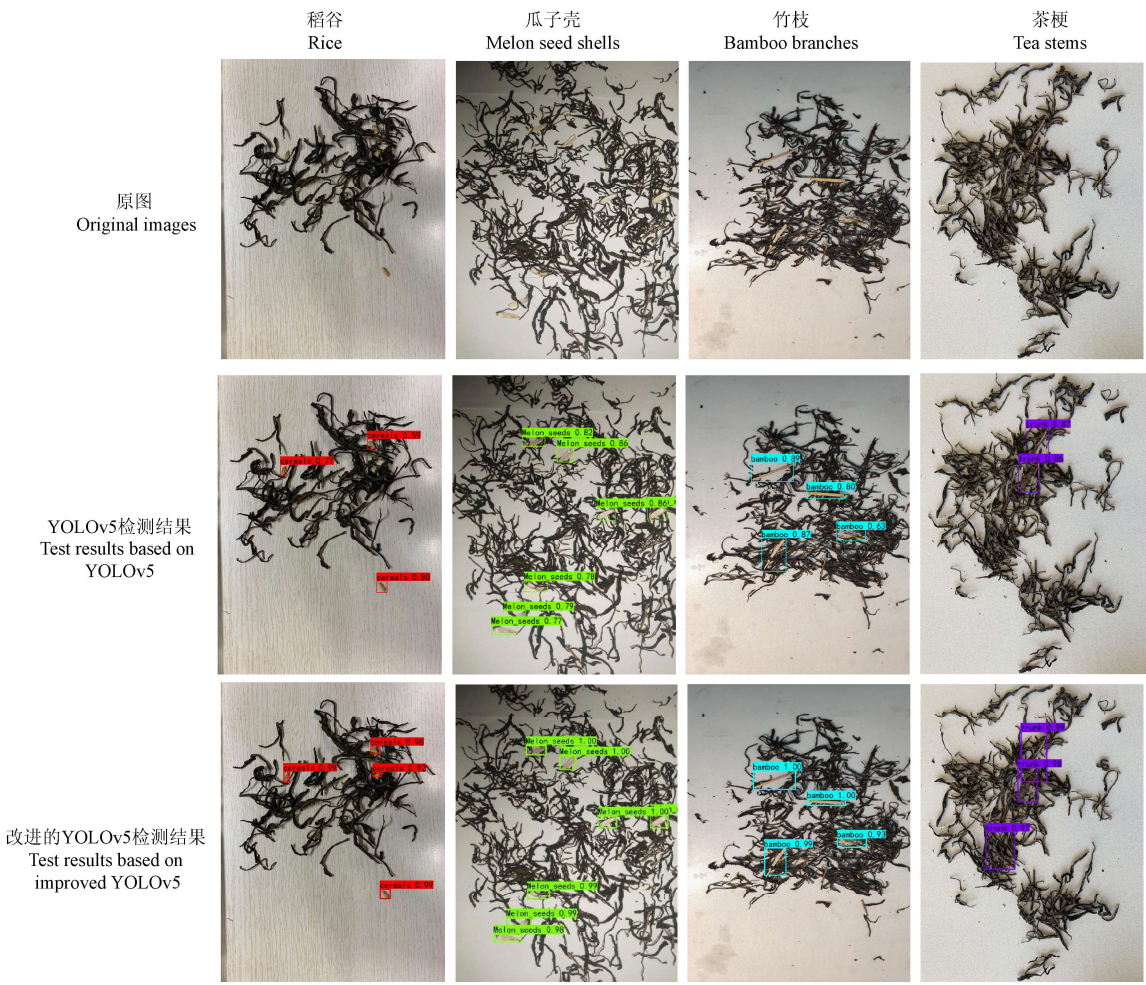


图 3 损失值变化曲线

Fig.3 Loss value change curve

3.4 检测试验对比

将常规 YOLOv5 和改进的 YOLOv5 在测试集上进行试验对比，且保证光照环境相等，每种杂质采样 1 000 张图像，各自随机选择 1 张检测结果展示，如图 4 所示。其中，I、II、III和IV分别代表茶叶中含有稻谷、瓜子壳、竹枝和茶梗杂质图像。



注：图中锚框代表检测出的杂质。
Note: The anchor frames represent the detected impurities.

