

DOI: 10.13733/j.jcam.issn.2095-5553.2025.07.023

杨义, 吴怡婧, 蒋学芹, 等. 基于深度学习的观光农业中的桃子采摘识别[J]. 中国农机化学报, 2025, 46(7): 153-163

Yang Yi, Wu Yijing, Jiang Xueqin, et al. Peach picking recognition in agritourism based on deep learning [J]. Journal of Chinese Agricultural Mechanization, 2025, 46(7): 153-163

基于深度学习的观光农业中的桃子采摘识别^{*}

杨义¹, 吴怡婧¹, 蒋学芹¹, 张洁¹, 万雪芬^{2, 3}

(1. 东华大学信息科学与技术学院, 上海市, 201620; 2. 华北科技学院计算机科学与工程学院, 河北廊坊, 065201; 3. 河北省物联网监控工程技术研究中心, 河北廊坊, 065201)

摘要: 针对桃子采摘园智慧化管理的需求, 提出一种基于深度学习的采摘识别方法。利用机器视觉与深度学习技术, 在轻量级人体姿态估计算法 Lightweight OpenPose、目标检测算法 YOLOv5s、目标跟踪算法 DeepSORT 的基础上, 提出桃子采摘行为检测方法。该方法按照功能顺序可分为基于人体关节角度的采摘姿态判定方法、基于最近邻检索的采摘目标确定方法及其优化、基于设定状态标志的采摘目标检测失效解决方法 3 个功能步骤。基于实际桃子采摘视频数据建立数据集, 进行相关性能测试。将基于人体关节角度方法与传统采用人体关节点外接矩形框的方法进行对比, 本方法对采摘举手动作的判定查准率 P 提高 16%。针对采摘目标判定问题, 基于最近邻检索的方法相比于传统的基于距离与参照物尺寸对比的方法、基于交并比 IoU 与阈值对比的方法, 查准率 P 至少提高 11%。基于设定状态标志的采摘目标检测失效方法, 较好地解决手部遮挡对检测结果的影响, 查准率 P 提高 39%。在此基础上, 设计试验系统, 在真实情境下对本方法进行测试。结果表明, 提出的桃子采摘识别方法能够在采摘桃园实际环境下完成对采摘动作的有效准确识别。

关键词: 智慧农业; 观光农业; 桃子; 采摘识别; 深度学习; 人体姿态

中图分类号: S24 **文献标识码:** A **文章编号:** 2095-5553 (2025) 07-0153-11

Peach picking recognition in agritourism based on deep learning

Yang Yi¹, Wu Yijing¹, Jiang Xueqin¹, Zhang Jie¹, Wan Xuefen^{2, 3}

(1. College of Information Science and Technology, Donghua University, Shanghai, 201620, China;
2. College of Computer Science and Engineering, North China Institute of Science and Technology, Langfang, 065201, China; 3. Hebei IoT Monitoring Engineering Technology Research Center, Langfang, 065201, China)

Abstract: For the needs of intelligent management of peach picking tourism orchards, a deep learning-based picking recognition method is proposed. The method uses machine vision and deep learning technologies to integrate a lightweight human posture estimation algorithm Lightweight OpenPose, a target detection algorithm YOLOv5s, and a target tracking algorithm DeepSORT to develop a peach picking behavior detection approach. It can be divided into three steps according to the functional order: the picking posture determination method based on the human body joint angles, the picking target determination method based on nearest neighbor retrieval and its optimization, and the picking target detection failure solution method based on the set status flags. A dataset is established based on the actual peach picking videos for performance tests. Comparing the method based on the angle of human joints proposed in this paper with the traditional method of using bounding boxes enclosing human joints, the method in this paper can improve the precision of determination (P) rate of hand-raising action by 16%. For the problem of determining the picking target, the nearest neighbor retrieval approach outperforms both the traditional method based on the comparison of distance and reference size and the method based on the comparison of IoU and thresholds, with an increased P rate by at least 11%. The picking target detection failure solution method based on set status flags effectively solves the influence of hand occlusion on

收稿日期: 2023 年 10 月 9 日 修回日期: 2023 年 11 月 22 日

^{*} 基金项目: 国家自然科学基金(12272092); 河北省物联网监控工程技术研究中心项目(3142018055, 3142016020); 廊坊市科学技术研究与发展计划资助项目(2021011035); 秦皇岛市科学技术研究与发展计划项目(201805A016); 东华大学 2025 年本科实践教学改革与建设项目

第一作者: 杨义, 男, 1978 年生, 成都人, 博士, 副教授; 研究方向为物联网技术及智慧农业。E-mail: yiyang@dhru.edu.cn

通讯作者: 万雪芬, 女, 1979 年生, 江苏连云港人, 硕士, 副教授; 研究方向为物联网技术及智慧农业。E-mail: calmerd@ncist.edu.cn

the detection results, substantially improving the P rate by 39%. On this basis, an experimental system is designed to test the proposed method under real-world conditions. The results show that the proposed peach picking recognition method achieves effective and accurate recognition of picking actions in actual orchard environments.

Keywords: smart agriculture; agritourism; peach; picking recognition; deep learning; human postures

0 引言

桃子是我国重要的果树品种,也是农业采摘园中常见的果树之一^[1]。相比于大棚环境下的果树种植,桃树通常采取露天栽培方式。在采摘园区内,游客的分布呈现较为分散的特点。游客的采摘行为往往与其采摘兴趣等个人偏好密切相关,通过对游客采摘动作的监督与分析,可以间接评估其采摘兴趣等。然而,近年来随着人工成本不断攀升等因素影响,依靠大量人力对游客行为进行监督分析的策略已变得不切实际。因此,亟须采用新的农业技术对桃子采摘果园进行科学高效管理。

深度学习技术在家禽家畜个体与姿态识别、农作物及杂草识别、农作物病虫害检测等农业领域应用广泛^[2, 3]。如李菊霞等^[4]采用 YOLOv4 模型对猪只饮食行为进行检测,试验结果表明,该方法在不同视角、不同遮挡程度以及不同光照下均能实时有效地检测猪只饮食行为。胡志伟等^[5]在特征金字塔网络 FPN 中引入融合通道注意力与空间注意力的双重注意力单元,实现对复杂场景中生猪实例的精确检测。彭明霞等^[6]将特征金字塔网络 FPN 融入 Faster R-CNN,并选用残差卷积网络提取图像特征,使模型在田间复杂条件下快速、准确识别棉花与杂草。刘莫尘等^[7]采用融合带色彩恢复的多尺度视网膜增强算法对 YOLOv4—tiny 模型进行改进,实现在复杂实际场景中对玉米苗和杂草的高效准确分类。鲍文霞等^[8]采用基于多路卷积神经网络的方法,对大田环境下麦穗赤霉病的识别精度高达 100%。甘雨等^[9]引入注意力模块 CA 对 EfficientNet 网络进行改进,构造出 CA—EfficientNet 模型。该模型有效提高了作物害虫的识别准确率,并大幅减少模型参数量,使其易于部署到移动端。Dyrmann 等^[10]提出了一种基于 DetectNet 网络实现复杂环境条件下对麦田中杂草的检测方法。Amara 等^[11]对 LeNet 架构进行改进,实现对现实场景中的香蕉树叶片进行高鲁棒性疾病检测。Veeraballi 等^[12]提出了一种改进的 ResNet50 网络对木瓜植物的叶片疾病进行识别和分类。上述研究表明,利用深度学习技术可以较好地农业领域的动、植物等生物对象及其行为进行识别。

