

基于手持式视频流及YOLOv8n的目标跟踪柑橘计数估产方法研究

斯乃瑾^{1,2}, 蔡健荣^{1*}, 王 陈¹, 朱文辉¹, 潘丙科¹, 石治国¹

(¹江苏大学食品与生物工程学院, 江苏镇江 212000; ²江苏大学农业工程学部, 江苏镇江 212000)

摘要:【目的】构建一套基于视频流的自动估产模型, 系统评估光照条件与拍摄角度对模型识别准确率的影响。【方法】以果实膨大期柑橘树为对象, 使用环绕旋转方式采集视频数据。利用改进的YOLOv8n-SC检测模型, 增加小目标检测头、引入CBAM注意力机制及优化损失函数, 提升对小尺寸和低对比度果实的识别能力; 结合DeepSORT算法实现果实的持续跟踪与计数, 并对比不同光照(自然光/人工光源)和拍摄角度(俯拍/平摄)条件下的模型性能。【结果】改进的YOLOv8n-SC模型在检测任务中召回率与mAP₅₀分别达到97.6%和93.7%; 在人工光源结合平摄条件下模型表现最优, 多目标跟踪准确度与跟踪精度分别为90.8%和94.0%, 平均计数精度(ACP)达89.46%, 可有效实现单株柑橘果实的自动化统计。【结论】本研究提出的基于环绕式采集视频流数据的估产系统能够适应复杂果园环境, 实现高精度柑橘计数。人工光源与平摄相结合为最优数据采集策略, 为柑橘产量自动化预测提供了可靠的技术支撑。

关键词: 柑橘; 果实膨大期; 旋转拍摄; 果实估产; YOLOv8n-SC; DeepSORT

中图分类号: S666

文献标志码: A

文章编号: 1009-9980(2026)05-1283-14

Citrus yield estimation based on multi-object tracking in video streams using an improved YOLOv8n model

SI Naijin^{1,2}, CAI Jianrong^{1*}, WANG Chen¹, ZHU Wenhui¹, PAN Bingke¹, SHI Zhiguo¹

(¹School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212000, Jiangsu, China; ²School of Agricultural Equipment Engineering, Jiangsu University, Zhenjiang 212000, Jiangsu, China)

Abstract: 【Objective】Accurate citrus yield estimation is a critical component of precision agricultural management, essential for optimizing production and ensuring economic benefits for growers. Traditional manual sampling methods are labor-intensive, time-consuming, and prone to subjective errors, failing to meet the demands of modern horticulture. While deep learning-based object detection has shown promise in fruit recognition from static images, its application in real-world orchard environments is hindered by challenges such as fruit occlusion, varying illumination, and perspective limitations, which lead to counting inaccuracies. Most existing video-based methods rely on linear, inter-row scanning, a technique suitable for trellis-trained orchards but ineffective for citrus trees, which typically grow in a standalone, three-dimensional crown shape. To address these limitations, this study aimed to develop and validate a robust and automated yield estimation system for individual citrus trees during the fruit expansion period. The primary objective was to establish an end-to-end pipeline that combines a novel data acquisition strategy—360° rotational video capture—with an improved object detection model (YOLOv8n-SC) and a multi-object tracking algorithm (DeepSORT). The research scope included a systematic investigation into the influence of different environmental factors, specifically lighting conditions (sunny, overcast, and artificial light) and camera perspectives (high-angle, straight-angle, and low-

收稿日期: 2025-09-10

接受日期: 2025-11-05

基金项目: 国家自然科学基金项目(51975259); 国家重点研发计划(2022YFD2002204)

作者简介: 斯乃瑾, 女, 在读硕士研究生, 研究方向为机器学习。E-mail: sinaijin_765@163.com

*通信作者 Author for correspondence. E-mail: jrcai@ujs.edu.cn

angle), on the overall accuracy and stability of the yield estimation system. **【Methods】** The methodology encompassed data acquisition, model development, and a multi-stage experimental evaluation. For data acquisition, a handheld imaging system was used to capture 360° rotational video streams around 20 individual citrus trees in a high-density orchard under three distinct lighting conditions and three different vertical camera angles, resulting in 180 video clips. A comprehensive dataset of 9000 images was created by extracting frames, which were then meticulously annotated and partitioned into training (80%) and testing (20%) sets. For model development, the baseline YOLOv8n was enhanced to create the YOLOv8n-SC model by implementing three key improvements: integrating a fourth, high-resolution (160×160) detection head to improve sensitivity to small targets; embedding the Convolutional Block Attention Module (CBAM) within the backbone network to enhance feature distinction between fruits and leaves; and replacing the original CIoU loss function with the Wise-IoU (WIoU) v3 loss function to accelerate convergence and improve bounding box precision. For tracking and counting, the DeepSORT algorithm was employed to process detection outputs from the video stream. Its workflow involves a Kalman filter for motion prediction, a cascaded matching strategy combining motion and appearance features for reliable identity association, and a track management system to assign unique IDs. The final fruit count was determined by the maximum unique ID assigned during the video sequence. **【Results】** The proposed YOLOv8n-SC model significantly outperformed other contemporary models. On the comprehensive dataset, it achieved a recall of 97.6% and a mean Average Precision at an IoU threshold of 0.5 (mAP_{50}) of 93.7%, marking improvements of 2.0 and 3.5 percentage points over the baseline YOLOv8n, respectively. It also showed clear advantages over Faster R-CNN and the lightweight YOLOv10n. A detailed ablation study confirmed that each of the three modifications—the additional detection head, the CBAM module, and the WIoU loss function—contributed positively to the model's overall performance. Lighting proved to be a critical factor, as the artificial light source yielded the best detection results with an mAP_{50} of 98.2%, markedly superior to overcast (96.1%) and sunny (83.5%) conditions. The camera angle also had a significant impact, with the straight-angle perspective providing the most balanced and comprehensive view, achieving the highest Multi-Object Tracking Accuracy (MOTA) of 90.8% and a high Multi-Object Tracking Precision (MOTP) of 94.0%. When comparing tracking algorithms, the combination of YOLOv8n-SC with DeepSORT provided the most accurate counts. For the optimal configuration (straight-angle video under artificial light), the system achieved an Average Counting Precision (ACP) of 89.46%, with a Root Mean Square Error (RMSE) of 10.83 and a Mean Absolute Error (MAE) of 9.98. This level of accuracy was superior to that achieved when pairing YOLOv8n-SC with either the SORT algorithm (ACP of 84.81%) or the Tracktor algorithm (ACP of 86.37%). **【Conclusion】** This study successfully demonstrates the viability of an automated citrus yield estimation system based on rotational video streams and deep learning. The specifically developed YOLOv8n-SC object detector proves highly effective for identifying green citrus fruits in complex orchard environments. The complete pipeline, integrating this detector with the DeepSORT tracking algorithm, provides reliable fruit counts for individual trees. Key findings indicate that data acquisition conditions are paramount; a stable, artificial light source is optimal for detection, and a straight-angle camera trajectory is superior for comprehensive tracking. The optimized system achieved a final counting accuracy of 89.46%, providing a strong technical foundation and a practical reference for the implementation of precision yield forecasting in modern citrus cultivation. This research makes two primary contributions to the field. First, it introduces a novel data acquisition paradigm—360° rotational video capture—tailored specifically for fruit trees with complex 3D crown structures, offering a

more comprehensive alternative to conventional linear scanning methods. Second, it presents a purpose-built detector, YOLOv8n-SC, whose targeted enhancements are systematically validated to address the specific challenges of detecting small, low-contrast fruits in visually cluttered agricultural settings.

