

基于视频跟踪算法的果园猕猴桃产量实时预估

郭明月¹ 刘雅晨¹ 李伟夫¹ 陈洪¹ 李善军² 陈耀晖²

(1. 华中农业大学信息学院, 武汉 430070; 2. 华中农业大学工学院, 武汉 430070)

摘要: 对猕猴桃产量的准确预估有利于合理安排后续采摘与运输工序,因此开发智能化的产量实时预估工具非常重要。针对大棚培育的猕猴桃矮化密植、分布范围广等特点,本研究利用果园履带小车采集视频,结合人工标注,建立猕猴桃检测和跟踪的数据集。考虑到自制数据集中猕猴桃占比小及密集分布的特点,本文提出使用 YOLO v7 模型加上 Soft-NMS 来检测每一帧图像内的猕猴桃。在卡尔曼滤波器预测的结果上,引入 VGG16 网络对猕猴桃进行特征提取,并结合匈牙利算法完成帧间目标的匹配。最后采用基于 YOLO v7 + DeepSort 跟踪算法的 ID 计数方法对猕猴桃进行产量估计。实验结果表明,改进的 YOLO v7 模型在猕猴桃检测数据集上表现良好,检测的 F1 值为 90.09%。猕猴桃跟踪数据集中使用的跟踪算法平均准确率为 89.87%,每个目标正确匹配的精确率为 82.34%,大型视频跟踪速度为 20.19 f/s。在环境影响较小的条件下,ID 计数准确率为 97.49%。该方法可为猕猴桃果园智能化生产管理中的估产、采收规划等提供技术支撑。

关键词: 猕猴桃; 产量预估; 视频跟踪; YOLO v7; Soft-NMS; DeepSort

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2023)06-0178-08 **OSID:** 

Real-time Production Prediction of Kiwifruit in Orchard Based on Video Tracking Algorithm

GUO Mingyue¹ LIU Yachen¹ LI Weifu¹ CHEN Hong¹ LI Shanjun² CHEN Yaohui²

(1. College of Informatics, Huazhong Agricultural University, Wuhan 430070, China

2. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: The use of machine vision to quickly and accurately estimate fruit yield is of great significance for the development of smart agriculture. In view of the characteristics of dwarf dense planting and wide distribution of kiwifruit cultivated in greenhouses, orchard crawler trolleys were used to shoot and obtain videos of kiwifruit orchards, and a dataset of kiwifruit detection and tracking was established combined with artificial labeling. Considering the small proportion and dense distribution of kiwifruit in the self-made dataset, the YOLO v7 model and Soft-NMS were proposed to detect kiwifruit in each frame. Based on the prediction results of the Kalman filter, the VGG16 network was introduced to extract the features of kiwifruit, and the Hungarian algorithm was used to complete the target matching of the before and after frames. Finally, the ID counting method based on YOLO v7 + DeepSort tracking algorithm was used to realize kiwifruit yield estimation. The experimental results showed that the improved YOLO v7 model performed well on the kiwifruit detection dataset, with an F1 score of 90.09%. The average accuracy of the adopted tracking algorithm on the kiwifruit tracking dataset was 89.87%, the precision of each target can be correctly matched was 82.34% and a large video tracking speed of 20.19 f/s. Under the condition of low environmental impact, the ID counting accuracy was 97.49%. This method can provide technical support for yield estimation and harvest planning in the intelligent management of kiwifruit orchards.

Key words: kiwifruit; yield estimation; video tracking; YOLO v7; Soft-NMS; DeepSort

收稿日期: 2023-03-30 修回日期: 2023-05-05
基金项目: 中央高校基本科研业务费专项资金项目(2662020LXQD002)
作者简介: 郭明月(1979—),女,讲师,主要从事应用统计、多元统计和时间序列分析研究,E-mail: moon@mail.hzau.edu.cn
通信作者: 陈耀晖(1993—),男,副研究员,博士,主要从事机器视觉和深度/强化学习研究,E-mail: yaohui.chen@mail.hzau.edu.cn

