

基于轻量化 YOLO 11n - CLL 和 BoT - SORT 的百香果计数方法

涂淑琴¹ 郭云峰¹ 毛亮² 张宇³ 张磊⁴ 谭柏洋¹

(1. 华南农业大学数学与信息学院, 广州 510642; 2. 深圳职业技术学院人工智能学院, 深圳 518055;
3. 华南农业大学人工智能学院, 广州 510642; 4. 广东技术师范大学计算机科学学院, 广州 510665)

摘要: 百香果轻量化检测、跟踪与计数是智慧农业中实现智能化采摘的关键技术。在真实果园场景中, 百香果与枝叶的重叠遮挡与环境光照变化会导致目标出现漏检、误检及重复计数等问题。针对上述问题, 本文提出了一种 YOLO 11n - CLL 融合 BoT - SORT 的轻量化计数方法。在目标检测中, YOLO 11n - CLL 利用 YOLO 11n 作为基线模型, 通过引入轻量化上下文引导模块, 增强上下文感知能力; 融合大核可分离注意力机制, 提升多尺度特征建模能力; 使用细节增强检测头, 提升遮挡果实的检测能力。在目标跟踪部分, BoT - SORT 跟踪器利用相机运动补偿模块, 抑制抖动干扰, 提升跟踪精度。在果实计数任务中, 选择最佳的区域计数方法, 实现果实准确计数。试验结果表明, YOLO 11n - CLL 平均准确率为 87.0%, 模型内存占用量为 5.0 MB; BoT - SORT 跟踪器 HOTA (Higher order tracking accuracy) 和 MOTA (Multiple object tracking accuracy) 分别为 62.4% 和 68.4%; 采用区域计数法对果实进行计数, 其平均计数准确率达到 92.8%, 比划线计数法和 ID 累加计数法分别高 8.8、29.9 个百分点。试验结果表明本文方法能实现百香果检测、跟踪与计数, 为百香果园智能化采摘提供技术支撑。

关键词: 百香果; YOLO 11n; 多目标跟踪; BoT - SORT; 果实计数

中图分类号: S126 文献标识码: A 文章编号: 1000-1298(2026)10-0208-11

OSID:



Passion Fruit Counting Method Based on Lightweight YOLO 11n - CLL and BoT - SORT

TU Shuqin¹ GUO Yunfeng¹ MAO Liang² ZHANG Yu³ ZHANG Lei⁴ TAN Baiyang¹

(1. College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China
2. School of Artificial Intelligence, Shenzhen Polytechnic University, Shenzhen 518055, China
3. College of Artificial Intelligence, South China Agricultural University, Guangzhou 510642, China
4. College of Computer Science, Guangdong Polytechnic Normal University, Guangzhou 510665, China)

Abstract: Lightweight detection, tracking, and counting passion fruit are crucial technologies for achieving intelligent harvesting in smart agriculture. In real orchard environments, factors such as overlapping occlusion with leaves and branches, along with lighting variations, often result in missed detections, false detections and repeated counting. To address these challenges, a lightweight counting framework was proposed based on YOLO 11n - CLL and BoT - SORT. In object detection, YOLO 11n - CLL was built upon YOLO 11n as its baseline model. Firstly, a lightweight context guided block was introduced to enhance contextual perception capability. Next, the large separable kernel attention mechanism was incorporated to improve multi-scale feature modeling. Finally, the lightweight shared detail-enhanced convolutional detection head was employed to effectively reconstruct the texture features of occluded fruits. For object tracking, the BoT - SORT tracker was adopted, which incorporated a camera-motion-compensation module to mitigate camera-shake interference and improve tracking accuracy. In fruit counting task, a region-based counting method was adopted to achieve accurate counting of passion fruits. Results showed that YOLO 11n - CLL achieved mAP@0.5 of 87.0% with a compact 5.0 MB model size. The BoT - SORT tracker achieved a HOTA of 62.4% and a MOTA of

收稿日期: 2025-11-28 修回日期: 2025-12-30

基金项目: 广东省普通高校重点领域项目(2025ZDZX4100)和广州市科技规划项目(2024E04J1245)

作者简介: 涂淑琴(1978—),女,讲师,博士,主要从事智慧农业研究,E-mail: tushuqin@163.com

通信作者: 毛亮(1983—),男,副研究员,博士,主要从事智慧农业、计算机视觉与深度学习研究,E-mail: maoliangscnu@szpu.edu.cn

68.4%。The region-based counting method yielded an average counting accuracy of 92.8%, outperforming the line-based counting and ID-based counting method by 8.8 percentage points and 29.9 percentage points. The results demonstrated that the proposed framework enabled detection, tracking, and counting of passion fruits and provided technical support for intelligent harvesting in passion fruit orchards.

Key words: passion fruit; YOLO 11n; multi-object tracking; BoT - SORT; fruit counting

0 引言

百香果作为一种具有独特风味与高营养价值的热带水果,已在多地广泛种植并深受消费者青睐^[1]。然而,传统的果园管理模式仍依赖人工操作,存在效率低、劳动强度大及果实计数精度不足等问题,不仅制约生产决策的科学性,也影响产业的可持续发展 and 经济效益。因此,亟需引入快速且高精度的人工智能计数方法,对果实进行自动化检测与计数,以支持百香果产业的智能化管理与高质量发展。

近年来,随着人工智能与深度学习(Deep learning, DL)^[4]的快速发展,基于 DL 的视觉检测方法正逐步替代传统人工方式,成为智能果园管理的重要技术支撑。文献[6 - 7]在提升水果检测准确性与效率方面取得了显著成效,为智慧农业的发展提供了切实可行的解决方案。然而,当前大多研究仍聚焦于果实检测单一环节,尚未形成系统化的智能感知体系^[8]。在实际应用中,精确的果实检测只是智能果园系统的基础,更关键的是依赖多目标跟踪(Multi-object tracking, MOT)技术以实现对果实个体的持续追踪与动态分析^[9],从而支撑后续的生长监测、产量评估与自动化采摘等核心任务。

多目标跟踪技术可实现对果园中多个果实目标的持续识别与轨迹分析,为果实生长监测、产量预测以及智能采摘规划等精细化农业管理任务提供关键数据支持。当前主流的 MOT 方法多基于“检测-跟踪”框架(Tracking-by-detection),如 DeepSORT^[12]、ByteTrack^[13]等。近年来,MOT 算法在农业场景中取得了一系列实用进展。文献[14 - 15]说明采用轻量化检测融合 MOT 技术可以实现果实的精准计数,但在果园自然场景中,面对百香果与枝叶的重叠遮挡与环境光照变化,这些技术仍面临着诸多严峻挑战^[16]。

在农业田间的实际应用场景中,风力扰动、机械振动等外部因素常导致视频图像产生抖动,严重干扰目标的运动轨迹预测,从而限制了 MOT 在果实跟踪任务中的性能^[17 - 18]。为应对此类挑战,AHARON 等^[19]提出了 BoT - SORT 算法,利用相机运动补偿机制将运动轨迹与外观特征融合,扩展卡尔曼滤波

状态向量,显著增强了目标在复杂环境下的跟踪稳定性。这一技术为农业场景中长时序、高精度目标跟踪任务提供了有效方法支撑,适用于智能化采摘等需持续追踪的农业生产管理应用,为精准农业实践奠定了技术基础^[20]。

