

引用格式:

周岳淮, 李震, 左嘉明, 龚琬蓉, 吕石磊, 温威, 黄莺. 基于改进 YOLOv5s 模型的山地果园单轨运输机搭载柑橘的检测[J]. 湖南农业大学学报(自然科学版), 2023, 49(4): 491–496.

ZHOU Y H, LI Z, ZUO J M, GONG W R, LYU S L, WEN W, HUANG Y. Detection of citrus carried by mountainous orchard monorail transporter based on improved YOLOv5s[J]. Journal of Hunan Agricultural University(Natural Sciences), 2023, 49(4): 491–496.

投稿网址: <http://xb.hunau.edu.cn>



## 基于改进 YOLOv5s 模型的山地果园单轨运输机 搭载柑橘的检测

周岳淮<sup>1</sup>, 李震<sup>1,3</sup>, 左嘉明<sup>1</sup>, 龚琬蓉<sup>1</sup>, 吕石磊<sup>1,3</sup>, 温威<sup>3,4</sup>, 黄莺<sup>1,2\*</sup>

(1.华南农业大学电子工程学院(人工智能学院), 广东 广州 510642; 2.柳州铁道职业技术学院自动控制学院, 广西 柳州 545616; 3.国家现代农业(柑橘)产业技术体系机械化研究室, 广东 广州 510642; 4.华南农业大学工程基础教学与训练中心, 广东 广州 510642)

**摘要:** 由于山地果园运输机立地条件差, 实时作业信息的获取、反馈、集中化管理较为困难, 为了解 7SYDD-200 型山地果园单轨运输机搭载货物情况, 合理调度运输装备, 建立了基于改进的 YOLOv5s 模型的运输机搭载柑橘果筐的检测方法: 在果园自然光环境下使用 RGB 相机(HSK-200)采集运输机搭载柑橘果筐的图像数据; 建立和优化 YOLOv5s 模型, 部署至嵌入式设备, 实现对搭载过程中的“空果筐”“柑橘”“满果筐”状态的检测。在模型的颈部网络引入 CBAM 注意力机制, 加强模型提取语义信息的能力, 解决检测过程中出现的“双重标签”的问题, 使用批归一化(BN)层稀疏的尺度因子衡量各通道对模型的表征能力, 并对表征能力弱的通道进行剪枝压缩, 以克服基模型 YOLOv5s 检测速度慢的问题, 通过多尺度训练策略对模型进行微调, 提高模型检测准确率。试验结果表明: 改进 YOLOv5s 模型的检测方法在柑橘搭载数据集上平均精度均值(mAP)为 93.3%; 模型的浮点数运算量和大小分别为 9.9 GFLOPs 和 3.5 MB, 比 YOLOv5s 的提高 60.3%和 21.3%; 在嵌入式平台 Jetson Nano 部署, 其检测速度为 78 ms/帧。

**关键词:** 山地果园单轨运输机; 目标检测; 剪枝压缩; CBAM 注意力机制; 改进 YOLOv5s

中图分类号: S229

文献标志码: A

文章编号: 1007-1032(2023)04-0491-06

## Detection of citrus carried by mountainous orchard monorail transporter based on improved YOLOv5s

ZHOU Yuehuai<sup>1</sup>, LI Zhen<sup>1,3</sup>, ZUO Jiaming<sup>1</sup>, GONG Wanrong<sup>1</sup>, LYU Shilei<sup>1,3</sup>, WEN Wei<sup>3,4</sup>, HUANG Ying<sup>1,2\*</sup>

(1.College of Electronic Engineering(College of Artificial Intelligence), South China Agricultural University, Guangzhou, Guangdong 510642, China; 2Automatic Control School, Liuzhou Railway Vocational Technical College, Liuzhou, Guangxi 545616, China; 3.Mechanization Laboratory of National Modern Agriculture(Citrus) Industrial Technology System, Guangzhou, Guangdong 510642, China; 4.Engineering Fundamental Teaching and Training Center, South China Agricultural University, Guangzhou, Guangdong 510642, China)