0 引言

我国是全球猕猴桃产量最高的国家,猕猴桃产业对我国农业的健康发展意义重大^[1]。目前猕猴桃的种植往往借助栽培架开展,果实悬垂于栽培架下并覆盖整个园区^[2]。猕猴桃产量预估有助于合理规划劳动力、采摘设备以及仓储空间,但是采用人工计数费时费力,准确率也难以保证^[3]。因此,为了有效降低作业成本并提高产业利润空间,开发快速、高效的猕猴桃产量预估工具非常重要。

传统农业中往往采用统计学方法对大面积种植作物进行产量估计。杨钰莹^[4]比较了简单随机抽样、按概率比例抽样以及回归估计^[5]对谷物产量的估计效果,在 95% 的置信度下,发现按概率比例抽样估计精度最高,达到 98.55%。黄凯等^[6]通过多源数据回归分析方法,采用无人机采集甘蔗不同时期的表型数据进行产量估计,得到估计结果与真实产量的 R^2 达到 0.87。吴菲菲等^[7]通过收集不同时期玉米生长情况数据,使用统计分析、方差分析以及遗传估计对玉米产量性状进行评估。尽管采用人工统计方法能对采集的数据进行预估,但是数据获取和评估过程需要耗费大量人力物力,且难以建立统一模型,因此不易推广^[8]。

随着深度学习技术在计算机视觉领域的快速发展^[9],一系列新的农作物产量预估方法被相继提出。其中一类工作主要基于单幅图像进行检测计数,从而获得视野内果实的产量。穆龙涛等^[10]提出的改进型 AlexNet 网络对果园环境下的猕猴桃果实进行检测,其识别率达 96.33%,单个果实检测时间为 0.029 s。REDMON 等^[11]提出的 YOLO 算法将候选区域生成与目标检测过程合并,大大提升了网络的检测速度。赵德安等^[12]采用 YOLO v3 对复杂背景下的苹果目标进行检测,准确率为 97.00%,单幅图像检测时间为 0.017 s。TIAN 等^[13]利用改进的 YOLO v3 网络对不同生长阶段的苹果果实进行检测,测试集上的 F1 值为 82%,单幅图像检测时长为 0.304 s。但受到视角变化、路径规划限制,难以获得较为准确的大区域训练拼接图像,来对果园的总产量进行估计。

另一类工作则对整个果园区域拍摄视频,通过视频跟踪的计数方式对果园总产量进行快速预估^[14]。这类方法一般先对每一帧内的果实目标进行定位检测,然后再进行目标匹配,避免了目标的重复计数。在跟踪算法中,卡尔曼滤波算法具有对未知状态进行估计的能力^[15-16],且算法本身占用内存资源小,因此广泛应用于处理连续变化的任务^[17]。

LIU 等^[18]通过使用卡尔曼滤波器算法,对单目摄像头拍摄下的芒果果实进行跟踪,通过获取目标编号 (Identity document, ID) 实现了芒果产量计数,预测产量与真实产量的 R^2 达到 0.88。WANG 等^[19]使用基于 YOLO 系列网络的水果检测模型,对芒果图像进行果实检测,并通过卡尔曼滤波预测视频中果实的位置,从而实现了实时芒果果实计数,与人工计数相比,该方法仅存在 7.3% 的计数误差。然而,这些研究仅对目标运动特征进行估计,缺失果实外观特征匹配的过程,因此在复杂场景中精度难以保证。

针对猕猴桃果园面积大、果实目标小且密度高、户外环境复杂的问题^[20],本研究拟开发基于视频跟踪的果实总产量预估方法,并在跟踪过程中引入运动特征与外观特征匹配,增加跟踪精度。首先以在小目标检测领域有显著优势的 YOLO v7^[21] 模型为基础,对猕猴桃果实进行检测。YOLO v7 模型通过引入更多特征融合^[22],能够减小目标特征丢失的情况,具有较高的检测精度与检测速度。同时,将 YOLO v7 模型与柔性非极大值抑制 (Soft non maximum suppression, Soft - NMS) 结合,减少检测框的误删。最后,通过 DeepSort 算法^[23]对猕猴桃进行跟踪与实现自动计数。其中,DeepSort 算法采用 VGG16 网络对果实特征进行提取,并结合运动特征,基于匈牙利算法匹配得出每个果实的唯一 ID,最终通过 ID 计数来估算猕猴桃的产量。