果实计数作为果园智能管理系统的关键环节,对于产量预测、采收调度与资源优化配置具有重要意义^[21]。现有计算机视觉计数方法主要包括 3 类:ID 累加计数法、划线计数法以及区域计数法,前两类方法在近年来被广泛研究。例如,王金鹏等^[22]利用 ID 累加计数法实现油茶果实的跟踪与计数,计数精度达到 81.9%;PARICO 等^[23]对 ID 累加计数法与划线计数法进行了比较,结果显示,F1 分数为 97.85%、72.94%。为提升计数稳定性与适应性,部分研究开始探索区域计数法。TU 等^[24]提出百香果区域计数方法,利用预定义计数区域避免重复计数,平均计数准确率达到 87.0%;AN 等^[25]引入双重区域计数法,使计数准确率达到 90.43%。总体而言,区域计数法在精度和适应性方面展现出更加强大的潜力,值得在真实果园场景中进一步验证和深入探索。

为解决自然果园环境中因枝叶遮挡、光照变化与摄像头抖动导致的检测与跟踪不稳定问题,本文提出一种基于轻量化 YOLO 11n - CLL 与 BoT - SORT 的百香果计数方法。对 YOLO 11n 进行轻量化改进以提升检测能力;将该模型与 BoT - SORT 算法相结合以增强跟踪稳定性;采用区域计数法以实现精确果实计数。

1 数据采集与预处理

1.1 数据采集

百香果数据集采集于广东省广州市花都区 and 从化区的标准百香果种植园,采集时间为 2021 年 10 月与 2022 年 10 月。视频使用同一型号手机(HUAWEI P50)采集,采集过程中,为模拟采摘机器人典型作业视角,将手机固定于距地面高度为 1.5 m,与目标之间保持 1 m 物距。在此配置下,每段视频可有效覆盖约 15 m²的种植区域,涵盖多种典型场景。如图 1 所示,数据集包括果实被枝叶严重遮挡的密集遮挡场景(图 1a)、强烈日照导致特征过

曝的强光场景(图 1b)、设备移动造成图像模糊的动态场景(图 1c),以及作为基准参照的理想光照场景(图 1d) 4 类典型场景。



图 1 百香果数据集样本示例
Fig. 1 Examples of dataset

1.2 数据预处理及数据集制作

视频数据集具体信息如表 1 所示,利用 DarkLabel 软件对 21 个视频进行标注,视频集共包含 37 800 帧图像,每帧包含 2 ~ 15 个百香果果实(平均果实数为 10)。视频分辨率为 1 080 像素 × 1 920 像素,长度为 1 min,帧率为 30 f/s。这些视频包含自然场景下不同光照下的果实场景,体现了数据多样性和复杂性。

表 1 数据集概述与描述

参数	数值
视频总数/个	21
视频时长/s	60 ± 5
视频帧率/(f·s ⁻¹)	30
视频像素尺寸/(像素 × 像素)	1 080 × 1 920
视频总图像数/f	37 800
检测训练集图像数/f	4 158
检测测试集图像数/f	1 040
MOT 训练视频数/个	13(随机选取)
跟踪-计数测试视频数/个	8(随机选取)

在目标检测中,从 13 个训练 MOT 视频中随机选择 5 198 帧图像。按照比例 4:1 将数据集分为检测训练集和测试集,进行目标检测模型训练和评估。在 MOT 和计数任务中,选择 13 个视频作为训练集,8 个视频作为 MOT 跟踪与计数的测试集。

2 算法原理

2.1 总体流程

本文所提百香果检测、跟踪与计数方法流程如图 2 所示,包含目标检测、跟踪和计数 3 个阶段。在检测阶段,YOLO 11n-CLL 在 YOLO 11n 基础上,进行轻量化改进,用于逐帧检测视频中每个百香果。在跟踪阶段,BoT-SORT 跟踪器以检测结果作为输入,通过双阶段关联与运动相机补偿策略,结合卡尔曼运动预测与匈牙利匹配算法实现稳定跟踪。最后,通过利用 ID 累加计数法、划线计数法和区域计数法实现果实计数,并从定量与定性 2 个维度,全面评估 3 种方法在真实农业环境中的实用性与适应性。所提方法能实现果园低成本自动化,满足边缘部署轻量化要求。

2.2 YOLO 11n-CLL 目标检测方法

为实现检测模型在精度与轻量化水平之间的高效平衡,提出了一种轻量化 YOLO 11n-CLL 检测模型,其整体结构如图 3 所示。

2.2.1 C3k2_CG 模块

C3k2_CG 模块在主干网络的 C3k2 模块中引入上下文引导模块(Context guided block, CGBlock),以替代原始的 C3k2 结构。通过将 CGBlock 集成进 C3k2 模块,能够有效增强模型对复杂背景和局部信息的感知,从而提升目标检测精度,整体结构如图 4 所示。

CG 模块由 4 个子模块组成,分别为局部特征提取器 $f_{loc}(\ast)$ 、周围上下文提取器 $f_{sur}(\ast)$ 、联合特征提取器 $f_{joi}(\ast)$ 和全局特征提取器 $f_{glo}(\ast)$ 。模块通过 1 × 1 卷积对输入特征图进行通道压缩,降低计算开销并保留关键语义信息。压缩后的特征被送入 2 个并行分支,局部分支和上下文分支。局部分支使用标准 3 × 3 卷积提取果实的局部细节,上下文分支使用 3 × 3 的深度卷积提取背景区域的上下文信息。2 个分支的输出在通道维度上拼接后输入通道注意力子模块 $f_{glo}(\ast)$,该模块通过全局平均池化、全连接层和激活函数生成通道权重向量,

