

改进 YOLOv5 识别复杂环境下棉花顶芽

彭 炫^{1,2}, 周建平^{1,2*}, 许 燕^{1,2}, 席光泽¹

(1. 新疆大学机械工程学院, 乌鲁木齐 830017; 2. 新疆维吾尔自治区农牧机器人及智能装备工程研究中心, 乌鲁木齐 830017)

摘 要: 为提高复杂环境下棉花顶芽识别的精确率, 提出了一种基于 YOLOv5s 的改进顶芽识别模型。建立了田间复杂环境下棉花顶芽数据集, 在原有模型结构上增加目标检测层, 提高了浅层与深层的特征融合率, 避免信息丢失。同时加入 CPP-CBAM 注意力机制与 SIOU 边界框回归损失函数, 进一步加快模型预测框回归, 提升棉花顶芽检测准确率。将改进后的目标检测模型部署在 Jetson nano 开发板上, 并使用 TensorRT 对检测模型加速, 测试结果显示, 改进后的模型对棉花顶芽识别平均准确率达到 92.8%。与 Fast R-CNN、YOLOv3、YOLOv5s、YOLOv6 等算法相比, 平均准确率分别提升了 2.1、3.3、2、2.4 个百分点, 该检测模型适用于复杂环境下棉花顶芽的精准识别, 为后续棉花机械化精准打顶作业提供技术理论支持。

关键词: 作物; 图像识别; 小目标检测; YOLOv5s; SIOU 损失函数; CPP-CBAM 注意力机制; 棉花打顶

doi: 10.11975/j.issn.1002-6819.202305026

中图分类号: S24

文献标志码: A

文章编号: 1002-6819(2023)-16-0191-07

彭炫, 周建平, 许燕, 等. 改进 YOLOv5 识别复杂环境下棉花顶芽[J]. 农业工程学报, 2023, 39(16): 191-197.

doi: 10.11975/j.issn.1002-6819.202305026 <http://www.tcsae.org>

PENG Xuan, ZHOU Jianping, XU Yan, et al. Cotton top bud recognition method based on YOLOv5-CPP in complex environment[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(16): 191-197. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202305026 <http://www.tcsae.org>

0 引 言

棉花是中国仅次于粮食的第二大作物, 是关系到国计民生的重要资源。近年来, 棉花种植主要依靠人工, 劳动力短缺、机械智能化生产程度低等问题仍需解决, 因此中国的棉花产量仍有很大提升空间^[1-5]。棉花打顶是棉田作业过程中的重要环节, 脱离人工打顶促进智能机械化打顶是降低成本、提升棉花生产效率的重要手段。对于棉花打顶机器人的设计, 首要任务是解决复杂棉田环境中不损伤棉花蕾, 精准驱动激光去除顶芽是其重中之重, 因此精准识别定位棉花顶芽是关键环节。

近年来, 随着 Fast R-CNN、YOLO 系列等^[6-9]经典目标检测算法不断迭代更新, 其迅速在各个领域得到了广泛应用。国内外研究学者针对目标检测技术在农业领域的应用进行了广泛的研究。KOIRALA 等^[10]为验证芒果发育阶段影响开花数量, 使用 YOLOv3 算法在芒果发育阶段自动检测圆锥花序数量, 但 YOLOv3 检测物体的实时性较慢。刘海涛等^[11]为实现非接触、低成本、精准识别棉花顶芽, 提出一种基于 YOLOv4 模型的复杂环境下棉花顶芽识别方法, 然而该设计对顶芽了解不够全面,

并没有对顶芯进行识别, 造成顶芽识别率过低, 打顶效果不佳。JIA 等^[12-13]改进了 Mask RCNN 算法, 提升了树叶遮挡等复杂环境下苹果识别的准确度, 但是 Mask RCNN 模型为一阶段检测模型, 较比 YOLO 模型检测效率较低。朱旭等^[14]采用 Fast R-CNN 模型算法在恶劣环境下对蓝莓进行识别, 目标检测准确率和效率明显提高, 但是 Fast R-CNN 算法需要采集大量的数据训练模型, 在样本不足的时检测准确度明显降低。李天华等^[15-16]设计了一种基于 YOLOv4 的改进算法, 实现了对复杂环境下的番茄识别定位, 能够有效解决果实遮挡、光照影响下番茄识别准确率低的问题, 但训练成本过大, 检测帧数过低, 模型计算量过大, 不利于模型轻量化运行。CHEN 等^[17]为提升茶蕾识别定位精度, 保证茶蕾采摘的完整率, 设计了一种基于 YOLOv3 的改进茶蕾识别算法模型, 该模型能在光照变化和部分遮挡等复杂环境下识别茶蕾并定位采摘点。茶蕾与棉花顶芽同样是小目标农作物, 但 YOLOv3 检测算法对小目标提取效果并不理想, 该算法识别的准确率有待加强。薛月菊等^[18]针对果园场景下光照的多样性、背景的复杂性及芒果与树叶颜色的高度相似性, 特别是树叶和枝干对果实遮挡及果实重叠等问题, 提出了改进 YOLOv2 检测方法, 但由于 YOLOv2 模型过大, 识别效果和检测速度不够理想。樊湘鹏等^[19]为建立高效、准确的葡萄叶部病害检测系统, 引入迁移学习机制, 利用大型公开数据集对 VGG16 模型预训练, 保持模型前端 13 个层等参数和权重不变, 对全连接层和分类层改进后利用新数据集微调训练模型, 虽然能够完成对葡萄叶部病虫害识别, 但训练成本过大, 检测速度较慢。