具体到果园场景中,深度学习的应用主要集中在

对果树病虫害识别以及对果树各个部位的识别(如识别果实,进而进行果实计数、果实目标跟踪等问题的研究)。彭红星等^[13]将经典 SSD 深度学习模型中的 VGG16 输入模型替换为 ResNet—101 模型,解决了重叠和受遮挡果实的目标检测问题。穆龙涛等^[14]采用改进的 AlexNet 作为 Faster R-CNN 的特征提取层,降低了网络复杂度,减少了计算量,并且具有较高的检测精度。刘芳等^[15]将其设计的含有残差模块的 DarkNet—20 网络替换 YOLO 的主干网络,并融合多尺度检测模块,更好地实现特征提取。Rahmehoonfar 等^[16]提出一种基于改进 Inception—ResNet 架构的方法获取多尺度特征,用于实现果实的准确计数以及产量估计,实现了 91% 的准确性,而原始 Inception—ResNet 的准确性为 76%。Koirala 等^[17]基于 YOLO 检测框架设计了 MangoYOLO,该架构比 YOLOv1 更深,但比 YOLOv3 更浅,用于芒果果实检测任务中的内存和速度优化,并且模型对水果遮挡及可变照明条件等问题具有鲁棒性。Kestur 等^[18]提出了深度全卷积神经网络 MangoNet 来分割图像中的芒果果实,通过连接对象对图像中的果实进行计数,该网络的 $F1$ 达 0.84,较传统目标检测网络有显著提高。

本文重点研究在观光采摘果园中识别游客等对象的采摘行为,该采摘行为不仅涉及对人体姿态的检测,也需要结合对果实空间位置变化的检测。采摘过程中,人体姿态动态变化,果实位置也会发生较大位移。以往的研究中,对果园情境中人体姿态识别的研究较为不足,多侧重于对果实及其位置进行识别与跟踪的研究工作,目的大多是为了进行产量估计,检测过程中果实位移量不大。此外,采摘行为的识别还需要对人体与果实接触情况进行判定,目前关于这方面的研究也有所不足。

考虑到对树上果实的采摘行为进行识别的任务涉及对人体动作以及果实位置两个方面的判断,因此,需要引入人体姿态估计算法和目标检测算法作为采摘识别任务的基础算法。由于采摘行为是一个动态过程,因而还需要结合目标跟踪算法对不同帧之间的果实进行关联。因此,本文融合基于人体姿态估计算法、目标检测算法和目标跟踪算法,针对现实场景中的具体问题对上述算法模型进行优化,构建基于深度学习的果树采摘识别架构,从而实现对果实采摘结果的准确判定。

1 果实采摘行为检测网络架构

1.1 网络架构 YOLOv5

YOLOv5 是一种单阶段目标检测算法,该算法在前几代 YOLO 算法的基础上进行改进,使其在检测速度与精度方面得到较大提升^[19]。YOLOv5 在主干网络中采用原创的 Focus 模块,将信息从宽高维度集中到通道维度,提高每个点的感受野,有效减少参数数量和计算量,提升推理速度。并且采用 C3 模块替换原来主干网络中的 BottleneckCSP 模块,使模型在性能不下降的情况下精简网络结构,减少计算量,降低模型推理时间。此外,将原本 SPP 模块中 3 个并行的最大池化层替换为串行结构的 SPPF 模块,大大减少前向计算和反向计算的时间。由于 YOLOv5 集成了各种高效模块,使其获得高精度和高实时性的检测效果。因此,若将其应用于果实采摘识别任务中,可望实现对果实检测目标的快速准确定位,为采摘目标判断以及果实采摘位移过程中跟踪果实目标提供稳定准确的位置信息。相比于 YOLOv5 其他版本,YOLOv5s 结构最为简洁,模型小、计算量小、速度快,特别适用于本文的应用情景。

1.2 Lightweight OpenPose 网络

Lightweight OpenPose 是在 OpenPose 基础上提出的一种轻量版^[20]。OpenPose 网络可根据输入图片生成一个部分置信图(PCM)和一个部分亲和场(PAF)。前者生成用来预测关键点位置的热力图(Heatmap);后者是在关键点之间建立的一个向量场,描述肢体的方向。获得热力图和部分亲和场后,使用二分图最大权匹配算法来对关键点进行局部关联,并用匈牙利算法(Hungarian Algorithm)求得相连关键点最优匹配,从而组合成人体骨架,描述人体姿态。Lightweight OpenPose 的流程与 OpenPose 基本一致,但在 OpenPose 的基础上做了以下改进:将 OpenPose 的主干网络 VGG19 替换为轻量型网络 Mobile v1,大幅减少模型参数量。在此基础上,把生成 keypoints 和 PAFs 的两个网络合并成一个,并将 7×7 卷积换成 3×3 和 1×1 的卷积。这些操作共同作用使得 Lightweight OpenPose 在检测速度方面的性能大幅提升,并且在参数量只有 OpenPose 的 15% 的情况下,保持较高的检测精度。因此,若将其应用于采摘识别模型当中,可以使其减少对高性能计算机的依赖,扩大模型的应用范围,并且可以保持较好的实时性。

1.3 DeepSORT 网络

DeepSORT 的前身是 SORT 算法^[21]。SORT 算法通过目标检测器在第一帧检测到的结果创建对应的轨迹。将卡尔曼滤波的运动变量初始化,通过卡尔曼滤波

获得轨迹在下一帧的预测框。在下一帧中将目标检测框和轨迹预测框进行 IoU 匹配,再通过匹配结果计算其代价矩阵。将其输入匈牙利算法,可以得到 3 种不同的线性匹配结果。第 1 种是 Tracks(代表轨迹信息)失配,将失配的 Tracks 删除;第 2 种是 Detections(代表通过目标检测算法获得的检测框)失配,将其初始化为一个新的 Track;第 3 种是目标检测框和卡尔曼滤波预测框匹配成功,说明前一帧和后一帧追踪成功,将其对应的 Detections 通过卡尔曼滤波更新到对应的 Tracks 变量。然而 SORT 算法在物体发生遮挡时,存在容易丢失 ID 的问题。DeepSORT 算法是 SORT 算法的改进版本,在 SORT 算法的基础上增加新轨迹的状态确认和级联匹配。

在新轨迹的状态确认方面,将 Tracks 分为确认态(Confirmed)和不确认态(Unconfirmed)。新产生的 Tracks 属于不确认态;不确认态的 Tracks 必须和 Detections 连续匹配一定的次数(默认 3 次)才可以转化成确认态。确认态的 Tracks 必须和 Detections 连续失配一定次数(默认 30 次),才会被删除。在级联匹配方面,根据卷积神经网络提取到的目标物体的外观特征和运动特征对确认态的 Tracks 与 Detections 进行级联匹配。通过计算前后两帧目标之间的匹配程度,实现轨迹关联,并为每个追踪到的目标分配 ID。通过 DeepSORT,可以高效准确地通过视频数据追踪果园场景中的对象,实现对采摘过程中的目标追踪检测。