Key words: Citrus; Fruits expanding-period; Rotational imaging; Fruit yield estimation; YOLOv8n-SC; DeepSORT

中国是世界主要的柑橘生产国,对中国的柑橘进行准确的产量预测是实现精准化农业管理、保障果农经济效益的关键环节^[1]。传统的人工抽样估产方法主观性强、效率低下,难以满足现代柑橘产业对精准化管理的需求^[2]。因此,研究自动化、高精度的估产技术具有重要的现实意义。

目标检测算法能在复杂环境中准确地识别、定位水果,是实现准确估产的基础^[3]。传统的目标检测算法包括图像分割、颜色分割等^[4],易受自然环境中遮挡、光照等的干扰^[5],通用性差。近年来,以深度学习为基础的目标检测算法^[6],特别是YOLO系列模型,检测精度高,已在静态图像的果实识别中得到广泛应用^[7];吕强等^[8]提出改进的YOLOv5s算法,实现柑橘果实的快速、准确检测,准确率达到97.2%,推理时长减少85.8%;Wu等^[9]基于YOLOv4-tiny的目标检测算法,结合阈值分割算法消除地面与后排的水果,将检测精度从89.1%提高到93.3%。

基于深度学习的目标检测算法研究仍多集中于单帧静态图像的果实识别,然而,在面对田间复杂的实际环境时,难以解决因果实间相互遮挡、拍摄视角变化导致的重复计数或漏检问题,从而限制了其在精准计数应用中的潜力。为克服此局限性,本研究采用基于视频流的多目标跟踪技术^[2]。传统的目标跟踪包括光流^[10]和帧差^[11],无法应对光照变化和背景干扰;基于激光雷达的目标跟踪方法采集高精度的三维空间数据,能够捕捉目标的形状和位置^[12],但其依赖昂贵的硬件设备,且存在计算复杂、处理周期长等局限,难以部署。基于深度学习的多目标跟踪技术,得益于深度学习本身泛化能力强、高效准确、易部署的特点,在农业领域展现出巨大潜力,不仅能够实现自动化估产,还能应用于成熟度判别、机械化采摘等智能化场景^[13]。张润池等^[14]为了稳定跟踪视频中的番茄目标,提出了超深掩蔽处理,将平均精度提高了12.63%;Li等^[15]基于ROI感兴趣区域,结合改进的ByteTrack算法对火龙果花、绿果、红果分别进行统计,平均识别精度为95.0%。

现有的视频数据采集以行间平移拍摄为主^[14-16],适用于篱笆式种植的果树,但对于普遍呈独立冠状生长的柑橘树,因固有视角的局限性仍无法有效解决遮挡问题。本文提出一种新颖的环绕单株果树旋转拍摄的数据采集方法,并在此基础上构建了一套结合深度学习目标检测与多目标跟踪的估产模型,旨在实现柑橘的自动化精准计数。

1 材料和方法

1.1 数据来源

试验于9月下旬在江苏省镇江市红美人柑橘园(119°58' E, 32°21' N)进行。由于江苏地区不属于理想的柑橘种植带,气候偏寒,所以较晚进入果实膨大期。果树采用高密度种植,行距2.5~3.0 m,株距2.0~2.5 m^[17]。单株果树高约2.0 m,冠幅饱满、叶幕稠密,其内部枝条结构复杂,叶片与果树空间分布高度重叠,形成了复杂的多层冠层结构。

1.2 数据获取

将运动相机和大功率LED光源安装在伸缩杆(最大1.5 m)上,构成手持式摄像系统,选用大疆OSMO Action4自动防抖的运动相机获取视频流原始数据,选用白激光光源,光通量约1200流明,色温约为6000 K,使用强光稳态照明模式,如图1-a所示。将手持式摄像系统绕树旋转360°拍摄,实现果树的全表面图像采集。试验条件包括不同光照条件(晴天、阴天、人工光源^[18])和拍摄角度(俯摄、平摄、仰摄)。拍摄过程中,以树冠平均高度H为参照,将相机与树干中心点的距离(旋转半径)控制在1.2~1.5 m之间,具体数值根据树冠宽度与株距动态调整。基于预试验的结果分析,最终选定30°为夹角,这在避免冠层遮挡和实现手持式操作性之间取得最佳平衡。基于H高度,水平拍摄称为平摄;以30°为夹角的俯视拍摄称为俯摄;以30°为夹角的仰视拍摄称为仰摄(如图1-b所示)^[19]。拍摄前在地面标记圆形路径,预先将相机与伸缩杆的夹角调节至对应角度,并通过云台固定,操作人员以标准化持握方式沿

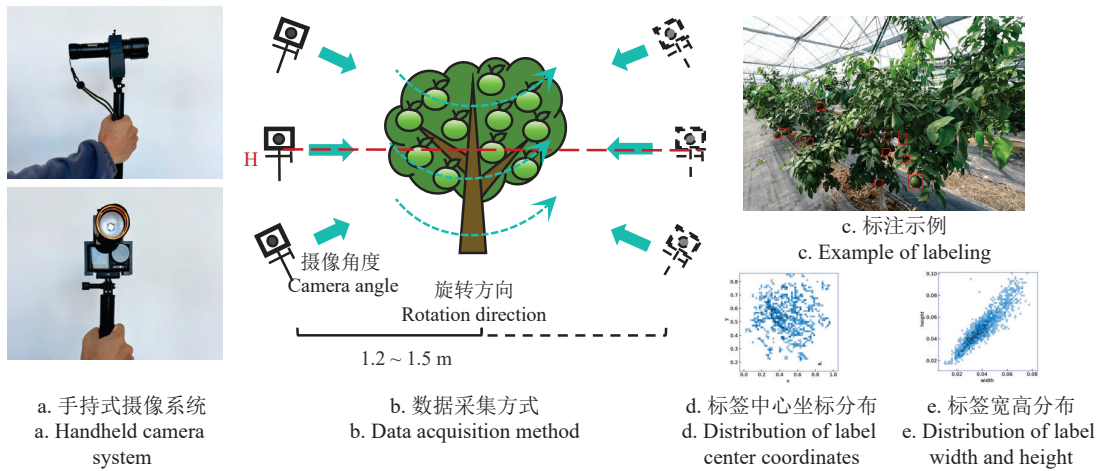


图 1 数据采集与标注
Fig. 1 Data acquisition and labeling

该路径匀速移动,实时微调保证杆身与地面垂直。由此可有效控制位置与角度的波动,将系统误差维持在可容错范围内。

根据不同光照、视角,共形成9组排列组合,如图2所示。20棵果树共计视频180个,以MP4格式保存,分辨率为 3840×2880 ,视频帧率为 $30 \text{ 帧} \cdot \text{s}^{-1}$ 。经过视频处理去除拍摄重复部分,每个视频占时约30 s。数据集既要保证图像间独立性,也要提取足

够数量的有效帧,因此,每隔18帧自动提取一幅图像,共获得柑橘图像9000幅。晴天、阴天、人工光源数据集各含3000幅图像。利用AnyLabeling标注工具对数据集进行人工标注,在标注过程中仅对当前树上人工可识别或可推测出果实轮廓的绿色柑橘进行标注,对于掉落果实、后排果实、图像边缘果实、树上裂果不进行标注^[20],标注示例如图1-c所示。在标注完成后,进行多轮审核,确保标注的准确性。图1-



图 2 9组数据示例
Fig. 2 Data examples of nine groups

d和图1-e分别给出了柑橘目标的标签中心坐标分布以及标签宽高分布,用以表征数据集中目标位置与尺度的整体分布情况。将数据集按照4:1的比例随机划分为训练集和测试集。

1.3 基于改进YOLOv8n的果实检测

YOLO(you only look once)系列的模型尺寸较小、灵活性强且准确率高,能够在单次前向传播中完成目标检测任务^[21]。然而,受遮挡、光照影响时仍会出现漏检或错检。为克服这些不足,提出基于YOLOv8n的小目标检测算法YOLOv8n-SC。

1.3.1 增加检测头 在YOLOv8n模型中,头部输出端(Head)负责目标检测和分类任务,默认有3个检测头(Detect),分别用于大型目标、中型目标和小型目标检测^[22]。柑橘果实在图像中的占比很小,其标签框面积与图像面积之比常低于0.1%。由于微小目标本身检测难度大,且易受遮挡、模糊及颜色失真等因素干扰,导致原有检测头对果实的识别性能不佳。为了解决这些问题,新增的检测头配备独特的锚框和预测层,针对微小目标的特征进行优化。如图3所示,增加检测特征图大小为160×160的小

目标输出层,该检测头与原有的3个检测头协同工作,形成1个更为强大的检测模块^[23]。由于采用分辨率更高的特征图,小目标在特征图上占据的像素比例增大,配合多个检测头并行进行柑橘目标检测,可在一定程度上减轻下采样过程中细节信息的损失,降低柑橘小目标的漏检率。