图 4 基于 YOLOv5 与改进 YOLOv5 的茶叶杂质视觉检测结果对比
Fig.4 Comparison of visual test results of tea impurities based on YOLOv5 and improved YOLOv5

从图 4 可以看出，由于稻谷杂质目标较小，混入茶叶中不易发现，常规 YOLOv5 对于部分被遮挡的稻谷出现了漏检现象，而改进的 YOLOv5 只是置信度略有降低。对于有茶叶遮挡的竹枝，常规 YOLOv5 的置信度分数大大低于改进的 YOLOv5。由于茶梗和茶叶特征相似，且颜色相近，常规 YOLOv5 中出现了漏检现象。从 4 种图像检测结果对比图中可看出，改进的 YOLOv5 相较于常

规 YOLOv5 具有更高的置信度分数，且定位更为准确，没有出现漏检现象，尤其对于有遮挡情况的杂质表现更为明显。同时可以发现，目标检测杂质被遮挡部分越多，则检测的置信度值越低。

通过视觉检测的定性分析后，还需对 YOLOv5 改进的每个策略进行消融试验对比的定量分析，如表 3 所示。

表 3 不同策略消融试验结果
Table 3 Results of ablation experiments with different strategies

检测方法 Test method	AP/%				mAP/%	FPS/(帧·s ⁻¹)
	稻谷 Rice	瓜子壳 Melon seed shells	竹枝 Bamboo branches	茶梗 Tea stems		
1	93.02	94.53	93.14	87.87	92.14	53
2	93.09	94.62	93.19	88.98	92.47	53
3	94.52	94.91	94.03	91.02	93.62	51
4	95.34	96.83	96.32	93.39	95.47	62
5	96.81	97.64	96.35	93.40	96.05	62

注：方法 1 表示常规 YOLOv5，方法 2 表示改进锚框尺寸，方法 3 表示改进锚框尺寸并引入注意力机制，方法 4 表示改进锚框尺寸、引入注意力机制并对网络整体结构进行改进，方法 5 表示改进锚框尺寸、引入注意力机制、改进网络整体结构和平衡不同尺度损失。
Notes: Method 1 represents Regular YOLOv5. Method 2 indicates improved anchor frame size. Method 3 means to improve the size of anchor frame and introduce attention mechanism. Method 4 means to improve the size of anchor frame, introduce attention mechanism and improve the overall network structure. Method 5 represents improving the size of anchor frame, introducing attention mechanism, improving the overall network structure and balancing losses of different scales.

从表 3 可知, 方法 2 比方法 1 检测精度略有提升, 表明利用 K-means 聚类得到的锚框尺寸更符合茶叶杂质数据集特征。方法 3 检测精度优于常规 YOLOv5, 但检测速度略有减慢, 表明引入注意力机制有利于网络提取茶叶杂质的关键特征, 但同时也增加了少量的模型复杂度。方法 4 的检测精度和速度都有一定提升, 表明在 FPN 中添加空间金字塔池化 (SPP) 模块有助于提取多尺度的特征, 检测精度得到提升, 且深度可分离卷积的引入节省了网络运算量, 实现网络轻量化, 提升了检测速度。本文采用的方法 5 的 mAP 和 FPS 达到了 96.05% 和 62 帧/s, 优于以上方法, 针对稻谷和瓜子壳小目标的 mAP 得到较大提升, 表明增大小目标损失权重, 对于提高小目标检测精度有很大帮助。

3.5 不同目标检测网络试验对比

选用 Faster R-CNN、SSD、YOLOv3、YOLOv4、YOLOv5 等主流目标检测网络和改进 YOLOv5 进行试验对比, 结果如表 4 所示。

表 4 不同网络试验结果

Table 5 Experimental results based on different network

网络 Network	mAP/%	FPS/(帧·s ⁻¹)
Faster R-CNN	90.92	15
SSD	78.59	42
YOLOv3	87.52	52
YOLOv4	90.52	46
YOLOv5	92.14	53
改进 YOLOv5	96.05	62

可以看出, Faster R-CNN 的检测精度较高, 但其采用双阶段检测方式, 结构复杂且网络参数多, 检测速度最慢, 无法满足实时性要求。SSD 参数量较少, 检测速度较快, 但特征提取能力弱于 YOLO 系列算法, mAP 最低。YOLOv3 和 YOLOv4 检测效果都低于常规 YOLOv5。改进后的 YOLOv5 相对于 Faster R-CNN、SSD、YOLOv3、YOLOv4 和常规 YOLOv5 检测精度和速度均最优。