2 果实采摘行为检测方法研究

2.1 果实采摘行为检测方法

对树上果实的采摘行为定义为举起手、手接触果实、摘下果实 3 个环节的一个过程,因此,将果树采摘行为的识别等效为对以上 3 个环节进行判定。若采用深度学习方法对这 3 个环节进行判定,即涉及目标检测技术、人体姿态估计技术和目标跟踪技术的应用。然而直接采用原始算法,难以满足实际场景的需求,所以在前述算法的基础上针对采摘识别情景做出 3 处重要适配改进。算法结构如图 1 所示。

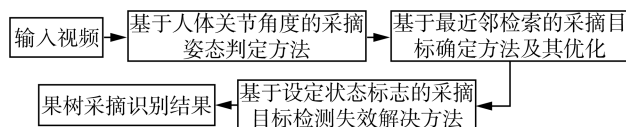


图1 果树采摘行为检测方法整体结构

Fig. 1 Overall structure of fruit tree picking behavior detection method

2.2 基于人体关节角度的采摘姿态判定方法

对采摘举手的判定涉及人体姿态,而与采摘时的举手动作最紧密相关的部位是人的前臂。人的前臂是由手肘关节和手腕关节相连,可以通过对手肘关节和手腕

关节的位置检测,分析前臂的活动情况。对于举起手这个动作,可以通过前臂与水平线之间的夹角 θ 进行判定。如图 2 所示,以左手采摘动作为例,分析夹角 θ 与左手肘和左手腕关节点坐标之间的关系。手肘和手腕关节点坐标可以通过 Lightweight OpenPose 模型给出。

通过图 2 中对应关系可以得出夹角 θ 的正弦值以及 θ 的角度大小,如式(1)和式(2)所示。

$$\sin\theta = \frac{y_{kpt6} - y_{kpt7}}{\sqrt{(x_{kpt6} - x_{kpt7})^2 + (y_{kpt6} - y_{kpt7})^2}} \tag{1}$$

$$\theta = \sin^{-1}\left(\frac{y_{kpt6} - y_{kpt7}}{\sqrt{(x_{kpt6} - x_{kpt7})^2 + (y_{kpt6} - y_{kpt7})^2}}\right) \tag{2}$$



图 2 夹角 θ 与手肘和手腕关节点坐标之间的关系

Fig. 2 Relationship between the angle θ and the coordinates of the elbow and wrist joint points

由于在采摘场景中还存在许多抬起前臂但并未进行采摘的动作,如图 3 所示。因此,不能简单地根据前臂抬起(即 $\sin\theta > 0$)这一条件来判定采摘时的举手动作,需要进一步区分采摘动作与采摘干扰动作。



图 3 干扰动作示例

Fig. 3 Interference action example

通过对数据集的观察,发现采摘动作与其他干扰动作前臂抬起的幅度大小有明显差异。因此,对数据集中不同动作的前臂抬起幅度进行统计。采用 $\sin\theta$ 的值来衡量前臂抬起幅度大小。对视频数据集中不同动作的 $\sin\theta(\theta > 0)$ 进行统计得到表 1。

表 1 数据统计表

Tab. 1 Statistical table of data

参数	采摘动作	走路姿态	整理东西	其他动作
$\max(\max(\sin\theta_i))$	0.96	0.34	0.47	0.54
$\min(\max(\sin\theta_i))$	0.67	0.18	0.23	0.18

注: $\max(\sin\theta_i)(i \in n, n$ 为一个视频包含的帧数)表示某一类动作在某一个视频中的最大 $\sin\theta_i$ 值,代表视频中此类动作的角度正弦峰值。 $\max(\max(\sin\theta_i))(j \in m, m$ 为数据集中此类视频的数量)与 $\min(\max(\sin\theta_i))$ 表示在数据集所有视频中该类动作的角度最大正弦值波动范围。

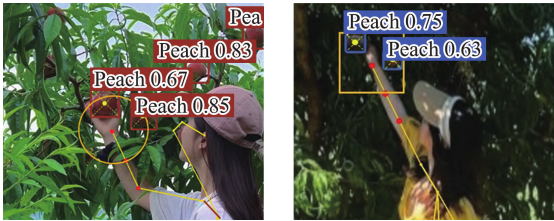
由表 1 可知,可以通过设置阈值为 0.6,将采摘举手动作和干扰动作区分。据此划分不同区间,并通过设置采摘举手标志 PickPutup 对不同区间进行表示。当前臂摆动幅度满足 $\sin\theta > 0.6$ 时,视为采摘举手动作,以此作为采摘举手的判定条件。此时可以继续后续采摘相关检测。而对摘下果实的判定条件为:若在手放下状态时采摘目标果实的检测框中心点位置比手腕关节点低,即视为果实被摘下。

2.3 基于最近邻检索的采摘目标确定方法及其优化

在果实采摘识别的环节中还有一个是对手接触果实的判定。传统对手接触物体的判定方法有两种:(1)当手腕关节点与目标物体中心的距离小于预设参照物的尺寸时,判定人手接触该物体。预设参照物一般取人手的长度,以往多采用前臂长度的 1/2 代表人手的长度;(2)利用交并比(IoU)算法,以手腕关节点为中心自设矩形框,边长为手长的 2 倍。当手腕矩形框与目标物体的交并比值超过设定阈值时可判定人手接触该物体。

由于对手是否接触果实进行判定是为了确定采摘目标,以便在下一步的摘下果实判定环节对相应果实的位置进行比对。说明判定手是否接触果实问题可以进一步转化为确定采摘目标问题,但是需要保证结果的唯一性。

关于基于手腕与目标物体之间的距离或交并比这两种方法,当有多个果实距离人手较近时,满足条件的果实可能不唯一,如图 4 所示,这样就难以保证确定采摘目标的唯一性。



(a) 判定方法 1 (b) 判定方法 2

图 4 传统方法检测效果示例

Fig. 4 Example of the effect of traditional methods of detection

利用最近邻检索方法找出距离手腕关节点最近的果实目标中心点。其对应的果实目标具有唯一性,可用于进行采摘目标的判定。图 5 为某一帧中查找距离手腕关节点最近果实的流程。

此处确定的果实目标只是在某一时刻距离手腕最近,显然不足以被判定为采摘目标。采摘目标对应的果实必然是在一段时间内多次距离手腕关节点最近的果实,那么便需要解决此处时间长短的问题。在视频中一般用连续的帧数来反映一段时间的长度,因此,问

题就转化为选取合适的检测帧数。

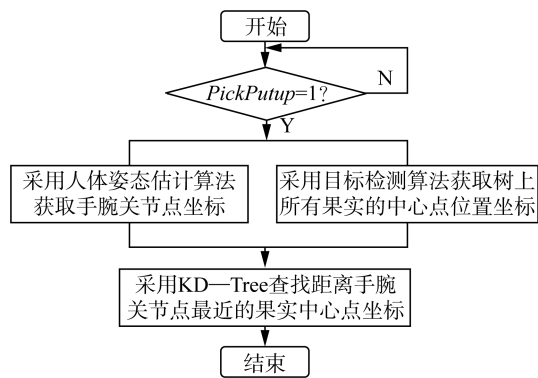


图 5 某一帧中查找距离手腕关节点最近果实的流程图
Fig. 5 Flowchart for finding the closest fruit to a wrist node in a given frame

