

基于改进 YOLO 12s 的甜瓜雌雄花识别与其边缘设备部署测试

张佐经¹ 王 晶¹ 刘明星¹ 刘幸韦¹ 杨昊东¹ 张 昭² 宋怀波¹

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100; 2. 宝鸡文理学院电子电气工程学院, 宝鸡 721016)

摘要: 授粉是甜瓜种植的关键环节, 机械授粉因其高效性成为重要的发展方向, 雌雄花朵精准识别是机械授粉作业的重要保障。本研究提出了一种基于改进的 YOLO 12s 目标检测网络的甜瓜雌雄花朵识别方法并进行了嵌入式系统平台部署。通过将原模型的 Backbone 部分替换为 ShuffleNetV2 结构降低模型复杂度与计算量, 以提升检测速度, 使其适于边缘设备上的部署, 并通过知识蒸馏进一步优化检测算法性能。对温室中拍摄得到的甜瓜雌雄花朵进行标注, 利用改进的 YOLO 12s 网络进行模型训练, 并在相同试验条件下, 与当前主流算法进行对比, 分别用精确率、召回率、平均精度均值(mean Average Precision, mAP)、模型内存占用量、检测速度 5 个指标对检测模型进行了性能评估。经测试, 模型精确率为 79.93%, 召回率为 87.65%, mAP 为 86.54%, 模型内存占用量为 4.10 MB, 检测速度为 153.46 f/s, 与 YOLO 12s、YOLO v10s、YOLO v8s、YOLO v5s、SSD 和 Faster R-CNN 模型相比, 其检测速度分别提升 79.32%、75.11%、45.14%、38.23%、746.44%、1 168.26%, 且在光照变化、部分遮挡及图像模糊等复杂场景下仍能实现精准检测。迁移部署于 Jetson Xavier NX 嵌入式平台, 模型 mAP 为 86.62%, 检测速度为 26.37 f/s, 其中, 检测速度分别比 YOLO 12s 和 YOLO v5s 高 90.40%、15.51%, 可满足实际授粉作业对精确度和实时性的需求。结果表明, 经轻量化处理后的模型可准确、快速地实现甜瓜雌雄花识别, 且在精度损失较小的前提下显著减小模型内存占用量, 有利于网络在边缘设备上的部署, 该研究结果可为甜瓜机械授粉提供技术支持。

关键词: 甜瓜花朵; 目标检测; YOLO 12s; 轻量化模型; Jetson Xavier NX

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2026)10-0219-11

OSID:



Performance Evaluation of Muskmelon Flower Gender Identification Based on Improved YOLO 12s and Its Deployment on Edge Devices

ZHANG Zuoqing¹ WANG Jing¹ LIU Mingxing¹ LIU Xingwei¹ YANG Haodong¹

ZHANG Zhao² SONG Huaibo¹

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China)

2. College of Electronic and Electrical Engineering, Baoji University of Arts and Sciences, Baoji 721016, China)

Abstract: Pollination is a critical stage in muskmelon cultivation, and mechanical pollination has emerged as a significant development direction due to its efficiency benefits. Accurate identification of female and male flowers is essential for the operation of pollination robots. For identifying melon female and male flowers, an improved YOLO 12s target detection network was applied, and the systematic deployment was implemented on edge device. By replacing the Backbone of the original model with the ShuffleNetV2 structure, the model's complexity and computational load were reduced, and the purpose of improving detection speed was achieved. Additionally, the knowledge distillation was also employed to further improve the performance of the improved model. A total of 3 597 muskmelon images were annotated in the greenhouse environment and then sent to the improved YOLO 12s object detection network for model training. Under identical experimental conditions, a comprehensive comparison was conducted with current mainstream algorithms. The melon flower detection model was evaluated across

收稿日期: 2025-10-19 修回日期: 2025-12-19

基金项目: 国家自然科学基金项目(32472964)、陕西省重点研发计划项目(2024NC-YBXM)和大学生创新创业训练计划项目(S202510712657)

作者简介: 张佐经(1982—),男,讲师,主要从事智能化检测技术研究,E-mail: zhangzuoqing@126.com

通信作者: 宋怀波(1980—),男,教授,博士生导师,主要从事图像处理研究,E-mail: songhuaibo@nwsuaf.edu.cn

five metrics: precision, recall, mean average precision, model size, and detection speed. Test results showed that the model achieved a precision of 79.93%, recall of 87.65%, mAP of 86.54%, a model size of 4.10 MB, and a detection speed of 153.46 f/s. Compared with YOLO 12s, YOLO v10s, YOLO v8s, YOLO v5s, SSD, and Faster R-CNN models, the detection speed was improved by 79.32%, 75.11%, 45.14%, 38.23%, 746.44% and 1 168.26%, respectively. Furthermore, the model maintained accurate detection performance under complex conditions such as varying lighting, partial occlusion, and image blur. When deployed on the Jetson Xavier NX embedded platform, this model achieved a mAP of 86.62%, the detection speed was 26.37 f/s. Compared with YOLO 12s and YOLO v5s, the detection speed was 90.40% and 15.51% faster, respectively. This performance was sufficient to meet the practical requirements for both accuracy and real-time processing in pollination operations. The results demonstrated that the lightweight model can accurately and efficiently identify male and female melon flowers. With only a minor loss in accuracy, it significantly reduced memory consumption, facilitating its deployment on edge devices. The research result can provide technical support for mechanical pollination of melons.

Key words: muskmelon flowers; object detection; YOLO 12s; lightweight model; Jetson Xavier NX

0 引言

甜瓜是典型的雌雄同株、异花授粉作物,温室大棚种植模式是提升其价值的重要手段^[1]。为提高温室的空间利用率,甜瓜多采用吊蔓种植,花朵分布密集,大量簇生花朵、叶片以及花朵层叠遮挡现象普遍,花朵检测难度较大,是花朵机械授粉效率不高的主要原因之一。开展设施场景下甜瓜雌雄花朵的高效识别,对提升授粉机器人的作业效率具有重要意义。

目前常用的设施甜瓜授粉方式主要包括人工授粉、化学辅助授粉和自然授粉 3 种^[2]。人工授粉作为目前最普遍的授粉方式,费工费时且授粉时机难以把握;化学辅助授粉易影响果实品质;自然授粉虽省时省工,但大棚高温易造成昆虫死亡、甜瓜植株早衰及裂瓜率增加,导致授粉率降低^[1]。

人工智能等技术大幅推动了设施甜瓜授粉智能化发展,机器视觉与图像处理技术可实现甜瓜花朵的实时检测与授粉,然而设施甜瓜生长环境复杂,花朵重叠、遮挡、姿态变化及光线条件复杂等问题都会影响花朵识别与检测。赵春江等^[3]提出了一种基于级联卷积神经网络(Convolutional neural networks, CNN)的番茄花期识别检测方法,将此方法部署到设施番茄授粉机器人上,识别准确率为 76.67%。GAO 等^[4]开发了一种用于果园中猕猴桃花精确授粉机器人,田间试验平均授粉成功率为 99.3%,平均果实坐果率为 88.5%。

花朵目标的准确检测是授粉机器人的核心技术之一。传统手工特征匹配方法识别精度有限、效率较低^[5]。随着人工神经网络技术的发展,深度学习方法已成为主流目标检测方法。目标检测方面,如 Faster R-CNN、YOLO 系列,能够自动提取花朵特征

