

DOI: 10.13733/j.jcam.issn.2095-5553.2024.04.030

张永宏, 李宇超, 秦夏洋, 等. 非结构化环境下番茄采摘机器人目标识别与检测[J]. 中国农机化学报, 2024, 45(4): 205-213

Zhang Yonghong, Li Yuchao, Qin Xiayang, et al. Target identification and detection for tomato harvesting robot in unstructured environments [J]. Journal of Chinese Agricultural Mechanization, 2024, 45(4): 205-213

非结构化环境下番茄采摘机器人目标识别与检测*

张永宏¹, 李宇超¹, 董天天¹, 秦夏洋¹, 刘云平¹, 曹景兴²

(1. 南京信息工程大学自动化学院, 南京市, 210044; 2. 无锡新松机器人自动化有限公司, 江苏无锡, 214000)

摘要:针对采摘机器人收获技术中的识别技术受限于非结构化环境中复杂背景干扰的问题,采用改进模型后处理的研究路线,提出一种改进 YOLOv5 算法。首先将果实目标的中心点距离、预测框宽高实际差值与面积交并比三者共同考虑为损失项,提升预测框实际尺寸精度,再利用中心点距离作为惩罚项加权面积交并比得分,提升密集目标的识别能力,最后通过设置辅助训练头,提供更多的梯度信息以防止过拟合现象。通过多种损失函数损失值对比与模型改进精度对比试验证明改进有效性,部署至机器人验证可行性。结果表明,改进后的算法模型识别平均精度 95.6%,召回率达到 90.1%,相较于改进前全类精度提升 0.4 个百分点,召回率提升 0.4 个百分点,满足采摘机器人识别需求。

关键词:非结构化;番茄果实;目标识别;损失函数优化;YOLOv5 算法

中图分类号:TP249; TP391.4 **文献标识码:**A **文章编号:**2095-5553 (2024) 04-0205-09

Target identification and detection for tomato harvesting robot in unstructured environments

Zhang Yonghong¹, Li Yuchao¹, Dong Tiantian¹, Qin Xiayang¹, Liu Yunping¹, Cao Jingxing²

(1. School of Automation, Nanjing University of Information Science and Technology, Nanjing, 210044, China;
2. Wuxi SIASUN Robot & Automation Co., Ltd, Wuxi, 210000, China)

Abstract: Aiming at the problem that the recognition technology of harvesting robots in crop picking was limited by complex background interference in unstructured environments, especially due to occlusion by foliage and the overlapping of fruits, resulting in lower accuracy in identification, an improved YOLOv5 algorithm was proposed based on the improved research approach involving post-processing of the model. Initially, the centroid distance of fruit targets, the actual difference in predicted box width and height, and the intersection-over-union of areas were collectively considered as loss terms. This was aimed at enhancing the accuracy of predicted box sizes. Furthermore, the centroid distance was utilized as a penalty term weighted by the intersection-over-union score to improve the recognition capability for densely clustered targets. Subsequently, auxiliary training heads were incorporated to provide additional gradient information, thereby preventing overfitting. Through comparative analysis of loss values using multiple loss functions and assessing the model improvement accuracy, the effectiveness of the enhancements was experimentally validated. Finally, the deployment onto the robot confirmed the feasibility of the proposed improvements. The results indicated that the improved algorithm model achieved an average accuracy of 95.6%, with a recall rate of 90.1%. Compared to the pre-improvement overall class accuracy, there was an increase of 0.4 percentage points in both accuracy and recall rate, meeting the recognition requirements for harvesting robots.

Keywords: unstructured; tomato; target recognition; loss function optimization; improved YOLOv5

0 引言

非结构化设施农业采摘机器人是指适用于不同设

施农业场景下的采摘机器人。设施农业生产环境相对整齐^[1],不仅为农作物提供良好的生长环境,也为采摘机器人提供稳定的工作环境,这一定程度上降低采摘

收稿日期:2023 年 9 月 25 日 修回日期:2023 年 11 月 26 日

*基金项目:江苏省现代农业机械装备与技术示范推广项目(NJ2022-02)