1.3.2 CBAM注意力机制 CBAM(convolutional block attention module)是一个轻量级通用模块,可以无缝集成到任何卷积神经网络模型的架构中,开销可以忽略不计^[24]。为解决膨大期柑橘果实与绿色叶片颜色相近、特征不易区分的难题,在骨干网络的C2f模块中引入CBAM注意力机制,其结构如图4所示。

CBAM包含通道注意力模块(channel attention module, CAM)和空间注意力模块(spatial attention module, SAM)两部分,分别关注目标的特征信息和位置信息^[25]。在YOLOv8n模型中,随着层次的深入,特征通道数逐渐增加,通道拼接时,会产生信息冗余。CAM通过动态加权,强化与果实检测高度相关的特征通道(色度、纹理),同时抑制与背景叶片相

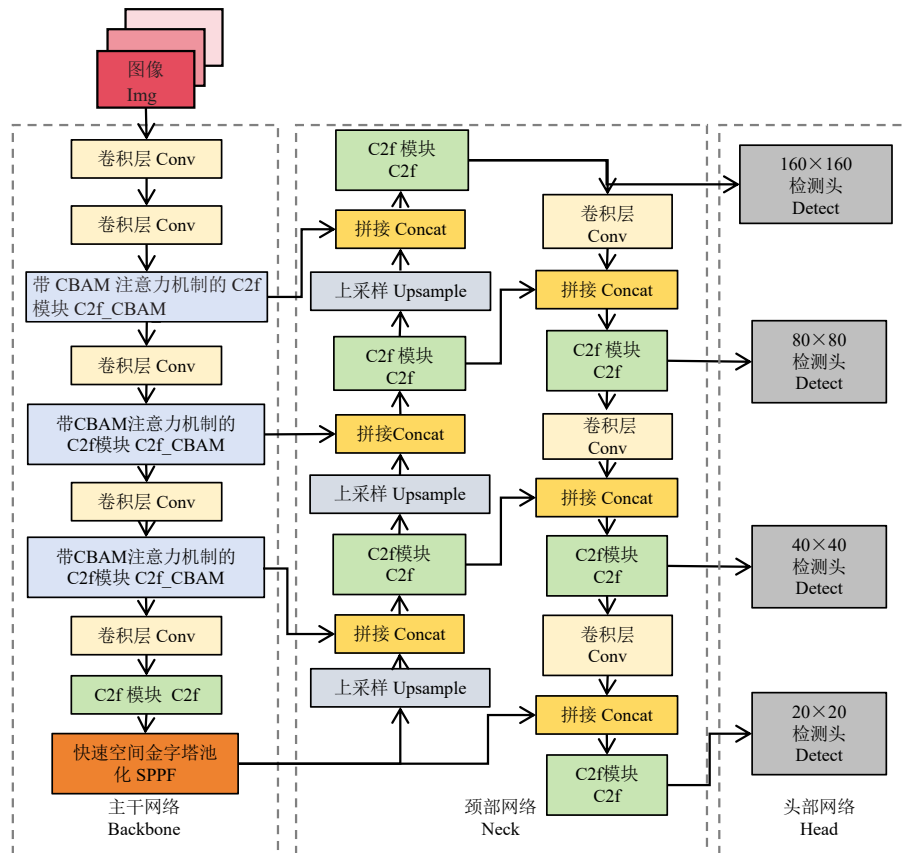


图3 YOLOv8n-SC 结构模型

Fig. 3 The structure of YOLOv8n-SC model

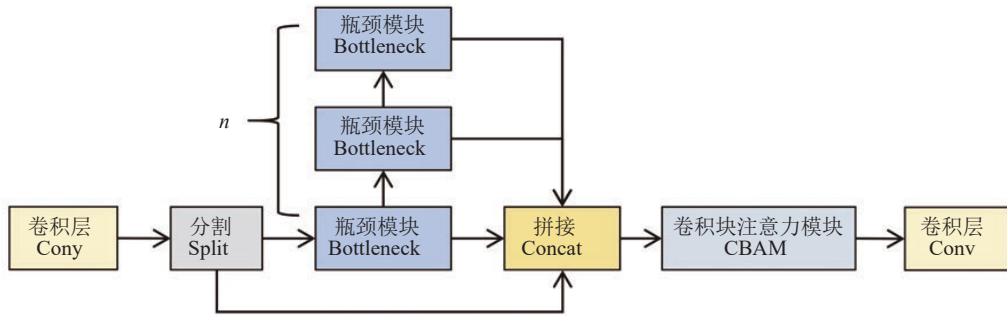


图 4 C2f_CBAM 结构

Fig. 4 The structure of C2f_CBAM

关或信息量低的通道,从而增强模型对关键特征的选择性感知^[24]。SAM通过生成空间权重图,更关注关键的局部信息,使模型聚焦于果实所在的局部空间位置,尤其是果实的边缘、轮廓等可鉴别性区域,有效抑制叶片背景等无关区域的噪声干扰,增强模型对果实目标的精确定位能力^[24]。

1.3.3 WIoU 损失函数 在 YOLOv8n 网络预测边框坐标损失 L_{CIoU} 采用的是完全交并比^[26](complete intersection over union, CIoU),通过将预测框与真实框的距离和尺寸引入到惩罚项中,使得模型收敛速度有所改善,如式(1)。

$$L_{\text{CIoU}} = L_{\text{IoU}} + \frac{(x-x_{\text{gt}})^2 + (y-y_{\text{gt}})^2}{W_g^2 + H_g^2} + av. \quad (1)$$

其中:

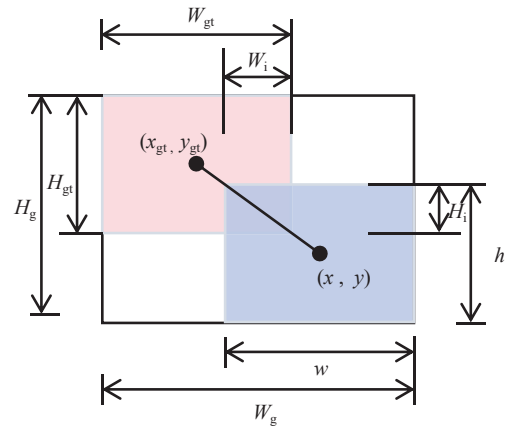
$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{\text{gt}}}{h_{\text{gt}}} - \arctan \frac{w}{h} \right)^2, \quad (2)$$

$$a = \frac{v}{(1-\text{IoU}) + v}, \quad (3)$$

$$L_{\text{IoU}} = 1 - \text{IoU} = 1 - \frac{W_i H_i}{wh + w_{\text{gt}} h_{\text{gt}} - W_i H_i}. \quad (4)$$

式中, L_{CIoU} 表示完全交并比损失函数, L_{IoU} 为交并比损失项; IoU 表示交并比(intersection over union),如图5所示; a 为权重参数,用于调节宽高比一致性项在损失计算中的贡献; v 为宽高比一致性度量项,用于衡量预测框与真实框长宽比之间的差异。

在检测过程中,柑橘与背景颜色相似,遮挡严重,识别困难。为了提高精度并加快边界框回归速度,采用动态非单调聚焦机制的加权交并比(weighted intersection over union, WIoU)替换原损失函数。其损失函数如式(5),定义离群度参数 β 以描述锚框的异常程度;参数 α 与 δ 调节样本在损失计算中的相对权重^[27];非单调聚焦系数 r 通过聚焦于普通质量的锚框,降低高质量样本对损失值的贡献率,



$w, h, (x, y)$ 分别表示预测框的宽、高尺寸和中心坐标; $W_{\text{gt}}, H_{\text{gt}}, (x_{\text{gt}}, y_{\text{gt}})$ 分别表示真实框的宽、高尺寸和中心坐标; W_i, H_i 分别表示交集宽、高尺寸; W_g, H_g 分别表示最小边框宽、高尺寸。

w, h , and (x, y) denote the width, height, and center coordinates of the predicted bounding box, respectively; $W_{\text{gt}}, H_{\text{gt}}$, and $(x_{\text{gt}}, y_{\text{gt}})$ denote the width, height, and center coordinates of the ground-truth bounding box, respectively; W_i and H_i denote the width and height of the intersection region, respectively; W_g and H_g denote the width and height of the minimum enclosing box, respectively.