并进行精准定位,被广泛用于花朵检测研究^[6-9];图像分割方面,研究者采用 K-means 聚类和峰值聚类等方法,实现花朵分类识别与花簇分簇,识别成功率与聚类指标均表现良好^[10-11]。

农业场景对目标检测实时性要求高且硬件资源有限,深层卷积神经网络模型因参数多、依赖高性能服务器、时效性不足难以适配,模型轻量化成关键趋势^[12-14]。为此研究人员提出了多种轻量化改进策略,包括优化网络结构、引入注意力机制及设计轻量级骨干网络等^[15-18]。

基于 CNN 的花朵检测虽取得了较好效果,但仍存在以下问题:①检测过程耗时较长,难以满足实时性需求。②已有检测模型针对设施内的甜瓜花朵检测尚不完善。③现有花朵检测模型大多存在参数量庞大、结构复杂的问题,边缘设备部署后的性能有待验证。基于以上问题,本文提出一种基于改进 YOLO 12s 的甜瓜雌雄花朵检测方法,将温室室内拍摄得到的甜瓜花朵进行标注并送入目标检测网络,在快速检测甜瓜花朵的同时进行雌雄花分类,并通过对模型进行边缘端部署测试与验证,以期实现一种轻量化且精度更高的花朵检测模型,为授粉机械视觉系统提供技术支持。

1 材料与方法

1.1 材料

由于当前公开数据集中包含设施环境下甜瓜花朵的数据信息较少,因此所使用甜瓜花朵图像为自建数据集。主要来源于陕西省西安市阎良区富秦果蔬专业合作社,甜瓜种植方式为温室种植,种植高度为 2 m,拍摄设备为一加 11(OnePlus 11)智能手机,相机焦距为 22 mm,分辨率为 4 096 像素×3 072 像素,拍摄时间为 2025 年 1 月 3 日,时间段为 11:00—

15:00。如图 1 所示,拍摄方式为手持设备形式,设备距离植株 20 ~ 40 cm,共采集图像 3 410 幅。

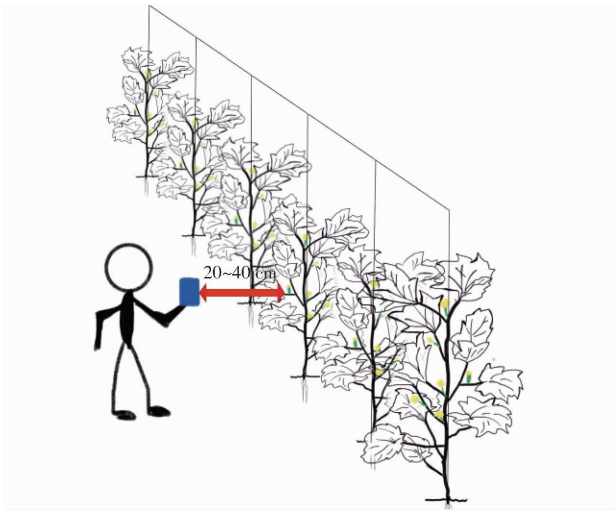


图 1 图像采集过程示意图

Fig. 1 Schematic of acquisition process of images

除此之外,本文还增加来自中国植物图像库的甜瓜数据集 187 幅,将获取到的 3 597 幅图像利用 Labellmg 标注工具进行人工标注,标注对象为目标授粉植株,标注框选择花朵最小外接矩形,如图 2 所示,将待检甜瓜花朵分为 2 类,包括雌性花朵 (female_flower)、雄性花朵 (male_flower)。

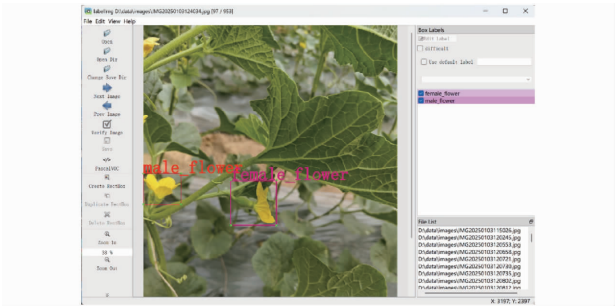


图 2 标注过程图

Fig. 2 Annotation process diagram

依据花朵下方有无膨大子房判别其花性,如图 3a、3b 所示。在目标检测任务中,花朵姿态朝向的多样性显著增加了识别难度。由于甜瓜花朵藤蔓攀援生长特性,部分花朵侧面朝向镜头,暴露其雌花子房或雄花花药,易于区分;部分正面朝向镜头,遮挡了辨别花朵雌雄的关键特征(雌花子房或雄花花药),如图 3c 红圈所示。仅依据图像中花瓣与花蕊结构的细微差异难以准确识别花朵的雌雄性,需要专业人员帮助,且误判率高。本研究标注策略基于甜瓜的实际生产需求,在“单蔓单果”的生产模式下,核心授粉任务在于识别出任意雌花,且雌花数量充足允许模型存在少量漏检。基于此,针对这类难以区分雌雄的花朵,均标注为无需授粉雄花,以提高授粉准确率。

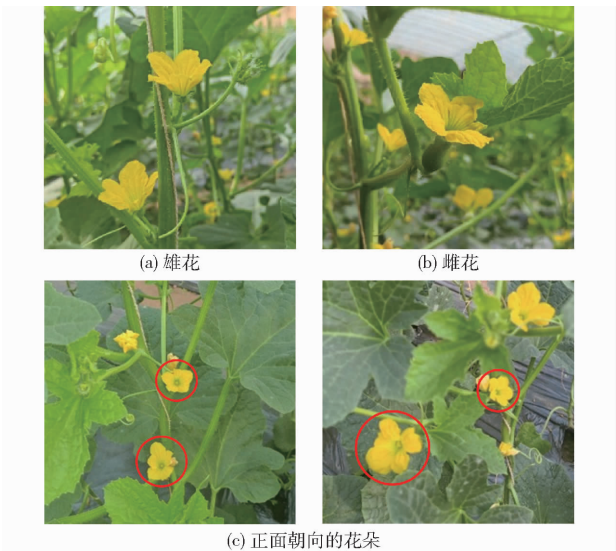


图 3 花朵数据采集示意图

Fig. 3 Flower data acquisition diagrams

由于拍摄位置和角度差异,甜瓜花朵光照强度不同,导致甜瓜花朵图像在整体亮度以及部分花朵局部亮度不同,如图 4 所示,为探究光照对花朵检测的影响,同时采集了甜瓜花朵在顺光和逆光情况下图像。

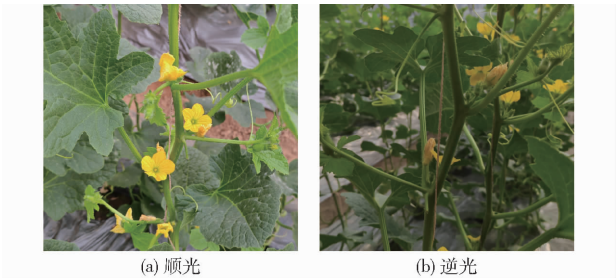


图 4 不同光照情况示例

Fig. 4 Examples of different lighting conditions

由于甜瓜花朵数量较多且多生长于叶腋处,其叶片宽大,藤蔓密集,导致花朵存在遮挡重叠现象,如图 5 所示,同时采集无遮挡、花朵遮挡、叶片遮挡和藤蔓遮挡 4 种情况花朵图像。