为此,在数据集上进行多轮试验与验证,发现当取 20 帧作为连续检测帧数的范围时,本方法检测采摘目标的准确性最高。但是依然存在小部分没有准确检测出采摘目标的情况。通过对检测失败的视频进行分析,发现检测错误的情况主要分为两种:一种是如图 6(a)所示的采摘目标与其他果实到手腕关节点的距离差不多的情况;另一种是如图 6(b)所示,手越过近处果实未采摘而采摘较远处的果实。



(a) 检测错误情况 1 (b) 检测错误情况 2

图 6 未准确检测出采摘目标的情况示例

Fig. 6 Example of a situation where the picking target was not accurately detected

读取图 6 中的采摘目标错检情况所对应的 20 帧数据,结果如表 2 所示。在图 6 所示的情况中,虽然被错检为采摘目标是连续 20 帧中最多次距离手腕关节点最近的果实,但并非是在检测范围的最后几帧中检测到最近的果实。说明此时手腕关节点已经远离该果实,该果实不应该被判断为采摘目标。

由此说明,根据前述条件并不足以确保对采摘目标的准确检测。因此,需要对确定采摘目标的条件进行优化,对其补充条件进行进一步约束。

通过观察数据集上的采摘动作,可以发现在手接触到采摘目标前的一小段时间内,采摘目标一直是距离手腕关节点最近的果实。将其转化为以下约束条件:对连续 20 帧这一检测范围进行划分,找出前 n 帧中与手腕关节点距离最近的频率最高的果实,将其视为潜在采摘目标。在后面的 $20-n$ 帧中,当其连续满

足成为潜在采摘目标的条件,则判定为采摘目标。采摘目标的检测流程如图 7 所示。

表 2 图 6 所示场景的检测数据

Tab. 2 Detection data for the scenario shown in figure 6

图 6(a)对应数据				图 6(b)对应数据			
第 n 帧	最近果实 id	第 n 帧	最近果实 id	第 n 帧	最近果实 id	第 n 帧	最近果实 id
156	1	166	2	97	1	107	1
157	1	167	2	98	1	108	2
158	1	168	2	99	1	109	2
159	1	169	2	100	1	110	2
160	2	170	2	101	1	111	2
161	2	171	1	102	1	112	2
162	2	172	1	103	1	113	2
163	2	173	1	104	1	114	2
164	2	174	1	105	1	115	2
165	2	175	1	106	1	116	2

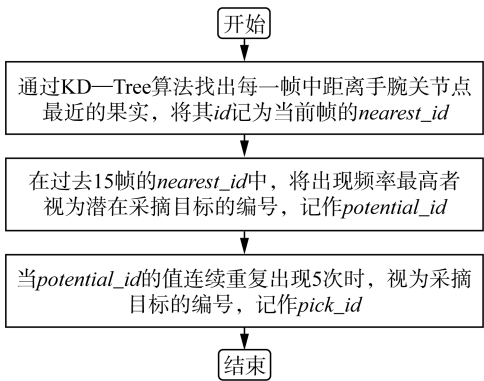


图 7 采摘目标的检测流程

Fig. 7 Flow chart for detection of picking targets

2.4 基于设定状态标志的采摘目标检测失效解决方法

通过数据集上的视频数据可以发现,在采摘过程中,当人手触及果实时往往会对果实造成遮挡,使目标检测算法无法检测出采摘目标,如图 8(a)所示,进而导致丢失对采摘目标的跟踪,如图 8(b)所示。

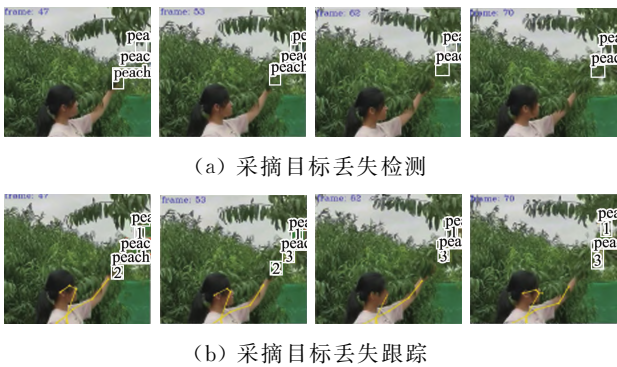


图 8 采摘目标丢失检测与跟踪示例

Fig. 8 Example of picking target loss detection and tracking
如果继续采用前述算法进行采摘目标的判断,那

此时检测到的距离手腕关节最近的果实并非采摘目标,这将导致采摘目标检测错误,并将影响到后续对摘下果实进行判定的准确性。

通过对模型功能需求的分析,发现在后续环节中只需用到采摘目标 id 这一参数,因此,没有必要在确定采摘目标后继续对其进行检测。针对这一问题,设定不进行采摘目标检测的标志 $Nodetect$,以避免模型对采摘目标的误检测。 $Nodetect$ 值为 0 时,表示当前需要对采摘目标 id (即 $pick_id$) 进行检测,值为 1 时则相反。将 $Nodetect$ 初始值设为 0。当检测出 $pick_id$ 后,将 $Nodetect$ 的值置为 1,表示在本次采摘动作过程中对 $pick_id$ 的检测结束,不再进行检测。当一次采摘动作结束,则将 $Nodetect$ 的值置 0,恢复到初始状态,为新一次的采摘动作检测做准备。图 9 为在一次采摘行为过程中 $Nodetect$ 取值的变化情况。

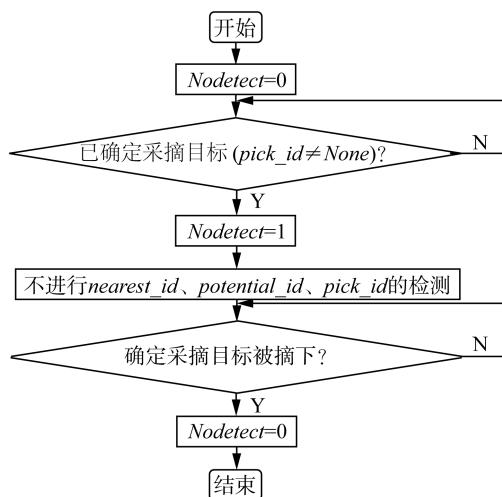


图 9 一次采摘行为过程中 $Nodetect$ 取值的变化情况

Fig. 9 Changes in $Nodetect$ values during one picking behavior

3 试验结果与分析

3.1 数据集制作

由于果实采摘识别主要涉及对人体动作以及果实位置变化的分析,因此,为满足模型训练的需要,不仅需要建立用于采摘动作检测的游客行为视频数据集,还需要构建用于果实目标检测的果实图像数据集。

试验果园位于上海市闵行区浦江镇,以当地广泛种植的桃子作为研究对象,采集游客在园中进行的采摘等行为活动的视频数据。在此基础上截取桃子图像,制作桃子图像数据集。

1) 采摘园中游客行为视频数据集。试验中,在园中选取不同合适位置设立支架,搭载摄像机拍摄以采摘桃子为主的园中游客行为活动的视频。共拍摄 120 段视频,其中包含采摘行为的视频有 78 段。每段视频时长 30 s 左右,视频帧率为 30 帧/s,保存为