1 材料与方法

本研究完整工作流程如图 1 所示,通过在猕猴桃果园进行视频拍摄,分别制作猕猴桃目标检测数据集与目标跟踪数据集;然后,建立目标检测与目标跟踪模型,并使用建立的数据集进行评估验证;最后结合目标检测与跟踪算法,进行猕猴桃产量估计,并对所提出方法进行分析与验证。



图 1 实验流程

Fig. 1 Experimental flowchart

1.1 实验数据采集

采集的猕猴桃数据来自于湖北省武汉市华中农业大学园艺标本园的猕猴桃种植区 (东经 114.35°, 北纬 30.47°), 通过履带小车搭载摄像机以 1 m/s 的速度, 采用 S 形路线对果园进行拍摄, 为保证获取植株完整视野, 采用 1 200 万像素长焦摄像头相机, 距离地面最低高度为 1.32 m, 距离两侧植株最少

1.5 m,以 45°仰角进行拍摄。具体拍摄范围与拍摄路线如图 2 所示,最终获得时长为 63 s、帧率 (Frames per second, FPS) 为 60.05 f/s 的视频数据。

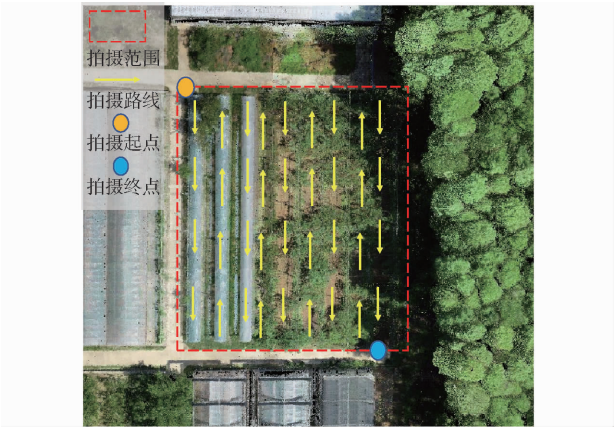


图 2 数据采集线路图
Fig. 2 Data acquisition circuit diagram

1.2 实验数据集制作

1.2.1 猕猴桃检测数据集

由于高帧率视频中相邻帧的相似程度较高,本研究首先采用 OpenCV 库对视频每间隔 5 帧抽取 1 帧,最终获得 534 幅分辨率为 3 840 像素×2 160 像素的彩色图像。同时,采用 Labelme 图像标注软件对每幅图像中的每个猕猴桃进行标注,并按照 8:1:1 的比例划分为训练集、测试集和验证集。

1.2.2 猕猴桃跟踪数据集

由于采集视频的前段、中段和后段中猕猴桃的种植密度及光照条件显著不同,跟踪难度差异较大。前段视频中,猕猴桃的密度中等,果实颜色与叶片颜色的差别明显,光线充足,属于“简单”模式;中段视频中猕猴桃果实密度较低,且果实与叶片颜色较为接近,属于“正常”模式;后段视频拍摄于猕猴桃果实的密集区域,相邻果实的重叠度较高,光照条件较

差,属于“困难”模式,如图 3 所示。

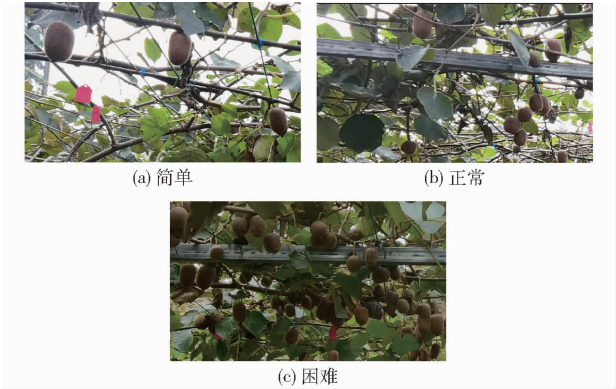


图 3 3 种模式比较
Fig. 3 Comparison of three modes

为衡量网络整体的跟踪精度,采用 Darklabel 软件截取连续 600 帧视频进行手动标记,并在 3 种模式下验证网络计数精度。

1.3 视频跟踪算法搭建

图 4 展示了基于视频追踪的猕猴桃产量预估算法框架。首先通过 YOLO v7 检测器对当前帧中的猕猴桃进行目标检测,并使用基于 VGG16 的特征提取网络对目标进行外观特征提取。然后,利用卡尔曼滤波器结合当前帧和前两帧图像获得目标猕猴桃的运动特征与外观特征,并预测下一帧目标的预测框及其相应的外观特征。最终,通过计算当前目标外观特征与下一帧外观特征的相似度及预测框与检测框的交并比 (Intersection over union, IoU),共同构建代价矩阵,并进行 ID 匹配。