由于甜瓜图像采集设备为智能手机,拍摄图像时存在一定程度虚化现象,且聚焦点各不相同,导致图像中目标植株清晰程度不一,如图 6 所示,图像清晰程度会干扰花朵细节特征的提取,从而影响检测效果。基于此,将采集到的甜瓜花朵图像划分为清晰和模糊 2 类,探究网络针对不同清晰程度花朵检测效果。

顺光和逆光条件下图像数量分别为 2 395、1 202 幅,有遮挡和无遮挡情况下图像数量分别为 1 197、2 400 幅,清晰和模糊情况下图像数量分别为 2 229、1 368 幅。如表 1 所示,将采集到的 3 597 幅图像按照比例 6:2:2 分为训练集(2 157 幅)、验证集(720 幅)和测试集(720 幅)进行模型训练与测试。



图 5 不同遮挡情况示例
Fig. 5 Examples of various occlusion conditions

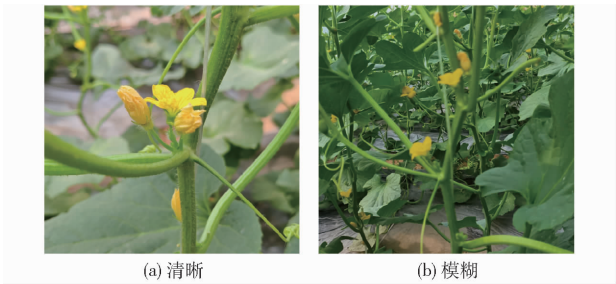


图 6 清晰程度
Fig. 6 Clarity level

表 1 拍摄图像详细信息

Tab.1 Detailed information of captured images		
影响因素	拍摄条件	图像数量
光照情况	顺光	2 395
	逆光	1 202
遮挡情况	有遮挡	1 197
	无遮挡	2 400
清晰程度	清晰	2 229
	模糊	1 368

1.2 试验平台

试验在 Windows 11 操作系统进行,处理器型号为 AMD Ryzen 5 6600H with Radeon Graphics,显卡为 NVIDIA GeForce RTX 3050。深度学习框架采用 PyTorch 2.5.1,编程平台为 PyCharm,编程语言为 Python 3.11.11,在训练过程中,输入图像分辨率为 640 像素×640 像素。模型参数设置如下:迭代次数(Epoch)为 300,批量大小(Batch size)为 16,初始学习率为 0.01,使用 SGD (Stochastic gradient descent) 优化器。所有算法均在相同环境下进行测试。

1.3 方法

1.3.1 改进 YOLO 12s 模型

YOLO 作为一种单阶段实时目标检测网络,具

备简洁的网络结构,具有速度快、精度高等优点,能够在复杂环境下完成目标检测任务,已广泛应用于诸多领域,随着 YOLO 版本的持续迭代更新,其性能和灵活性也不断提升^[19]。

相较于其他 YOLO 版本,YOLO 12 目标检测算法从根本上重构了以传统 CNN 为主导的架构,以注意力机制为核心,克服了传统注意力机制在实时性和计算效率方面不足等问题,能够实现快速推理速度和更高检测精度的实时目标检测^[20]。本研究选用性能更优,推理速度更快,计算复杂度更低的 YOLO 12s 作为基本网络框架,以确保模型具有较少参数量和较快检测速度,为边缘设备高效运行目标检测功能提供有力支持。

YOLO 12s 的 Backbone 首次采用 A2C2f 结构,该结构以残差高效层聚合网络(R-ELAN)为核心,优化了特征聚合能力,并沿用了 YOLO 11 的 C3k2 模块用于辅助特征提取;YOLO 12s Neck 部分用于聚合和细化多尺度特征,并通过上采样和拼接操作,整合不同层次的特征信息,增强特征的表达,同时引用了 FlashAttention 加速的区域注意力机制,克服注意力内存访问问题;Head 部分简化了多尺度检测路径,由多个检测层组成,负责最终的目标检测任务,输出检测目标的类别和位置信息^[21]。

为保证原有模型在目标检测中的性能,并在保持原有检测精度的同时减少网络模型参数,以 YOLO 12s 框架为基础进行结构改进。将原有模型中 Backbone 部分替换为 ShuffleNetV2 结构,用于降低网络部署和训练过程中的参数量,实现网络轻量化设计。此外,通过知识蒸馏方法,弥补轻量化带来的精度损失,以提升模型性能。将改进的网络结构命名为 YOLO 12s - ShuffleNetV2 - D,模型结构如图 7(YOLO 12s - ShuffleNetV2 - D 网络分为 Backbone、Neck、Head 3 个部分,每个框表示 1 个部分。ShuffleNetV2 为轻量级网络架构,返回 4 个不同尺度的网络特征图,Conv 为卷积操作,Upsample 为上采样操作,Concat 为特征融合函数,A2C2f 模块用于融合不同层次的特征,Detect 为检测头,Split 为分割操作)所示。

1.3.2 ShuffleNetV2 轻量级主干网络

ShuffleNetV2^[22]是一种高效的轻量级卷积神经网络架构,旨在通过优化卷积操作和通道设计,提高计算效率和模型精度。相较于 ShuffleNetV1,ShufflenetV2 摒弃了耗内存的 1×1 分组卷积,采用更高效的通道分割与通道混洗机制,并融合了多尺度特征以增强模型表达能力,使其在保持低参数与低计算量的同时,实现了精度与速度之间的平衡,适