**Abstract:** Due to the poor site conditions of mountainous orchard monorail transporter, it is difficult to obtain, feedback and centralized management of real-time operation information. In order to monitor the proceeding of deliveries by 7SYDD-200 mountainous orchard monorail transporter and reasonably dispatch transportation equipment, the detection method of citrus fruit basket carried by the transporter is established based on the improved YOLOv5s model. Images of the citrus fruit baskets carried by the transporter were collected by the RGB camera of HSK-200 under the natural light

收稿日期: 2022-05-16

修回日期: 2023-04-20

基金项目: 国家重点研发计划子课题(2020YFD1000107); 国家现代农业产业技术体系(CARS-26); 国家自然科学基金项目(31971797、61601189); 广东省省级乡村振兴战略专项(粤财农[2021] 37号)

作者简介: 周岳淮(1997—), 男, 广东深圳人, 硕士研究生, 主要从事图像处理研究, [scau\\_zyh@stu.scau.edu.cn](mailto:scau_zyh@stu.scau.edu.cn); \*通信作者, 黄莺, 教授, 主要从事智能检测与控制技术研究, [huangying800816@163.com](mailto:huangying800816@163.com)

environment of mountainous orchards. The YOLOv5s model was established and optimized, which was deployed into the embedded device to detect the states of “empty fruit basket”, “citrus” and “full fruit basket” during the loading process. convolutional block attention module(CBAM) is introduced into neck network of the model to strengthen the ability to extract semantic information and solve the problem of “double labels” in the detection process. The sparse scale factor of the batch normalization(BN) layer was used to measure the representation ability of each channel of the model. The channels with weak representation ability were pruned and compressed to overcome the problem of slow detection speed of the model based on YOLOv5s. The multi-scale training strategy is used to fine-tune the model to improve the detection accuracy. The test results show that the mean average precision of the improved detection method is 93.3% on the fruit dataset. The floating point operation and the size of the improved models were 9.9 G and 3.5 M, respectively, which were 60.3% and 21.3% higher than that of YOLOv5s. The detection speed of the improved model was 78 ms/img, when it was deployed into the Jetson Nano embedded platform.

**Keywords:** mountainous orchard monorail transporter; object detection; pruning and compressing; convolutional block attention module(CBAM); improved YOLOv5s

南方山地果园立地条件多数为丘陵和山地, 果园坡度大, 地形复杂, 较难形成完善的交通运输网络, 山地果园单轨运输机的推广应用提高了山地运输的经济效益, 促进了果业的可持续发展<sup>[1-3]</sup>。

山地果园单轨运输机主要运输水果、化肥、小型农机具等, 装载能力较强, 可靠性高, 且轨道铺设灵活, 但工作环境较为复杂, 实时作业信息的集中化管理难度较大, 若不按额定搭载量行驶, 容易造成安全事故, 影响运输效率<sup>[4-5]</sup>。及时了解和掌握运输机的搭载状况, 对提高运输装备的资源配置、提升作业安全性和作业效率具有重要意义。机器视觉技术在识别运送装备搭载情况时具有较好的应用前景。张毅等<sup>[6]</sup>利用 Harris 角点特征和特征点匹配方法识别码垛作业中的烟包, 在特定环境识别准确率高, 但识别效果受光照影响较大。HUSSIN 等<sup>[7]</sup>通过圆形 Hough 变换对挂果芒果进行目标检测, 在芒果重叠和遮蔽的情况下, 检测位置偏差较大, 准确率较低。传统视觉算法在自然光变化较大时检测的鲁棒性较低, 不适合光线变化较大的果园环境检测单轨运输机搭载状况, 而深度学习有望解决传统图像处理算法存在的效率低和表征能力不足等问题<sup>[8]</sup>。金秋等<sup>[9]</sup>通过微调两阶段目标检测网络 Faster-RCNN, 实现在仓储环境下对货物和叉车的识别, 识别准确率高, 受环境因素影响较小, 但检测速度较慢。王晨等<sup>[10]</sup>基于 YOLOv4 网络, 将其骨干网络更换为更轻量级的 MobileNetV3, 实现了仓储货物的目标实时检测。汤寓麟等<sup>[11]</sup>使用改进的 YOLOv5 算法实现了海底沉船目标的检测, 检测速度比 YOLOv3 提升将近 1 倍。说明单阶段目标检测网络(如 YOLO 等)比两阶段目标检测网络(如 Faster-RCNN)的检测速度更快, 更适合部署于嵌入式平台, 通过迁移学习和模型微调保证检测精度,