第一作者、通讯作者:张永宏,男,1974 年生,山东莒南人,博士,教授;研究方向为方向模式识别与智能系统。E-mail: zyh@nuist.edu.cn

机器人的开发难度、增强机器人运行时的稳定性^[2, 3]。在早期作物识别试验中,大多采用机器学习的方法^[4]。利用数字图像处理的方式,采用卷积进行滤波特征提取,使用支持向量机、BP 神经网络等机器学习模型,实现特征的分类,从而达到识别的效果^[5]。但由于图像采集环境因素多变且特征向量选取存在多样性,导致最终识别效果通用性与可移植能力较弱^[6]。从 20 世纪中期开始,机器视觉不断发展,从构建三维向量到特征识别,从浅层神经网络到深度学习,人们逐渐拓宽机器视觉应用领域,提升农业作物识别任务的精度^[7-10]。

吕志远等^[11]通过先分类识别再进行目标检测的级联网络,实现检测小目标与密集型目标的任务,模型 mAP 达到 92.35%,对比改进前 YOLOX 网络提升 2.38 个百分点。李天华等^[12]采用 HSV 色域分割方法,对成熟番茄识别精度达到 94.77%,较改进前 YOLOv4 模型提升 4.3 个百分点。杨坚等^[13]采用 CBAM 注意力模块与 YOLOv4-tiny 骨干网络结合,加强全局特征融合,识别平均精度达到 97.9%,较 YOLOv3 提升 30.9 个百分点。何斌等^[14]改进 YOLOv5 的损失函数,通过使用交并比的目标位置损失函数,识别夜间温室番茄,综合平均精度达到 97.6%。朱智惟^[15]通过对原始 YOLOv5 网络结构与 NMS 极大值抑制进行改进,增加跨层连接,更好地实现不同层级间特征融合,使识别精度达到 97.62%,相较于改进前提升 2.38%,抓取绝对误差不超过 ± 3 mm。

现有的机器视觉技术侧重目标检测中识别率与识别精度的提升,但对于摄像头放置于机械手上的识别采摘实际应用而言,检测算法更倾向于提升局部范围内当前目标检测识别预选框的精度,即目标的坐标信息精度^[16, 17]。在此需求下,应寻找一种面向提升局部画面,确定目标坐标信息的高精度模型。本文以番茄果实为研究对象,采用改进损失函数和极大抑制相结合的方法,提出一种基于 YOLOv5 的番茄采摘机器人的果实目标识别算法。通过融合中心点距离、预测框宽高实际差值与面积交并比三类损失值,提升预测框实际尺寸精度;再引入中心点距离作为惩罚项加权面积交并比得分,提升小范围内密集目标下模型的极大值抑制能力;最后通过设置辅助分类头,降低模型过拟合风险,提升预测框中心点坐标精度,为非结构环境下果实自动化采摘提供可靠的视觉识别支持。

1 番茄机器人系统

1.1 采摘机器人系统结构

番茄采摘系统的设计主要方法:(1)通过深度学习研究方法研究复杂环境下番茄检测与识别。(2)通过深度

相机确定番茄目标位置。(3)通过远程监控实时返回检测画面及监测数据。(4)通过六自由度机械臂及末端执行器完成采摘任务。针对上述需求,所设计的番茄识别定位采摘系统主要包括采摘机器人、深度相机、通信模块,控制模块,机械臂及末端执行器,整体系统结构如图 1 所示。

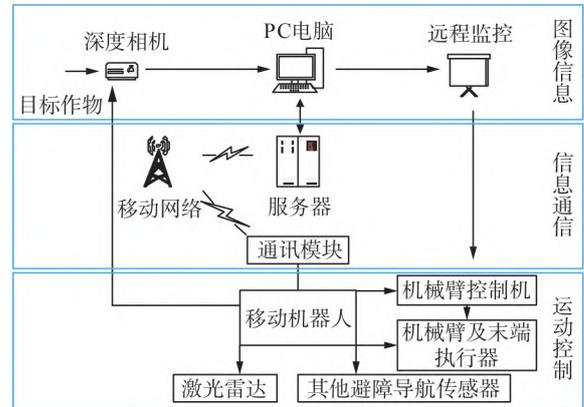


图 1 采摘机器人系统结构图

Fig. 1 Structure of the harvesting robot system

采摘机器人移动平台采用四轮差动复合机器人作为载体,安装避障传感器和激光雷达,以满足运动精度要求。移动平台搭载六自由度机械臂完成果实采摘。深度相机采用 Realsense D435i,由一对立体红外传感器、一个红外激光发射器和一个彩色相机组成,其中 RGB 图像最大帧分辨率为 $1\ 920$ 像素 \times $1\ 080$ 像素,帧率 30 FPS;深度图像分辨率为 $1\ 280$ 像素 \times 720 像素,最大帧率 90 FPS,最小深度距离 10 cm。本次采摘试验兼顾深度信息与色彩信息,因此设置设备参数为分辨率 $1\ 280$ 像素 \times 720 像素,帧率 30 FPS。采摘机器人技术参数主要如表 1 所示。