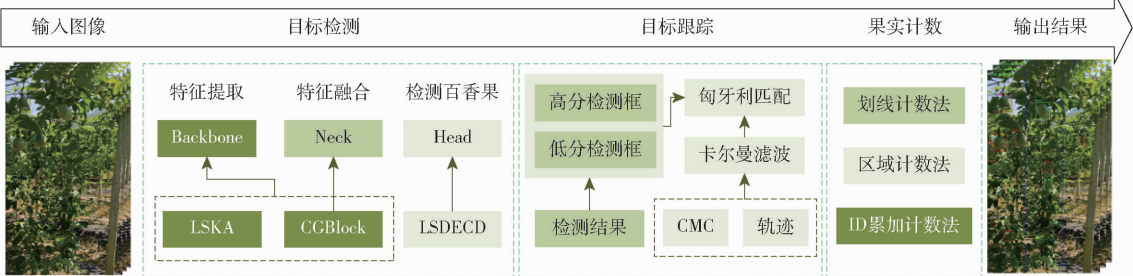


图 2 算法整体流程图
Fig. 2 Overall process flowchart

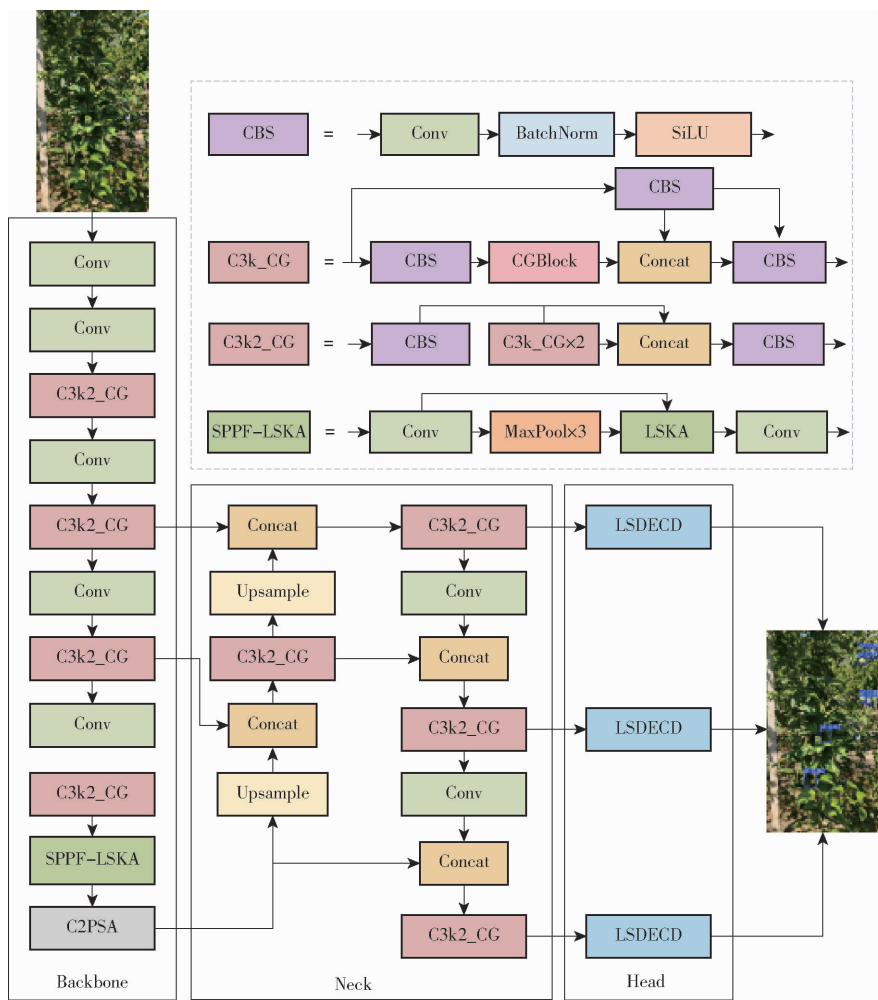


图 3 YOLO 11n-CLL 模型结构

Fig.3 Structure of YOLO 11n-CLL model

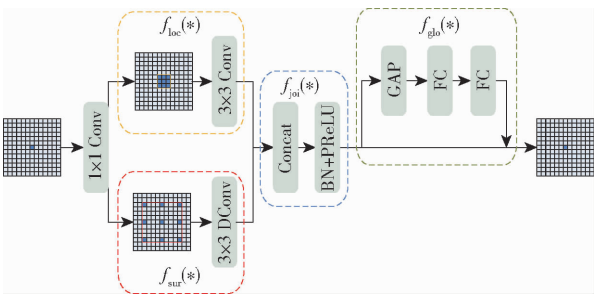


图 4 CG 模块结构

Fig.4 Structure of context guided block module

并对拼接特征进行加权,从而实现对关键特征的自适应增强。

引入 CG 模块后,模型在遮挡和复杂背景下展现出更强的判别能力和鲁棒性,同时保持良好的轻量化特性,使其适用于边缘计算等资源受限场景。

2.2.2 SPPF-LSKA 模块

为减轻遮挡与复杂背景对目标检测性能的影响,设计 SPPF-LSKA 模块,在 YOLO 11n 骨干网络中融合 LAU 等^[26]提出的轻量级空间核自适应注意力模块(Large separable kernel attention, LSKA),以实现果实在复杂场景下的特征提取。通过 LSKA 机

制,模型能够自适应地调整特征图的权重,显著提升了在光照变化下的特征表征稳定性。

LSKA 模块结构如图 5 所示,首先将输入特征图在通道维度上均匀划分为多个组,每组特征经过标准卷积 Conv 提取基础局部特征;再利用空洞卷积 D-Conv 捕获更广泛的上下文信息;最后将各组输出的特征,在通道维度上重新拼接,通过逐点卷积(P-Conv)完成跨通道融合,从而形成具备丰富多尺度信息的特征图。LSKA 模块通过深度卷积(DW-Conv)和空洞深度卷积(DW-D-Conv)联合扩展有效感受野,从而在遮挡、背景干扰和光照变化等自然场景中,具备更强的上下文理解能力与区域辨识能力。

SPPF-LSKA 模块在多尺度信息聚合、动态感受野调控和轻量化注意力建模方面形成互补优势,显著提升了模型在复杂场景中的检测性能,减轻了计算负担,为轻量化目标检测网络在实际应用中的部署提供了可靠解决方案。

2.2.3 LSDECD 模块

LSDECD (Lightweight shared detail-enhanced

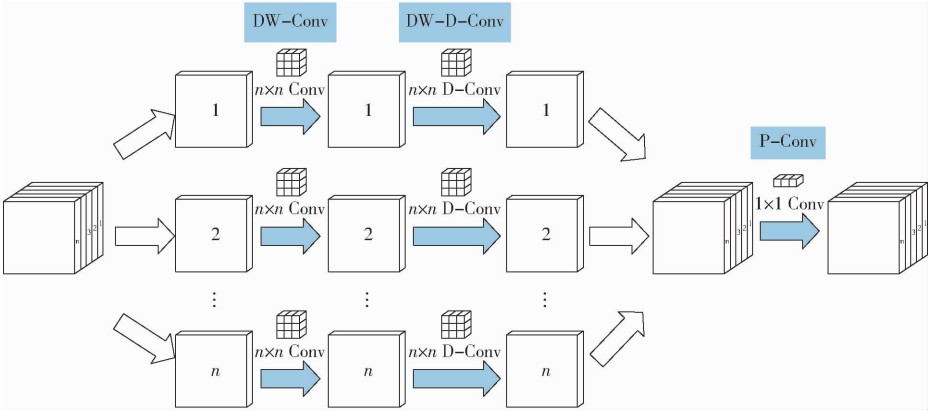


图 5 LSKA 模块结构

Fig. 5 LSKA structure

convolutional detection head) 是一种轻量级共享细节增强检测头。该模块基于 WANG 等^[27] 提出的轻量化检测头 LSCD 结构,以细节增强卷积 (Detail-enhanced convolution, DEConv) 取代其中的标准卷积,实现在保持轻量化的同时,显著提升局部差异建模能力。该模块特别适用于小目标和被遮挡区域的检测,能够显著提升检测精度。

LSDECD 模块由特征压缩模块 Conv_GN 和共享细节增强模块 DEConv_GN 组成,结构如图 6 所示。来自骨干网络的输入特征输入到 Conv_GN,实现通道压缩与基础特征整合。随后,获得的特征输入 2 个 DEConv_GN 模块中,获得更稳定的特征。DEConv_GN 为增强模型对不同方向局部变化的敏感性,引入 4 种方向感知的差分卷积操作:中心差分卷积 (Conv2d_cd)、水平差分卷积 (Conv2d_hd)、垂直差分卷积 (Conv2d_vd) 和反对称差分卷积 (Conv2d_ad),4 种方向感知的差分卷积操作具体功能见表 2。最后,整体输出经过 GroupNorm 处理,进一步提升特征分布的一致性和训练过程的稳定性。