满足使用需求。

为了解山地果园单轨运输机搭载柑橘情况, 建立了一种果园环境中单轨运输机柑橘搭载情况人工智能分析方法: 采集运输机柑橘搭载图像后, 优化 YOLOv5s 目标模型, 在颈部网络引入 CBAM 注意力机制, 解决检测过程中出现“双重标签”的问题; 通过剪枝压缩减少模型的参数量和计算量, 提高检测速度, 并使用多尺度训练策略, 提高检测精度, 以实现柑橘搭载过程中“空果筐”“柑橘”“满果筐”的实时检测。

## 1 数据处理

根据 7SYDD-200 型山地果园单轨运输机常见的搭载 1 层和 2 层果筐和柑橘, 采集运载拖车中柑橘搭载情况图像。为满足自然光下柑橘搭载情况的检测需求, 在华南农业大学的山地果园单轨运输机测试平台采集数据。考虑到果园中树木遮挡会使光照发生较大变化, 采集时间涵盖了晴天、阴天和多云等多种天气情况。在搭载过程中, 果园运输机搭载的果品、农资会环绕山岭任意点装卸, 并且动力电池需能够为检测系统提供电源, 因此视觉传感器安装于运输机载货车斗前方, 位置高于车斗前方横杆 40 cm, 与车斗平面夹角为 45°, 从而得到完整的载货车斗图像。视觉传感器安装如图 1-a 所示。数据采集设备为 HSK-200(鸿视康), 分辨率 1920 像素 × 1080 像素。

2021 年 5 月至 6 月, 采集原始图像共 872 张, 共有 3 种类型, 分别为“空果筐”“满果筐”“柑橘”, 样本标注数分别为 1046、688、6720 个, 如图 1-b 所示。考虑到山地果园中树木遮挡导致光线变化较大, 在原始数据集分别加入高斯噪声和随机调整图像亮度进行数据增强, 增强后的数据集共 2616 张。