MP4 格式。为弥补场地的单一性对人的行为活动多样性造成的限制,在网上搜集合适的视频对数据集进行补充。在网上下载 281 个在采摘果园环境中的包含游客的视频。通过人工挑选,筛除镜头切换太频繁导致单一个体出镜时间太短无法判断其行为以及个体双手全被遮挡无法判断手部活动等情况的视频。保留符合要求的有效视频数据 89 个。此视频数据集共包含 209 个视频,部分视频示例如图 10 所示,其中包含采摘行为的有 126 个,采摘动作共计 278 次。其他不包含采摘行为的视频大致分类及对应数量如下:走路视频 32 个,整理东西视频 21 个,其他行为视频 30 个。



图 10 视频数据集中部分视频示例

Fig. 10 Selected video examples from the video dataset

2) 桃子图像数据集。首先,从以上视频数据集中截取包含桃子的图像,得到果园环境中桃子图像 1 268 幅。然后,使用 LabelImg 软件对这些图像中的桃子果实位置进行标注。标注时,对于肉眼无法清楚判断是否存在果实的模糊像素点不进行标注。标注后得到的 xml 文件包含桃子果实的中心坐标、宽、高信息。为增加图像的多样性,避免模型在训练时发生过拟合的情况,采用数据增强的方式对图像进行扩增。采用水平翻转、旋转角度、随机颜色、对比度增强和颜色增强的数据增强方法,效果如图 11 所示。

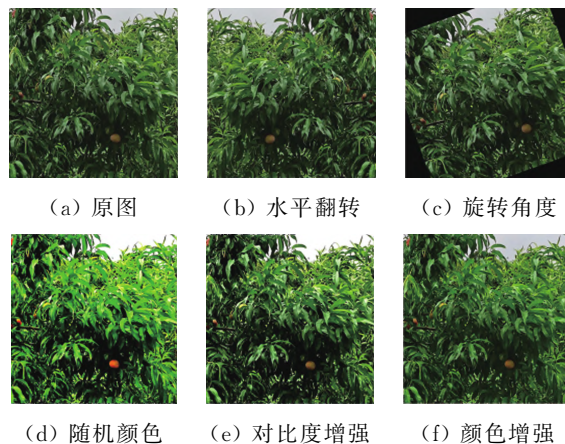


图 11 数据集数据增强示例

Fig. 11 Example of dataset data enhancement

采用数据增强方式扩增后得到 7 608 幅图像及其对应的 xml 文件。由于 YOLO 系列算法在训练时需要 txt 格式的标注文件,因此,通过 python 编程将 xml 文件转化为 txt 文件。将其与对应的图像作为本试验的桃子图像数据集,并按照 8 : 2 的比例将其划分成训练集与测试集。

3.2 试验平台与评价指标

试验硬件环境:CPU 为 i7-12700H,GPU 为 NVIDIA RTX 3060。操作系统为 Windows 10,基于 PyTorch 1.7 平台实现对模型的构建与训练。软件工具为 PyCharm 2021,编程语言为 Python 3.7。并在 Anaconda 虚拟环境中安装 OpenCV 等依赖库,搭建试验环境。

对采摘行为识别模型最重要的评价标准是算法识别的准确性,并且针对已有模型的改进方法也是为了 提高对采摘行为识别的准确性。因此,在试验中以算法识别的准确性作为主要评价标准,评估提出方法的有效性。在采摘识别问题中可以用查准率 P 、查全率 R 两个指标来衡量识别结果的准确性。查准率 P 和查全率 R 计算如式(3)和式(4)所示。

$$P = \frac{TP}{TP + FP} \times 100\%$$

(3)

$$R = \frac{TP}{TP + FN} \times 100\%$$

(4)

式中: TP ——实际为正样本且被检测为正样本的数量;

FP ——实际为负样本但被检测为正样本的数量;

FN ——实际为正样本但被检测为负样本的数量。

提出的方法按照功能顺序分别为基于人体关节角度的采摘姿态判定方法(用于采摘姿态识别)、基于最近邻检索的采摘目标确定方法及其优化(用于判定采摘目标)、基于设定状态标志的采摘目标检测失效解决方法(人手遮挡误检测抑制)3 个功能步骤。为验证提出的方法在实际采摘果园应用场景中的使用效果,按照功能步骤。

3.3 识别采摘姿态的不同方法比较

试验对采用不同方法识别采摘姿态的准确性进行对比。传统上对人体姿态进行识别是根据人体关节点外接矩形框的宽高比值范围对人体姿态进行判定。基于自制数据集对采摘姿态的人体框宽高比值进行统计,发现对于采摘姿态该值范围为(0.38,0.45),宽高比算法可以达到最佳效果,因而,在试验中以此作为宽高比算法采摘姿态的判定条件。提出的方法为采用基于关节点角度的方法对采摘姿态进行判定,当前臂与水平线

夹角的正弦值 $\sin\theta > 0.6$ 时,则判定为采摘姿态。

采用视频数据集的全部视频进行试验,以查准率 P 和查全率 R 作为衡量不同方法识别准确性的指标。在试验中将 TP 定义为实际为采摘姿态且被识别为采摘姿态的次数, FP 定义为实际为其他姿态但被识别为采摘姿态的次数, FN 定义为实际为采摘姿态但未被识别为采摘姿态的次数。结果如表 3 所示。

表 3 采用不同方法识别采摘姿态的结果
Tab. 3 Results of using different methods to identify picking gestures

方法	TP	FP	FN	P	R
人体框宽高比方法	249	56	29	0.82	0.90
关节点角度方法	278	5	0	0.98	1

由表 3 可知,采用人体框宽高比方法识别采摘姿态的查准率 P 显著低于本文方法。结合式(3)与表 3 可知,此处查准率 P 较低,主要是由 FP 较高引起的。而 FP 较高,说明该方法将许多其他姿态误识别为采摘姿态。图 12(a)~图 12(c)展示了采用人体框宽高比方法而误将其他姿态判定为采摘姿态的例子。从这 3 个错误示例可以看出,仅根据人体关节点外接矩形的宽高比这一条件不足以界定采摘姿态,因为还有许多其他姿态也可以对应到采摘姿态的人体关节点外接矩形宽高比范围中。说明该方法的判定条件难以有效区分采摘动作与其他动作。另外,现实情况下往往存在设备未能完整采集到人体图像的问题,如图 12(c)所示,严重干扰到此种方法的检测结果,说明该方法难以应用到复杂的实际场景中。



图 12 不同方法对采摘姿态的识别结果对比
Fig. 12 Comparison of the recognition results of different methods for picking gesture

图 12(d)~图 12(f)展示本方法在相应视频上的识别结果。结果显示,对于采用人体框宽高比方法未能准

确判定的姿态,本方法均可正确判定。相比之下,本文基于关节点角度的判定方法与采摘姿态的契合度更强,因而,在查准率和查全率指标上的表现更佳。

3.4 判定采摘目标的不同方法比较

试验对采用不同方法进行采摘目标判定的准确性进行对比。采摘目标的判定问题实质上是判定人手是否接触物体问题的延伸。以往对人手接触物体的判定方法主要有基于手腕与目标物体之间的距离与交并比这两种。将其应用到判定人手是否接触果实的问题上,以上两种方法可以进一步描述:前者是当手腕关节点与目标物体中心的距离小于预设参照物的尺寸时,判定人手接触该物体;后者以手腕关节点为中心自设矩形框,当手腕矩形框与目标物体的交并比值超过设定阈值时,判定人手接触该物体。所提出的方法是基于最近邻检索算法找出距离手腕关节点最近的果实,在此基础上进行人手接触果实的判定。