图 5 IoU 损失函数

Fig. 5 IoU loss function

动态调整边界框梯度增益^[28],在训练后期减少低质量锚框产生的有害梯度,最终提高模型定位能力。

$$L_{\text{WIoU}} = rgR_{\text{WIoU}}L_{\text{IoU}}, R_{\text{WIoU}} \in [1, e], L_{\text{IoU}} \in [0, 1]. \quad (5)$$

$$\text{其中: } r = \frac{\beta}{\delta \alpha^{\beta-a}}, \quad (6)$$

$$R_{\text{WIoU}} = \exp\left(\frac{(x-x_{\text{gt}})^2 + (y-y_{\text{gt}})^2}{(W_g^2 + H_g^2)^{\gamma}}\right). \quad (7)$$

L_{WIoU} 表示加权交并比损失函数, R_{WIoU} 为基于中心点距离的距离权重项,式(7)中“*”表示停止梯度处理。WIoU旨在去除CIoU中的纵横比惩罚项,平衡高质量和低质量锚框对模型回归的影响,从而增强模型的泛化能力并提高整体性能。

1.4 基于DeepSORT的果实跟踪与计数

1.4.1 状态估计与果实匹配 随着相机的绕圈旋转运动,果实目标会频繁出现、消失或位置剧烈变化,使得相邻帧可检测到的果实数量存在较大波动。因

此,引入多目标跟踪算法DeepSORT,实现在复杂环境中柑橘的实时跟踪与计数,其框架结构如图6所示。

卡尔曼滤波器的引入为处理上述不确定性提供

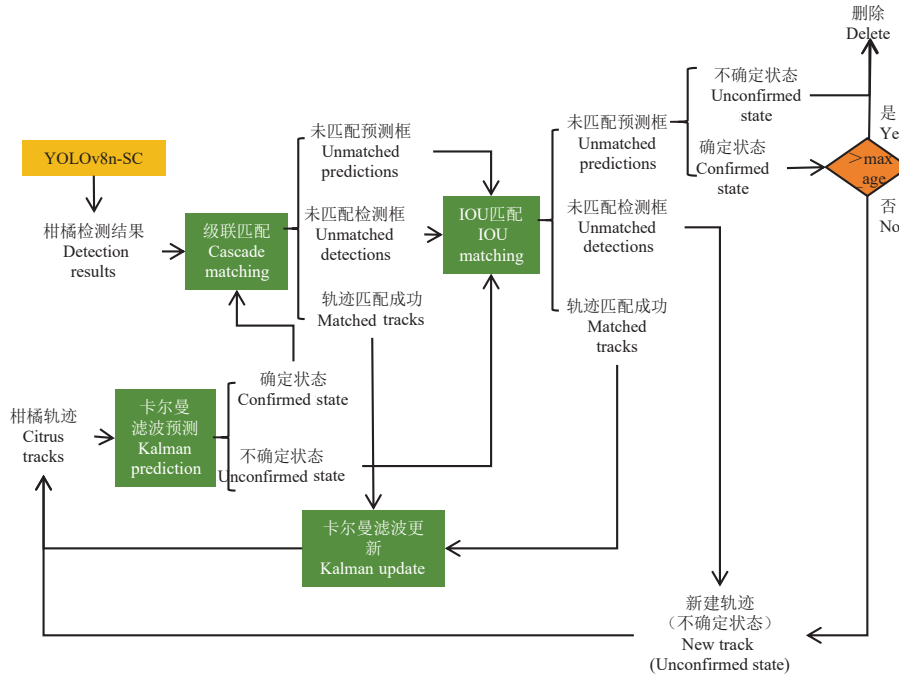


图6 DeepSORT 柑橘跟踪框架结构图

Fig. 6 The structure of DeepSORT citrus tracking frame

了理论依据。由于视频序列帧间果实目标的位置变化较小,可以近似视为匀速运动。卡尔曼滤波通过其“预测-更新”的递归框架,为此模型提供了最优状态估计^[29]:在预测阶段,利用上一个时刻的状态和控制输入来初步估计当前状态,这一机制使得在目标因短暂遮挡而未被检测器发现时,跟踪器仍能维持其轨迹,显著降低了因漏检导致的跟踪中断概率;在更新阶段,结合实际观测、修正预测,有效平滑了由检测框抖动带来的位置噪声,从而输出更稳定、精确的运动轨迹。

关联度计算是将检测结果与预测轨迹状态进行匹配的重要步骤,DeepSORT的优化主要基于匈牙利算法里的代价矩阵。在IoU匹配之前进行一次额外的级联匹配,结合了外观特征和马氏距离,适用于动态环境中的目标检测。在目标跟踪时,通过计算马氏距离,可有效衡量预测位置与真实位置之间的偏差。当目标运动模式变化频繁或被遮挡时,引入余弦距离来捕捉外观特征,以保持一定的匹配能力。这两种度量方式从不同方面服务目标匹配,互

为补充。

1.4.2 跟踪与计数 DeepSORT算法对每条轨迹都设置一个初始值为0的计时器(age),每当未成功匹配时计数加1,成功匹配时则重置为0。其值表示该轨迹离上次成功匹配的时间,设定最大时间(max_age)为90帧,若超过90帧仍未成功重新匹配,则将该轨迹删除。图6所示的算法详细过程如下:(1)为视频帧中的每个柑橘果实创建其对应的轨迹信息,将卡尔曼滤波的运动变量初始化,预测其对应的检测框信息。(2)将步骤1中的轨迹与本轮中YOLOv8n-SC检测出的目标一起送入级联匹配中,柑橘目标得到三种状态,分别为未匹配检测、未匹配轨迹和已匹配轨迹^[30]。(3)由于步骤2中存在遗漏,将其与步骤1中未确定的轨迹合并,用IoU匹配再次匹配,也可以得到三种状态。(4)将未匹配轨迹中的未确认和已确认但age>90的状态删除。(5)将已匹配轨迹合并,进行卡尔曼滤波更新后输出轨迹;步骤3中未匹配目标作为新建的轨迹;步骤4中已确认但age≤90。此三类都作为新一轮的轨迹。(6)重复循