按照 8 : 2 的比例将数据集划分为训练集与验证集, 其中训练集为 2092 张, 验证集为 524 张。

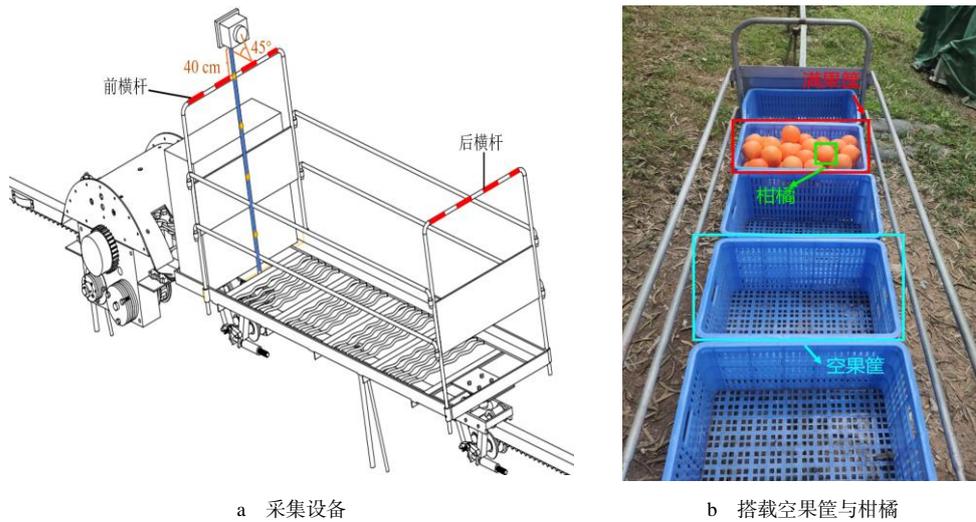


图 1 山地果园单轨运输机搭载柑橘及果筐图像的采集

Fig.1 Image acquisition of citrus and fruit basket carried by mountainous orchard monorail transporter

## 2 改进的 YOLOv5s 网络模型

### 2.1 基模型的选择

YOLOv5s 网络主要由骨干网络、颈部网络和预测网络构成<sup>[12]</sup>。骨干网络中的 Focus 模块能够避免输入图像下采样造成特征信息的丢失, 保留图像更丰富的特征。颈部网络由特征金字塔和路径聚合网络组成, 能将深层语义信息和骨干网络中浅层空间信息进行融合, 提高模型表达能力<sup>[13]</sup>。预测网络部分, YOLOv5s 能够输出大、中、小 3 种尺寸的特征

图, 更好地预测不同尺寸的目标。

### 2.2 算法改进

#### 2.2.1 CBAM 混合域注意力机制

应用 YOLOv5s 模型识别“空果筐”和“满果筐”时会出现“双重标签”的问题, 为提高识别准确率, 在 YOLOv5s 颈部网络的 C3 模块后引入 CBAM 注意力机制模块<sup>[14-15]</sup>, 如图 2 所示。

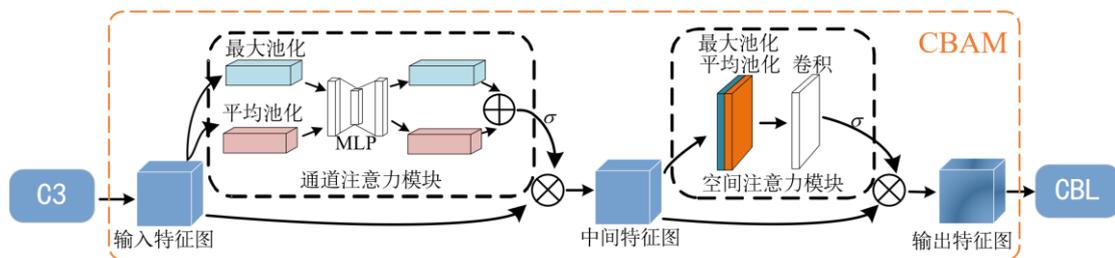


图 2 改进的 YOLOv5s 子结构

Fig.2 Substructure of improved YOLOv5s

首先对尺寸为  $C \times H \times W$  的输入特征图在空间维度上使用最大值池化和平均值池化, 变换为  $C \times 1 \times 1$  并经过 MLP、加和操作和 Sigmoid 激活, 得到具有通道注意力的特征图, 丰富其深层特征。再将具有通道注意力的特征图输入空间注意力模块在通道维度上进行特征提取, 关注目标的位置信息, 从而提高模型的位置和语义信息的能力。

#### 2.2.2 模型的剪枝压缩

融合 CBAM 模块的 YOLOv5s 在嵌入式平台 Jetson Nano 上检测速度较慢, 每帧耗时约为 184 ms。为提高模型推理速度, 减少网络冗余参数量, 基于 YOLOv5s 卷积层后普遍连接 BN 层的模型结构, 通过稀疏化 BN 层尺度因子, 将表征能力弱的通道进行剪枝, 剪枝流程如图 3 所示。