采用两种传统方法进行试验时,需要对其涉及的参数设定合适参数值。采用距离与参照物尺寸进行对比的方法需要选取参照物并确定其尺寸,沿用以往方法中以人手作为参照物并取前臂长度的一半来代表人手长度的做法。采用交并比值与阈值对比方法需要选取自设矩形框的边长和交并比的阈值,自设矩形框的边长可沿用以往方法选定为手长的 2 倍,而交并比的阈值则需要结合实际统计来进行选取。为此,通过在数据集上进行试验与统计,发现在采摘行为中,以手腕关节点为中心的自设矩形框与目标果实的检测框二者的交并比值一般大于 0.9。因此,选取 0.9 作为该方法的交并比阈值。

以上讨论的是判定手是否接触果实问题的 3 种方法,应进一步转化使其可应用于解决采摘目标判定问题。采用 3 种方法进行采摘目标判定的具体试验方案:以发生采摘举手动作为采摘行为的开端,在此基础上进行采摘目标的确定。以发生采摘举手动作后的连续 20 帧作为采摘目标检测范围,3 种方法各自对应采摘目标的判定条件如表 4 所示。

表 4 不同方法对采摘目标的判定条件
Tab. 4 Judgment conditions of different methods for picking targets

方法	判别条件
方法 1: 距离与参照物尺寸对比方法	符合与手腕关节点的距离小于手长度的条件最多次的果实视为采摘目标
方法 2: IoU 与阈值对比方法	符合与手腕矩形框交并比大于 0.9 的条件最多次的果实视为采摘目标
方法 3: 本文的最近邻检索方法	符合距离手腕关节点最近的果实视为采摘目标

由表 5 可知,采用前两种方法对采摘目标的判定都存在查全率高而查准率低的问题。通过追溯采用这两种方法而未被正确判定采摘目标的视频,对查准率低的原因进行分析。图 13 选取试验采取的不同方法在 2 个典型场景下的检测结果进行展示。这 2 个场景的最大差别在于果实目标大小以及人的手臂长短。在第 1 个场景中,采用方法 1,有两个果实满足判定条件,这是因为这两个果实距离手腕关节点都比较近;而采用方法 2 和方法 3 都只确定了 1 个果实目标。在第 2 个场景中,采用方法 2,有 2 个果实满足判定条件,这是由于在此场景中果实目标较小而且人的手臂较长,使交并比的值较大,易达到设定阈值。而采用方法 1 和方法 3 都只确定 1 个果实目标。

表 5 采用不同方法判定采摘目标的试验结果
Tab. 5 Experimental results of using different methods to determine picking targets

方法	TP	FP	FN	P	R
距离与参照物尺寸对比方法	265	35	13	0.88	0.95
IoU 与阈值对比方法	259	47	19	0.85	0.93
本文的最近邻检索方法	276	2	2	0.99	0.99

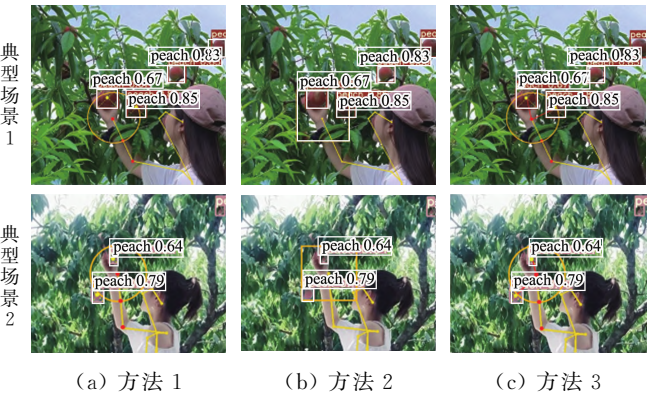


图 13 不同方法在两种典型场景下的试验结果
Fig. 13 Experimental results of different methods in two typical scenarios

通过以上分析可知,若采用前两种方法,无论如何选取参照物与阈值,在真实场景下都可能存在不止 1 个果实满足此条件。结合到本文具体问题,这样就容易将其他果实误判为采摘目标。说明手接触果实的判定问题在转化为采摘目标的确定问题时,还需要考虑检测结果的唯一性这一因素。相比之下,本文提出的基于最近邻检索的方法,使每帧满足条件的果实对象具有唯一性,有效弥补了前述方法的不足,更有利于准确检测采摘目标。表 5 从数据方面验证本文方法的检测准确性较高,在查准率和查全率上的表现较好。

3.5 模型 Nodetect 状态标志对采摘目标检测的影响
设定 Nodetect 状态标志是为避免人手遮挡导致

模型对采摘目标的误检测。为验证该方法的有效性，在试验中对 Nodetect 状态标志设定与否对模型检测采摘目标的准确性进行对比。

由于试验要考察的也是采摘目标检测的准确性，因此，试验的视频范围、参数定义以及评价指标皆与 3.4 节中的试验设定相同。由表 6 可知，若不设定 Nodetect 状态标志，即使模型可以检测出每个视频中正确的采摘目标，但也会将许多非采摘对象的果实误识别为采摘目标。即在一次采摘过程中，模型可能会检测出多个采摘目标，将严重影响到后续果实摘下判定环节的正确执行。

表 6 设定状态标志与否对采摘目标检测的影响
Tab. 6 Effect of setting status flags or not on picking target detection

方法	TP	FP	FN	P	R
方法 4: 不设定 Nodetect 状态标志方法	278	189	2	0.60	0.99
方法 5: 设定 Nodetect 状态标志方法	278	2	2	0.99	0.99

图 14 展示以上 2 种方法在采摘目标受遮挡的同一视频上的试验结果。图 14(a)反映采用不设定 Nodetect 状态标志方法时，当 2 号采摘目标受到人手遮挡丢失目标检测与目标跟踪后，模型将 3 号果实错判为采摘目标。图 14(b)反映采用设定 Nodetect 状态标志的方法时，当检测出采摘目标后就不再对其进行检测，可以有效避免对采摘目标误判的情况，确保采摘行为识别后续环节可以读取到正确的采摘目标信息以进行下一步操作。

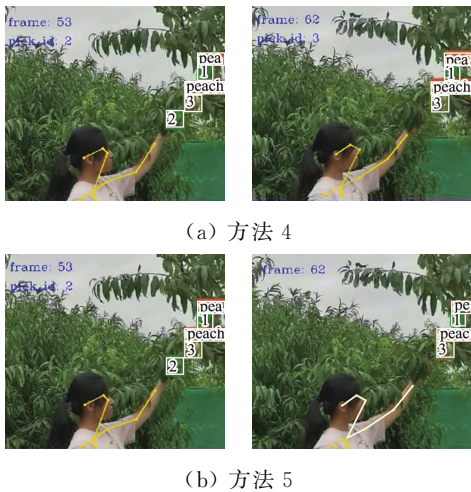


图 14 设定状态标志与否的试验结果对比
Fig. 14 Comparison of experimental results with and without setting status flags