1.3.1 果实检测网络

(1) YOLO v7 检测网络

YOLO v7 为 YOLO 系列目标检测网络的最新版本,显著提升了小目标检测性能^[24],因此选择其作为猕猴桃的检测网络。

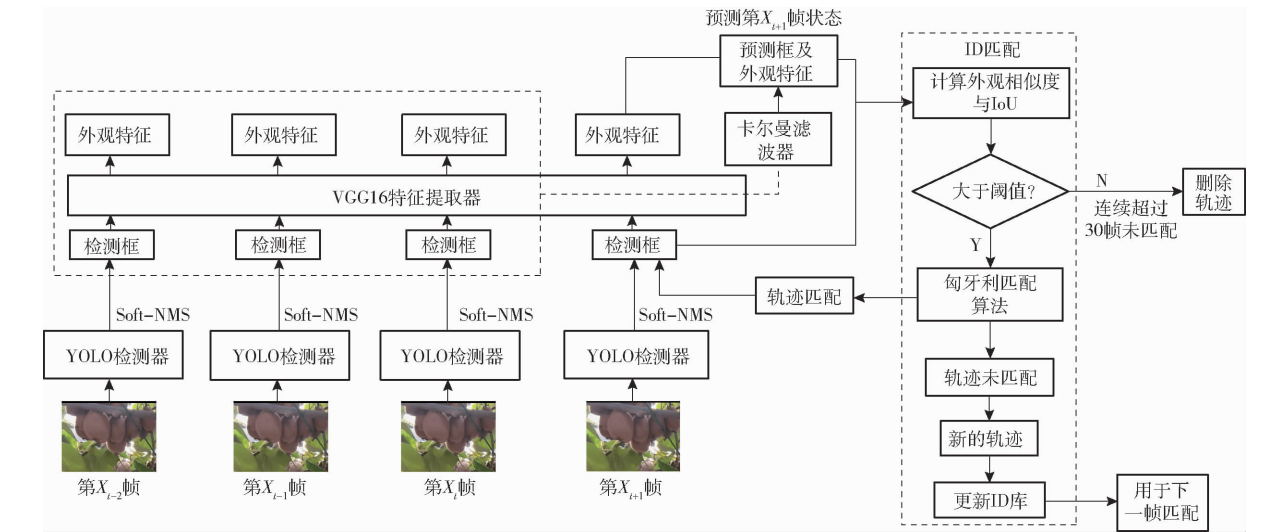


图 4 视频跟踪算法框图
Fig. 4 Architecture of video tracking algorithm

YOLO v7 结构如图 5 所示,输入图像经过数据增强等操作后,根据网络输入要求对图像进行缩放并送入 backbone 网络进行特征提取,backbone 网络通过高效的网络聚合来提高模型特征提取性能。提取后的特征分别通过双向融合网络 (Path aggregation network, PAN) 和特征金字塔网络 (Feature pyramid network, FPN) 进行融合,得到 3 种不同尺度的特征图后,融合的特征被送入检测头进行检测。由于网络会重复检测每个特征点,因此使用非极大抑制 (NMS) 来选择置信度 (con) 最高的预测框作为输出。

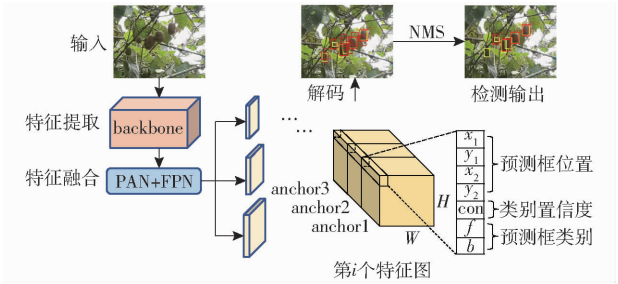


图 5 YOLO v7 结构图

Fig. 5 Architecture diagram of YOLO v7

(2) Soft - NMS 检测框筛选

由于视频中猕猴桃数据量多,重叠严重,NMS 操作可能会错误地删除相邻检测框。为了解决此类问题,本研究使用 Soft - NMS 方法^[25],通过对重叠检测框使用高斯加权来修改重叠检测框置信度,从而减少由 NMS 操作引起的重叠检测框被误删,当前检测框置信度得分计算公式为