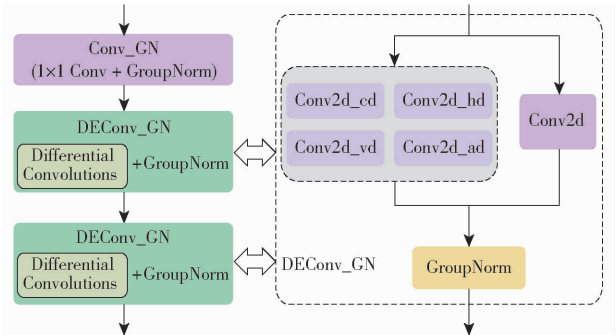


图 6 LSDECD 模块结构

Fig. 6 LSDECD structure

总体而言,LSDECD 能在不增加计算开销的前提下,有效增强了特征表达能力。与传统 YOLO 11n 检测头对比,该模块在复杂背景和小目标场景下均能展现出更优的检测性能。

表 2 不同差分卷积功能对比

Tab. 2 Functional comparison of differential convolutions

名称	功能
Conv2d_cd	提取中心像素与周围像素之间的特征差异,增强局部敏感性
Conv2d_hd	计算水平方向相邻像素的特征变化,建模水平细节结构
Conv2d_vd	计算垂直方向相邻像素的特征变化,建模垂直细节结构
Conv2d_ad	计算邻近像素之间的特征差异,捕获更自由方向的局部变化
Conv2d	提取整体局部特征,不特定于某一方向

2.3 BoT - SORT 多目标跟踪

在农业场景中,百香果在拍摄过程中,目标常因相机快速移动、果实遮挡和视角变化等呈现复杂的运动轨迹,使传统跟踪方法难以保持目标轨迹连续性与稳定性,导致跟踪性能下降。

针对上述问题,利用 BoT - SORT 跟踪器,采用相机运动补偿特点降低相机抖动的影响。同时,通过校正卡尔曼滤波器的预测结果,提高了目标定位的稳定性与跟踪精度。BoT - SORT 的整体流程如图 7 所示,具体流程包括 5 个阶段:①目标检测与特征提取。BoT - SORT 使用 YOLO 11n - CLL 完成果实目标检测,获得高分和低分检测框。对于高置信度检测框,通过外观特征提取器,提取外观特征,提供外观匹配信息。低置信度检测框暂时保留,用于后续第 2 阶段匹配。②相机运动补偿与目标预测。BoT - SORT 使用 CMC 模块提取图像中目标关键点,通过卡尔曼滤波器结合轨迹对目标位置进行预测,得到当前帧的预测边界框。③第 1 次匹配操作。将高分检测框与已有轨迹进行第 1 次特征和外观的匈牙利匹配。将匹配的轨迹和检测框放入轨迹管理模块,非匹配的轨迹输入第 2 次匹配模块。④第 2 次匹配操作。将第 1 次匹配失败的轨迹,利用 IoU 匹配方式,和剩余的低置信度检测框进行第 2 次匈

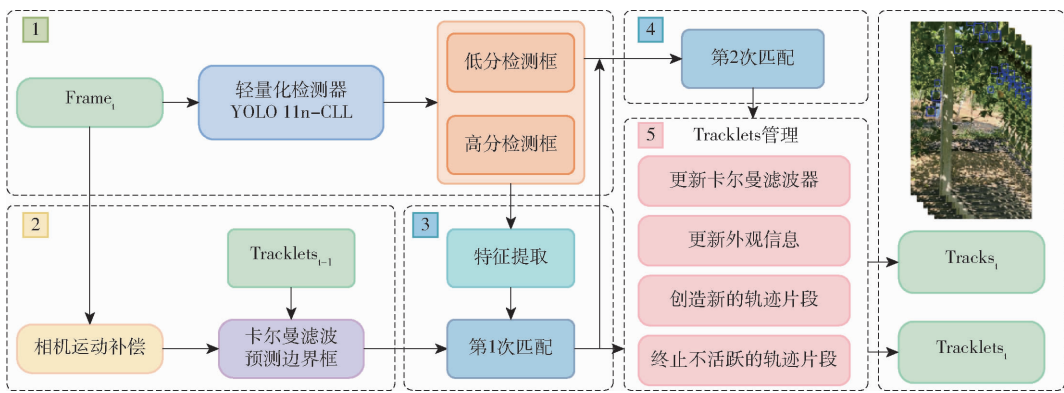


图 7 BoT-SORT 跟踪器结构

Fig. 7 BoT-SORT structure

匈牙利匹配。将匹配的轨迹输入轨迹管理模块进行更新,未匹配的轨迹符合删除条件则删除。⑤轨迹管理和后处理。对 2 次匹配的轨迹,进行卡尔曼滤波器和外观特征更新;对未匹配的高分检测框,创建新的轨迹;对超过 30 帧未匹配的轨迹进行删除操作,实现轨迹的管理。

BoT-SORT 跟踪器通过稀疏光流估计相邻帧的相对运动,修正卡尔曼滤波器的预测结果,并通过相机运动补偿机制,有效缓解相机震动、目标快速运动或局部遮挡带来的累积误差,能显著提升跟踪的稳定性与连续性。

2.4 果实计数方法

为准确评估不同计数方法在实际果园环境中的适用性,选取 ID 累加计数法、划线计数法与区域计数法 3 种方法进行果实计数,如图 8 所示。

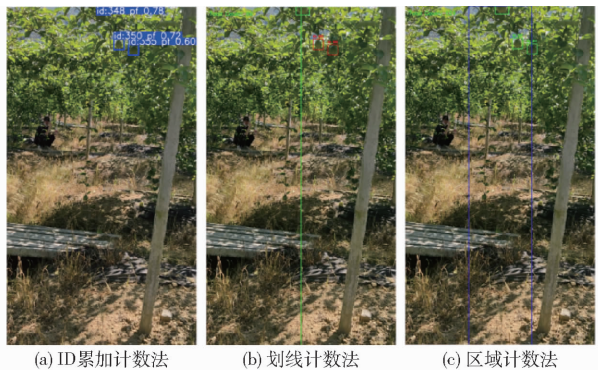


图 8 不同计数法计数示意图

Fig. 8 Diagram of different counting methods

图 8a 为 ID 累加计数法,其通过累加跟踪过程中所分配的最大 ID 值来估算果实总量,例如该帧中最大 ID 值为 355,说明当前累计检测到的果实数量为 355。图 8b 为划线计数法,其在视频画面中轴线设定 1 条绿色虚拟检测线,依据果实质心穿越该线的有效方向进行计数触发,绿色虚拟线左边的检测框表示已计数果实,右边的红色检测框则表示未计数果实,画面左上角显示当前累计数量为 98。图 8c 为区域计数法,其在视频帧中心定义 1 个矩形感兴