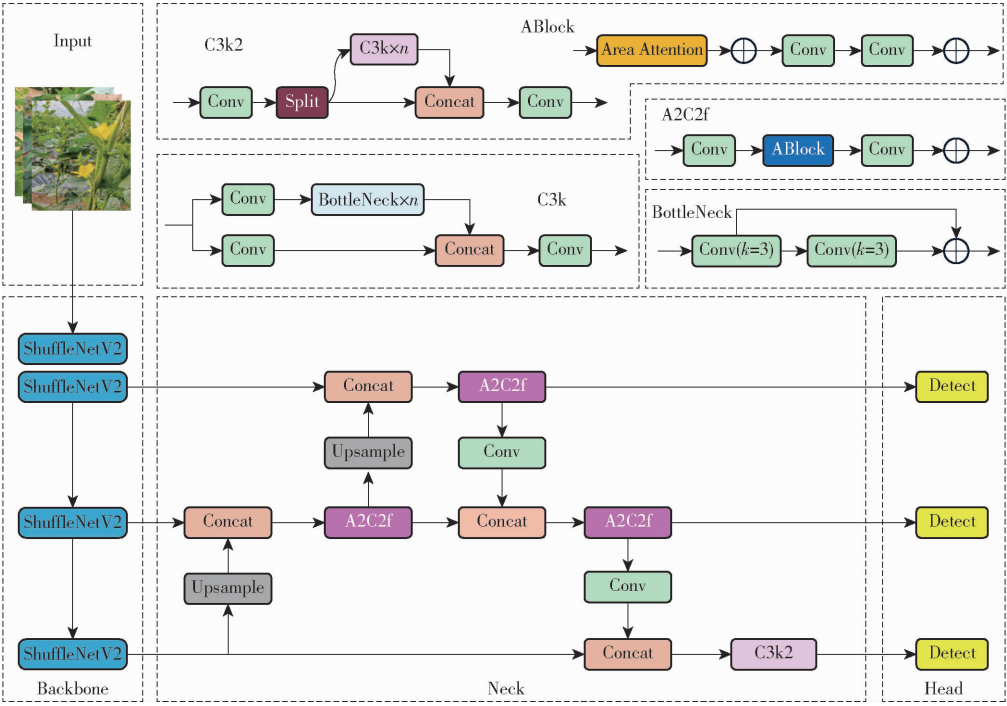


图 7 YOLO 12s - ShuffleNetV2 - D 模型结构

Fig. 7 Model structure of YOLO 12s - ShuffleNetV2 - D

用于移动端等资源受限设备。

如图 8 所示,该模型在基本单元中引入通道分割 (Channel Split) 操作,将输入特征图拆分为 2 个分支,1 个分支进行恒等映射以保留原始信息并降低碎片化,另 1 个分支通过一系列卷积进行特征变换,确保输入输出通道数相等以最小化内存访问开销,并对 2 个分支的输出结果进行 Concat 操作,随后通过通道混洗 (Channel Shuffle) 操作促进特征信息的交融,打破分组卷积可能造成的信息隔离,并辅以多尺度特征融合,使得 ShuffleNetV2 在参数量、计算速

度和模型精度之间达到最优权衡。

为选择合适的轻量化骨干网络,分别利用 MoblieNetV3^[23]、GhostNetV2^[24] 和 LSNet^[25] 轻量化网络,构建相同的网络结构,在相同环境下利用相同数据集进行对比试验,结果表明,ShuffleNetV2 保持较高检测精度的同时,在模型轻量化方面表现最优。基于此,将 ShuffleNetV2 模块引入 YOLO 12s 网络模型,旨在实现一种轻量化设计,显著提升其在移动端及边缘设备上的部署能力。

1.3.3 知识蒸馏

知识蒸馏^[26]是提高轻量化模型性能的常用方法,通过将复杂教师模型的知识迁移至轻量学生模型,减少模型的计算成本和内存占用量,以适配资源受限环境下的高精度目标识别任务需求。相较于传统的逻辑蒸馏,基于特征的蒸馏通过让学生模型学习教师模型中间层的特征表示来传递知识,使学生模型能捕获更多结构特征与细节信息,进而增强学生模型的特征表达和泛化能力。

CWD (Channel-wise knowledge dissillation)^[27]是通过通道维度上实现软对齐来进行知识迁移的蒸馏方法。将教师模型与学生模型对应通道的激活图转换为概率分布,并利用 KL 散度最小化二者差异,使蒸馏过程专注于学习每一通道的显著区域,从而实现通道级知识的有效迁移。CWD 损失一般形式可表示为

$$\varphi(\phi(y^t), \phi(y^s)) = \varphi(\phi(y_c^t), \phi(y_c^s)) \quad (1)$$

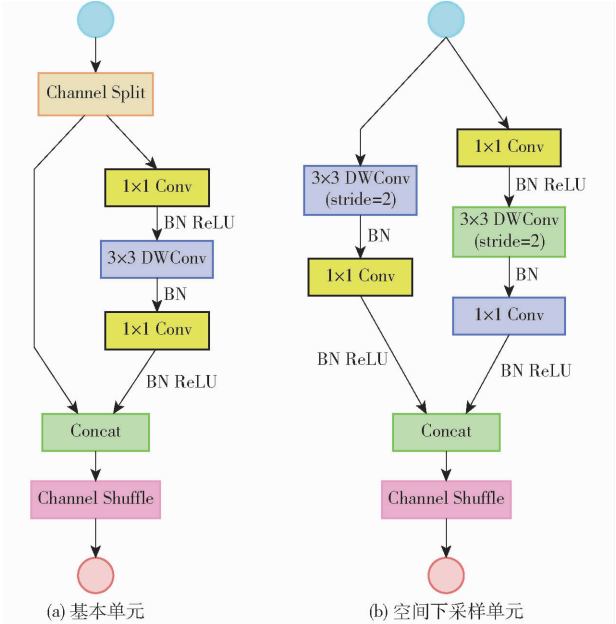


图 8 ShuffleNetV2 结构

Fig. 8 ShuffleNetV2 frame structure

其中

$$\phi(y_c) = \frac{\exp\left(\frac{y_{c,i}}{T}\right)}{\sum_{i=1}^{WH} \exp\left(\frac{y_{c,i}}{T}\right)}$$

(2)

式中 y^t ——教师网络激活映射
 y^s ——学生网络激活映射
 $\phi()$ ——将激活值转换为概率分布函数,即使用 Softmax 归一化,将每个通道的特征映射转换为概率映射
 c ——通道索引
 i ——通道内的空间索引
 $W、H$ ——特征图宽度、高度
 T ——温度超参数

利用 KL 散度评估教师与学生模型之间通道分布的差异,教师与学生特征蒸馏损失值计算式为

$$\varphi(y^t, y^s) = \frac{T^2}{C} \sum_{c=1}^C \sum_{i=1}^{WH} \phi(y_{c,i}^t) \ln\left(\frac{\phi(y_{c,i}^t)}{\phi(y_{c,i}^s)}\right)$$

(3)

式中 C ——特征图总通道数

KL 散度是一种非对称度量,其优化过程赋予教师网络中不同区域以不同的重要性。由式(3)可知,若 $\phi(y_{c,i}^t)$ 很大,那么 $y_{c,i}^s$ 也应相应增大,才能使 KL 散度最小。当 $\phi(y_{c,i}^t)$ 非常小时,KL 散度对最小化 $y_{c,i}^s$ 的关注较少。因此,学生网络趋向于在前景显著性上与教师网络产生近似激活分布,而教师网络的背景区域对应的激活对学习的影响较小。

1.3.4 评价指标

采用精确率 (Precision, P)、召回率 (Recall, R)、平均精度均值 (mean Average Precision, mAP)、模型内存占用量、检测速度对甜瓜花朵检测模型进行评价。

2 结果与分析

2.1 模型训练

为弥补轻量化模型带来的精度损失,进一步提高改进模型的性能,利用 CWD 知识蒸馏技术,将大型教师模型的知识迁移到小型学生模型,以提高甜瓜花朵识别过程中的检测精度。

在 YOLO 12 系列中,YOLO 12x 具有更深的网络结构和更宽的特征通道,其生成的多尺度特征图包含丰富的细节信息,表现出更强的鲁棒性和稳定性。因此,选取 YOLO 12x 作为教师模型,改进的轻量化网络 YOLO 12s - ShuffleNetV2 作为学生模型,通过知识蒸馏技术,实现模型精度与速度的平衡。YOLO 12s - ShuffleNetV2 - D 模型训练结果如图 9 所示,在模型训练初期,模型学习效率较高,目标检