$$s_i^* = s_i e^{-\frac{IoU(M, b_i)^2}{\delta}} \quad (1)$$

式中 s_i ——当前检测框置信度得分

M ——当前检测框

s_i^* ——加权后检测框置信

δ ——算法标准差 b_i ——重叠框

在相同置信度下,重叠度越高时,检测框置信度得分越低。如图 6 所示,A、B、C 为满足阈值条件 (置信度大于 0.8,交并比大于 0.5) 的 3 个检测框,为得到最终检测结果,NMS 处理通常保留置信度最高的检测框 A,而直接删除其邻近的检测框 B、C。该方法虽然得到了目标的唯一检测框,但是错误地删除了检测框 C。采用 Soft - NMS 时,通过对置信度加权的方式降低了邻近目标的检测框置信度,从而保留了检测框 C,提高了检测精度。

1.3.2 果实跟踪算法

由于视频中猕猴桃外观较为相似且重叠较多,为了提高追踪精度,采用结合运动轨迹与外观特征的追踪策略。其中,外观特征由 VGG16^[26] 网络进行

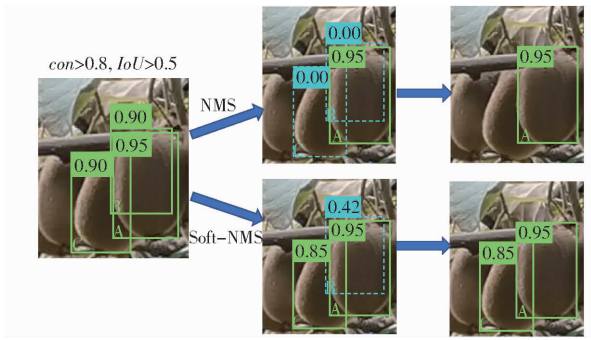


图 6 Soft - NMS 示意图

Fig. 6 Schematic of Soft - NMS

提取,运动轨迹特征提取及匹配由 DeepSort 算法实现。

(1) VGG16 特征提取网络

对于检测到的猕猴桃,为保留原始目标的外观特征信息,并同时满足网络输入尺度要求,对图像中不满足尺寸要求的部分进行补零处理,获得统一尺寸的 RGB 图像作为 VGG16 网络的输入。第 1 次卷积后,特征图尺寸不变,通道数变为 64,第 1 次池化后,特征图尺寸变为原来的 1/2,之后经过 4 次卷积与池化处理,得到 512 通道、尺寸为 7×7 的特征图,3 次全连接后的特征向量作为最终输出,得到猕猴桃目标的特征编码。

(2) DeepSort 目标跟踪

对于视频帧 t 进行检测后获得检测框坐标 d_1^t 、 d_2^t 、 $\dots$ 、 d_n^t ,其中 n 为当前视频帧中检测框的个数。VGG16 网络对检测框中的目标进行特征提取,获得目标外观特征 f_1^t 、 f_2^t 、 $\dots$ 、 f_n^t 。卡尔曼滤波器则根据当前帧 t 和前两帧 $t-1$ 、 $t-2$ 对检测到的目标进行状态预测,得到 $t+1$ 帧目标位置 d_1^{t+1} 、 d_2^{t+1} 、 $\dots$ 、 d_n^{t+1} 。将预测的运动特征与提取外观特征同时进行匹配,匹配成功即可判断为同一个猕猴桃目标。

1.3.3 基于 ID 计数的产量估计

猕猴桃产量估计需要获得视频中出现的猕猴桃总数,如图 7 所示。由于在目标跟踪算法运行时,将从数字 1 开始为每个目标赋予不同的 ID,并保证相邻视频序列中相同目标的 ID 不变,因此仅需要获取视频中 ID 最大值便可获得产量估计结果。