基于此, 上述研究学者在目标检测技术存在小目标检测精度低、实时检测慢、模型计算量大等问题, 本文

收稿日期: 2023-05-06 修订日期: 2023-07-22

基金项目: 2022 新疆大学优秀博士研究生科研创新项目 (XJU2022BS086); 新疆维吾尔自治区自然科学基金项目: 复杂环境下棉花顶芽识别定位与激光去除技术研究 (No.2022D01C67); 新疆农机研发制造推广应用一体化项目: 棉花激光打顶机研发制造推广应用 (No.2022D14002)

作者简介: 彭炫, 博士, 研究方向为计算机视觉技术农业领域应用。

Email: xjdxpx@126.com

*通信作者: 周建平, 教授, 博士生导师, 研究方向为机电一体化、农业机械化与信息化。Email: linkzhou@163.com

对复杂环境下对棉花顶芯识别为研究对象, 本文选用 YOLOv5s 模型, 针对上述问题进行优化, 设计一种基于 YOLOv5s 的改进棉花顶芽识别模型, 为棉花机械化打顶提供视觉识别研究基础。

1 试验数据获取

棉花顶芽数据采集地于新疆生产建设兵团第八师 143 团棉花基地, 品种为新路早 50, 采集设备为索尼 A7M3 全高清摄像机, 为采集真实环境的复杂棉花顶芽数据, 不同光照和不同天气条件下均分为单株顶芽、多株顶芽、遮挡顶芽 3 种类别进行拍摄, 共采集到 1 360 张顶芽图像, 经筛选后得到 506 张清晰顶芽图像 (其中不同光照 156 张, 不同天气 150 张, 不同遮挡顶芽面积 200 张), 经数据增强 (平移、翻转、剪切、降噪、缩放比例、亮度增强) 扩充到 3 103 张图像, 随机抽取 2 663 张作为训练集, 240 张作为验证集, 200 张作为测试集, 图 1 为棉花顶芽样本数据。

2 YOLOv5s 检测算法改进

2.1 YOLOv5-CPP 模型

使用 YOLOv5s 目标检测算法对棉花顶芽数据集进行训练, 在复杂环境下棉花顶芽识别效果并不理想。随着网络层的加深, 小目标信息逐渐丢失, 同时由于 YOLOv5s 算法中原有 3 个检测头, 对多尺度目标特征信息提取效果不佳, 尤其对棉花顶芽小目标检测识别会出现错检漏检的问题。为了实现顶芽较好的检测效果, 以 YOLOv5s 目标检测算法为基础, 将主干网络得第 2 层引出第 4 个目标检测头, 因浅层含有特征信息更加丰富, 为避免随着网络加深信息丢失, 将浅层与深层提取到的

特征信息进行融合, 解决小目标特征信息随着网络结构加深信息丢失问题, 避免特征提取不充分。由于在改进的 YOLOv5s 结构中增加目标检测层会使模型的计算量加大, 因此在原有目标检测网络中, 需要设计一种既能解决模型参数量问题又能解决对小目标提取效果好的模块来代替 C3 模块。经试验分析, CPP-CBAM 注意力机制为最佳模块, 将 CPP-CBAM 加入到第 13 层得到棉花顶芽最佳目标检测模型, 图 2 为 YOLOv5-CPP 目标检测模型改进结构图。