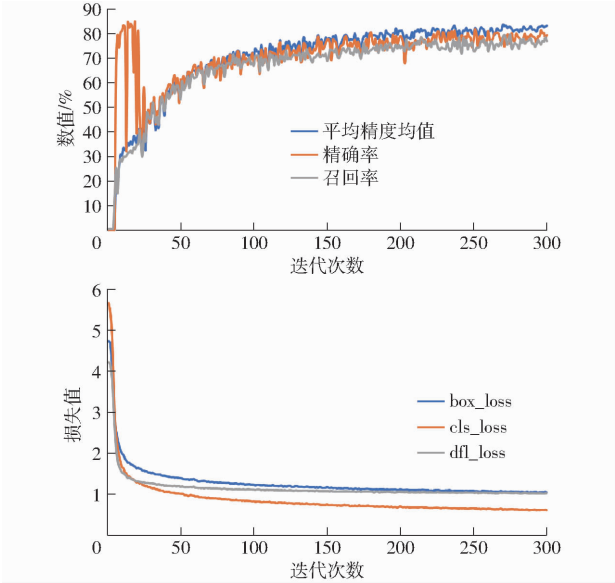


图 9 训练结果

Fig. 9 Training results

测损失 (box_loss)、分类损失 (cls_loss) 及分布焦点损失 (dfl_loss) 均快速收敛,当迭代至 225 轮时,模型学习效率趋于饱和。

模型测试结果如表 2 所示。由表 2 可知,通过知识蒸馏后的模型 YOLO 12s - ShuffleNetV2 - D 在保持轻量化优势的基础上,相较于 YOLO 12s - ShuffleNetV2 模型,其精确率、召回率、mAP 分别提高 1.46、6.87、2.33 个百分点;相较于 YOLO 12s 模型,其 mAP 低 1.77 个百分点,但模型内存占用量仅为 YOLO 12s 模型的 22.69%,且检测速度快 79.32%。结果表明,YOLO 12s - ShuffleNetV2 - D 在精度与轻量化间取得了较好平衡,对模型的改进是有效的。

表 2 模型测试结果

Tab. 2 Results of model test

算法	精确率/%	召回率/%	mAP/%	模型内存占用量/MB	检测速度/(f·s ⁻¹)
YOLO 12s	86.86	82.50	88.31	18.07	85.58
YOLO 12s - ShuffleNetV2	78.47	80.78	84.21	4.10	148.76
YOLO 12s - ShuffleNetV2 - D	79.93	87.65	86.54	4.10	153.46

2.2 检测结果

为验证 YOLO 12s - ShuffleNetV2 - D 网络检测甜瓜花朵的有效性,对测试集中 720 幅甜瓜花朵图像进行测试,结果表明,该方法检测甜瓜花朵精确率为 79.93%,召回率为 87.65%,mAP 为 86.54%,模型内存占用量为 4.10 MB,检测速度为 153.46 f/s。YOLO 12s - ShuffleNetV2 - D 目标检测算法检测甜瓜花朵效果如图 10 所示。

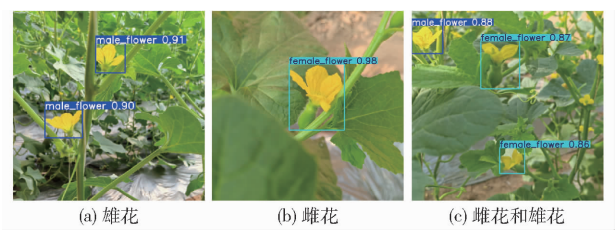


图 10 甜瓜花朵识别效果

Fig. 10 Melon flower recognition rendering

检测置信度在检测框上方表示,目标植株的每个雌花和雄花均可准确识别,且置信度均高于 0.85,表明该算法可以有效地检测并识别出雌花和雄花,且检测的置信度较高。如图 11 所示,在花朵数量较多、背景复杂及遮挡等复杂场景中,该算法亦可实现准确检测。

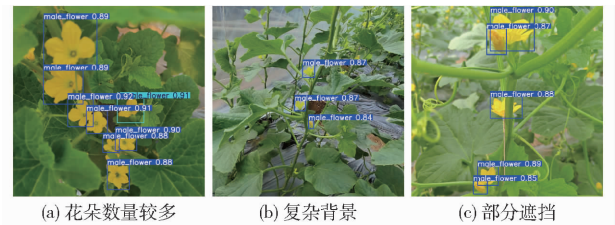


图 11 复杂情况下的花朵识别效果

Fig. 11 Flower recognition renderings in complex situations

2.3 不同目标检测算法性能测试

为了评估 YOLO 12s – ShuffleNetV2 – D 网络对甜瓜花朵的检测效果,将其与主流目标检测算法进行了对比,涵盖双阶段检测算法 Faster R – CNN^[28]与单阶段目标检测算法 SSD^[29]、YOLO v5s^[30]、YOLO v8s^[31]、YOLO v10s^[32]和 YOLO 12s,对甜瓜花朵数据集在相同条件下进行训练并验证,再利用测试集对各检测算法的性能进行评估,不同模型检测性能如表 3 所示。

表 3 7 种目标检测算法性能测试结果

Tab.3 Performance test results of seven target detection algorithms

算法	精确率/%	召回率/%	mAP/%	模型内存占用量/MB	检测速度/(f·s ⁻¹)
YOLO 12s – ShuffleNetV2 – D	79.93	87.65	86.54	4.10	153.46
YOLO 12s	86.86	82.50	88.31	18.07	85.58
YOLO v10s	83.29	80.32	87.13	15.80	90.49
YOLO v8s	84.20	80.71	86.17	21.50	105.73
YOLO v5s	86.89	82.12	87.76	13.73	111.02
SSD	78.50	79.34	64.10	59.26	18.13
Faster R – CNN	66.46	82.68	70.52	314.30	12.10

由表 3 可知,YOLO 12s – ShuffleNetV2 – D 的精确率为 79.93%, 分别比 YOLO 12s、YOLO v10s、YOLO v8s 和 YOLO v5s 低 6.93、3.36、4.27、6.96 个

百分点,分别比 SSD 和 Faster R – CNN 高 1.43、13.47 个百分点;其召回率为 87.65%, 分别比 YOLO 12s、YOLO v10s、YOLO v8s、YOLO v5s、SSD 和 Faster R – CNN 高 5.15、7.33、6.94、5.53、8.31、4.97 个百分点;其 mAP 为 86.54%, 分别比 YOLO 12s、YOLO v10s 和 YOLO v5s 低 1.77、0.59、1.22 个百分点,但分别比 YOLO v8s、SSD 和 Faster R – CNN 高 0.37、22.44、16.02 个百分点;其模型内存占用量为 4.10 MB,分别比 YOLO 12s、YOLO v10s、YOLO v8s、YOLO v5s、SSD 和 Faster R – CNN 小 77.31%、74.05%、80.93%、70.14%、93.08%、98.70%;其检测速度为 153.46 f/s, 分别比 YOLO 12s、YOLO v10s、YOLO v8s、YOLO v5s、SSD 和 Faster R – CNN 快 79.32%、75.11%、45.14%、38.23%、746.44%、1 168.26%。

YOLO 12s – ShuffleNetV2 – D 模型的精确率虽略低于 YOLO 12s、YOLO v10s、YOLO v8s 和 YOLO v5s,但显著高于 SSD 和 Faster R – CNN,且召回率均高于所有对比算法,验证其漏检控制能力最优。其 mAP 略低于 YOLO 12s、YOLO v10s 和 YOLO v5s,但高于 YOLO v8s、SSD 和 Faster R – CNN。在模型内存占用量和检测速度方面,YOLO 12s – ShuffleNetV2 – D 于 7 种模型中表现最佳。在能基本满足检测精度的情况下,YOLO 12s – ShuffleNetV2 – D 的模型内存占用量更小、检测速度更快,其轻量化特性更适合于在边缘设备上的部署,为在计算资源有限的设备上实现甜瓜花朵的实时检测提供了有效的技术方案。