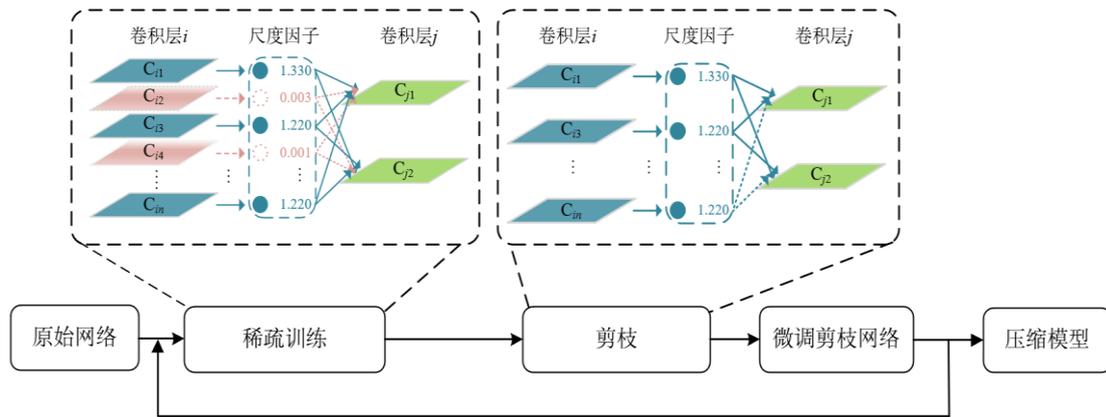


图 3 模型剪枝流程

Fig.3 The flowchart of model pruning

根据 BN 层变换公式(1)、公式(2),  $z_{in}$  与  $z_{out}$  为 BN 层的输入与输出,  $\mu_B$  和  $\sigma$  表示当前批量在激活输入的平均值和标准差值。由于尺度因子  $\gamma$  与上一层卷积层输出相乘, 从而能够作为权重用于衡量当前通道重要程度, 因此通过稀疏训练对  $\gamma$  进行稀疏化, 并将  $\gamma$  较小(如图 3 中  $C_{i2}$  和  $C_{i4}$ )的通道进行剪枝, 达到模型压缩的效果。

$$\hat{z} = \frac{z_{in} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \tag{1}$$

$$z_{out} = \gamma \hat{z} + \beta \tag{2}$$

选取稀疏率 0.000 5 对模型进行稀疏训练, 并使用剪枝率 0.6 对模型进行压缩。稀疏训练前 BN 层  $\gamma$  值分布如图 4-a 所示,  $\gamma$  值大小, 在 1 附近呈正态分布, 不利于剪枝压缩; 而稀疏训练后的  $\gamma$  值如图 4-b 所示, 稀疏化较为明显, 部分层的  $\gamma$  值已经趋近于 0, 即该通道对模型表征能力影响较弱, 可以被裁剪。选用稀疏率 0.000 5 和剪枝率 0.6 对模型进行剪枝, 能够达到模型轻量化的效果, 提高模型检测速度。