图 1 样本数据集
Fig.1 Sample data set

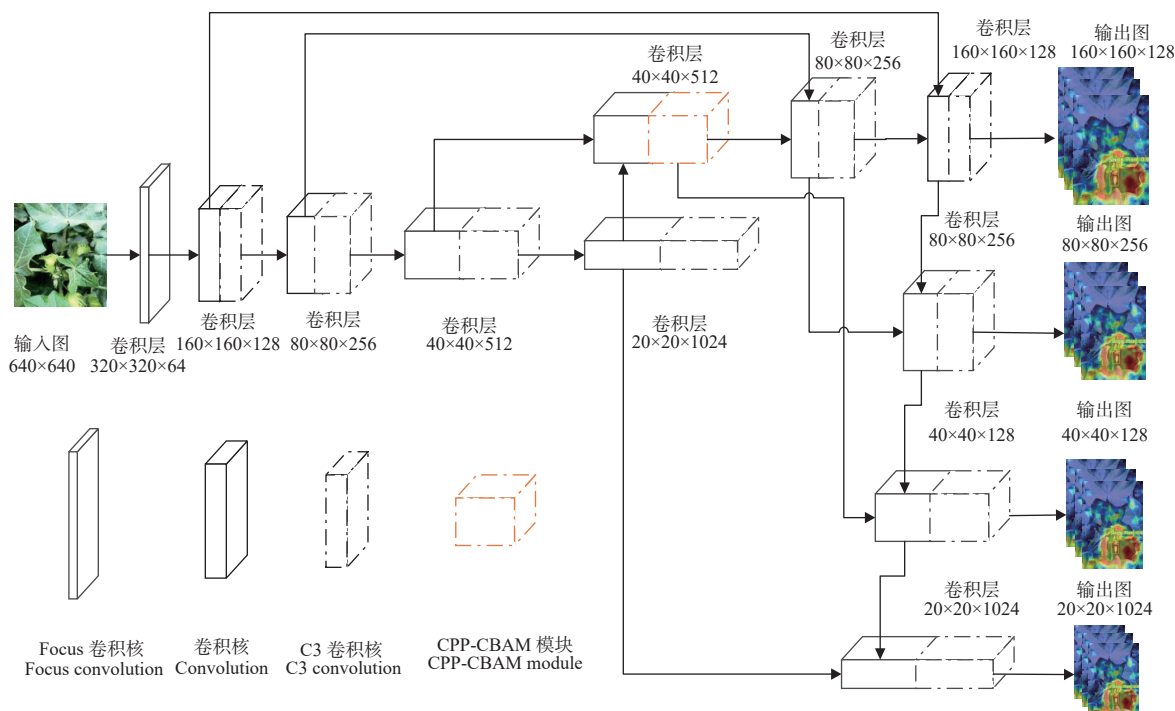


图 2 YOLOv5-CPP 模型结构图
Fig.2 The YOLOv5-CPP model structure figure

2.2 CPP-CBAM 注意力机制模块

CPP-CBAM 注意力机制模块如图 3 所示, C3 卷积核结构分为两个分支, 第一分支将特征图先升维后降维然后与第二分支网络进行特征融合, 特征融合后特征图再进入通道模块, 最大池化层 Maxpool 以及平均池化层 AvgPool 将特征图变成两个矩阵 $[C, 1, 1]$ 的权重向量, 两个拼接的权重向量经过 MLP 网络获取到通道的权重值, 增加原特征图的特征通道数, 通道数得到两个激活函数,

将两个输出结果的每个元素相加, 再由 Sigmoid 激活函数过滤掉冗余过多顶芽特征, 按权重重要程度先进行筛选特征图, 筛选出的特征图在再经过空间模块 Maxpool 最大池化以及平均池化层 AvgPool, 形成一个 $[1, H, W]$ 的权重向量, 由于 CBAM 模块多数经过 1×1 卷积核、最大池化层以及平均池化层操作, 能够压缩特征通道维度, 将复杂的模型特征通道数大大减少, 达到增加注意力机制后, 目标检测模型参数也减少的目的。