环步骤1~5,直至视频帧结束。按照果实在视频帧被检测到的顺序,为果实匹配从1开始的数字ID,实现视频中柑橘计数。

1.5 柑橘估产试验设计

试验分为三部分,分别为目标检测、目标跟踪和目标计数,具体处理如图7所示。

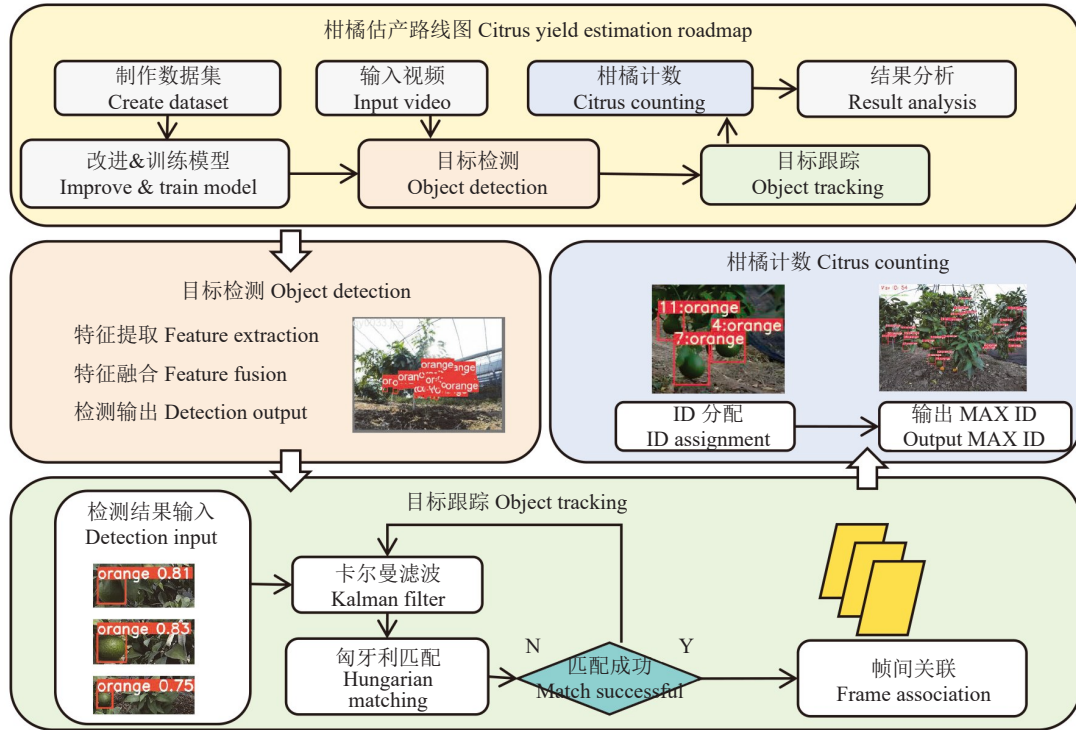


图7 柑橘估产路线图

Fig. 7 Citrus yield estimation roadmap

试验环境:处理器为 Intel(r) Xeon(r) Platinum P-8124 CPU @ 3.00GHz 3.00 GHz(双核),运行内存为 64 GB,显卡为 NVIDIA GeForce RTX 2080 Ti,显存大小为 11 G,软件环境为 Windows 10 专业版操作系统,在 PyCharm2022 软件中使用 Pytorch 深度学习框架进行模型训练,为了提高模型训练速度,使用 CUDA11.2 加速训练。模型训练时训练轮次(epochs)设置为 400,初始化学习率(lr)设置为 0.01。

数据获取方式:在采集视频数据的同时,采用传统的归拢树枝法由 5 名研究人员清点树上果实,将其平均处理后得到果实的果园真实值。对视频中的果实进行计数,将其平均处理后得到果实的视频真实值,同时记录 ID 变换、误检、漏检的个数。经过估产系统得到的计数结果称为算法预测值(MAX ID)。

1.6 试验评价指标

1.6.1 果实检测试验 采用准确率 P(precision)、召回率 R(recall)、平均精度均值 mAP(mean average

precision)、计算复杂度 FLOPs(floating point operations)^[31]等指标对训练的模型进行性能评价。其中, mAP 是一个综合评价指标,可以反映模型的检测能力,在 IoU 阈值分别为 0.5 和 0.5~0.95 时,平均精度值表示为 mAP₅₀ 和 mAP₅₀₋₉₅。

1.6.2 果实跟踪试验 使用多目标跟踪准确度 MOTA (multiple object tracking accuracy) 表示算法在保持跟踪时的性能;多目标跟踪精度 MOTP (multiple object tracking precision) 评估检测目标与真实目标之间的匹配精度^[20];跟踪速度 FPS (frame per second) 评价算法的实时性;果实 ID 切换率 IDSR (ID switch rate) 指视频中 ID 发生变换的果实占有计数果实的比值^[13]。

1.6.3 果实计数试验 使用均方根误差 RMSE (root mean square error) 和平均绝对误差 MAE (mean absolute error) 衡量预测值与实际观察值之间的差异^[13];采用 ACP (average counting precision) 值作为柑橘果实计数的精度指标。

2 结果与分析

2.1 果实图像检测结果

2.1.1 目标检测模型比较 为了验证改进后模型的有效性,将模型YOLOv8n-SC与YOLOv8n对比,同时增加常用的Faster RCNN模型以及新推出的轻量级模型YOLOv10n。均使用总数据集进行训练和测试,结果见表1。

由表1可知,YOLOv8n-SC与YOLOv8n相比,

在平均精度指标 mAP_{50-95} 上提高了1.5%。由于输出端新增了检测头,YOLOv8n-SC模型大小有小幅度增加,虽未减少模型参数量,但其计算量却略有下降,这表明模型改进是有效的。与传统的双阶段模型Faster R-CNN相比,YOLOv8n-SC在 mAP_{50-95} 指标上提高了4.1%,优势显著。与YOLOv10n相比,虽然模型大小与运算速度存在劣势,但 mAP_{50-95} 与召回率均比YOLOv10n提高了0.7%,表明YOLOv8n-SC在识别膨大期柑橘方面有较强的竞争力。

表1 不同模型的柑橘果实目标检测性能比较

Table 1 Comparison of citrus object detection performance among different models

模型 Model	精确率 Precision (P)/%	召回率 Recall (R)/%	平均精度均值 Mean average precision (mAP ₅₀)/%	平均精度均值 Mean average precision (mAP ₅₀₋₉₅)/%	参数量 Parameters/ M	浮点运算数 Float operation per second (FLOPs)/G
YOLOv8的轻量级版本 YOLOv8n	93.7	95.6	90.2	59.1	14.5	8.1
经典的双阶段目标检测模型 Faster RCNN	88.1	79.3	84.7	56.5	100.3	65.2
改进的YOLOv8n模型 YOLOv8n-SC	97.2	97.6	93.7	60.6	14.8	7.9
YOLOv10的轻量级版本 YOLOv10n	94.5	96.9	90.8	59.9	8.7	6.8

2.1.2 目标检测模型的消融试验 为了验证YOLOv8n模型改进后的性能,对各个改进的模型和原模型在总数据集和同一硬件环境下进行消融试验,如表2所示。

由表2可知,在YOLOv8n模型增加一个新的检测头后,柑橘果实的R、 mAP_{50-95} 分别提高了0.6%和1.1%,说明新的检测头可以提高微小柑橘目标的识别率,将C2f_CBAM替换C2f模块后,可提升模型检测性能,R、 mAP_{50-95} 分别提高了0.9%和0.7%,由于CBAM是一个轻量级的通用模块,对计算量影响较小。更改损失函数为WIoU,模型性能也有小幅度提升。综合比较4个卷积网络评估指标,YOLOv8n-

SC(第4行)性能最佳。

2.1.3 不同光照条件结果分析 在不同光照强度下,柑橘果实的颜色和纹理特征差异较大,使得果实的细节难以识别。为验证光照条件对柑橘目标检测准确率的影响,将YOLOv8n-SC模型分别在晴天、阴天、人工光源采集的柑橘数据下进行训练和测试,结果如表3所示。

光照条件是影响模型检测性能的关键因素。由表3可知,人工光源下的检测性能最佳,其 mAP_{50} 达到98.2%,显著优于阴天(96.1%)和晴天(83.5%)。这主要是因为人工光源提供了稳定、均匀的光照环境,有效提高了柑橘果实与背景的对比值,同时消除

表2 消融试验结果

Table 2 Results of ablation experiment

模型 Model	检测头 Detect	CBAM	WIoU	召回率 Recall (R)/%	平均精度均值 Mean average precision (mAP ₅₀)/%	平均精度均值 Mean average precision (mAP ₅₀₋₉₅)/%	浮点运算数 Float operation per second (FLOPs)/G
YOLOv8n模型 YOLOv8n model				95.6	90.2	59.1	8.1
YOLOv8n模型 YOLOv8n model	√			96.2	92.6	60.2	8.4
YOLOv8n模型 YOLOv8n model	√	√		96.5	93.3	59.8	8.4
YOLOv8n模型 YOLOv8n model	√	√	√	97.6	93.7	60.6	7.9

表 3 不同光照条件的柑橘果实检测结果比较

Table 3 Comparison of citrus detection results under different light conditions

数据集 Dataset	精确率 Precision (P)/%	召回率 Recall (R)/%	平均精度均值 Mean average precision (mAP ₅₀)/%	平均精度均值 Mean average precision (mAP ₅₀₋₉₅)/%
晴天 Sunny	86.1	81.3	83.5	45.1
阴天 Overcast	93.5	93.4	96.1	61.4
人工光源 Artificial light	97.2	96.5	98.2	66.9

了自然光下因强光反射和阴影变化所引入的图像噪声,从而大幅提升了模型的识别准确率。

2.2 视频流果实估产试验结果

2.2.1 不同视角视频流跟踪结果分析 为减少因柑橘树枝茂密和果实遮挡聚集等现象带来的估产误差,本课题采用在人工光源条件下分别采集的俯摄、平摄、仰摄三组视频。基于 YOLOv8n-SC 模型结合 DeepSORT 跟踪算法进行跟踪与计数训练,置信度设置为 0.7。跟踪结果如表 4 所示。