表 1 技术参数表

Tab. 1 Technical parameter table

模块	参数	数值
移动底盘	外形尺寸(长 \times 宽 \times 高) /(mm \times mm \times mm)	865 \times 650 \times 785
	激光雷达	速腾 RS-16
深度相机	最小深度距离/cm	10
	图像分辨率/(像素 \times 像素)	1 280 \times 720
六自由度机械臂	工作范围/mm	700
	定位精度/mm	± 0.01
	最大负载/kg	5

整个系统通过深度摄像头确定番茄位置坐标,移动机器人运动至合适位置,工控机计算机机械臂各关节角度并控制运动,最后由末端执行器完成番茄的采摘。同时通过远程监控实时返回摄像头视频输出,完成对采摘过程的可视化监控。摄像头安装如图 2 所示,采摘机器人

系统如图 3 所示。



图 2 深度摄像头安装示意图
Fig. 2 Depth camera installation diagram



图 3 采摘机器人系统
Fig. 3 Harvesting robot system

1.2 目标识别及定位算法

考虑到检测需求的实时性和快速性,番茄识别基于 YOLOv5 算法^[18]。YOLOv5 网络结构分为 3 部分,用于提取图像特征的骨干网络、用于特征融合的特征金字塔和用于对金字塔特征检测的目标检测头^[19]。

YOLOv5 使用 CSPDarknet53 作为其主干网络^[20],头部采用卷积层和池化层,用于预测目标的类别、位置和置信度。与 YOLOv4 相比,YOLOv5 的头部网络采用更大的感受野和更小的步幅,可以有效提高检测的准确性。此外采用 CIoU (Complete Intersection over Union) 损失函数,该函数不仅考虑目标框的位置和大小,还计算目标框的形状影响,可以更准确地衡量检测结果的好坏。与其他算法相比,YOLOv5 具有速度快、精度高、模型轻量级等特点,适合于实时处理和轻量级部署的场景^[21-23]。

2 改进 YOLOv5 算法原理和方法

针对在非结构化场景中番茄生长状态不一、枝叶遮挡等问题,本文主要从网络模型后处理角度出发,对 YOLOv5 算法进行针对性改进。

2.1 多因素损失优化

EIoU^[24] (Efficient Intersection over Union) 是一种改进的目标检测损失函数,是在原有的 CIoU 损失函数基础上进一步优化而来。EIoU 损失函数在计算真实框 GT(Ground Truth)与预测框 BBox(Bounding Box)之间的差异时,即引入距离惩罚项,并且又直接考虑边界框的宽高与其置信度的真实差异,避免纵横比对模型相似性的有效优化,从而更准确地评估预测框的位置和尺寸,原理如图 4 所示,红色部分为预测框,绿色为真实

框,蓝色为最小外接矩形。计算如式(1)所示。该损失函数包含重叠损失、中心距损失,宽高损失三部分。

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} \\ = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} \quad (1)$$

式中: L_{EIoU} —— EIoU 损失值;

L_{IoU} —— 预测框与真实框面积交并比损失;

L_{dis} —— 预测框与真实框中心点距离损失值;

L_{asp} —— 预测框与真实框方位损失;

IoU —— 面积交并比;

$\rho(\cdot)$ —— 预测框与真实框间欧式距离;

b, w, h —— 预测框的中心点和宽高;

b^{gt}, w^{gt}, h^{gt} —— 真实框中心点和宽高;

c, C_w, C_h —— 覆盖预测框与真实框的最小外接框的对角线长度、宽度和高度。

在计算过程中,首先计算真实框与预测框之间的交集面积和并集面积之间的比值,再加入真实框和预选框中心点之间的欧氏距离作为惩罚项,最后考虑真实框和预选框之间的中心点位置关系计算损失值。因此在模型后续沿损失值减小方向优化过程中,预测框将逐渐与真实框面积重合,中心点距离缩小,形状相同。