趣区域,仅在目标质心于该区域内持续停留多帧后方将其计入总数,画面左上角显示当前累计数量为 124。

2.5 模型训练

2.5.1 训练参数设置

试验在搭载 Intel 酷睿 i9-12900KF 处理器与 NVIDIA GeForce RTX 3090 显卡 (24 GB 显存) 的台式计算机上进行,操作系统采用 Ubuntu 20.04.6 LTS 版本。

在目标检测任务中,设置训练次数为 60,批处理大小为 16,同时选定 $conf_thres$ 为 0.2 与 iou_thres 为 0.5 的阈值组合。跟踪任务与计数任务直接继承检测阶段的置信度与交并比阈值配置,其余参数保持多目标跟踪算法的默认设定。

2.5.2 目标检测评价指标

为准确评估模型性能,采用精确率 (Precision)、召回率 (Recall)、平均精度均值 (mean Average Precision, mAP)、F1 分数、模型内存占用量对模型性能进行全面评价。其中,mAP 采用 $mAP@0.5$,表示当 IoU 阈值为 0.5 时的平均精度均值。

2.5.3 目标跟踪评价指标

在多目标跟踪评估方面,采用多目标跟踪领域广泛认可的 MOT 评价指标体系,包括跟踪精度 (Higher order tracking accuracy, HOTA)、跟踪准确率 (Multiple object tracking accuracy, MOTA)、跟踪调和平均数 (Identification F1 score, IDF1) 和 ID 切换数 (ID switch, IDSW) 等核心指标,以全面评估跟踪器的综合性能。

2.5.4 目标计数评价指标

在果实计数中,采用准确率 (Accuracy, Acc) 作为主要评估指标。

3 试验与结果

3.1 目标检测模型消融试验

为验证上下文引导模块 (CGBlock)、自适应注

意力模块 (LSKA) 与轻量级共享细节增强检测头 (LSDECD)对模型性能的协同影响,使用百香果数

据集对 YOLO 11n-CLL 进行消融试验,结果如表 3 所示。

表 3 消融试验结果

Tab.3 Results of ablation experiment

模型	CGBlock	LSKA	LSDECD	精确率/%	召回率/%	F1 分数/%	mAP@0.5/ %	模型内存占 用量/MB
YOLO 11n(基线)	×	×	×	80.0	79.7	79.8	85.9	5.5
YOLO 11n-C	√	×	×	80.4	80.4	80.4	86.2	4.7
YOLO 11n-L1	×	√	×	80.2	79.0	79.6	86.2	6.0
YOLO 11n-L2	×	×	√	80.0	80.9	80.4	86.6	5.2
YOLO 11n-CL	√	×	√	79.6	79.6	79.6	86.1	4.4
YOLO 11n-CLL	√	√	√	81.5	80.5	81.0	87.0	5.0

注:√表示采用此模块(方法),×表示不采用此模块(方法)。

由表 3 可知,引入 CGBlock 的 YOLO 11n-C,精确率、召回率、F1 分数和平均精度均值分别提升 0.4、0.7、0.6、0.3 个百分点,模型内存占用量降至 4.7 MB。引入 LSKA 的 YOLO 11n-L1 模型指标中,精确率和平均精度均值较基线模型分别提高 0.2、0.3 个百分点,但召回率和 F1 分数略有下降。引入 LSDECD 的 YOLO 11n-L2 模型,召回率较基线模型提升 1.2 个百分点,F1 分数提高 0.6 个百分点,平均精度均值提升 0.7 个百分点。将 CGBlock 与 LSDECD 结合后的 YOLO 11n-CL 模型,精确率、召回率和 F1 分数略有下降,但平均精度均值仍保持优化,模型内存占用量进一步降低,说明 LSDECD 在控制参数量的同时提升了特征的细粒度建模能力。将 3 种改进方法结合后的 YOLO 11n-CLL 模型,其精确率、F1 分数和平均精度均值在 6 个模型中最高(81.5%、81.0% 和 87.0%),证明了 YOLO 11n-CLL 模型在百香果识别任务中的准确性和综合性能均最优;模型内存占用量(5MB)虽较 YOLO 11n-CL 略大,但仍为 6 个模型中较少。消融试验结果证明 YOLO 11n-CLL 模型在百香果检测任务中性能最优,模型内存占用量较小,为未来轻量化部署提供了条件。

为验证改进后 YOLO 11n-CLL 模型在复杂环境下的检测性能,选取 3 幅具有代表性的百香果图像进行分析,结果如图 9 所示,其中图 9c 为光照不均、果实尺寸差异与遮挡的复合干扰场景。试验结果表明,在叶片遮挡条件下,YOLO 11n-CLL 模型仍能够较为准确地定位果实目标轮廓;在强光干扰场景中,模型有效抑制了背景噪声带来的误检现象;在多重干扰共存的复杂场景下,模型整体检测结果保持稳定,未出现明显漏检或误检情况。以上结果表明,该模型对复杂农业环境中的光照变化、目标尺度差异及遮挡干扰具有较强的适应能力。其稳定的检测能力,能为后续果园动态场景中跨帧目标关联

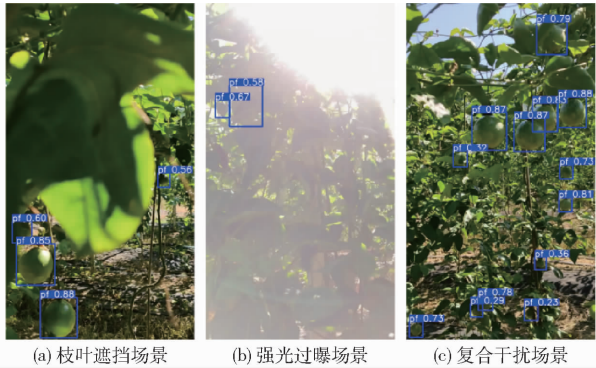


图 9 复杂果园环境下 YOLO 11n-CLL 检测可视化结果

Fig.9 Visualization of YOLO 11n-CLL in complex orchard environments

与果实跟踪计数提供可靠技术支撑。

3.2 目标检测模型对比试验

为验证改进模型性能,选用 YOLO v8n、YOLO v8n-P2、YOLO v10n、YOLO 11n 和 YOLO 11n-P2 主流检测模型与 YOLO 11n-CLL 在相同百香果检测数据集上进行了对比试验,结果如表 4 所示,其中,P2 版本为采用高分辨率特征层以提升小目标检测能力。

表 4 不同目标检测模型性能测试结果

Tab.4 Performance test results of different target detection models

模型	精确 率/%	召回 率/%	F1 分 数/%	mAP@ 0.5/%	模型内存 占用量/ MB
YOLO v8n	79.0	78.0	78.5	83.7	5.6
YOLO v8n-P2	77.4	78.2	77.8	82.6	5.9
YOLO v10n	79.0	78.6	78.8	84.7	5.7
YOLO 11n	80.0	79.7	79.8	85.9	5.5
YOLO 11n-P2	79.2	77.2	78.2	83.8	5.7
YOLO 11n-CLL	81.5	80.5	81.0	87.0	5.0