将测试集中 720 幅图像按照不同光照、不同遮挡情况和不同清晰程度进行分类,在相同试验条件下,利用上述 7 种目标检测算法分别对不同场景下的甜瓜花朵图像测试集进行了测试,表 4 列出了不同场景下 7 种目标检测算法的检测效果。

由表 4 可知,在不同光照、遮挡及清晰度场景下,YOLO 12s – ShuffleNetV2 – D 模型精确率整体上略低于 YOLO 12s、YOLO v8s 和 YOLO v5s 模型,但在不同清晰程度和逆光条件下,优于 YOLO v10s,在逆光条件下优于 YOLO v8s 模型,且整体上显著优于 SSD 和 Faster R – CNN 模型;其召回率整体上略低于其他对比模型,但在顺光和遮挡条件下表现较优;其 mAP 略低于 YOLO 12s、YOLO v10s、YOLO v8s 和 YOLO v5s 模型,但显著高于 SSD 和 Faster R – CNN 模型,在保持轻量化的同时,具有较强的场景适应能力。

结果表明,YOLO 12s – ShuffleNetV2 – D 在精度与效率间取得了较好平衡,核心优势在于其轻量

表 4 不同场景下 7 种目标检测算法的检测精度

Tab.4 Accuracy of seven algorithms in different scenarios

%					
算法	场景	类别	精确率	召回率	mAP
YOLO 12s - ShuffleNetV2 - D	光照情况	顺光	80.33	84.01	84.80
		逆光	84.59	79.77	85.73
	遮挡情况	遮挡	79.11	81.04	84.39
		无遮挡	84.53	82.30	87.09
	清晰程度	清晰	86.09	80.68	86.67
		模糊	82.05	77.84	84.31
YOLO 12s	光照情况	顺光	87.03	82.36	88.66
		逆光	85.76	83.06	88.04
	遮挡情况	遮挡	83.28	81.46	86.31
		无遮挡	89.71	85.00	90.76
	清晰程度	清晰	90.43	82.70	89.34
		模糊	82.94	83.44	87.48
YOLO v10s	光照情况	顺光	86.29	78.63	87.21
		逆光	80.39	82.98	86.98
	遮挡情况	遮挡	84.24	77.91	86.32
		无遮挡	84.53	81.72	88.29
	清晰程度	清晰	84.46	82.08	87.87
		模糊	83.18	78.31	86.55
YOLO v8s	光照情况	顺光	88.77	77.65	86.73
		逆光	83.55	81.50	85.80
	遮挡情况	遮挡	82.25	79.33	84.49
		无遮挡	90.10	80.70	88.47
	清晰程度	清晰	84.51	84.63	88.26
		模糊	84.01	77.88	84.79
YOLO v5s	光照情况	顺光	87.90	81.65	87.99
		逆光	86.80	82.54	89.29
	遮挡情况	遮挡	86.13	81.47	87.37
		无遮挡	90.33	83.34	90.71
	清晰程度	清晰	89.38	83.12	89.60
		模糊	85.49	81.89	88.14
SSD	光照情况	顺光	79.72	79.11	63.01
		逆光	77.33	79.57	67.63
	遮挡情况	遮挡	77.00	76.13	61.77
		无遮挡	80.66	84.22	72.49
	清晰程度	清晰	78.75	82.01	67.59
		模糊	78.28	77.04	63.21
Faster R - CNN	光照情况	顺光	68.11	82.68	69.44
		逆光	64.61	72.18	60.77
	遮挡情况	遮挡	63.02	79.58	64.59
		无遮挡	71.88	87.39	77.73
	清晰程度	清晰	68.79	84.26	71.98
		模糊	64.51	81.33	68.99

化与快速检测性能。虽在检测精度上略低于其他 YOLO 模型,但显著高于 SSD 和 Faster R - CNN,且置信度均在 0.80 以上。在能够基本完成甜瓜雌雄花的实际检测精度需求的情况下,该模型更适合在计算资源有限的边缘设备上部署,为甜瓜花朵的实时检测提供了高效的解决方案。

2.4 不同算法对模型迁移的影响

为验证 YOLO 12s - ShuffleNetV2 - D 目标检测网络的准确性与实时性,选取性能较好的 3 种模型 (YOLO 12s - ShuffleNetV2 - D、YOLO 12s 和 YOLO v5s) 迁移至 NVIDIA Jetson Xavier NX 开发板上。该开发板配置 384 个 CUDA 核心的 Volta 架构的 GPU 和六核 NVIDIA Carmel CPU、Ubuntu 18.04 操作系统、Python 3.8 版本、PyTorch 2.2.0 版本深度学习框架。在未优化的前提下,于嵌入式设备上使用本数据集的测试集对模型性能进行了测试,各模型 mAP 与检测速度如表 5 所示。

表 5 3 种目标检测算法测试结果

Tab.5 Test results of three target detection algorithms

算法	mAP/%	检测速度/(f·s ⁻¹)
YOLO 12s - ShuffleNetV2 - D	86.62	26.37
YOLO 12s	88.30	13.85
YOLO v5s	87.78	22.83

由表 5 可知,YOLO 12s - ShuffleNetV2 - D 的 mAP 分别比 YOLO 12s 和 YOLO v5s 低 1.68、1.16 个百分点,但检测速度分别比 YOLO 12s 和 YOLO v5s 高 90.40%、15.51%。

结果表明,YOLO 12s - ShuffleNetV2 - D 在精度上与迁移前的模型相当,但检测速度因硬件资源限制下降 82.27%,但仍显著高于 YOLO 12s 和 YOLO v5s。在计算资源受限的条件下,YOLO 12s -ShuffleNetV2 - D 凭借其轻量化特性与较快的检测速度,展现出良好的综合性能。虽其精度略低于 YOLO 12s 和 YOLO v5s,但 mAP 仍为 86.62%,基本满足甜瓜雌雄花的识别任务,且其快速的检测能力可较好地与移动式授粉设备结合,确保设备在行进间及时、准确地完成雌雄花朵识别与定位,提升授粉效率。因此,在实际授粉场景下,本研究模型可更好地平衡精度与速度,是更优的边缘部署选择。

2.5 不同环境下检测效果

为全面评估 YOLO 12s - ShuffleNetV2 - D 目标检测模型在温室环境下的鲁棒性与泛化能力,系统分析了其不同光照、遮挡及清晰度条件下甜瓜花朵的检测效果,以验证该模型应对环境变化的稳定性与可靠性。该网络在不同光照情况下花朵检测效果如图 12 所示,经测试,本方法检测顺光、逆光条件下甜瓜花朵精确率分别为 80.33%、84.59%,召回率分别为 84.01%、79.77%,mAP 分别为 84.80%、85.73%,结果表明,在不同光照情况下,YOLO 12s - ShuffleNetV2 - D 网络仍然可以准确识别出甜瓜雌