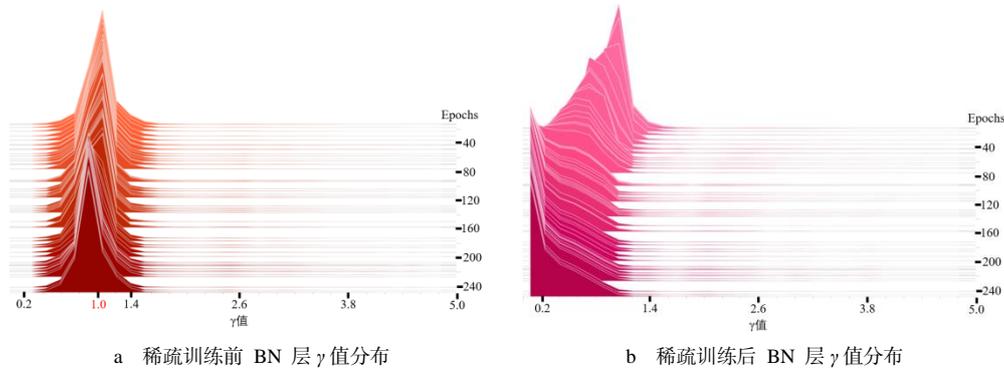


图 4 模型稀疏训练的评估

Fig.4 Evaluation of sparse training for the model

### 2.2.3 多尺度训练

为避免 YOLOv5s 网络使用单一尺寸的图片进行训练时容易出现识别目标漏检的情况<sup>[16]</sup>, 为提高模型的鲁棒性, 采取多尺度训练策略。改进的 YOLOv5s 网络最大下采样倍率为 32 倍, 且当图像输入尺寸为 640 像素×640 像素时检测精度较高, 但检测速度较慢; 当输入尺寸为 480 像素×480 像素时, 检测速度能满足需求。为达到检测精度和速度的要求, 选取{640、576、544、512、480}等 5 个尺度作

为多尺度训练输入图像尺寸大小, 使模型学习不同尺度输入的图片特征, 提高模型的检测精度, 且每隔 10 个轮次后改变图片大小, 再对模型进行训练。

## 3 试验环境与评价指标

### 3.1 试验设备和参数设置

模型训练的服务器配置为 Intel(R) Xeon(R) Gold 6142 CPU 处理器, 内存 29 GB, 显卡型号为 NVIDIA GeForce RTX 3080 10 GB。在 Ubuntu18.04 系统和 Pytorch 框架下实现 YOLO 算法。试验平台硬件配置

为 Intel(R) Core i5-11400 2.60 GHz 处理器, 内存 16 GB, 显卡为 GeForce RTX2060 6GB。嵌入式平台选用 Jetson Nano B01, 运行内存为 4 GB。

网络训练时初始学习率为 0.01, 初始动量为 0.937, 动量系数为 0.000 5, 采用 Adam 算法进行优化, 训练 250 个轮次。为提高单 GPU 训练效率, 使用 Mosaic 数据增强。

### 3.2 模型评价指标的选取

选择平均精度(AP)和平均精度均值(mAP)<sup>[17]</sup>作为准确度指标; 其他指标包括每帧耗时(ms)、浮点运算数 FLOPs 和模型大小(MB)。

### 3.3 消融对比试验

为验证改进的子结构和训练策略对模型的影响, 对改进机制和策略进行消融对比试验, 其中 YOLOv5sA 为带有 CBAM 注意力的 YOLOv5s 模型, YOLOv5sAP 为引入注意力和模型剪枝的 YOLOv5s 模型, 结果如表 1 所示。在基模型 YOLOv5s 加入注意力机制, 模型检测“满果筐”的平均精度为 96.9%, 比基模型提高了 1.3%, 说明注意力机制能够提高模型提取语义信息的能力; 对模型加入注意力机制和剪枝压缩, 并使用多尺度训练策略, 模型检测平均精度均值为 93.3%, 在 Jetson 上的检测速度为 78 ms/帧, 检测精度和检测速度相比于基模型都有所提升, 检测效果较好。

表 1 模型消融试验的性能指标

Table 1 Model performance in ablation experiments

模型	注意力	剪枝	多尺度训练	平均精度/%			mAP/%	模型大小/ MB	每帧耗时/ ms
				空果筐	柑橘	满果筐			
基模型				97.4	84.2	95.6	92.4	14.4	145
YOLOv5sA	√			97.5	85.1	96.9	93.2	14.6	184
YOLOv5sAP	√	√		97.3	84.6	96.4	92.8	3.5	78
改进 YOLOv5s	√	√	√	97.7	85.6	96.5	93.3	3.5	78

### 3.4 不同模型的检测效果

为对比与常见目标检测模型的效果, 将检测结果与典型的单阶段检测网络 RetinaNet-Res50 和 YOLOv3-tiny 进行对比, 且均以图像尺寸为 480 像素×480 像素输入网络进行训练, 采用算法默认超参数进行训练。