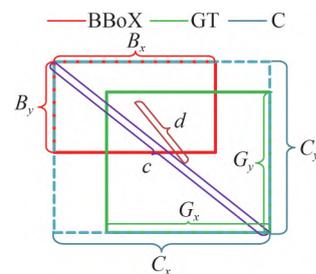


图 4 EIoU 损失函数原理图

Fig. 4 Principle of the EIoU loss function

2.2 多目标重叠抑制

非极大值抑制 NMS(Non-Maximum Suppression)是目标检测中常用的一种后处理方法,用于去除与同一真实框匹配的冗余的预测框。传统 NMS 方法采用不同预测框间面积交并比 IoU (Intersection over Union)作为衡量两个检测结果重叠程度的指标,即当两个检测框的 IoU 大于某个阈值时,将得分较低的框去除。

然而,NMS 极大值抑制过程中,若预测框中存在多个真实框目标,则该预测框和与不同目标真实框匹配的预测框间面积重叠值较大,正确的预测框可能会被 NMS 错误抑制。针对此问题,不仅需要考虑目标框之间的重叠情况,还需要考虑目标框之间的中心点距离,才可以更准确地选择最终的检测结果。因此,一组预测框中,与置信度最高的预测框中心点距离越近,则是冗余框的可能性越大,即引入相邻框的中心距参

数,创建新的影响参数项,提高同一预选框多目标情况下的召回率。在计算过程中,首先按照置信度对所有检测结果进行排序,从置信度最高的框开始,依次遍历每个框,将每个预测框与置信度最大的框对比,得分 C_{DIoU} 小于一定阈值的框标记为冗余框,并从结果列表中删除。将被保留下来的预测框作为新的种子框,继续进行遍历和删除操作,直到处理完毕。这种将 $DIoU$ ($Distance-IoU$) 的损失算法替换 NMS 中原有的 IoU 损失算法即为 $DIoU-NMS$ 。数学模型如(2)式所示。

$$S_i = \begin{cases} S_i & C_{DIoU} < \epsilon \\ 0 & C_{DIoU} \geq \epsilon \end{cases} \quad (2)$$

$$C_{DIoU} = IoU - \left(\frac{d^2}{c_1^2}\right)^\beta$$

式中: S_i ——遍历的第 i 个预选框;
 C_{DIoU} —— $DIoU$ 得分;
 ϵ ——去除阈值;
 d ——不同预测框的中心点距离;
 c_1 ——覆盖相邻两预测框最小外接矩形的对角线长度;
 β ——控制距离惩罚项的惩罚幅度。

这种方式在面对具有重叠目标的情况下,具有较好的效果,可以有效地减少误判和漏检,提高目标检测的精度和鲁棒性。如图 5 所示,绿色框为置信度最大的预测框,蓝色框为可能的冗余预测框,三种情况下冗余预测框与置信度最大的预测框面积交并比相同,无法通过 IoU 阈值去除三者中的冗余框,引入中心距项后, IoU 得分排序为 $B > A > C$,这时再合理设置阈值即可去除冗余预测框 C , AB 则表示所预测的物体与绿色预选框预测的物体重合。

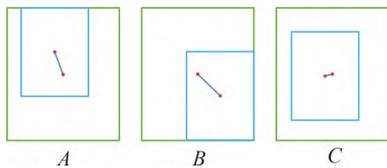


图 5 Diou-NMS 原理图
 Fig. 5 Principle diagram of Diou-NMS

2.3 多输出头辅助

在网络训练过程中,为了让中间层学到更多的信息,从而具有丰富的梯度信息,将网络训练得更好,提高模型的精度。本文选择在头部网络的倒数第二次和倒数第三层添加辅助训练头。这些概率值可以用特定的损失函数来计算分类损失,并且可以在反向传播中以一定的权重更新模型参数与主模型的损失函数共同训练模型。另外,每个辅助头对应一个不同的特征图尺寸,并负责识别不同大小的物体。这样辅助头可以为模型提供更多的信息,增强模型对同幅画面中大小不一的物