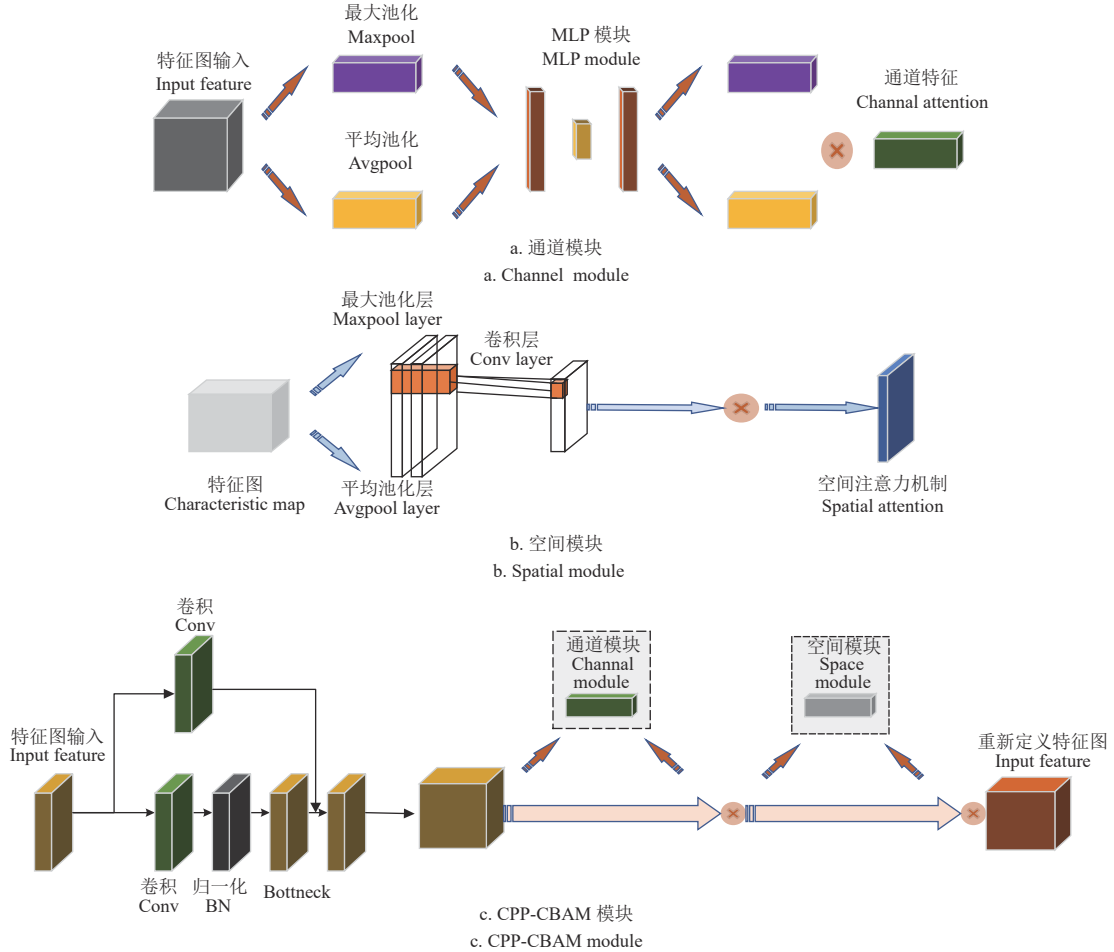


图 3 注意力机制图

Fig.3 Attention mechanism figure

2.3 分类 Focal Loss 损失函数

机械打顶作业过程中由于棉田环境复杂, 棉花顶芽被叶片严重遮挡, 给机器对顶芽识别带来困难, 为了使预测框回归到真实框速度加快, 因此考虑将 CIOU^[20] 损失函数替换为 SIOU^[21] 损失函数。GIOU^[22]、DIOU^[23]、CIOU 损失函数只考虑了预测框与真实框的宽高比, 并没有考虑到它们之间的方向, 易退化成 IOU 损失函数, 导致顶芽识别率精确度下降, 式 (1) SIOU 损失函数引入了真实框与预测框之间的向量角, 其包含式 (2) 角度损失函数式、式 (3) 距离损失函数式、式 (4) 形状损失函数。

$$L_{\text{loss}} = 1 - I_{\text{OU}} + \frac{\Delta + \Omega}{2} \quad (1)$$

$$\beta = \frac{\Pi}{2} - \alpha \quad (2)$$

$$\Delta = \sum_{t=x, y} (1 - e^{-\gamma \rho_t}) = 2 - e^{-\gamma \rho_x} - e^{-\gamma \rho_y} \quad (3)$$

$$\Omega = \sum_{t=w, h} (1 - e^{-wt})^\theta \quad (4)$$

式中 γ 为赋予时间的距离值, ρ_x 为真实框与预测框中心点和最小外接框宽比值的平方, ρ_y 为真实框与预测框中心点和最小外接框高比值的平方, 式 (3) θ 为控制对形状损失的关注程度, t 、 w 、 h 代表真实框与预测框的宽与高, 为预测框的宽与高, 接近 4 才能避免过度关注形状损失函数。

2.4 系统平台与评价指标

试验运行环境操作系统为 Windows 10 (64 位), 内存为 16 GB, 显卡驱动 RTX2060 SUPRER, 搭载 12 th Gen Intel(R) Core(TM) i5-9400F 2.90 GHz 处理器, 编程平台 Anaconda 4.12.0, CUDA 11.6, 开发环境 PyTorch,

编程使用 Python3.8, 通过精确率 (Precision, P), 召回率 (Recall, R)。

$$P = \frac{T_P}{T_P + F_P} \quad (5)$$

$$R = \frac{T_P}{T_P + F_N} \quad (6)$$

式中 T_P 表示正样本预测出来正确的框, F_P 表示负样本预测出正样本的框, F_N 表示从正样本预测出负样本的数量, $T_P + F_P$ 表示正样本预测出来正确的框加上负样本预测出正样本的框, 也就是表示预测框的个数。 $T_P + F_N$ 表示正样本预测出来正确的框加上正样本预测出负样本的数量也就是标注框的个数。

3 结果与分析

3.1 内部参数对模型的影响

为测试出 YOLOv5s 模型的最佳内部参数, 即不同批量大小对 YOLOv5s 的影响, 在训练时利用 github 官网提供的 YOLOv5.pt 预训练权重, 通过自制 YOLO 格式棉花顶芽数据集, 模型设置图像大小为 640×640 分辨率, 训练轮次设置为 150, 以及初始学习率为 0.01, 采用 Adam 优化器, 分别设置不同批量大小分别为 10、12、15、17、20、25 进行训练, 试验结果如表 1 所示。通过数据对比可以看出, 在相同迭代周期内, 批量 Batch size 设置为 20 时效果最好, 即使平均准确率 mAP 达到了 90.3%, 图片测试验证效果可看出, 未改进的目标检测模型仍有提升空间。

表 1 内部参数对模型的影响
Table 1 Effect of batch size on model

批量大小 Batch size	平均准确率 Mean average precision mAP%	精确率 Precision P/%	召回率 Recall R/%	置信度 Confidence/%
10	87.3	92.1	96.0	86.0
12	88.2	93.2	96.0	84.0
15	88.5	93.5	96.0	84.0
17	89.1	94.2	96.0	82.0
20	90.3	95.2	95.0	86.0
25	89.1	94.2	96.0	85.0