4 结 论

以常规 YOLOv5 网络为基础进行改进, 更高效地实现茶叶杂质检测。使用 K-Means 聚类获得适合茶叶杂质特征的锚框, 引入注意力机制, 提高主干网络特征提取能力, 添加空间金字塔池化 (Spatial Pyramid Pooling, SPP) 模块, 融合不同感受野多尺度特征, 提升检测精度, 采用深度可分离卷积降低网络复杂度, 加快检测速度, 平衡不同尺度损失, 提高对小目标检测精度。通过定性和定量试验结果对比, 验证了改进 YOLOv5 网络的有效性。改进 YOLOv5 的多类别平均精度 mAP 和每秒传输帧数 (Frame Per Second, FPS) 分别达到 96.05% 和 62 帧/s, 满足茶叶杂质检测的需求, 并且对小目标检测领域提供了一定的参考。

对于不同的背景、光照条件、茶叶平铺厚度以及图像分辨率等复杂条件对改进 YOLOv5 模型在茶叶杂质检测精度等方面产生的影响, 仍需要进一步的研究与试验。

[参 考 文 献]

- [1] 鹿瑶, 王伟, 钟斌, 等. 融合光电色选的皮带筛式油茶果壳籽分选机设计与试验[J]. 农业机械学报, 2020, 51(S1): 429-439.
- Lu Yao, Wang Wei, Zhong Bin, et al. Design and experiment of belt screen seed sorter for *Camellia oleifera* fused with photoelectric color sorting technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(S1): 429-439. (in Chinese with English abstract)
- [2] 李仿舟. 茶叶色选机的技术创新和发展趋势[J]. 农机质量与监督, 2018, 27(12): 18-19.
- [3] 丁勇, 徐奕鼎, 王辉军. 工夫红茶精制中色选技术应用研究[J]. 中国茶叶加工, 2009, 26(3): 22-23, 27.
- [4] 黄留锁, 宋艳. 茶叶茶梗的计算机在线识别技术及其应用[J]. 福建茶叶, 2016, 38(7): 161-162.
- [5] 陈培俊. 基于图像处理和模式分类的茶叶杂质识别研究[D]. 南京: 南京航空航天大学, 2014: 1-2.
- Chen Peijun. Research on Recognizing Tea-Leaves and Impurities Based on Image Processing and Pattern Classification[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2014: 1-2. (in Chinese with English abstract)
- [6] Cao S T, Jiang Y L. Combining attention mechanisms with RPN networks for target tracking[J]. International Core Journal of Engineering, 2021, 7(10): 438-443.
- [7] 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, Columbus, USA: IEEE, 2014: 580-587.
- [8] Ren S, He K M, 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, 2017, 39(6): 1137-1149.
- [9] 亢洁, 刘港, 郭国法. 基于多尺度融合模块和特征增强的杂草检测方法[J]. 农业机械学报, 2022, 53(4): 254-260.
- Kang Jie, Liu Gang, Guo Guofa. Weed detection based on multi-scale fusion module and feature enhancement[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 254-260. (in Chinese with English abstract)
- [10] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]//European Conference on Computer Vision, German: Springer, 2016: 21-37.