图 7 ID 计数示意图

Fig. 7 ID count diagram

1.4 评价指标

1.4.1 猕猴桃检测性能评估指标

为了评估网络对猕猴桃的检测性能,采用平均交并比(Mean intersection over union, mIoU)来衡量检测框与真实框之间的重合情况。在评估检测准确性方面,使用精确率(Precision)来衡量网络预测结果中是否为正样本;使用召回率(Recall)来衡量在所有正样本中网络预测为正样本的能力;使用F1值(F1 score)来衡量网络准确率和召回率的平均水平。

1.4.2 猕猴桃跟踪性能评估指标

采用多目标跟踪准确率(Multiple object tracking accuracy, MOTA)来衡量跟踪网络在每帧中正确跟踪的能力,使用多目标跟踪精确率(Multiple object tracking precision, MOTP)来衡量跟踪网络在每帧中跟踪框与实际框匹配的程度。其中MOTA与MOTP的具体公式为

$$M_{OTA} = 1 - \frac{\sum_t (F_{N_t} + F_{p_t} + I_{DSW_t})}{\sum_t G_{T_t}} \quad (2)$$

$$M_{OTP} = \frac{\sum_t \delta_{t,i}}{\sum_t c_i} \times 100\% \quad (3)$$

式中 G_{T_t} ——目标出现的总数
 M_{OTA} ——多目标跟踪准确率
 M_{OTP} ——多目标跟踪精确率
 F_{N_t} ——目标跟踪失败数
 F_{p_t} ——目标跟踪成功数
 I_{DSW_t} ——目标ID变化的次数
 $\delta_{t,i}$ ——跟踪框与检测框之间的重叠率
 c_i ——当前帧匹配成功的目标数
 t ——测试视频帧序列

1.4.3 猕猴桃计数指标

使用平均计数精度(Mean counting accuracy,

MCA)来衡量跟踪网络的计数精度,计算公式为

$$M_{CA} = \frac{1}{n} \sum_{i=1}^n \frac{P_i}{G_i} \times 100\% \quad (4)$$

式中 M_{CA} ——平均计数精度
 P_i ——检测到的目标数
 G_i ——真实目标个数
 n ——测试视频个数

2 实验结果与分析

实验环境为PC机(Windows 10操作系统,处理器为Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz,显卡为NVIDIA GeForce GTX 1660 SUPER,内存为16 GB,开发环境为Python 3.6.13)。

2.1 猕猴桃目标检测结果分析

YOLO系列网络在经过特征提取层后,通常生成 13×13 、 26×26 、 52×52 共3种不同尺度的特征图,其中尺寸越小的特征图适合越大的先验框。原始检测网络的先验框是在COCO数据集上获得,并不适用于本研究数据集,为了获得更加适合的先验框尺寸,采用K-means算法对人工标注的猕猴桃尺寸进行聚类,获得不同尺度下的先验框大小,结果见表1。

表 1 先验框聚类结果

Tab.1 Clustering results of prior boxes

特征图尺寸	感受野	先验框
13 × 13	大	10 × 18、11 × 25、16 × 32
26 × 26	中	6 × 12、7 × 16、8 × 21
52 × 52	小	3 × 7、5 × 9、4 × 12

聚类后所获得的锚框作为网络的先验框,用于目标检测。所用的数据集共包含260幅图像,10391个人工标注的检测框,作为不同检测网络的对比实验数据,多次训练的检测网络的平均测试结果见表2,可视化结果见图8。

表 2 不同网络检测结果对比

Tab.2 Comparison of detection results for different networks

模型	正检数	漏检数	误检数	精确率/%	召回率/%	F1值/%	帧速率/(f·s ⁻¹)
YOLO v3	9 165	1 226	1 123	89.08	88.20	88.64	22.78
YOLO v4	8 637	1 754	2 080	80.59	83.12	81.84	24.65
YOLO v5	8 794	1 597	2 477	78.02	84.63	81.19	25.84
YOLO v7	9 252	1 139	897	91.16	89.04	90.09	25.92
Faster R-CNN	7 832	2 559	4 422	63.91	75.37	69.17	13.24

实验结果显示,YOLO v7网络的检测结果明显优于其他网络,其精确率达到91.16%,F1值达到90.09%,每秒能够检测25.92帧图像,满足基本的视频实时检测需求,为视频跟踪提供了可靠的接口。