3.2 不同注意力机制对模型的影响

为测试不同注意力机制对模型的影响, 分别将不同注意力机制加入到 YOLOv5s 目标检测模型第 13 层, 将 C3 模块更换为 SE、BotNet^[24]、CotNet^[25]、C2f、CCA、CCP-CBAM 6 种注意力机制模块, YOLOv5-CPP 与其他 5 种模型对比结果如表 2 所示, YOLOv5-CPP 检测平均准确率 mAP 上升 2.5、1.6、1.7、0.8、1.4 个百分点, 而 YOLOv5-SE、YOLOv5-BotNet、YOLOv5-CotNet、YOLOv5-CCA 这 4 种模型检测单株顶芽的平均精度值分别降低了 1.1、0.1、0.3、0.4 个百分点, 说明在棉花顶芽数据集上 SE、BotNet、CotNet、CCA 注意力机制抑制了棉花顶芽的通道特征, 造成总体识别顶芽精度下降, 相反 YOLOv5-CPP 精度上升比例最大, 进一步说明了 CPP-CBAM 自注意力机制经过空间模块与通道模块, 对重要特征进行数值编码, 强化了各特征之间的关系, 降低了

对全局注意力计算的结果, 更加注重顶芽特征, 因此 CPP-CBAM 注意力机制在 3 种样本上效果均有所改善, 尤其是对遮挡顶芽这种数量多的样本改善效果更加明显, 因此得出 CPP-CBAM 注意力机制既能够解决在复杂情况下对小目标在复杂环境下的特征提取任务。

表 2 各个注意力机制模型对比
Table 2 Comparison of attention mechanism models

注意力机制 Attention mechanism	mAP%	P/%	R/%	检测速度 Detection speed/ (帧·s ⁻¹)	平均精度 Average accuracy AP/%		
					单株顶芽 Single plant bud	多株顶芽 Multiple plant bud	遮挡顶芽 Shelter plantbud
YOLOv5-SE	90.3	94.3	96.0	50.0	94.0	96.2	80.8
YOLOv5-BotNet	91.2	91.8	96.0	41.0	95.0	98.1	80.7
YOLOv5-CotNet	91.1	92.6	96.0	47.0	94.8	97.8	80.8
YOLOv5-C2f	92.0	92.2	96.0	50.0	95.3	97.5	83.2
YOLOv5-CCA	91.4	92.0	95.0	47.0	94.7	97.1	83.2
YOLOv5-CPP	92.8	92.2	96.0	43.0	95.1	98.4	85.1

3.3 注意力机制不同位置加入模型对比分析

为测试 CPP-CBAM 注意力机制对模型的影响, 本试验通过将注意力机制 CPP-CBAM 插入到不同网络层中, 测试出模型的检测速度与平均准确率, 试验效果见表 3, 第 13 层分别加入 CPP-CBAM 注意力机制之后, 较比插入到 1、3、5、9、20 层平均准确率 mAP 明显有所提升, 因为随着网络加深, 第 13 层 C3 模块对特征图处理效果较比 CPP-CBAM 注意力机制特征提取能力较弱, 浅层小目标信息较为丰富, C3 模型尚且能够对小目标信息特征进行提取, 但是随着网络加深小目标信息丢失, C3 模块逐渐对小目标的特征提取效果不理想, 因此可以推出第 13 层插入 CPP-CBAM 注意力机制效果最佳。

表 3 不同层加入注意力机制对模型的影响
Table 3 The effect of different layers adding attention mechanisms on the model

特征提取 层位置 Position in backbone	mAP%	P/%	R/%	检测速度 Detection speed/ (帧·s ⁻¹)	AP/%		
					单株顶芽 Single plant bud	多株顶芽 Multiple plant bud	遮挡顶芽 Shelter plant bud
Layer=1	90.8	95.8	94.9	43	93.2	95.2	80.2
Layer=3	90.6	95.2	95.1	43	95.1	97.1	80.6
Layer=5	91.2	96.3	95.6	43	94.7	97.7	80.7
Layer=9	90.4	95.3	94.1	43	94.4	97.3	82.0
Layer=13	90.1	95.2	94.1	43	95.1	97.1	82.9
Layer=20	91.2	95.1	94.7	43	95.4	97.2	83.0

3.4 与其他模型对比分析

与其他模型对比结果见表 4 所示,

表 4 各个检测模型对比
Table 4 Comparison of each loss function

模型 Model	模型参数量 Model parameter quantity/MB	mPA%	P/%	R/%	检测时间 Detection time/s	检测速度 Detection speed/(帧·s ⁻¹)
YOLOv3	61.0	89.5	94.7	96.0	0.04	24.0
YOLOv5s	7.20	90.8	95.8	96.0	0.03	37.0
YOLOv6s	37.8	90.4	95.1	96.0	0.03	35.0
Fast R-CNN	30.0	90.7	95.5	96.0	0.02	33.0
YOLOv5-CPP	3.70	92.8	92.2	96.0	0.01	36.0