由表 4 可知,YOLO 11n-CLL 模型在多项指标上均表现优异,mAP@0.5 可达 87.0%,较基准模型 YOLO 11n 提升 1.1 个百分点,较同系列模型 YOLO

v8n、YOLO v8n-P2、YOLO v10n 以及 YOLO 11n-P2 分别提升 3.3、4.4、2.3、3.2 个百分点；精确率达 81.5%，相较基准模型 YOLO 11n、模型 YOLO v8n、YOLO v8n-P2、YOLO v10n 以及模型 YOLO 11n-P2 分别提升 1.5、2.5、4.1、2.5、2.3 个百分点。此外，模型在保持 80.5% 召回率的同时，内存占用量仅为 5.0 MB。结果表明，YOLO 11n-CLL 模型在精度与轻量化之间取得了良好平衡，其较低的内存占用量与良好的检测性能显示出在边缘设备中部署的实际潜力，为果园环境下实时检测任务提供了可行方案。

为验证各模型在复杂果园环境中的表现，选取 1 幅包含显著光照差异、多尺度果实及叶片遮挡的百香果图像作为测试图像，使用对比试验中训练好的模型进行检测，各模型测试输出图像如图 10 所示，其中红框表示漏检或误检目标。

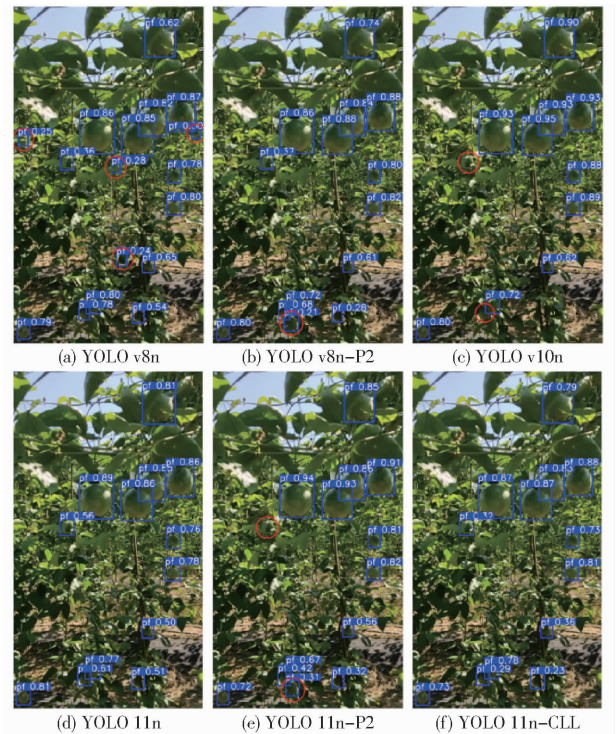


图 10 YOLO 11n-CLL 与其他检测器可视化结果
Fig. 10 Visualization results of YOLO 11n-CLL and other detectors

在所有模型中，YOLO 11n-CLL 取得最优检测效果，能够准确识别图像中全部果实。YOLO v8n 在样例图像左侧、右侧及中心区域均出现将叶片误识别为果实的情况；YOLO v8n-P2 在样例图像左下角仍存在误检问题；YOLO v10n 在样例图像左上与下侧区域分别漏检 1 个被遮挡果实和 1 个低光照区域果实。值得注意的是，YOLO 11n 与 YOLO 11n-CLL 模型在此测试样本中均未出现漏检或误检。然而，YOLO 11n-CLL 模型在保持高精度的同时具有

更小的内存占用量(5.0 MB)，使其在边缘端部署中更具优势。

3.3 百香果目标跟踪模型试验

采用 BoT-SORT^[19]、ByteTrack^[13]、OC-SORT^[28]、DeepOC-SORT^[29] 4 种方法进行 MOT 跟踪，结果如表 5 所示。

表 5 目标跟踪模型试验结果				
Tab. 5 Experiment results of target tracking models				
模型	HOTA/%	MOTA/%	IDF1/%	IDSW
BoT-SORT	62.4	68.4	73.0	1 102
ByteTrack	56.6	62.0	67.7	744
OC-SORT	60.8	66.0	70.8	2 105
DeepOC-SORT	60.2	65.6	71.2	403

由表 5 可知，BoT-SORT 算法在多项核心指标上均展现出最优性能，其 HOTA、MOTA、IDF1、IDSW 分别为 62.4%、68.4%、73.0%、1 102，前 3 项核心指标均显著领先于其他对比算法。

与 ByteTrack、OC-SORT、DeepOC-SORT 算法对比，HOTA 分别上升 5.8、1.6、2.2 个百分点，MOTA 分别上升 6.4、2.4、2.8 个百分点，IDF1 分别上升 5.3、2.2、1.8 个百分点，在 IDSW 上，比 OC-SORT 减少 1 003 次。综合评估表明，尽管 BoT-SORT 在 IDSW 指标上略逊于 ByteTrack 和 DeepOC-SORT，但在 HOTA、MOTA 和 IDF1 3 项关键跟踪指标上均优于所有对比算法，这充分证明了其在百香果园的目标跟踪应用中的优越性能。

为系统评估跟踪算法在复杂果园环境中的稳定性，选取了包含密集遮挡与光照不均等挑战的测试帧进行可视化对比试验。如图 11 所示，可视化对比图集包含人工标注真值与 4 种算法(BoT-SORT、ByteTrack、DeepOC-SORT、OC-SORT)在视频序列 10 第 289 帧与视频序列 21 第 874 帧的跟踪结果。可视化结果显示，BoT-SORT 在 2 个测试帧中均实现了全部百香果目标，表现出较强的抗遮挡与稳定性；ByteTrack 在 2 处测试均表现出较多漏检(在第 289 帧为 2 例，在第 874 帧为 4 例)，且伴随检测框尺寸不匹配的错误；DeepOC-SORT 在 2 个测试帧中均出现了 2 例因目标重叠导致的漏检；OC-SORT 在视频序列 10 第 289 帧同样完成了全部目标的识别，但在视频序列 21 第 874 帧受到颜色干扰，将叶片误检为果实；综上，BoT-SORT 的相机运动补偿机制使其在复杂果园环境中表现最为稳定。