4 系统设计及现场测试

为进一步在真实情境中对上述方法进行实际使用

效果测试，并实现对采摘园中的游客采摘行为的智能化监督管理，基于上述采摘识别模型，设计桃子采摘园智慧监管系统。

该系统由萤石监控摄像头(CS-H5)、萤石云平台和上位机三部分组成。监控摄像头负责采集采摘园现场图像数据，通过 Wi-Fi 或有线网络将图像数据输送到云平台；云平台可以实现摄像头与上位机端的图像数据传输；上位机端向云端发送请求获取视频数据，采用训练好的采摘识别算法对获取的视频图像进行检测。本系统的网络拓扑图如图 15 所示。当系统识别出采摘行为后，将采摘信息(如人数、采摘次数等)与检测后的视频图像展示在系统用户界面上，并将采摘信息数据存储到数据库中，以便管理人员进行数据分析。

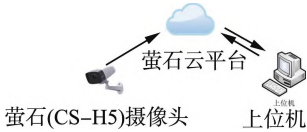


图 15 系统网络拓扑图
Fig. 15 System network topology diagram

由于现实场景中存在对历史和实时 2 种形式视频文件中采摘行为的检测需求，因此，该系统在工作流程方面提供 2 种选择，如图 16 所示。

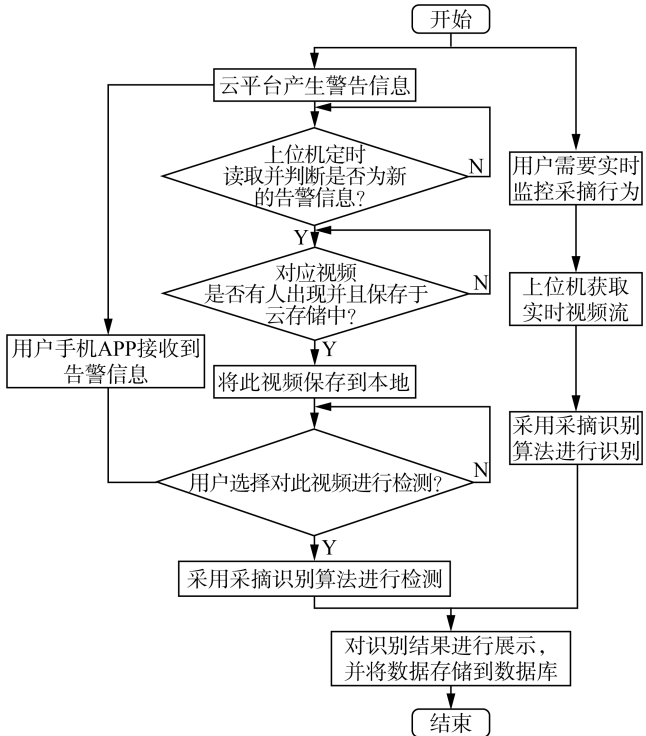


图 16 系统功能实现流程
Fig. 16 Flow chart of system function implementation

选择一，系统通过运用云平台的人像告警功能，使上位机端访问云平台获取告警视频。待用户需要时可选取相应的历史告警视频进行采摘识别检测。选择二，如果用户需要实时监控现场是否发生采摘行为，则上位

机系统通过访问摄像头在萤石云平台的直播源地址获取实时视频流,通过采摘识别算法对捕获到的视频流进行检测。系统将检测到的采摘行为信息展示在系统用户界面上,并连同其他相关信息一同存入数据库。

在上海市闵行区某桃子采摘园进行系统功能测试。首先,对两种不同视频输入源的系统检测效果进行测试。当监控区域有人出现时,云平台向用户端发送告警信息,并且对该监控区域录像。而录像视频的保存名称为视频录制的起始时间。当用户收到有人进入监控区域的告警消息后,用户可根据实际需求,在上位机应用上选择历史视频中相应的视频文件。然后,选择模型权重文件,点击界面的运行按钮,系统将展示检测后的视频图像、视频文件信息和检测结果。图 17 为历史视频的检测结果展示。如果用户想要实时监控现场是否发生采摘行为,则点击用户界面上的实时视频对想要查看的监控区域对应的摄像头编号进行选择,其他操作同上。图 18 为实时视频的检测结果。



图 17 历史视频检测结果展示

Fig. 17 Recorded video test results show



图 18 实时视频检测结果展示

Fig. 18 Real-time video detection results show

对基于果树采摘行为识别方法的采摘园智慧监管系统的识别准确性进行测试。在摄像头视场范围内进行 500 次随机动作的果树果实采摘行为及 200 次随机采摘干扰动作(如整理东西、随意走动等)。测试结果表明,除有 3 次因为镜头被飞虫等遮挡导致系统未能读取到清晰完整画面,因而未能识别采摘行为之外,剩余采摘行为都被成功识别。而 200 次采摘干扰动作都没有被误判为采摘动作。

5 结论

1) 将深度学习技术引入到农业环境中对人体姿态进行识别,基于深度学习技术的采摘行为识别中存在 3 个重要的判定环节,分别是采摘姿态判别、采摘目标的判定和果实摘下判定。

2) 传统对人体姿态判别方法的研究主要考虑人体骨架的整体变化情况。对于采摘果园场景中的采摘行为,对采摘姿态的判别应考虑选择更有针对性的约束条件。针对树上果实的采摘行为,采摘姿态明显不同于其他姿态,因此,应找到有效区分采摘姿态与其他姿态的判别方法。通过设定关节角度的方法,对游客不同活动的关节角度变化情况进行分析,提出一种通过关节角度判定采摘姿态的方法。通过与传统人体姿态判别方法的试验对比可知,基于人体关节角度的姿态判别方法能够更准确地识别采摘姿态,查准率提高 16%。

3) 采摘目标的判定实质上是一个对人手与果实位置关系进行考量的问题。传统对人手与物体接触情况的研究主要基于手腕与目标物体之间的距离或交并比的方法进行衡量。对于采摘识别场景中的果实目标,采摘目标的确定应着重考虑唯一性,因此,应找到一种可以在每一帧的检测中匹配到唯一果实的约束条件。针对采摘行为中人手与果实的位置关系,提出一种新型基于手腕与果实之间距离关系的采摘目标确定方法。针对现实场景中的手腕与果实的复杂位置关系,对该方法进行优化。通过与传统对人手接触物体对象的判别方法对比可知,优化后基于距离的新型方法能够更准确地判定采摘目标,查准率提高 11%。

4) 在以上采摘目标的判定环节中,还存在人手遮挡导致采摘目标错检的问题,该问题将影响后续果实摘下判定环节的正确判定。考虑通过设定状态标志的方法,避免由于果实目标检测失效影响到最后的采摘行为判定结果。试验结果表明,该状态标志的设定可有效避免对采摘目标的误判,查准率提高 39%。

参 考 文 献

- [1] 徐磊,陈超. 中国桃产业经济分析与发展趋势[J]. 果树学报, 2023, 40(1): 133—143.
- [2] 滕光辉,冀横溢,庄晏榕,等. 深度学习在猪只饲养过程的应用研究进展[J]. 农业工程学报, 2022, 38(14): 235—249. Teng Guanghui, Ji Hengyi, Zhuang Yanrong, et al. Research progress of deep learning in the process of pig feeding [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(14): 235—249.
- [3] 翟肇裕,曹益飞,徐焕良,等. 农作物病虫害识别关键技术研究综述[J]. 农业机械学报, 2021, 52(7): 1—18.