为验证 YOLOv5-CPP 模型在棉花顶芽数据集上检测效果, 本文分将 YOLOv3^[26]、YOLOv5s、YOLOv6^[27]、Fast-RCNN 这 4 种网络模型在棉花顶芽数据集上进行训

练, 在相同迭代次数内, YOLOv3、YOLOv5s、YOLOv6、Fast-RCNN 这 4 种网络模型对数据集进行提取特征的效果并不理想, 进一步说明网络模型中模型对顶芽部分特征提取不到位, 导致检测精度不高。将 CPP-CBAM 注意力机制加入到 YOLOv5s 模型中, 并改进边界框回归损失函数, 对比 4 种网络模型在棉花顶芽的训练情况, 得出改进 YOLOv5s 的 mAP 值比 YOLOv3、YOLOv5s、YOLOv6s、Faster-RCNN-tiny 模型分别高出 3.3、2、2.4、2.1 个百分点, 同时模型参数量分别减少 57.3、3.5、34.1、26.3 MB, 但检测速度并未达到实时检测棉花顶芽的帧率。

3.5 Jetson nano 部署试验分析

为测试 YOLOv5-CPP 模型移植到试验样机实时检测

顶芽的效果, 将 YOLOv5-CPP 与 YOLOv5s、YOLOv5-CotNet、YOLOv5-BotNet 3 种模型进行对比分析, 4 种模型检测棉田同一区域被遮挡顶芽, 如图 4 所示。通过测试顶芽图可以发现, YOLOv5s、YOLOv5-BotNet、YOLOv5-CotNet 模型识别遮挡的顶芽分别漏检顶芽个数为 5、5、2, 同时 YOLOv5-BotNet 错检 1 个顶芽, YOLOv5-CPP 检测模型 (图 4d) 较比 YOLOv5-CotNet 检测模型 (图 4c) 检测遮挡的顶芽, 平均准确度 mAP 分别提升 15、29 个百分点, YOLOv5-CPP 模型鲁棒性进一步加强, 模型识别准确度大幅度提升, 同时在 TensorRT 加速情况下检测速度达到了 52 帧/s, 满足实时检测的效果, 在遮挡情况下的顶芽基础上识别率、模型的鲁棒性进一步提升, 验证了对遮挡情况的下顶芽有提升的效果。

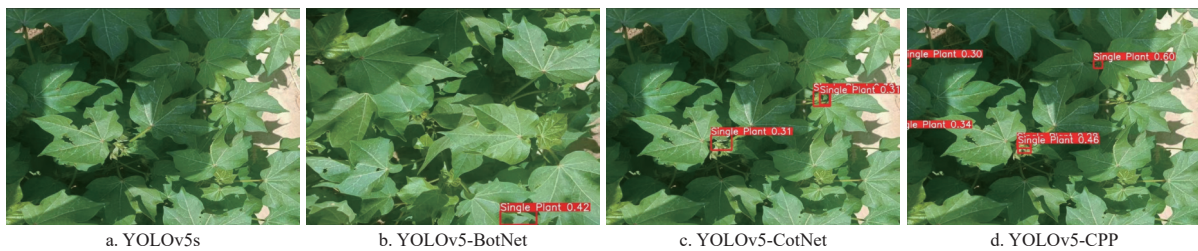


图 4 不同试验模型对比图

Fig.4 Comparison of different test models

4 结 论

本文在复杂环境识别遮挡下的棉花顶芽, 建立不同不同光照和不同天气条件下的棉花顶芽样本, 通过增加小目标检测层提高浅层与深层的特征融合, 提升检测精度通过增加小目标检测层, 并且添加注意力机制, 同时改进 Focss Loss 损失函数, 提升棉花顶芽样本识别率, 并通过棉花打顶平台验证不同检测模型对棉花顶芽识别, 得到以下几个结论:

1) 通过在模型特征提取层增加小目标检测层以及 YOLOv5-CPP 并与 YOLOv5-SE、YOLOv5-BotNet、YOLOv5-CotNet、YOLOv5-C2f、YOLOv5-CCA 5 种模型进行对比, 平均准确度 mAP 分别提高了 2.5、1.6、1.7、0.8、1.4 个百分点, 尤其在识别被遮挡顶芽平均精度 AP 分别提高了 4.3、4.4、4.3、1.9、1.9 个百分点, 充分验证在模型特征提取层增加 CPP-CBAM 注意力机制能够完成在复杂情况下的顶芽精确识别。

2) 本文通过对 YOLOv5 模型改进, YOLOv5-CPP 不均衡各类样本检测精度 mAP 达到了 92.8%, 比 YOLOv5s 模型提高了 2 个百分点, 与 YOLOv3、YOLOv6、Fast RCNN 相比, mAP 分别提高了 3.3、2.4、2.1 个百分点, 分析得出 YOLOv5-CPP 对复杂情况下样本检测优势明显。