RetinaNet-Res50、YOLOv3-tiny 和改进的 YOLOv5s 的性能指标如表 2 所示。在自定义数据集上, 改进的 YOLOv5s 的 mAP 最高(93.3%), 比

RetinaNet-Res50 高 10.2%。此外, RetinaNet-Res50 模型较大, 不适用于嵌入式设备。在模型大小相近的情况下, YOLOv3-tiny 比本模型的 mAP 低 12.4%, 2 个模型检测柑橘的平均精度分别为 53.1% 和 85.6%, 检测精度有明显差异。这是由于 YOLOv3-tiny 模型输出只有 2 种不同尺度的输出特征图, 对存在粘连且较小的目标检测效果较差, 易出现柑橘漏检的情况, 因此在检测速度和精度上有更好的性能。

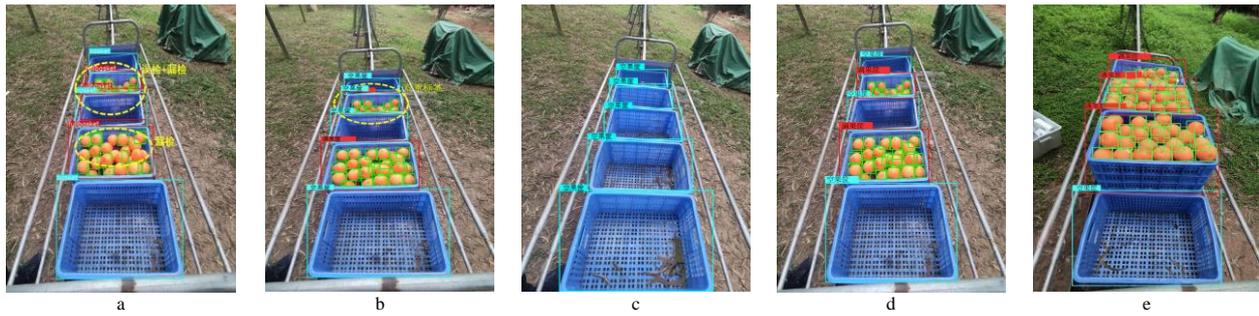
表 2 不同模型的检测性能对比

Table 2 Comparison of the detection performance for the different models

模型	AP/%			mAP/%	浮点运算量/ GFLOPs	模型大小 /MB
	空果筐	柑橘	满果筐			
RetinaNet-Res50	91.5	73.5	84.2	83.1	156.0	145.8
YOLOv3-tiny	95.5	53.1	94.2	80.9	12.9	17.4
改进 YOLOv5s	97.7	85.6	96.5	93.3	9.9	3.5

YOLOv3-tiny 和 YOLOv5s 的检测结果如图 5 所示。仅拥有 2 种尺寸特征图输出的 YOLOv3-tiny 模型在检测“柑橘”时, 会出现较严重的漏检, 与表 2 中的结果相符, 不能很好地检测存在粘连且尺寸较小的目标; 用 YOLOv3-tiny 和 YOLOv5s 的检测时均出现了“多重标签”的情况, 即同时标有“空果筐”

和“满果筐”的标签, 这是由于模型不能准确地提取图像的语义信息, 模型存在特征表达能力不足的问题。改进后模型的检测效果如图 5-c 所示。以运输机典型的搭载情况进行检测分析, 当模型检测搭载 1 层“空果筐”、混载 1 层“空果筐”和“柑橘”以及搭载 2 层“空果筐”和“柑橘”时, 不会出现漏检和误检。