3.4 果实计数方法试验

表 6 为 3 种计数方法在测试视频中的准确率结果。由表 6 可知，区域计数法在计数准确率方面

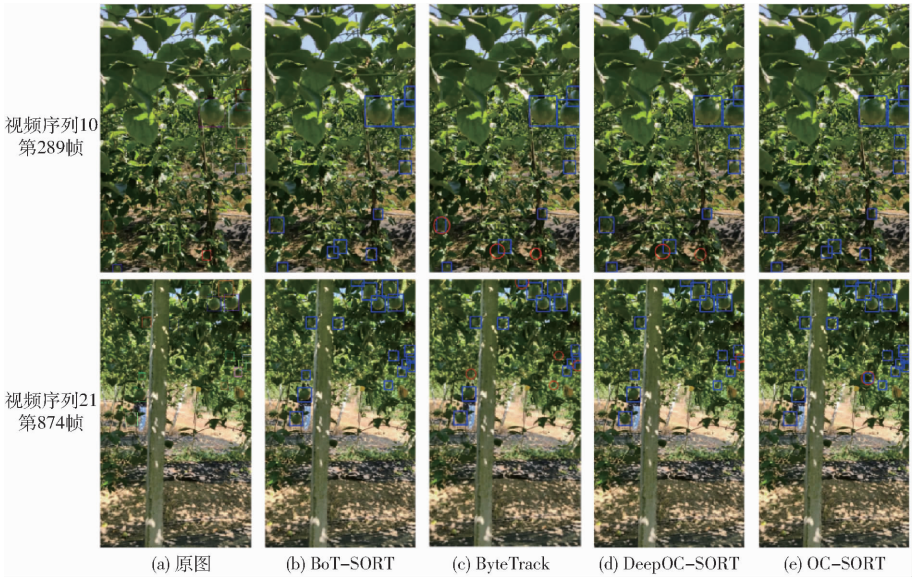


图 11 BoT-SORT 与其他跟踪器可视化结果

Fig. 11 Visualization of BoT-SORT and other tracker

表 6 果实计数方法试验结果

Tab. 6 Experiment results of fruit counting methods

				%
视频序列	区域计数法	划线计数法	ID 累加计数法	
10	90. 4	88. 1		65. 7
21	98. 9	78. 6		74. 6
41	93. 8	84. 3		67. 8
51	94. 0	89. 3		69. 9
61	89. 4	70. 6		74. 8
71	88. 0	85. 2		42. 9
80	95. 4	89. 3		50. 3
90	92. 1	86. 6		56. 8
所有	92. 8	84. 0		62. 9

始终优于其他 2 种方法,在所有测试序列中平均准确率为 92.8%,较划线计数法提升 8.8 个百分点,较 ID 累加计数法提升 29.9 个百分点。此外,区域计数法在每个单独视频中均取得最高准确率。

为了对比分析划线计数法与区域计数法的性能差异,对视频序列 10 第 229~234 帧进行可视化,结果如图 12 所示,其中红色箭头标识的是 ID 号为 75 的百香果目标。区域计数法在这 3 帧中都取得正确结果,如图 12 第 2 行。划线计数法出现错误计数,如图 12 第 1 行,其 ID75 目标在第 229 帧时因短时遮挡导致跟踪中断,第 230 帧重新捕获后,划线计数法因目标已穿越检测边界未触发单次计数。相较之下,区域计数法通过持续监测机制实现稳健计数。结果表明,区域计数法通过引入时间连续性约束条件,有效缓解了因目标短时丢失产生的计数误差,更适用于目标遮挡频发的自然农业场景。

3.5 百香果检测、跟踪与计数可视化系统

为验证该方法在资源受限场景中的潜在适用性,

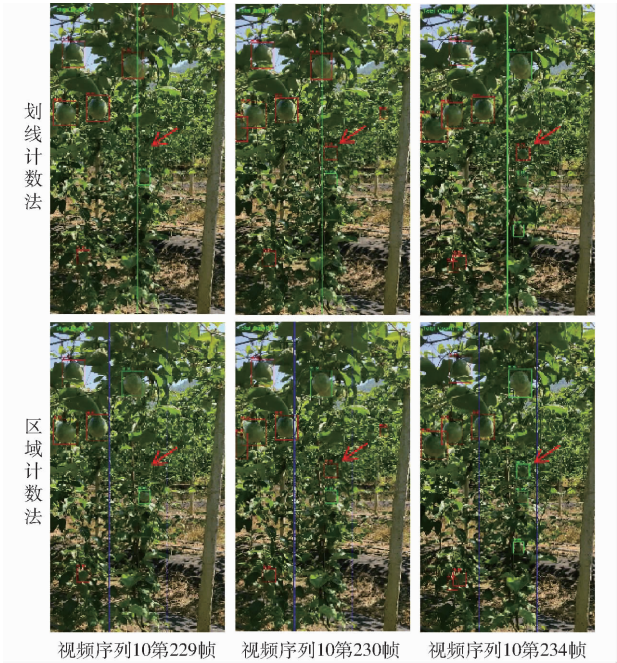


图 12 区域计数法与划线计数法可视化结果

Fig. 12 Visualization results of region based counting method and line based counting method

在边缘端设备上构建百香果检测、跟踪和计数的应用系统。边缘端设备的配置为 AMD Ryzen 3 2200U 处理器与 NVIDIA GeForce 940MX 显卡(2 GB 显存),硬件整体计算性能较低,用于模拟资源受限的边缘计算环境。该应用系统基于改进的 YOLO 11n-CLL 检测模型与 BoT-SORT 多目标跟踪算法,对输入视频序列中的百香果目标进行多目标跟踪,利用基于区域的果实计数策略,实现对果实的稳定统计,获得准确的果实数量。模型处理完成后,系统以结构化形式展示果实总数及对应的性能指标。通过该系统,用户能够直观观察果实的跟踪与计数过程,并对模型精度与运

行效率进行综合分析。

果实自动计数系统计数的可视化界面如图 13 所示。界面左侧为计数完成后的百香果园视频(以视频序列 80 为例),右侧为计数操作按钮及结果展示区域。试验结果表明,在该视频中共检测到 406 个百香果,系统平均运行帧率为 19.6 f/s,总处理时长为 26.1 s,峰值显存占用约为 192 MB。上述结果表明,所提方法在资源受限的实际边缘端计算环境下仍能保持较好的运行效率与较低的资源消耗,具备良好的应用部署可行性。



图 13 百香果视频检测、跟踪与计数系统可视化界面

Fig. 13 Visualization interface for passion fruit video detection, tracking, and counting

4 结论

(1)提出了一种轻量化的 YOLO 11n-CLL 目标检测方法。该方法对 YOLO 11n 模型在骨干网

络、特征融合与检测头 3 个层面进行协同优化,通过骨干网络引入 CGBlock,增强了上下文感知能力;同时利用 LSKA 模块提高了光照特征鲁棒性;在头部网络采用 LSDECD 模块增强了对小目标和遮挡目标的识别精度。YOLO 11n-CLL 方法提升了模型在复杂果园环境中的性能。

(2)在百香果果实检测任务中,试验结果表明, YOLO 11n-CLL 模型精确率、召回率、F1 分数和 mAP@ 0.5 分别为 81.5%、80.5%、81.0% 和 87.0%,与原始 YOLO 11n 模型相比,精确率、召回率、F1 分数和 mAP@ 0.5 各提升 1.5、0.8、1.2、1.1 个百分点,模型内存占用量为 5.0MB。在保持精度的同时降低了模型对硬件资源的需求。

(3)提出 BoT-SORT 跟踪器和区域计数法算法实现果实的稳定跟踪和计数。BoT-SORT 跟踪器利用运动补偿机制,改善了目标轨迹的稳定性,其 HOTA、MOTA 和 IDF1 分别为 62.4%、68.4%、73.0%。对跟踪结果,采用 ID 累加计数法、划线计数法和区域计数法 3 种方法实现视频的果实自动计数,通过 3 种方法试验结果验证对比,区域计数法优于其他计数方法,本文选择采用区域计数法进行果实自动计数,其平均计数准确率为 92.8%,具备在实际果园环境中部署应用的潜力,为果实采摘智能化管理提供了可行的技术方案。

参 考 文 献

[1] BISWAS S, MISHRA R, BIST A S. Passion to profession; a review of passion fruit processing[J]. Aptisi Transactions on Technopreneurship (ATT), 2021, 3(1): 48-57.