由表 4 可知,由于平摄减少了视频盲区,兼顾了

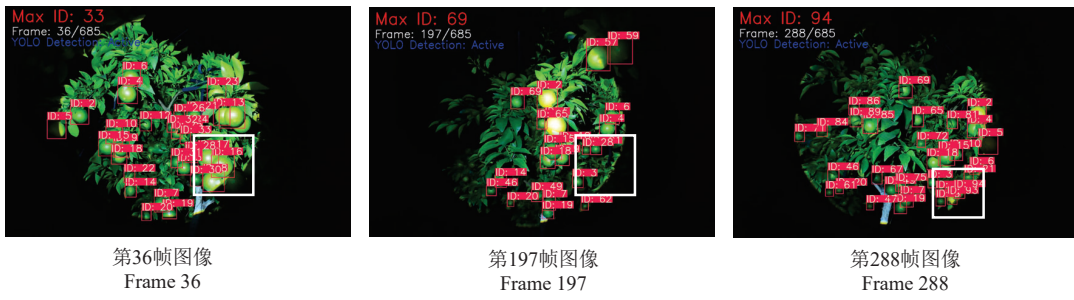
树冠俯仰角度的果实,因此跟踪效果最好, MOTA 达到 90.8%, MOTP 高达 94.0%, 误差的主要原因可能是相互遮挡导致 ID 的变化率较高, 为 11.8%。俯摄检测效果次之, 多目标跟踪精度高, ID 变换表现稳定, 检测精度低于平摄的原因可能是果树下部果实被遮挡严重而导致遗漏。仰摄检测效果最低, 因果实的大部分着生在冠层区域, 仰摄检测时的遮挡降低了检测率。

图 8 为结果示例, 图 8-a 的仰摄视频中 16、17、30 号果实, 在 34 帧时处于连续跟踪的状态, 但在 197 帧

表 4 不同视角的柑橘果实跟踪结果比较

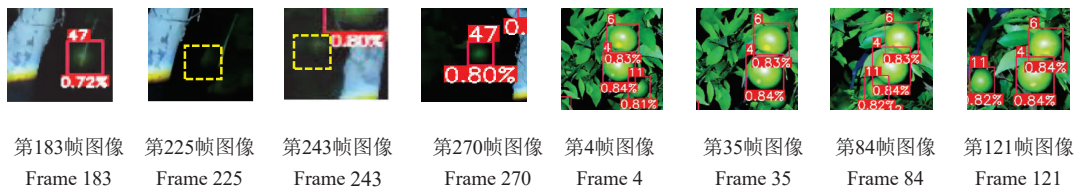
Table 4 Comparison of citrus tracking results from different shooting angles

拍摄角度 Camera angle	多目标跟踪准确度 Multiple object tracking accuracy (MOTA)/%	多目标跟踪精度 Multiple object tracking precision (MOTP)/%	跟踪速度 Tracking speed (FPS)/s	ID 切换率 ID switch rate (IDSR)/%
俯摄 High-angle shot	84.9	95.7	21.74	5.6
平摄 Straight angle shot	90.8	94.0	21.74	11.8
仰摄 Low-angle shot	85.8	87.9	21.74	27.4



a. 仰视角的 ID 变动情况示例

a. Example of ID variation from a low-angle perspective



b. 检测丢失后重新匹配情况示例

b. Example of re-matching after detection loss

c. 跳帧检验果实跟踪情况示例

c. Example of fruit tracking verification with frame skipping

图 8 不同情况下的结果示例

Fig. 8 Examples of fruit tracking results under different detection conditions

时丢失,消失的目标未在90帧(max_age)内重新匹配,因此在第288帧目标重新出现时,就会错误地分配新的ID:92、93、94。如图8-b所示,47号果实在运动过程中被树干遮挡,且在置信度小于0.7时不予匹配,但经历了87帧的目标丢失后,仍匹配了原ID。如图8-c所示,在旋转视角的过程中,虽然4、6、11号果实的相对位置发生变化,果实在匹配过程中未出现差错,这表明所用算法具有较好的跟踪能力。

2.2.2 多角度视频流计数结果分析 根据获取的果园真实值、视频真实值、算法预测值,得到计数精度结果,如表5所示。

从表5数据可知,由于膨大期的柑橘果实质量增大呈下垂状态,俯摄时,被叶片遮挡导致视频真实

值与果园真实值差距过大,对柑橘估产没有意义。平摄时, RMSE与MAE值比仰摄分别低11.33和10.03, ACP值比仰摄高8.43%,说明在人工光源下,视频流平摄估产更接近于果园的真实果树数量。

2.2.3 跟踪算法间比较 为进一步验证算法的有效性,采用YOLOv8n-SC模型,分别结合DeepSORT、SORT、Tracktor算法跟踪器,组合出多种柑橘计数方法,检测结果见表6。

由表6可知,DeepSORT跟踪算法具有最高的ACP计数精度,分别比SORT和Tracktor跟踪器高了4.65%和3.09%, RMSE和MAE分别为10.83和9.98。但由于需提取和计算每一帧的外观特征,导致计算复杂度较高, FPS最低。

表5 不同视角的柑橘果实计数精度比较

Table 5 Comparison of citrus counting accuracy from different shooting angles

拍摄角度 Camera angle	果园真实值 Orchard ground truth	视频真实值 Video ground truth	算法预测值 Algorithm prediction value	均方根误差 Root mean square error (RMSE)	平均绝对误差 Mean absolute error (MAE)	平均计数精度 Average counting precision (ACP)/%
俯摄 High-angle shot	2195	697	539	10.35	8.97	89.87
平摄 Straight angle shot		2103	2298	10.83	9.98	89.46
仰摄 Low-angle shot		2091	2484	22.16	20.01	81.03

注:所需计算的数据都基于视频真实值,果园真实值仅作参考。

Note: All data required for calculation are based on the ground-truth values from the video, while the orchard ground-truth values are for reference only.

表6 不同跟踪算法的柑橘果实计数精度比较

Table 6 Comparison of citrus counting accuracy of different tracking algorithms

模型 Model	均方根误差 Root mean square error/RMSE	平均绝对误差 Mean absolute error/MAE	跟踪速度 FPS/s	平均计数精度 Average counting precision ACP/%
YOLOv8n-SC结合SORT跟踪算法 YOLOv8n-SC+SORT	17.50	16.33	31.84	84.81
YOLOv8n-SC结合Tracktor跟踪算法 YOLOv8n-SC+Tracktor	11.21	10.67	25.25	86.37
YOLOv8n-SC结合DeepSORT跟踪算法 YOLOv8n-SC+DeepSORT	10.83	9.98	21.74	89.46

3 讨论

本研究验证了基于360°旋转视频流、改进的YOLOv8n-SC模型及DeepSORT算法在单株柑橘估产任务中的可行性与有效性。试验结果表明,该方法在特定条件下能够实现较高的计数精度。然而,本研究所提出的方法在展现其应用价值的同时,也暴露出当前技术体系存在的局限性,这为后续研究的深化与拓展指明了方向。

3.1 模型在复杂光照与遮挡条件下的局限性分析

试验结果显示,模型性能对环境光照条件高度敏感,晴天条件下mAP₅₀降至83.5%,这凸显了当前

算法在真实田间环境中的适应性不足。尽管人工光源可优化检测效果,但该方法对理想照明条件的依赖限制了其在实际生产环境中的推广应用。因此,提升模型在复杂光照条件下的鲁棒性应成为后续研究的重点。在跟踪环节中,平视角下ID切换率为11.8%,这表明DeepSORT算法在面对由视角转变引起的目标间相对位置快速变化时,存在适应性不足。当一个果实因遮挡消失超过max_age阈值后被重新识别,算法机制必然会视其为“新目标”并分配新ID,导致重复计数。旋转拍摄虽能提供比行间扫描更完整的视角,但视频计数结果与真实产量间仍存在差异。本质上其仍是将三维冠层

结构投影至二维图像序列,无法有效探测被茂密枝叶遮挡的冠层内部果实,这表明基于二维视觉的采集策略可能已接近其性能上限^[32]。要解决空间遮挡下的计数问题,核心突破口可能不在于微调跟踪参数,而在于引入能够理解空间关系的更高维度信息。