YOLOv5-CPP 模型在复杂背景下识别精度、鲁棒性大幅度提升, 模型的检测速度虽然降低, 但 YOLOv5-CPP 模型通过 TensorRT 加速后经过实地验证后可以满足棉花打顶实时识别检测要求, 而且对于棉花打顶机而言, 在保证识别准确率的前提下, 提升速度指标, 是保障机

械作业效率的前提, 为模型在棉花打顶机械上的迁移部署提供技术支持。

[参 考 文 献]

- [1] 熊伟, 秋黎凤, 白雪. 棉花发展和纺纱进步的回顾与展望[J]. *纺织器材*, 2023, 50(1): 9-14.
XIONG Wei, QIU Lifeng, BAI Xue. Review and prospect of cotton development and spinning progress[J]. *Textile Equipment*, 2023, 50(1): 9-14. (in Chinese with English abstract)
- [2] 侯慧杰. 新疆棉花生育期气象灾害对棉花生长的影响及应对措施研究[J]. *农业灾害研究*, 2023, 13(1): 22-24.
HOU Huijie. Effects of meteorological disasters on cotton growth in Xinjiang and countermeasures[J]. *Agricultural Disaster Research*, 2019, 13(1): 22-24. (in Chinese with English abstract)
- [3] 徐爱武, 梁燕, 尚斌. 我国棉花专业仓储的现状问题及对策[J]. *中国棉花加工*, 2022(6): 14-16.
- [4] 马乾. “十四五”期间新疆棉花公证检验事业与棉花产业的发展趋势探讨[J]. *中国棉花加工*, 2022(6): 30-31
- [5] 陈常兵. 我国内地棉区棉花生产现状与发展对策[J]. *中国农技推广*, 2022, 38(10): 10-13.
- [6] GIRSHICK R. Fast r-cnn[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, 2015: 1440-1448
- [7] REDMON J, DIVVALA S, GIRSHICK R, You only look once: Unified, real-time object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV,

- USA, 2016: 779-788.
- [8] Redmon J, Farhadi A. YOLOv3: An incremental improvement[EB/OL]. (2018-04-08)[2022-11-19]<https://arxiv.org/abs/1804.02767>.
- [9] WANG C, BOCHKOVSKIY A, LIAO H Y M. Faster-RCNN: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[EB/OL]. (2022-07-06)[2022-11-19]<https://arxiv.org/abs/2207.02696>.
- [10] KOIRALA A, WALSH K B, WANG Z, et al. Deep learning for mango (*Mangifera indica*) panicle stage classification[J]. *Agronomy*, 2020, 10(1): 143.
- [11] 刘海涛, 韩鑫, 兰玉彬, 等. 基于 YOLOv4 网络的棉花顶芽精准识别方法[J]. 中国农业科技导报, 2022, 24(8): 99-108.
- LIU Haitao, Han Xin, LAN Yubin, et al. Precision identification of cotton terminal bud based on YOLOv4 network[J]. *Journal of Agricultural Science and Technology of China*, 2022, 24(8): 99-108. (in Chinese with English abstract)
- [12] JIA W, TIAN Y, LUO R. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot[J]. *Computers and Electronics in Agriculture*, 2020, 172: 105380.
- [13] 张振国, 吕全贵, 任杰宇, 等. 旋转剪切式红花花丝采摘机械关键部件的设计[J]. 中国农机化学报, 2019, 40(7): 1-6.
- ZHANG Zhenguo, Lü Quanguai, REN Jieyu. Design of critical components for safflower harvesting machinery by rotary shear[J]. *Journal of Chinese Agricultural Mechanization*, 2019, 40(7): 1-6. (in Chinese with English abstract)
- [14] 朱旭, 马湜, 姬江涛, 等. 基于 Faster R-CNN 的蓝莓冠层果实检测识别分析[J]. 南方农业学报, 2020, 51(6): 1493-1501.
- ZHU Xu, MA Hao, JI Jiangtao, et al. Detecting and identifying blueberry canopy fruits based on Faster R-CNN[J]. *Journal of Southern Agriculture*, 2020, 51(6): 1493-1501. (in Chinese with English abstract)
- [15] 李天华, 孙萌, 丁小明, 等. 基于 YOLOv4+HSV 的成熟期番茄识别方法[J]. 农业工程学报, 2021, 37(21): 183-190.
- LI Tianhua, SUN Meng, DING Xiaoming, et al. Tomato recognition method at the ripening stage based on YOLO v4 and HSV[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2021, 37(21): 183-190.
- [16] 杨坚, 钱振, 张燕军, 等. 采用改进 YOLOv4-tiny 的复杂环境下番茄实时识别[J]. 农业工程学报, 2022, 38(9): 215-221.
- YANG Jian, QIAN Zhen, ZHANG Yanjun, et al. Real-time recognition of tomatoes in complex environments based on improved YOLOv4-tiny[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(9): 215-221. (in Chinese with English abstract)
- [17] CHEN C, LU J, ZHOU M, A. YOLOv3-based computer vision system for identification of tea buds and the picking point[J]. *Computers and Electronics in Agriculture*, 2022, 198: 107116.
- [18] 薛月菊, 黄宁, 涂淑琴, 等. 未成熟芒果的改进 YOLOv2 识别方法[J]. 农业工程学报, 2018, 34(7): 173-179.
- XUE Yueju, HUANG Ning, TU Shuqin, et al. Improved YOLOv2 identification method of immature Mango[J]. *Transactions of the Chinese Society of Agricultural Engineering, (Transactions of the CSAE)* 2018, 34(7): 173-179. (in Chinese with English abstract)
- [19] 樊湘鹏, 许燕, 周建平, 等. 基于迁移学习和改进 CNN 的葡萄叶部病害检测系统[J]. 农业工程学报, 2021, 37(6): 151-159.
- FAN Xiangpeng, XU Yan, ZHOU Jianping, et al. Grape leaf disease detection system based on transfer learning and improved CNN[J]. *Transactions of the Chinese Society of Agricultural Engineering, (Transactions of the CSAE)* 2021, 37(6): 151-159. (in Chinese with English abstract)
- [20] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*, New York, USA, 2020: 12993-13000.
- [21] GEVORGYAN Z. Siou loss: More powerful learning for bounding box regression[C]//*Computer Vision and Pattern Recognition*. New Orleans, 2022: 1-12
- [22] HAMID R, NATHAN T, JUN Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Long Beach, 2019: 658-666.
- [23] ZHENG Z, WANG P, LIU W, et al. DistanceIoU loss: Faster and better learning for bounding box regression[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto, 2020: 1-7.
- [24] Srinivas A, Lin Tian Y, Parmar N, Bottleneck transformers for visual recognition[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville, TN, USA, 2021: 16519-16529.
- [25] LI Y, YAO T, PAN Y, et al. Contextual transformer networks for visual recognition[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New Orleans, IEEE, 2022: 1489-1500.
- [26] Redmon J, Farhadi A. YOLOv3: An incremental improvement[EB/OL]. (2018-04-08)[2020-07-07]. <https://arxiv.org/abs/1804.02767>.
- [27] LI Chuyi, LI Lulu, JIANG Hongliang, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. 2022[2023-05-06]. <https://arxiv.org/abs/2209.02976>