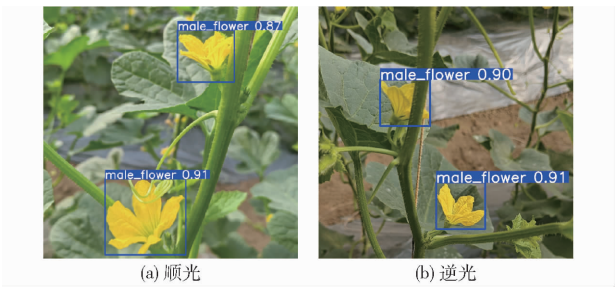


图 12 不同光照情况下花朵检测效果

Fig. 12 Recognition results under different illumination conditions

雄花朵且检测置信度较高。

为验证模型对不同遮挡情况下的花朵识别效果,利用 YOLO 12s-ShuffleNetV2-D 目标检测算法对测试集中不同遮挡情况的花朵图像进行了测试,检测结果如图 13 所示,经测试,本方法对无遮挡和遮挡花朵的精确率分别为 84.53%、79.11%,召回率分别为 82.30%、81.04%,mAP 分别为 87.09%、84.39%,结果表明,该目标检测算法可以准确识别出不同遮挡情况的甜瓜雌雄花朵且检测的置信度较高。

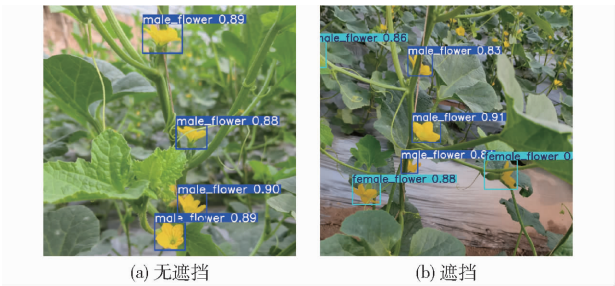


图 13 不同遮挡情况下花朵检测效果

Fig. 13 Recognition results under different occlusion conditions

不同清晰程度甜瓜花朵检测效果如图 14 所示,经测试,本方法检测清晰和模糊情况下甜瓜花朵精确率为 86.09%、82.05%,召回率为 80.68%、77.84%,mAP 为 86.67%、84.31%,结果表明,对于不同的清晰程度,YOLO 12s-ShuffleNetV2-D 目标检测算法均可以准确识别甜瓜花朵且置信度较高。

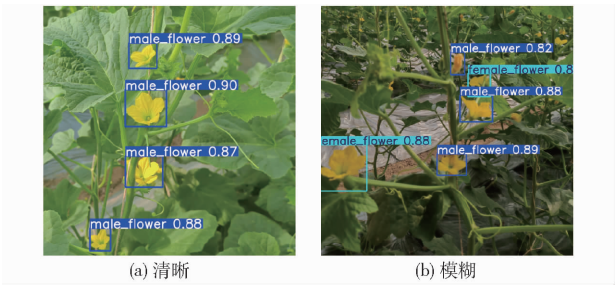


图 14 不同清晰程度下花朵检测效果

Fig. 14 Recognition results at different levels of clarity

3 讨论

图 15 为部分花朵被漏检和误检的情况,部分甜瓜花朵被漏检(图 15a~15c 箭头所示),同时也存在雌、雄花识别误检(图 15d、15e 箭头所示)等情况。分析出现漏检和误检的可能原因如下:

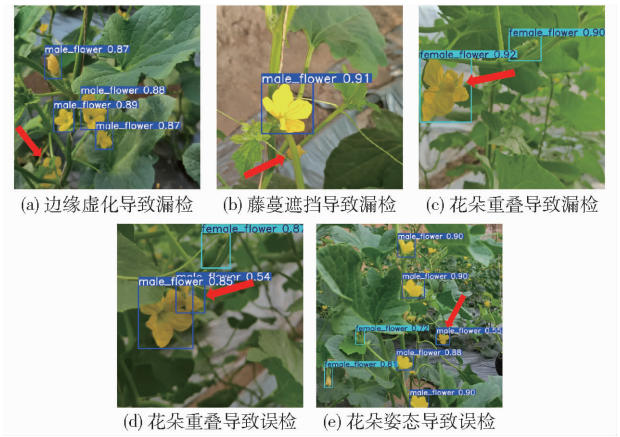


图 15 漏检和误检效果

Fig. 15 Renderings of missed and false detection

(1)甜瓜花朵图像采集设备为智能手机,拍摄图像时会存在一定程度虚化,且温室大棚膜、金属架等可能会产生反光,干扰图像质量,导致处于图像中心的花朵细节明显而处于图像边缘的花朵细节模糊,如图 15a 箭头所示。甜瓜花朵检测模型在处理图像边缘的花朵时难以提取其特征,导致检测时出现部分图像边缘模糊花朵漏检。

(2)甜瓜植株具有叶片硕大且藤蔓密集生长的特性,上层的藤蔓或大型叶片常对下层花朵形成物理遮挡。如图 15b 箭头所示,目标检测模型在处理此类场景时难以提取下层花朵特征,导致检测时出现仅上层花朵成功检出,而下层花朵无法检出的情况,造成漏检。

(3)甜瓜花朵呈簇状密集分布,相邻花朵易相互重叠,处于上层的花朵会严重遮挡下层花朵的特征信息,使得上层花朵特征清晰易辨而下层花朵的特征难以提取。此外,在生长过程中花朵姿态各异,且互相挤压导致形状不规则,子房作为区分雌花与雄花的关键特征,易被自身花瓣或邻近花朵遮挡,如图 15c~15e 箭头所示。当目标花朵的关键特征在预测框内无法捕捉时,模型对该花朵识别会出现偏差甚至失败,造成误检与漏检。

4 结论

(1)基于改进 YOLO 12s 模型实现了甜瓜花朵高效检测与识别,所用算法对花朵检测的精确率为 79.93%,召回率为 87.65%,mAP 为 86.54%,模型内存

占用量为 4.10 MB,检测速度为 153.46 f/s。与 YOLO 12s、YOLO v10s、YOLO v8s、YOLO v5s、SSD 和 Faster R-CNN 目标检测算法进行对比,其模型内存占用量分别减少 77.31%、74.05%、80.93%、70.14%、93.08%、98.70%,检测速度分别提升 79.32%、75.11%、45.14%、38.23%、746.44%、1 168.26%。

(2)针对不同光照、遮挡及清晰度条件下的甜瓜花朵,提出的甜瓜花朵检测方法均能较好地检测并识别甜瓜雌雄花朵,说明基于 YOLO 12s -

ShuffleNetV2 - D 的甜瓜花朵检测方法具有较好的鲁棒性。

(3)将 YOLO 12s - ShuffleNetV2 - D、YOLO 12s 和 YOLO v5s 3 种模型部署于 Jetson Xavier NX 开发板上进行实测,YOLO 12s - ShuffleNetV2 - D 的 mAP 为 86.62%,检测速度为 26.37 f/s,其检测速度分别比 YOLO 12s 和 YOLO v5s 提高 90.40%、15.51%,同时在精度方面也能满足实际需求,是面向实际授粉场景的更优边缘部署选择。

参 考 文 献

[1] KIATOKO N, POZO M, VAN OYSTAEYEN A, et al. Effective pollination of greenhouse Galia musk melon (*Cucumis melo* L. var. *reticulatus* ser.) by afrotropical stingless bee species[J]. *Journal of Apicultural Research*, 2022, 61(5): 664 - 674.

[2] HE L L, LIU X J, DING Y S, et al. Advancements in artificial pollination of crops: from manual to autonomous[J]. *Computers and Electronics in Agriculture*, 2025, 231: 110067.