体检测准确性。总的来说,辅助头可以提高模型的精度,使得模型更好地适应不同的数据集和场景。

为了更好地介绍辅助头技术路线,将原本的检测头称为引导头。引导头导向标签器是根据引导头的预测结果与真实标签进行损失计算,通过优化生成软标签。由于标签更能代表源数据与目标数据的相关性,此时生成的软标签将作为辅助头和引导头后续的目标训练模型。此外,因为辅助头学习能力低,为了避免丢失信息,在检测任务中重点提升辅助头的召回率,引导头将从召回的结果中选择高精度结果作为输出结果。在此过程中,辅助头和引导头的权重比需要额外注意,本试验设置成 $1 : 0.25$ 。其损失融合的数学模型如式(3)所示。

$$l = l_p + \lambda \times l_r \quad (3)$$

式中: l ——软标签最终损失值;
 λ ——粗标签权重系数;
 l_p ——细标签损失值;
 l_r ——粗标签损失值。

根据改进方法,提高 YOLOv5 算法在实际应用中的鲁棒性,更加适合于番茄生长状态不一、枝叶遮挡等复杂场景下的目标检测任务,改进后整体网络模型如图 6 所示。

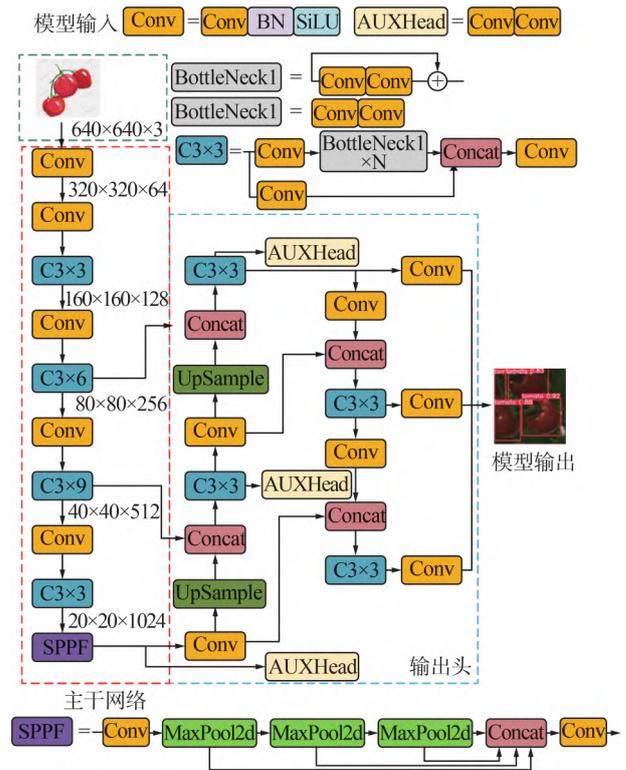


图 6 改进 YOLOv5 模型网络结构图

Fig. 6 Improved YOLOv5 model network

注:Conv 卷积池化激活一体模块, $C3 \times N$ 多个 $C3$ 特征提取模块, BottleNeck 瓶颈模块, SPPF 空间金字塔池化模块, MaxPool2d 池化模块, UpSample 上采样模块, Concat 拼接模块, AUXHead 辅助头模块

3 试验结果与分析

3.1 识别试验准备

模型训练的数据集采用不同遮挡程度、不同识别角度、不同成熟度的 800 张的番茄图像集,使用 Labelimg 进行手动标注成 YOLO 所需 txt 格式的标签,制作而成。之后将数据集中的图像和标签按照 8:2 方式随机划分为训练集和验证集。

本试验示意如图 7 所示,采摘机器人依次经过四个采摘点。拟抓取试验场地如图 8 所示,设置四株番茄树采摘点,田间宽度 1.8 m,作物生长高度 0.8~1.5 m,每棵树上包含遮挡程度不同的目标 8~15 颗,模拟如图 9 所示的实际采摘效果。本试验软件环境采用 python3.8 版本、pytorch1.10.0 深度学习框架,CUDA10.2 版本。超参数设置为:预训练权重使用 yolov5s.pt;最大迭代次数 100 次;批次尺寸(batchsize)设置为 16 张;学习率为 0.001。

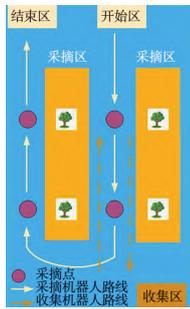


图 7 试验示意图
Fig. 7 Workflow



图 8 试验场地搭建图
Fig. 8 Experimental site setup



图 9 实际场地采摘示意图
Fig. 9 Actual site harvesting diagram

由于在图片读取后,主干网络经过 focus 模块及一个卷积操作后又通过 4 个 C3 模块由浅到深逐层提

取特征图,每次特征提取操作将图片宽高尺寸各减小原来的一半,最小图像减少至原图 1/32,因此主干网络输入图像尺寸应为 32 的整数倍,默认为 640 像素×640 像素。通过长度宽度等比缩放再填充计算实现,长度 1 280 像素×640 像素/1 280 像素=640 像素,宽度 720 像素×640 像素/1 280 像素=360 像素,之后宽度由 360 像素填充至 640 像素,形成 640 像素×640 像素尺寸的图片。