3.2 二维与三维视觉方案对比分析

本研究提出的二维视觉方案在多个关键维度上相较于三维感知方法展现出显著优势。在技术扩展性方面,该框架可快速迁移至苹果、梨等具有相似冠层结构的乔木果树,仅需通过迁移学习更新少量模型参数即可完成适配,而三维方法通常需要复杂的点云配准和重新标定^[33]。在硬件成本方面,本方案仅需配置普通运动相机和常规计算设备,其成本远低于激光雷达系统,更适合在中小规模果园中实现规模化部署。在处理效率方面,经过优化的轻量级网络可在边缘设备上实现近实时处理,单株果树的完整分析可在数分钟内完成,而基于点云处理的三维方法需要消耗大量计算资源。

三维视觉方法在解决空间定位精度、三维重建方面具有独特优势。通过激光雷达、立体视觉方法可以实现对果树冠层结构的高精度三维建模。此外,基于点云分割、三维检测和表面重建的方法能够提供更丰富的植株参数,为精准农业中的树形分析和生长监测提供更完整的数据基础^[34]。因此,在需要高精度空间信息的应用场景中三维方法仍有不可替代的价值。

3.3 本研究的局限性

本研究在取得预期成果的同时,也存在若干需要正视的技术局限性,这些局限主要源于数据采集方式、算法架构及感知维度三个层面,对未来实际应用提出了一定挑战。

首先,数据采集过程中采用的手持拍摄方式可能存在运动不稳定性,导致图像序列存在运动模糊与帧间稳定性不足。这不仅降低了单帧图像的检测质量,更破坏了多目标跟踪算法中基于匀速运动的动力学假设,成为系统误差的重要来源。其次,在算法性能层面,DeepSORT算法通过复杂的匹配策略维持较高的跟踪精度,但相应地抑制了计算效率,导致处理帧率难以满足实时性要求较高的应用场景。最后,在感知维度层面,基于二维视觉的感知框架本质上无法解析三维空间中的深度信息与立体结构关

系,在面对多层次叶片遮蔽、果实空间重叠等复杂场景时,该方法存在固有的感知盲区。

3.4 未来的研究方向

基于本研究存在的局限性,未来的研究方向同样从三个维度展开。在数据采集标准化层面,研发机载式稳定采集系统,可通过预设标准化运动轨迹与采集参数,建立稳定的多视角数据获取系统,从根本上消除手持操作引起的运动模糊与帧间不稳定问题。在算法层面,可采用轻量化重识别网络架构,结合模型剪枝与知识蒸馏技术,实现精度与效率的平衡。同时,重点研究基于叶幕密度影响的修正模型,建立基于生态特征的检测结果校正机制^[35]。在多维感知方面,探索将本研究的框架与轻量级三维感知方法结合的可能性,在保持低成本、高效率、可扩展的基础上,通过空间特征增强提升对遮挡目标的感知能力,最终形成适用于复杂田间环境的多维感知解决方案。

4 结 论

本试验利用YOLOv8n-SC模型结合DeepSORT算法对膨大期柑橘进行检测及计数,探索了不同光照条件、不同视角下估产的准确性,取得主要结论如下:

1)经改进的YOLOv8n-SC模型综合检测性能优于YOLOv8n、Faster RCNN和YOLOv10n模型,表明了模型改进的有效性和对膨大期柑橘果实估产的能力。

2)不同的光照环境对柑橘估产有显著影响。晴天的强光造成的反射和阴影,降低了检测的准确率;阴天的光照较为均匀,各项检测指标均有明显提升;人工光源条件下,能突出果实、降低背景干扰,检测性能最佳,mAP₅₀达到98.2%。

3)不同的视频拍摄角度对柑橘估产精度有显著影响。平摄视频流的跟踪性能最佳,其MOTA达到90.8%,MOTP高达94.0%;俯摄视频流因视角受限导致果实被遮挡而可见目标过少,影响估产计数的真实性;仰摄视频流则因严重的ID切换问题和较低的估产能力不适合柑橘估产。

4)对比柑橘果实多目标跟踪算法,YOLOv8n-SC+DeepSORT的模型组合的估产精度最高,计数精度为89.46%,较YOLOv8n-SC搭载SORT或Tracker分别提高4.65%和3.09%。

参考文献 References:

- [1] 李馨雨,阳文娟,何丽波.我国柑橘产业布局变迁及其影响因素研究[J].农业科技与信息,2024(5):135-139.
LI Xinyu, YANG Wenjuan, HE Libo. Research on the change of citrus industry layout and its influencing factors in China[J]. Agricultural Science- Technology and Information, 2024(5): 135-139.
- [2] 程洪.面向园艺应用的机器视觉目标辨识方法创新[D].北京:中国农业大学,2015.
CHENG Hong. The innovation of machine vision target identification method[D]. Beijing:China Agricultural University, 2015.
- [3] 高芳芳,武振超,索睿,周忠贤,李瑞,傅隆生,张昭.基于深度学习与目标跟踪的苹果检测与视频计数方法[J].农业工程学报,2021,37(21):217-224.
GAO Fangfang, WU Zhenchao, SUO Rui, ZHOU Zhongxian, LI Rui, FU Longsheng, ZHANG Zhao. Apple detection and counting using real-time video based on deep learning and object tracking[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(21):217-224.
- [4] LI H, LEE W S, WANG K. Immature green citrus fruit detection and counting based on fast normalized cross correlation (FNCC) using natural outdoor colour images[J]. Precision Agriculture, 2016, 17(6):678-697.
- [5] 毕松,张璐.自然环境下的柑橘采摘点识别方法研究[J].计算机仿真,2021,38(12):227-231.
BI Song, ZHANG Lu. Research on method of citrus picking point recognition in natural environment[J]. Computer Simulation, 2021, 38(12):227-231.
- [6] 曾风,滕开良,张磊.基于改进YOLOv8n模型的咖啡果实成熟度检测算法研究[J/OL].智能计算机与应用,1-6(2024-12-02). <https://doi.org/10.20169/j.issn.2095-2163.24110506>.
ZENG Feng, TENG Kai-liang, ZHANG Lei. Coffee fruit maturity detection algorithm based on YOLOv8n model[J/OL]. Intelligent Computers and Applications, 1-6(2024-12-02). <https://doi.org/10.20169/j.issn.2095-2163.24110506>.
- [7] 熊俊涛,刘振,汤林越,林睿,卜榕彬,彭红星.自然环境下绿色柑橘视觉检测技术研究[J].农业机械学报,2018,49(4):45-52.
XIONG Juntao, LIU Zhen, TANG Linyue, LIN Rui, BU Rongbin, PENG Hongxing. Visual detection technology of green citrus under natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(4):45-52.
- [8] 吕强,林刚,蒋杰,王明之,张皓杨,易时来.基于改进YOLOv5s模型的自然场景中绿色柑橘果实检测[J].农业工程学报,2024,40(18):147-154.
LÜ Qiang, LIN Gang, JIANG Jie, WANG Mingzhi, ZHANG Haoyang, YI Shilai. Detecting green citrus fruit in natural scenes using improved YOLOv5s model[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(18):147-154.
- [9] WU Z C, SUN X M, JIANG H H, MAO W L, LI R, ANDRIYANOV N, SOLOVIEV V, FU L S. NDMFCS: An automatic fruit counting system in modern apple orchard using abatement of abnormal fruit detection[J]. Computers and Electronics in Agriculture, 2023, 211:108036.
- [10] 李艳萍,林建辉,杨宁学.一种基于SIFT特征光流的运动目标跟踪算法[J].计算机科学,2015,42(11):305-309.
LI Yanping, LIN Jianhui, YANG Ningxue. Algorithm of moving target tracking based on SIFT feature optical flow[J]. Computer Science, 2015, 42(11):305-309.
- [11] 李萌,郑娟毅,门瑜.基于帧差法的运动车辆阴影去除算法[J].电视技术,2016,40(10):11-14.
LI Meng, ZHENG Juanyi, MEN Yu. Algorithm of moving vehicle shadow elimination based on frame difference[J]. Video Engineering, 2016, 40(10):11-14.
- [12] 魏超,吴西涛,朱耿霆,舒用杰,李路兴,随淑鑫.基于视觉相机和激光雷达融合的无人车障碍物检测与跟踪研究[J/OL].机械工程学报,1-15(2024-03-05). <https://link.cnki.net/urlid/11.2187.TH.20240301.1406.005>.
WEI Chao, WU Xitao, ZHU Gengting, SHU Yongjie, LI Luxing, SUI Shuxin. Research on obstacle detection and tracking of autonomous vehicles based on the fusion of vision camera and LiDAR[J/OL]. Journal of Mechanical Engineering, 1-15(2024-03-05). <https://link.cnki.net/urlid/11.2187.TH.20240301.1406.005>.
- [13] 沈磊.基于机器视觉的葡萄果穗跟踪计数方法研究[D].杨凌:西北农林科技大学,2023.
SHEN Lei. Grape clusters tracking and counting method based on machine vision[D]. Yangling: Northwest A & F University, 2023.
- [14] 张润池,周云成,侯玉涵,刘泽钰,赵鸿舸,赵昱涵.基于超深掩蔽与改进YOLOv8的不同成熟度番茄计数方法[J].农业工程学报,2024,40(24):146-156.
ZHANG Runchi, ZHOU Yuncheng, HOU Yuhan, LIU Zeyu, ZHAO Hongge, ZHAO Yuhan. Counting tomatoes with different maturities using ultra-depth masking and improved YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(24):146-156.
- [15] LI X H, WANG X, ONG P, YI Z R, DING L, HAN C. Fast recognition and counting method of dragon fruit flowers and fruits based on video stream[J]. Sensors, 2023, 23(20):8444.
- [16] 殷慧军,王宝丽,景运革,李菊霞,王鹏岭,权高翔,孙婷婷.基于改进YOLOv7的无人机航拍视频西瓜计数方法[J].农业工程学报,2024,40(19):124-134.
YIN Huijun, WANG Baoli, JING Yunge, LI Juxia, WANG Pengling, QUAN Gaoxiang, SUN Tingting. Improved YOLOv7 method for counting watermelons in UAV aerial videos[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(19):124-134.
- [17] 毛妮妮,苏西娅,任俊鹏,毕秀芝,刘书仁.‘红美人’柑橘在江