[3] 赵春江,文朝武,林森,等. 基于级联卷积神经网络的番茄花期识别检测方法[J]. *农业工程学报*, 2020, 36(24): 143 - 152. ZHAO Chunjiang, WEN Chaowu, LIN Sen, et al. Tomato florescence recognition and detection method based on cascaded neural network[J]. *Transactions of the CSAE*, 2020, 36(24): 143 - 152. (in Chinese)

[4] GAO C Q, HE L L, FANG W T, et al. A novel pollination robot for kiwifruit flower based on preferential flowers selection and precisely target[J]. *Computers and Electronics in Agriculture*, 2023, 207: 107762.

[5] 王宇. 基于深度学习的目标检测算法综述[J]. *科技资讯*, 2025, 23(2): 64 - 66. WANG Yu. An overview of object detection algorithm based on deep learning[J]. *Science and Technology Information*, 2025, 23(2): 64 - 66. (in Chinese)

[6] FARJON G, KRIKEB O, BAR HILLER A, et al. Detection and counting of flowers on apple trees for better chemical thinning decisions[J]. *Precision Agriculture*, 2020, 21(3): 503 - 521.

[7] 文朝武,龙洁花,张宇,等. 基于 3D 视觉的番茄授粉花朵定位方法[J]. *农业机械学报*, 2022, 53(8): 320 - 328. WEN Chaowu, LONG Jiehua, ZHANG Yu, et al. Positioning method of tomato pollination flowers based on 3D vision[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(8): 320 - 328. (in Chinese)

[8] 肖克辉,杨宏,苏章顺,等. 基于改进 YOLOv5 的蝴蝶兰花朵识别与计数[J]. *中国农机化学报*, 2023, 44(11): 155 - 161, 209. XIAO Kehui, YANG Hong, SU Zhangshun, et al. Identification and counting of phalaenopsis flowers based on improved YOLOv5[J]. *Journal of Chinese Agricultural Mechanization*, 2023, 44(11): 155 - 161, 209. (in Chinese)

[9] 任龙龙,杜永辉,李玉强,等. 基于 CEG - YOLO 的轻量化设施环境樱花检测检测方法[J]. *农业机械学报*, 2025, 56(9): 648 - 656. REN Longlong, DU Yonghui, LI Yuqiang, et al. Lightweight cherry blossom detection method in facility environments based on CEG - YOLO[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025, 56(9): 648 - 656. (in Chinese)

[10] 刘浩洲,陈礼鹏,穆龙涛,等. 基于 K - means 聚类的猕猴桃花朵识别方法[J]. *农机化研究*, 2020, 42(2): 22 - 26. LIU Haozhou, CHEN Lipeng, MU Longtao, et al. A recognition method of kiwifruit flowers based on K - means clustering[J]. *Journal of Agricultural Mechanization Research*, 2020, 42(2): 22 - 26. (in Chinese)

[11] GENG M Y, SHANG Y Y, XIANG S Y, et al. Using improved density peak clustering algorithm for flower cluster identification and apple central and peripheral flower detection[J]. *Computers and Electronics in Agriculture*, 2025, 232: 110095.

[12] 宋怀波,尚钰莹,何东健. 果实目标深度学习识别技术研究进展[J]. *农业机械学报*, 2023, 54(1): 1 - 19. SONG Huaibo, SHANG Yuying, HE Dongjian. Review on deep learning technology for fruit target recognition[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(1): 1 - 19. (in Chinese)

[13] 于萍. 面向边缘设备的智慧农业病虫害图像识别关键技术研究[D]. 长春: 吉林大学, 2025. YU Ping. Research on key technologies for image recognition of pest and disease in smart agriculture for edge devices[D]. Changchun: Jilin University, 2025. (in Chinese)

[14] ZAIDI S S A, ANSARI M S, ASLAM A, et al. A survey of modern deep learning based object detection models[J]. *Digital Signal Processing*, 2022, 126: 103514.

[15] 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 onimproved YOLOv4 - tiny[J]. *Transactions of the CSAE*, 2022, 38(9): 215 - 221. (in Chinese)

[16] YANG G L, WANG J X, NIE Z L, et al. A lightweight YOLOv8 tomato detection algorithm combining feature enhancement and attention[J]. *Agronomy-Basel*, 2023, 13(7): 1824.

[17] 赵博, 柳苏纯, 张巍朋, 等. 面向樱桃番茄采摘识别的轻量化 Transformer 架构优化研究[J]. *农业机械学报*, 2024, 55(10): 62 – 71, 105.
ZHAO Bo, LIU Suchun, ZHANG Weipeng, et al. Performance optimization of lightweight Transformer architecture for cherry tomato picking[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(10): 62 – 71, 105. (in Chinese)

[18] 温毓晨, 赵永杰, 蒲六如, 等. 基于改进 YOLO v5s 模型的奶山羊乳房区域热红外图像检测方法[J]. *农业机械学报*, 2024, 55(6): 237 – 245.
WEN Yuchen, ZHAO Yongjie, PU Liuru, et al. Thermal infrared image detection method of dairy goat breast region based on improved YOLO v5s model[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(6): 237 – 245. (in Chinese)

[19] SAPKOTA R, FLORES-CALERO M, QURESHI R, et al. YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series[J]. *Artificial Intelligence Review*, 2025, 58(9): 274.

[20] GE T, NING B, XIE Y. YOLO – AFR: an improved YOLOv12 – based model for accurate and real-time dangerous driving behavior detection[J]. *Applied Sciences-Basel*, 2025, 15(11): 6090.

[21] MA J, ZHOU Y, ZHOU Z, et al. Toward smart ocean monitoring: real-time detection of marine litter using YOLOv12 in support of pollution mitigation[J]. *Marine Pollution Bulletin*, 2025, 217: 118136.

[22] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design [C] // *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 116 – 131.

[23] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C] // *Proceedings of IEEE/CVF International Conference on Computer Vision*, 2019: 1314 – 1324.

[24] TANG Y, HAN K, GUO J, et al. GhostNetV2: enhance cheap operation with long-range attention[J]. *arXiv Preprint, arXiv: 2211.12905*, 2022.

[25] WANG A, CHEN H, LIN Z, et al. LSNet: see large, focus small [C] // *Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025: 9718 – 9729.

[26] GOU J P, YU B S, MAYBANK S J, et al. Knowledge distillation: a survey[J]. *International Journal of Computer Vision*, 2021, 129(6): 1789 – 1819.

[27] LI Z H, XU P F, CHANG X J, et al. When object detection meets knowledge distillation: a survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(8): 10555 – 10579.

[28] REN S Q, 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.

[29] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] // *Proceedings of the Computer Vision-ECCV 2016; 14th European Conference*, 2016: 21 – 37.

[30] YAN B, FAN P, LEI X Y, et al. A real-time apple targets detection method for picking robot based on improved YOLOv5 [J]. *Remote Sensing*, 2021, 13(9): 1619.

[31] LI Y T, FAN Q S, HUANG H S, et al. A modified YOLOv8 detection network for UAV aerial image recognition[J]. *Drones*, 2023, 7(5): 304.

[32] MAO M, LEE A, HONG M. Efficient fabric classification and object detection using YOLOv10 [J]. *Electronics*, 2024, 13(19): 3804.