同时,为了验证在不同定位精度损失函数下,损失值的收敛情况。本试验同时对比 $CIoU$ (Complete Intersection over Union)^[25]、 $SIoU$ (Scylla Intersection over Union)^[26]、 $GIoU$ (Generalized Intersection over Union)^[27]、 $EIoU$ (Efficient Intersection over Union)、 $Focal EIoU$ (Focal and Efficient Intersection over Union)、 $WIoU$ (Wise Intersection over Union)^[28] 六种损失函数在相同数据集上的情况,主要对比参数为损失值、收敛时间、召回率与全类平均精度。其中图 10 损失值为整体回归损失值,损失值内容包括边框回归损失,类别损失与置信度损失,分别代表预测框定位精度与预测框分类精度。

$$Loss = L_{localization} + L_{Class} + C_{prediction}$$

式中: $Loss$ —— 总损失值;

$L_{localization}$ —— 定位损失;

L_{Class} —— 类别损失;

$C_{prediction}$ —— 精度损失。

召回率为被精准识别出的番茄与全部番茄的数量比,计算如式(4)所示。

$$R = \frac{T_p}{T_p + F_n} \quad (4)$$

式中: R —— 召回率;

T_p —— 标签为番茄的目标被识别成番茄是样本数量;

F_n —— 标签为番茄的目标没有识别成番茄的样本数量。

全类平均精度为模型不同类别上的精度进行综合评估的指标,计算如式(5)所示。

$$mAP = \frac{1}{m} \sum_{i=0}^{m-1} \left[\sum_{j=1}^{n-1} \left[\max[P(R_{j+1}), P(R_j)] \right] \right] \times (R_{j+1} - R_j) \quad (5)$$

$$P(R) = \frac{T_p}{T_p + F_p}$$

式中: F_p —— 将背景识别为番茄的预测框数量;

$P(R)$ —— 在召回率为 R 的情况下的识别精度;

R_j —— 为对召回率等距取 n 个插值中,第 j 个召回率的数值;

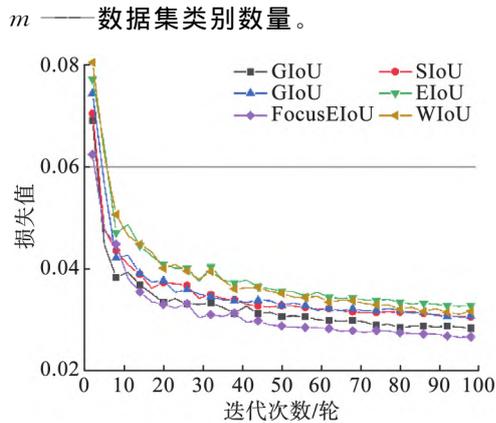


图 10 损失值对比图

Fig. 10 Loss value comparison chart

3.2 模型训练结果分析

将不同角度和遮挡程度的番茄按照有无遮挡和单果多果将数据集划分四类,其模型训练结果分别如图 11 所示,当单果无遮挡时,识别置信度最高,当多果且有遮挡时,识别置信度最低,即多果有遮挡识别难度最高。



图 11 模型训练结果

Fig. 11 Model training results

经过 YOLOv5s 网络模型训练后,不同损失函数的损失值对比如图 11 所示。由图可得在整个迭代过程中损失值有小幅波动,但总体仍呈下降趋势,说明各类损失函数都可使模型收敛。在前 20 次迭代过程中,各类损失函数中损失值均下降迅速,损失值范围减小至 0.04 以下。在后 80 次迭代过程中,各损失函数的损失值下降较为平缓,损失值均缩小至 0.035 以下。

具体来看,EIoU 最初计算出的损失值最大,其次是 GIoU,而 SIoU、CIoU、FocusEIoU 三者初始损失值大小相差无几均为最小。后续迭代过程中 EIoU 损失值下降最多为 0.071,CIoU 损失值下降最少为 0.04。综上所述,EIoU 在损失计算过程中考虑因素较多初始计算出