- 苏南京的引种表现及生态栽培技术[J]. 园艺与种苗, 2024, 44(3):3-4.
- MAO Nini, SU Xiya, REN Junpeng, BI Xiuzhi, LIU Shuren. Introduction performance and ecological cultivation techniques of *Citrus reticulata* 'Hong Mei Ren' in Nanjing of Jiangsu Province[J]. Horticulture & Seed, 2024, 44(3):3-4.
- [18] 熊俊涛, 郑镇辉, 梁嘉恩, 钟灼, 刘柏林, 孙宝霞. 基于改进 YOLO v3 网络的夜间环境柑橘识别方法[J]. 农业机械学报, 2020, 51(4):199-206.
- XIONG Juntao, ZHENG Zhenhui, LIANG Jia'en, ZHONG Zhuo, LIU Bolin, SUN Baoxia. Citrus detection method in night environment based on improved YOLO v3 network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(4):199-206.
- [19] AL RIZA D F, MUSAHADA L C, AUFA R I, HERMANTO M B, NUGROHO H, HENDRAWAN Y. Comparative study of citrus fruits (*Citrus reticulata* Blanco cv. Batu 55) detection and counting with single and double labels based on convolutional neural network using YOLOv7[J]. Smart Agricultural Technology, 2025, 10:100763.
- [20] 高芳芳. 基于 YOLO 轻型网络和树干跟踪的苹果检测与计数方法研究[D]. 杨凌:西北农林科技大学, 2022.
- GAO Fangfang. Apple detection and counting method based on YOLO light network and trunk tracking[D]. Yangling: Northwest A & F University, 2022.
- [21] 蔡健荣, 朱文辉, 乔宇, 李奇洋, 梁小祥, 杨笑难, 潘丙科. 基于 YOLOv8 的疫苗胚蛋活性视觉检测[J]. 农业工程学报, 2024, 40(21):274-284.
- CAI Jianrong, ZHU Wenhui, QIAO Yu, LI Qiyang, LIANG Xiaoxiang, YANG Xiaonan, PAN Bingke. Visual detection method for vaccine embryo vitality based on YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(21):274-284.
- [22] 李茂, 肖洋秩, 宗望远, 宋宝. 基于改进 YOLOv8 模型的轻量化板栗果实识别方法[J]. 农业工程学报, 2024, 40(1):201-209.
- LI Mao, XIAO Yangyi, ZONG Wangyuan, SONG Bao. Detecting chestnuts using improved lightweight YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(1):201-209.
- [23] 苏宇锋, 边锋, 张玉堂. 基于改进 YOLOv8s 算法的胸环靶弹孔检测技术[J]. 郑州大学学报(工学版), 2024, 45(5):16-22.
- SU Yufeng, BIAN Feng, ZHANG Yutang. Bullet hole detection technology of chest bitmap based on improved YOLOv8s algorithm[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(5):16-22.
- [24] WOO S, PARK J, LEE J Y, KWEON I S. CBAM: Convolutional block attention module[M]//BERTINO E, GAO W, STEFFEN B, YUNG M. Computer Vision- ECCV 2018, Lecture Notes in Computer Science. Switzerland: Springer International Publishing, 2018:3-19.
- [25] 王欣, 江涛, 马珍, 魏玉梅. 基于改进 YOLOv3-Tiny 的遥感舰船小目标检测方法[J]. 计算机时代, 2023(3):111-115.
- WANG Xin, JIANG Tao, MA Zhen, WEI Yumei. Small target detection of remote sensing ship based on improved YOLOv3-Tiny algorithm[J]. Computer Era, 2023(3):111-115.
- [26] ZHENG Z H, WANG P, REN D W, LIU W, YE R G, HU Q H, ZUO W M. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2022, 52(8):8574-8586.
- [27] ZHANG H, XU C, ZHANG S J. Inner-IoU: More effective intersection over union loss with auxiliary bounding box[PP/OL]. V4. arXiv (2023-11-14)[2025-09-10]. <https://doi.org/10.48550/arXiv.2311.02877>
- [28] MA D Q, LI S J, DANG B, ZANG H Y, DONG X Q. Fostc3net: A lightweight YOLOv5 based on the network structure optimization[J/OL]. Journal of Physics: Conference Series, 2024, 2824(1):012004. doi:10.1088/1742-6596/2824/1/012004.
- [29] TU S Q, HUANG Y F, HUANG Q, LIU H X, CAI Y F, LEI H. Estimation of passion fruit yield based on YOLOv8n + OC-SORT + CRCM algorithm[J/OL]. Computers and Electronics in Agriculture, 2025, 229:109727. <https://doi.org/10.1016/j.compag.2024.109727>
- [30] HE L Y, WU F D, DU X Q, ZHANG G F. Cascade-SORT: A robust fruit counting approach using multiple features cascade matching[J/OL]. Computers and Electronics in Agriculture, 2022, 200:107223. <https://doi.org/10.1016/j.compag.2022.107223>.
- [31] WU Z C, SUN X M, JIANG H H, GAO F F, LI R, FU L S, ZHANG D, FOUNTAS S. Twice matched fruit counting system: An automatic fruit counting pipeline in modern apple orchard using mutual and secondary matches[J]. Biosystems Engineering, 2023, 234:140-155.
- [32] KAMILARIS A, PRENAFETA-BOLDÚ F X. Deep learning in agriculture: A survey[J]. Computers and Electronics in Agriculture, 2018, 147:70-90.
- [33] GENÉ-MOLA J, SANZ-CORTIELLA R, ROSELL-POLO J R, MORROS J R, RUIZ-HIDALGO J, VILAPLANA V, GREGORIO E. Fruit detection and 3D location using instance segmentation neural networks and structure-from-motion photogrammetry[J]. Computers and Electronics in Agriculture, 2020, 169:105165. DOI:10.1016/j.compag.2019.105165.
- [34] LI J, TANG L. Developing a low-cost 3D plant morphological traits characterization system[J]. Computers and Electronics in Agriculture, 2017, 143:1-13.
- [35] NIKLAS K J. Plant allometry: The scaling of form and process[M]. Chicago:University of Chicago Press, 1994.