图8给出了不同网络检测结果的可视化对比。为了进一步评估不同网络的检测效果,本实验选取4组不同环境拍摄的猕猴桃图像数据进行检测。检测结果由检测框和置信度两部分组成,检测框内的



图 9 目标跟踪结果可视化

Fig. 9 Visualization of network tracking results

表 4 网络计数结果比较

Tab.4 Comparison of network count results

模式	视频 ID	人工统计数/个	ID 计数/个	MCA/%
困难	v1	721	609	84. 47
	v2	598	516	86. 29
	v3	475	421	88. 63
	合计	1 794	1 546	86. 18
简单	v4	439	428	97. 49
	v5	357	340	95. 24
	合计	796	768	96. 48
正常	v6	69	65	94. 20
	v7	58	55	94. 83
	合计	127	120	94. 49

量估计。采用最新的 YOLO v7 目标检测框架和 Soft-NMS 算法,得到 F1 值为 90.09% 的检测精度以及 25.92 f/s 的检测速度。目标跟踪阶段,结合训练好的 YOLO v7 检测网络、VGG16 特征提取器以及卡尔曼滤波器,通过代价矩阵和匈牙利算法实现了前后帧的目标跟踪,平均跟踪准确率为 89.87%,精确率为 82.34%,跟踪速率为 20.19 f/s。最后,通过 ID 计数实现了猕猴桃产量估计,平均计数精度达到 91.59%,简单模式下,计数精度高达 97.49%,近似于人工统计结果,困难模式下也能保持在 88.63%,进一步证明了本文算法的可行性。

参 考 文 献

[1] 王菊琴,王国军,赵汉红. 汉中猕猴桃高产高效栽培技术[J]. 果农之友, 2022(9): 36-38.
WANG Juqin, WANG Guojun, ZHAO Hanhong. High-yield and high-efficiency cultivation technology of kiwifruit in Hanzhong [J]. Friends of Fruit Farmers, 2022(9): 36-38. (in Chinese)

[2] 崔永杰,苏帅,王霞霞,等. 基于机器视觉的自然环境中猕猴桃识别与特征提取[J]. 农业机械学报,2013,44(5):247-252.
CUI Yongjie, SU Shuai, WANG Xiaxia, et al. Kiwifruit recognition and feature extraction in natural environment based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(5): 247-252. (in Chinese)

[3] 王卓,王健,王泉雄,等. 基于改进 YOLO v4 的自然环境苹果轻量级检测方法[J]. 农业机械学报, 2022, 53(8): 294-302.
WANG Zhuo, WANG Jian, WANG Xiaoxiong, et al. Lightweight apple real-time detection method based on improved YOLO v4 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8): 294-302. (in Chinese)

[4] 杨钰莹. 不同抽样方法在我国谷物产量估计中的应用[J]. 统计与管理, 2018, 33(3): 46-48.
YANG Yuying. Application of different sampling methods in grain yield estimation in China[J]. Statistics and Management, 2018, 33(3): 46-48. (in Chinese)

[5] 金勇进,杜子芳,蒋妍. 抽样技术[M]. 北京:中国人民大学出版社, 2002.

[6] 黄凯,查元源,史良胜. 基于多源数据回归分析的糖料蔗产量估计[J]. 节水灌溉, 2020, 45(6): 24-28.
HUANG Kai, ZHA Yuanyuan, SHI Liangsheng. Estimation of sugar cane yield based on multi-source data regression analysis [J]. Water-saving Irrigation, 2020, 45(6): 24-28. (in Chinese)

[7] 吴菲菲,滕延福. 玉米产量性状的统计分析及遗传力估计[J]. 农业装备技术, 2022, 48(4): 27-28, 31.
WU Feifei, TENG Yanfu. Statistical analysis and heritability estimation of yield traits of maize[J]. Agricultural Equipment Technology, 2022, 48(4): 27-28, 31. (in Chinese)

[8] 程祥云,宋欣. 果蔬采摘机器人视觉系统研究综述[J]. 浙江农业科学,2019,60(3): 490-493.
CHENG Xiangyun, SONG Xin. Review of visual system of fruit and vegetable picking robot [J]. Zhejiang Agricultural Sciences, 2019, 60(3): 490-493. (in Chinese)