的损失值最大,同时收敛性较强,损失下降差值最大。

本试验采用检测时间 t ;召回率 $Recall$;全类平均精度 mAP 作为采用不同损失函数模型的主要的评价指标。其分别代表检测单张图像的平均耗时,实际目标中被认为是目标概率及被认为是目标中实际目标的概率。由表 2 可知,检测速度最快为 SIoU 和 FocusEIoU,均为每幅图 9.6 ms,检测时间最慢的为 CIoU,为每幅图 11.5 ms。精度最高为 EIoU (95.6%)与 CIoU、SIoU、GIoU、FocalEIoU、WIoU 函数相比精度提升 0.4、0.8、1.4、1.4 个百分点。召回率最高为 FocusEIoU,但是相较于 EIoU 其精度却下降 0.9 个百分点。说明对于本试验的数据集,虽然 FocusEIoU 让精度最高的 EIoU 的检测速度和召回率有一定提升,但代价却是降低 1 个百分点。因此综合各方面分析,采用 EIoU 时,运行速度较快且召回率也较高,识别精度最高,对于番茄数据目标预测率提高,最终使得精度提高。

表 2 损失函数检测指标

Tab. 2 Detection metrics for loss functions

损失函数	检测时间/ms	Recall/%	mAP/%
CIoU	11.5	89.7	95.2
SIoU	9.6	86.3	94.8
GIoU	9.9	89.5	94.2
EIoU	9.9	89.8	95.6
FocusEIoU	9.6	90.1	94.6
WIoU	9.7	86.5	94.2

之后对模型的极大值抑制方法进行改进添加辅助训练头机制,改进前后识别效果如图 12 所示。

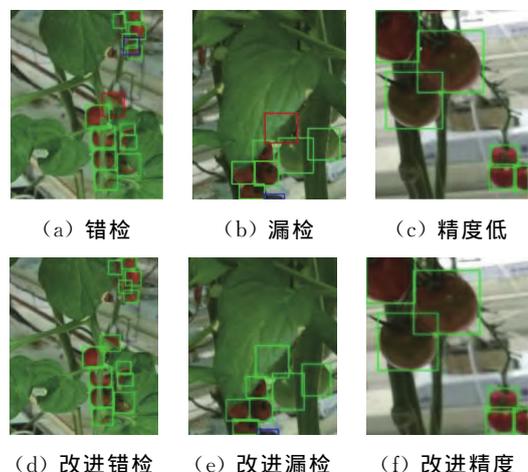


图 12 模型改进结果对比

Fig. 12 Recognition results comparison

绿色边框为识别正确的边框,蓝色为错误边框,红色为漏检的边框。在红色边框处,原本番茄的枝叶遮挡率过高或者果实重叠无法识别,在提取特征数量不

变的情况下,通过添加辅助训练头和改进损失函数方式,增强模型的特征分类处理能力,从而提高识别精度。另外在蓝色部分,其与周围预选框交并比值很大,无法被极大值抑制。使用距离极大值抑制方法(DIoU-NMS)后,引入不同预选框的中心点距离作为惩罚项,降低蓝色边框得分后便可成功将其抑制。最后,对比改进前后绿色边框,在改进后的边框面积更小,更加贴合番茄轮廓,在返回番茄中心点坐标时精确度更高,更利于番茄的采摘。

由表 3 可知,对比改进前后模型的检测指标可得:由于添加辅助头训练,单张图像预测时间增加 1.2 ms,召回率提升 0.4 个百分点,全类精度提升 0.4 个百分点。总体精度上升,模型拥有更好的鲁棒性,改进后模型精度如图 13 所示。