Cotton top bud recognition method based on YOLOv5-CPP in complex environment

PENG Xuan^{1,2}, ZHOU Jianping^{1,2*}, XU Yan^{1,2}, XI Guangze¹

(1. School of Mechanical Engineering, Xinjiang University, Urumqi 830017, China; 2. Engineering Research Center of Agricultural and Animal Husbandry Robots and Intelligent Equipment, Xinjiang Uygur Autonomous Region, Urumqi 830017, China)

Abstract: Cotton bud is one of the most important cash crops with many uses, such as the textile and cotton fabric raw material. However, manual cotton topping cannot fully meet the large-scale production in recent years, due to the efficiency and labor costs. Intelligent mechanical topping can be expected to serve as an inevitable trend during cotton topping. There is a high demand to accurately identify the cotton top bud under the complex environment (such as light and shadow), particularly for the various shape characteristics of cotton top core in the process of cotton cap removal. In this study the accurate recognition was proposed to locate the cotton top bud in the complex environment using an improved YOLOv5s. A dataset was also constructed to contain 3103 cotton top buds with different morphological characteristics under a complex environment. Three categories were divided: single, multi-plant and screened top bud under different occlusion areas, weather and light conditions. Systematic training and analysis were then performed on the data. An improved YOLOv5s model was proposed to reduce the error and leakage detection for the sufficient features of small targets. Multi-scale target features were also effectively extracted using only three detection heads in the original YOLOv5s. At the same time, the ideal extraction of the improved YOLOv5s was obtained with the deepening of the network layer. The object detection and CPP-CBAM attention mechanism were first added to the original structure, in order to improve the shallow and deep feature fusion rate. Secondly, the regression of prediction frame was enhanced to avoid the information loss of the real frame. The loss function SIOU boundary frame regression was added to further accelerate the prediction frame regression speed for the detection accuracy. The SIOU loss function was introduced by the vector angle between the real and predicted frame including angle, distance, and shape loss function. As such, the aspect ratio between the predicted and real frame was considered in the CIOU loss function with the direction between them in the SIOU. The high accuracy of recognition rate was obtained on the top bud, resulting in the less degrading into an IOU loss function. In addition, an accurate identification was performed on the cotton top bud that was seriously obscured by the leaves. The results show that the average accuracy (mAP) reached 92.8% for the recognition of cotton top bud by the improved model. Furthermore, the average accuracy was improved by 2.1, 3.3, 2, and 2.4 percentage points, respectively, compared with Fast R-CNN, YOLOv3, YOLOv5s, and YOLOv6. The improved detection model can be expected to accurately and efficiently identify the cotton top bud under a complex environment in real time. The small size of the model parameters and the high identification speed were suitable for the migration deployment of cotton topping machinery, and the operation of the laser topping robot. The finding can also provide technical and theoretical support to the subsequent operation of cotton laser topping machines, particularly for the mechanized real-time topping.

Keywords: crops; image recognition; small target detection; YOLOv5s; SIOU loss function; CPP-CBAM attention mechanism; cotton topping