- [9] 张凌栩, 韩锐, 李文明, 等. 大数据深度学习系统研究进展与典型农业应用[J]. 农业大数据学报, 2019, 1(2): 88–104.
ZHANG Lingxu, HAN Rui, LI Wenming, et al. Research progress and typical agricultural application of big data deep learning system[J]. Journal of Agricultural Big Data, 2019, 1(2): 88–104. (in Chinese)
- [10] 穆龙涛, 高宗斌, 崔永杰, 等. 基于改进 AlexNet 的广域复杂环境下遮挡猕猴桃目标识别[J]. 农业机械学报, 2019, 50(10): 24–34.
MU Longtao, GAO Zongbin, CUI Yongjie, et al. Occlusion kiwifruit target recognition in wide area complex environment based on improved AlexNet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(10): 24–34. (in Chinese)
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779–788.
- [12] 赵德安, 吴任迪, 刘晓洋, 等. 基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位[J]. 农业工程学报, 2019, 35(3): 164–173.
ZHAO Dean, WU Rendi, LIU Xiaoyang, et al. Robot apple picking positioning based on YOLO deep convolutional neural network in complex background[J]. Transactions of the CSAE, 2019, 35(3): 164–173. (in Chinese)
- [13] TIAN Y, YANG G, WANG Z, et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]. Computers and Electronics in Agriculture, 2019, 157: 417–426.
- [14] 高芳芳, 武振超, 索睿, 等. 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, et al. Apple detection and video counting method based on deep learning and object tracking[J]. Transactions of the CSAE, 2021, 37(21): 217–224. (in Chinese)
- [15] 何文乐. 无迹卡尔曼滤波 Cam-Shift 视频监控跟踪算法改进[J]. 无线电工程, 2021, 51(2): 97–103.
HE Wenle. Improvement of Cam-Shift video surveillance tracking algorithm for unscented Kalman filter[J]. Radio Engineering, 2021, 51(2): 97–103. (in Chinese)
- [16] 张宏鸣, 汪润, 董佩杰, 等. 基于 DeepSORT 算法的肉牛多目标跟踪方法[J]. 农业机械学报, 2021, 52(4): 248–256.
ZHANG Hongming, WANG Run, DONG Peijie, et al. Multi-target tracking method for beef cattle based on DeepSORT algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 248–256. (in Chinese)
- [17] LUKEZIC A, VOJIR T, CEHOVIN Z L, et al. Discriminative correlation filter tracker with channel and spatial reliability[J]. International Journal of Computer Vision, 2018, 126(7): 671–688.
- [18] LIU X, CHEN S W, LIU C H, et al. Monocular camera based fruit counting and mapping with semantic data association[J]. IEEE Robotics and Automation Letters, 2019, 4(3): 2296–2303.
- [19] WANG Q, ZHANG L, BERTINETTO L, et al. Fast online object tracking and segmentation: a unifying approach[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Long Beach, 2019: 1328–1338.
- [20] 李兆家. 猕猴桃种植技术分析 & 病虫害防治[J]. 农业开发与装备, 2023(1): 200–201.
LI Zhaojia. Analysis of kiwifruit planting technology and pest control[J]. Agricultural Development and Equipment, 2023(1): 200–201. (in Chinese)
- [21] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLO v7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. arXiv Preprint, arXiv:2207.02696, 2022.
- [22] 杜豫川, 岳光华, 刘成龙, 等. 探地雷达多特征融合的城市空洞自动识别方法[J]. 中国公路学报, 2023, 36(3): 108–119.
DU Yuchuan, YUE Guanghua, LIU Chenglong, et al. Automatic identification method of urban void fusion of ground penetrating radar[J]. China Journal of Highway and Transport, 2023, 36(3): 108–119. (in Chinese)
- [23] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//Proceedings of International Conference on Image Processing, 2018: 3645–3649.
- [24] 戚玲珑, 高建瓴. 基于改进 YOLO v7 的小目标检测[J]. 计算机工程, 2023, 49(1): 1–10.
QI Linglong, GAO Jianling. Small target detection based on improved YOLO v7[J]. Computer Engineering, 2023, 49(1): 1–10. (in Chinese)
- [25] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS-improving object detection with one line of code[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 5561–5569.
- [26] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of International Conference on Learning Representations, 2015: 1–14.