[2] DEDIEU B, CONTZEN S, NETTLE R, et al. The multiple influences on the future of work in agriculture: global perspectives[J]. Frontiers in Sustainable Food Systems, 2022, 6: 889508.

[3] 张孙健, 吴应海, 彭熙, 等. 从江县百香果产业发展存在问题与建议[J]. 耕作与栽培, 2023, 43(6): 145-147.

ZHANG Sunjian, WU Yinghai, PENG Xi, et al. Challenges and suggestions for the development of passiflora edulis industry in Congjiang County[J]. Tillage and Cultivation, 2023, 43(6): 145-147. (in Chinese)

[4] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[5] 陈媛媛, 游炯, 幸泽峰, 等. 世界主要国家精准农业发展概况及对中国的发展建议[J]. 农业工程学报, 2021, 37(11): 315-324.

CHEN Yuanyuan, YOU Jiong, XING Zefeng, et al. Review of precision agriculture development situations in the main countries in the world and suggestions for China[J]. Transactions of the CSAE, 2021, 37(11): 315-324. (in Chinese)

[6] LI H, GU Z, HE D, et al. A lightweight improved YOLOv5s model and its deployment for detecting pitaya fruits in daytime and nighttime light-supplement environments[J]. Computers and Electronics in Agriculture, 2024, 220: 108914.

[7] 宋怀波, 马宝玲, 尚钰莹, 等. 基于 YOLO v7-ECA 模型的苹果幼果检测[J]. 农业机械学报, 2022, 54(6): 233-242.

SONG Huaibo, MA Baoling, SHANG Yuying, et al. Detection of young apple fruits based on YOLO v7-ECA model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 54(6): 233-242. (in Chinese)

[8] XIAO F, WANG H, XU Y, et al. Fruit detection and recognition based on deep learning for automatic harvesting: an overview and review[J]. Agronomy, 2023, 13(6): 1625.

[9] ZHANG Y, LIANG Y, LENG J, et al. SCGTracker: spatio-temporal correlation and graph neural networks for multiple object tracking[J]. Pattern Recognition, 2024, 149: 110249.

[10] HERNANDEZ B, MEDEIROS H. Multi-object tracking in agricultural applications using a vision transformer for spatial association[J]. Computers and Electronics in Agriculture, 2024, 226: 109379.

[11] 李景顺, 刘美, 孟亚男, 等. 应用轻量化 FEB-YOLO 模型的荔枝果实动态识别计数方法[J]. 计算机测量与控制, 2025, 33(2): 229-237, 261.

LI Jingshun,LIU Mei,MENG Yanan,et al. Dynamic recognition and counting method for litchi fruit using lightweight FEB – YOLO model[J]. Computer Measurement & Control,2025,33(2):229 – 237,261. (in Chinese)

[12] WOJKE N,BEWLEY A,PAULUS D. Simple online and realtime tracking with a deep association metric[C] // 2017 IEEE International Conference on Image Processing (ICIP),2017:3645 – 3649.

[13] ZHANG Y,SUN P,JIANG Y,et al. Bytetrack: multi-object tracking by associating every detection box[C] // European Conference on Computer Vision,2022:1 – 21.

[14] TU S,HUANG Y,LIANG Y,et al. A passion fruit counting method based on the lightweight YOLOv5s and improved DeepSORT [J]. Precision Agriculture,2024,25(3):1731 – 1750.

[15] 黄成龙,张忠福,华向东,等. 基于改进 Faster R – CNN 和 Deep Sort 的棉铃跟踪计数[J]. 农业机械学报,2022,54(6):205 – 213.

HUANG Chenglong,ZHANG Zhongfu,HUA Xiangdong,et al. Cotton boll tracking and counting based on improved Faster R – CNN and Deep Sort[J]. Transactions of the Chinese Society for Agricultural Machinery,2022,54(6):205 – 213. (in Chinese)

[16] TANG Y, QIU J, ZHANG Y, et al. Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment;a review[J]. Precision Agriculture,2023,24(4):1183 – 1219.

[17] DONG L, CHEN L, WU Z C, et al. Video stabilization-based elimination of unintended jitter and vibration amplification in centrifugal pumps[J]. Mechanical Systems and Signal Processing,2025,229:112500.

[18] 杨洋,宋品德,钟春来,等. 无人机视角下基于深度学习的多目标跟踪研究进展[J]. 计算机工程与应用,2023,59(23):48 – 62.

YANG Yang,SONG Pinde,ZHONG Chunlai,et al. Research progress of multi-target tracking based on deep learning from perspective of UAV[J]. Computer Engineering and Applications,2023,59(23):48 – 62. (in Chinese)

[19] AHARON N,ORFAIG R,BOBROVSKY B Z. Bot-sort:robust associations multi-pedestrian tracking[J]. arXiv preprint arXiv:220614651,2022.

[20] 杨莹莹,唐朝,陈洁,等. 智慧果园发展现状及对策[J]. 农业工程,2025,15(2):44 – 48.

YANG Yingying,TANG Chao,CHEN Jie,et al. Development status and countermeasures of smart orchard [J]. Agricultural Engineering,2025,15(2):44 – 48. (in Chinese)

[21] AHMAD J,GUEAIEB W,ELSADDIK A,et al. Yield estimation and health assessment of temperate fruits;a modular framework [J]. Engineering Applications of Artificial Intelligence,2024,136:108871.

[22] 王金鹏,何萌,甄乾广,等. 基于改进 COF – YOLO v8n 的油茶果静态与动态检测计数方法[J]. 农业机械学报,2024,55(4):193 – 203.

WANG Jinpeng,HE Meng,ZHEN Qianguang,et al. *Camellia oleifera* fruit static and dynamic detection counting based on improved COF – YOLO v8n[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(4):193 – 203. (in Chinese)

[23] PARICO A I B,AHAMED T. Real time pear fruit detection and counting using YOLOv4 models and deep SORT[J]. Sensors,2021,21(14):4803.

[24] TU S,HUANG Y,HUANG Q,et al. Estimation of passion fruit yield based on YOLOv8n + OC – SORT + CRCM algorithm[J]. Computers and Electronics in Agriculture,2025,229:109727.

[25] AN Q,CUI Y,TONG W,et al. StraTracker:a dynamic counting method for growing strawberries based on multi-target tracking [J]. Computers and Electronics in Agriculture,2024,227:109564.

[26] LAU K W,PO L M,REHMAN Y A U. Large separable kernel attention:rethinking the large kernel attention design in CNN [J]. Expert Systems with Applications,2024,236:121352.

[27] WANG X,GAO H,JIA Z. Improved road defect detection algorithm based on YOLOv8[J]. International Applied Computational Electromagnetics Society Symposium,2024,13:2413.

[28] CAO J,PANG J,WENG X,et al. Observation-centric sort:rethinking SORT for robust multi-object tracking[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition,2023:9686 – 9696.

[29] MAGGIOLINO G,AHMAD A,CAO J,et al. Deep OC-SORT:multi-pedestrian tracking by adaptive re-identification[C] //2023 IEEE International Conference on Image Processing (ICIP),2023:3025 – 3029.