- Zhai Zhaoyu, Cao Yifei, Xu Huanliang, et al. Review of key techniques for crop disease and pest detection [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1—18.
- [4] 李菊霞, 李艳文, 牛帆, 等. 基于 YOLOv4 的猪只饮食行为检测方法[J]. 农业机械学报, 2021, 52(3): 251—256.
Li Juxia, Li Yanwen, Niu Fan, et al. Pig diet behavior detection method based on YOLOv4 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 251—256.
- [5] 胡志伟, 杨华, 娄甜田. 采用双重注意力特征金字塔网络检测群养生猪[J]. 农业工程学报, 2021, 37(5): 166—174.
Hu Zhiwei, Yang Hua, Lou Tiantian. Instance detection of group breeding pigs using a pyramid network with dual attention feature [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(5): 166—174.
- [6] 彭明霞, 夏俊芳, 彭辉. 融合 FPN 的 Faster R—CNN 复杂背景下棉田杂草高效识别方法[J]. 农业工程学报, 2019, 35(20): 202—209.
Peng Mingxia, Xia Junfang, Peng Hui. Efficient recognition of cotton and weed in field based on Faster R—CNN by integrating FPN [J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(20): 202—209.
- [7] 刘莫尘, 高甜甜, 马宗旭, 等. 基于 MSRCR—YOLOv4—tiny 的田间玉米杂草检测模型[J]. 农业机械学报, 2022, 53(2): 246—255, 335.
Liu Mochen, Gao Tiantian, Ma Zongxu, et al. Target detection model of corn weeds in field environment based on MSRCR algorithm and YOLOv4—tiny [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(2): 246—255, 335.
- [8] 鲍文霞, 孙庆, 胡根生, 等. 基于多路卷积神经网络的大田小麦赤霉病图像识别[J]. 农业工程学报, 2020, 36(11): 174—181.
Bao Wenxia, Sun Qing, Hu Gensheng, et al. Image recognition of field wheat scab based on multi-way convolutional neural network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(11): 174—181.
- [9] 甘雨, 郭庆文, 王春桃, 等. 基于改进 EfficientNet 模型的作物害虫识别[J]. 农业工程学报, 2022, 38(1): 203—211.
Gan Yu, Guo Qingwen, Wang Chuntao, et al. Recognizing crop pests using an improved EfficientNet model [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(1): 203—211.
- [10] Dyrmann M, Jørgensen R N, Midtby H S. Robo weed support detection of weed locations in leaf occluded cereal crops using a fully convolutional neural network [J]. Advances in Animal Biosciences, 2017, 8(2): 842—847.
- [11] Amara J, Bouaziz B, Algergawy A. A deep learning-based approach for banana leaf diseases classification [C]. Datenbanksysteme für Business, Technologie und Web (BTW 2017)—Workshopband. Gesellschaft für Informatik eV, 2017: 79—88.
- [12] Veeraballi R K, Nagugari M S, et al. Deep learning based approach for classification and detection of papaya leaf diseases [C]. 18th International Conference on Intelligent Systems Design and Applications, 2020.
- [13] 彭红星, 黄博, 邵园园, 等. 自然环境下多类水果采摘目标识别的通用改进 SSD 模型[J]. 农业工程学报, 2018, 34(16): 155—162.
Peng Hongxing, Huang Bo, Shao Yuanyuan, et al. General improved SSD model for picking object recognition of multiple fruits in natural environment [J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(16): 155—162.
- [14] 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. Kiwifruit detection of far-view and occluded fruit based on improved AlexNet [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(10): 24—34.
- [15] 刘芳, 刘玉坤, 林森, 等. 基于改进型 YOLO 的复杂环境下番茄果实快速识别方法[J]. 农业机械学报, 2020, 51(6): 229—237.
Liu Fang, Liu Yukun, Lin Sen, et al. Fast recognition method for tomatoes under complex environments based on improved YOLO [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6): 229—237.
- [16] Rahnmounfar M, Sheppard C. Deep count: Fruit counting based on deep simulated learning [J]. Sensors, 2017, 17(4): 905.
- [17] Koirala A, Walsh K B, Wang Z, et al. Deep learning for real-time fruit detection and orchard fruit load estimation: Benchmarking of ‘MangoYOLO’ [J]. Precision Agriculture, 2019, 20(6): 1107—1135.
- [18] Kestur R, Meduri A, Narasipura O. MangoNet: A deep semantic segmentation architecture for a method to detect and count mangoes in an open orchard [J]. Engineering Applications of Artificial Intelligence, 2019, 77: 59—69.
- [19] Redmon J, Farhadi A. YOLOv3: An incremental improvement [J]. Computer Science, 2018, 4(1): 1—6.
- [20] Cao Z, Hidalgo G, Simon T, et al. OpenPose: Realtime multi-person 2D pose estimation using part affinity fields [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(1): 172—186.
- [21] Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric [C]. IEEE International Conference on Image Processing, 2018: 3645—3649.

论文精选:基于深度学习的观光农业中的桃子采摘识别

研究概况

针对桃子采摘园智慧化管理的需求,提出一种基于深度学习的采摘识别方法。利用机器视觉与深度学习技术,在轻量级人体姿态估计算法 LightweightOpenPose、目标检测算法 YOLOv5s、目标跟踪算法 DeepSORT 的基础上,提出桃子采摘行为检测方法。基于实际桃子采摘视频数据建立数据集,进行相关性能测试。将基于人体关节角度方法与传统采用人体关节点外接矩形框的方法进行对比,该方法对采摘举手动作的判定查准率提高 16%。针对采摘目标判定问题,基于最近邻检索的方法相比于传统的基于距离与参照物尺寸对比的方法、基于交并比与阈值对比的方法,查准率至少提高 11%。基于设定状态标志的采摘目标检测失效方法,较好地解决手部遮挡对检测结果的影响,查准率提高 39%。提出的桃子采摘识别方法能够在采摘桃园实际环境下完成对采摘动作的有效准确识别。

主要技术指标

方法	TP	FP	FN	查准率 P	查全率 R
距离与参照物尺寸对比方法	265	35	13	0.88	0.95
交并比 IoU 与阈值对比方法	259	47	19	0.85	0.93
最近邻检索方法	276	2	2	0.99	0.99
设定 Nodetect 状态标志方法	278	2	2	0.99	0.99
不设定 Nodetect 状态标志方法	278	189	2	0.60	0.99



团队介绍

该成果由东华大学信息科学与技术学院农业物联网创新团队研发,团队负责人为杨义副教授。该研究由国家自然科学基金(12272092)、河北省物联网监控工程技术研究中心项目(3142018055,3142016020)、廊坊市人民政府科学技术研究与发展计划资助项目(2021011035)、秦皇岛市科学技术研究与发展计划项目(201805A016)等资助。团队主要致力于无人机农业图像处理、观光农业智慧管理措施及农业物联网智慧终端设备等方面的研究,取得了多项重要科技成果。

联系方式

联系人:杨义,电话:13817951737,邮箱:yiyang@dhu.edu.cn