- [11] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA: IEEE, 2016: 779-788.
- [12] 胡根生, 吴继甜, 鲍文霞, 等. 基于改进 YOLOv5 网络的复杂背景图像中茶尺蠖检测[J]. 农业工程学报, 2021, 37(21): 191-198.
- Hu Gensheng, Wu Jitian, Bao Wenxia, et al. Detection of Ectropis oblique in complex background images using improved YOLOv5[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(21): 191-198. (in Chinese with English abstract)
- [13] 尚钰莹, 张倩如, 宋怀波, 等. 基于 YOLOv5s 的深度学习在自然场景苹果花朵检测中的应用[J]. 农业工程学报, 2022, 38(9): 222-229.
- Shang Yuying, Zhang Qianru, Song Huaibo. Application of deep learning using YOLOv5s to apple flower detection in natural scenes[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(9): 222-229. (in Chinese with English abstract)
- [14] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, Seattle, USA: IEEE, 2020: 1-17.
- [15] 吕禾丰, 陆华才. 基于 YOLOv5 算法的交通标志识别技术研究[J]. 电子测量与仪器学报, 2021, 35(10): 137-144.
- Lyu Hefeng, Lu Huacai. Research on traffic sign recognition technology based on YOLOv5 algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(10): 137-144. (in Chinese with English abstract)
- [16] 吕军, 方梦瑞, 姚青, 等. 基于区域亮度自适应校正的茶叶嫩芽检测模型[J]. 农业工程学报, 2021, 37(22): 278-285.
- Lyu Jun, Fang Mengrui, Yao Qing, et al. Detection model for tea buds based on region brightness adaptive correction[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(22): 278-285. (in Chinese with English abstract)
- [17] 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, Montreal, Canada: IEEE, 2021: 2778-2788.
- [18] 徐凤如, 张昆明, 张武, 等. 一种基于改进 YOLOv4 算法的茶树芽叶采摘点识别及定位方法[J]. 复旦学报(自然科学版), 2022, 61(4): 460-471.
- Xu Fengru, Zhang Kunming, Zhang Wu, et al. Identification and localization method of tea bud leaf picking point based on improved YOLOv4 algorithm[J]. Journal of Fudan University (Natural Science), 2022, 61(4): 460-471. (in Chinese with English abstract)
- [19] 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.
- Peng Cheng, Zhang Qiaohong, Tang Zhaohui, et al. Research on mask wearing detection method based on YOLOv5 enhancement model[J]. Computer Engineering, 2022, 48 (4): 39-49. (in Chinese with English abstract)
- [20] 杨俊闯, 赵超. K-Means 聚类算法研究综述[J]. 计算机工程与应用, 2019, 55(23): 7-14.
- Yang Junchuang, Zhao Chao. Survey on K- Means clustering algorithm[J]. Computer Engineering and Applications, 2019, 55(23): 7-14. (in Chinese with English abstract)
- [21] 安晓飞, 王培, 罗长海, 等. 基于 K-means 聚类和分区寻优的秸秆覆盖率计算方法[J]. 农业机械学报, 2021, 52(10): 84-89.
- An Xiaofei, Wang Pei, Luo Changhai, et al. Corn straw coverage calculation algorithm based on K-means clustering and zoning optimization method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10): 84-89. (in Chinese with English abstract)
- [22] Darmawan E, Novantara P J, Suwanto G P, et al. The implementation of K-means algorithm to determine the quality of teak wood in image based on the texture[J]. Journal of Physics: Conference Series, 2021, 1933(1): 1-7.
- [23] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision, German: Springer, 2018: 3-19.
- [24] Selvaraju R R, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]//IEEE International Conference on Computer Vision, Venice, Italy: IEEE, 2017: 618-626.
- [25] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [26] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA: IEEE, 2017: 1251-1258.

Detecting the impurities in tea using an improved YOLOv5 model

Huang Shaohua, Liang Xifeng*

(College of Mechanical & Electrical Engineering, China Jiliang University, Hangzhou 310018, China)

Abstract: Tea sorting has been one of the most important links in tea production. Manual sorting has been often adopted to remove the excess impurities (such as branches and grains) from the collected fresh tea in traditional processing. However, the current sorting cannot fully meet the high requirement of taste and quality in the finished tea products after collection in recent years, due to the labor-intensive and high cost. Fortunately, machine vision has been gradually applied to tea impurity sorting, particularly for fully automatic sorting in the process of tea collection. Among them, the single-stage lightweight network (represented by YOLOv5 deep learning) can perform better performance for small targets with high detection speed and accuracy. However, the conventional YOLOv5 network cannot be used to extract the characteristics of tea impurities, due to the disorderly clusters, the generally small targets, the complex types of impurities, and the similar color to the tea. Particularly, the overlapping small targets can cause an inaccurate prediction box, leading to low accuracy or miss detection of the tea impurities. It is necessary to improve the conventional YOLOv5 network to meet the requirements of tea impurity detection. In this study, an improved YOLOv5 model was proposed to detect the tea impurity with a higher accuracy and detection speed than before. The YOLOv5 was taken as the baseline network. The K-Means clustering was applied to cluster the real boxes of impurities as the anchor frame size suitable for the characteristics of tea impurities. Convolutional Block Attention Module (CBAM) was introduced into the backbone feature extraction network (CSPDarkNet). The key features were obtained using the channel and spatial dimension of feature images. A Spatial Pyramid Pooling (SPP) module was added to the neck network, in order to integrate and extract the multi-scale features of different sensory fields. The deep separable convolution was updated to reduce the number of network parameters for the higher detection speed. The confidence loss weight of the small target prediction in the feature map was improved for the higher detection accuracy of the network for the small targets. The data set was taken as the Tieguanyin tea mixed with the rice, melon seed shell, bamboo branches, and tea stems. The results show that the improved YOLOv5 presented a higher confidence score than the conventional one, where the positioning was much more accurate without missing detection. The mAP and FPS of improved YOLOv5 reached 96.05% and 62 frames/s, respectively. The higher efficiency and robustness of the improved model were achieved to compare the mainstream target detections. The findings can provide a strong reference for the detection accuracy and speed of small target impurities in the tea production process.

Keywords: deep learning; algorithm; object detection; YOLOv5; tea impurities