a YOLOv3-tiny; b YOLOv5s; c 搭载1层果筐; d 搭载1层果筐和柑橘; e 搭载2层果筐和柑橘。

图5 改进前后模型的检测效果

Fig.5 Detection results of the model before and after improvement

## 4 结论

以果园山地单轨运输机柑橘搭载为研究场景,采集果园自然光环境下运输机搭载柑橘图像数据集;通过改进 YOLOv5s 模型,在颈部网络引入 CBAM 模块和使用多尺度训练策略,提升模型检测精度;使用通道级模型剪枝,实现模型轻量化,实现了对“空果筐”“满果筐”“柑橘”的目标检测。结果表明,所提出的检测方法检测精度达 93.3%,在 Jetson Nano 上检测速度为 78 ms/帧,适合部署于计算资源有限的嵌入式设备。后续将探索深度学习模型高效压缩的方法,在保障检测精度的前提下进一步提高模型的检测速度,并将该检测模型辅助三维数字建模,用于水果搭载情况分析。

### 参考文献:

- [1] LIU Y, HONG T S, LI Z. Influence of toothed rail parameters on impact vibration meshing of mountainous self-propelled electric monorail transporter[J]. *Sensors*, 2020, 20(20): 5880
- [2] 盛玲玲, 宋淑然, 洪添胜, 等. 广东省山地果园机械化现状与发展思考[J]. *农机化研究*, 2017, 39(11): 257-262.
- [3] 李震, 洪添胜, 吕石磊, 等. 山地果园自走式电动单轨运输装备的研究进展[J]. *现代农业装备*, 2020, 41(4): 2-9.
- [4] 吕石磊, 魏志威, 吴奔雷, 等. 果园单轨运输机在轨状态感知系统研制[J]. *农业工程学报*, 2020, 36(15): 56-64.
- [5] 吕石磊, 梁尹聪, 李震, 等. 基于超高频RFID双天线双标签对照的果园单轨运输机定位[J]. *农业工程学报*, 2018, 34(4): 71-79.
- [6] 张毅, 王彦博, 付华森, 等. 基于机器视觉的不规则烟包校对码垛系统[J]. *烟草科技*, 2019, 52(6): 105-111.
- [7] HUSSIN R, JUHARI M R, KANG N W, et al. Digital image processing techniques for object detection from complex background image[J]. *Procedia Engineering*, 2012, 41: 340-344.
- [8] 毕松, 高峰, 陈俊文, 等. 基于深度卷积神经网络的柑橘目标识别方法[J]. *农业机械学报*, 2019, 50(5): 181-186.
- [9] 金秋, 李天剑. 仓储环境下基于深度学习的物体识别方法研究[J]. *北京信息科技大学学报(自然科学版)*, 2018, 33(1): 60-65.
- [10] 王晨, 袁庆霓, 白欢, 等. 面向仓储货物的轻量化目标检测算法[J]. *激光与光电子学进展*, 2022, 59(24): 74-80.
- [11] 汤寓麟, 边少锋, 翟国君, 等. 侧扫声纳检测沉船目标的改进 YOLOv5 法[J/OL]. *武汉大学学报(信息科学版)*: 1-11(2021-09-01). <https://doi.org/10.13203/j.whugis20210353>.
- [12] BOCHKOVSKIY A, WANG C, LIAO H. YOLOv4: Optimal Speed and Accuracy of Object Detection[OL]. arXiv preprint arXiv: <https://arxiv.org/abs/2004.10934>, 2020.
- [13] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017: 6517-6525.
- [14] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision. Munich, Germany: IEEE, 2018: 3-19.
- [15] 郝帅, 马瑞泽, 赵新生, 等. 基于卷积块注意模型的 YOLOv3 输电线路故障检测方法[J]. *电网技术*, 2021, 45(8): 2979-2987.
- [16] 张锦, 屈佩琪, 孙程, 等. 基于改进 YOLOv5 的安全帽佩戴检测算法[J]. *计算机应用*, 2022, 42(4): 1292-1300.
- [17] 胡嘉沛, 李震, 黄河清, 等. 采用改进 YOLOv4-Tiny 模型的柑橘木虱识别[J]. *农业工程学报*, 2021, 37(17): 197-203.

责任编辑: 罗慧敏  
英文编辑: 吴志立