表 3 改进模型性能对比
Tab. 3 Comparison of the improved model

损失函数	Recall/%	mAP/%
YOLOv5	89.7	95.2
改进 YOLOv5	90.1	95.6

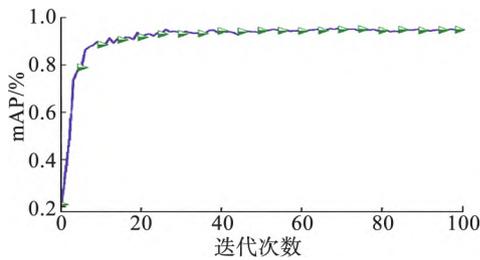


图 13 改进后模型全类精度值

Fig. 13 Overall class accuracy of the improved

3.3 实际应用结果分析

为防止模型训练过程中过拟合,先将训练好的模型用于识别实验室番茄,具体效果如图 14 所示,针对训练过程中提出的枝叶遮挡与果实重叠问题有较好地处理性。

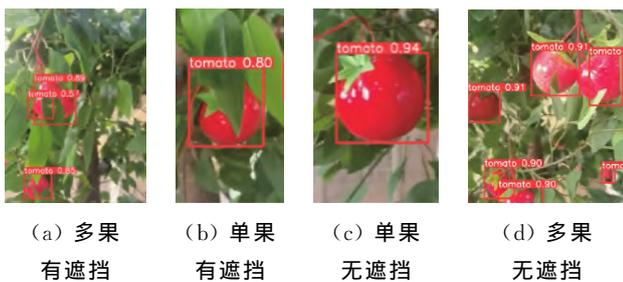


图 14 识别试验结果

Fig. 14 Recognition experiment results

图 15 展示了采摘机器人试验过程,上位机显示抓取画面如图 15(a)、图 15(b)所示,左上角显示识别信息,左下角显示机械臂抓取信息,右侧显示深度相机返

回图像信息。其中包含识别目标的三维坐标信息,具体坐标值再结合深度相机内置矩阵和像素坐标计算得出,抓取效果如图 15(c)、图 15(d)所示。

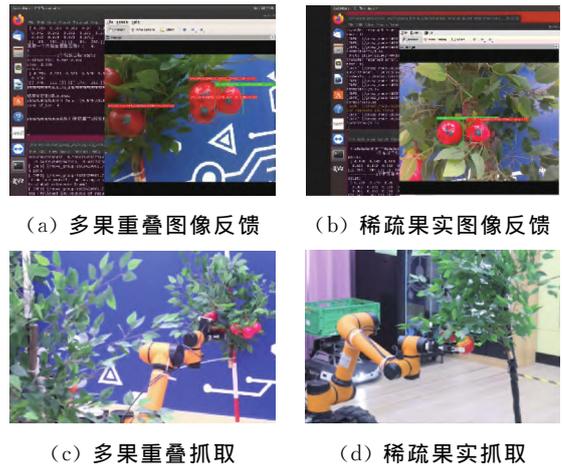


图 15 抓取采摘试验

Fig. 15 Grasping and harvesting experiment

3.4 结果误差分析

根据试验结果,在采摘过程中存在不完全抓取,抓取不牢的情况,图 16(a)表示抓取部位过少,果实弹出末端执行器受力范围,导致抓取失败,图 16(b)所示虽可正常采摘但是果实受力点分布不均,导致采摘时果实破损。经分析误差产生首先来源于机械臂及运动部分采摘开始时,所处的初始位置,机械臂及深度相机坐标较原始坐标有偏差,使得识别结果坐标转换为采摘坐标时出现偏移。其次算法部署时识别帧率不足,在抓取前由风吹等因素干扰果实识别位置时,抓取坐标实时更新性差。

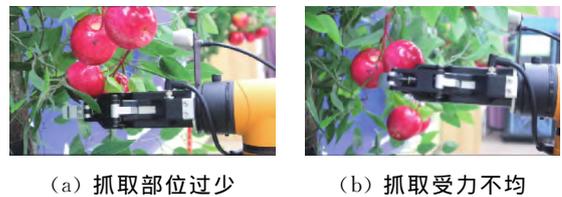


图 16 采摘误差现象

Fig. 16 Harvesting error phenomenon

4 结论

本文以非结构化环境下番茄果实为研究对象,对番茄采摘机器人目标果实进行图像识别,提出一种应用于番茄采摘机器人采摘技术的果实目标识别与检测方法,解决非结构环境下番茄果实目标识别精度低的问题,有效提高受枝叶遮挡,果实重叠等因素干扰的果实目标识别精度。

1) 针对番茄识别采摘存在的果实遮挡及识别定位不准的问题,本文提出改进 YOLOv5 识别算法,通过添加辅助训练头、改进损失函数与极大抑制方法,让模型

产生的预测框更加贴合番茄轮廓,解决成串番茄相互遮挡,预测框错误抑制的问题,提高番茄的位置信息精度。对比改进前后模型的性能指标,改进后的模型召回率 *Recall* 和全类精度 *mAP* 达到 90.1% 和 95.6%, 分别提升 0.4 和 0.4 个百分点,可以满足抓取需求。

2) 对比多种不同损失函数,采用 *EIoU* 损失情况的模型精度与速度综合属性最好,与采用 *CIoU* 损失的模型相比,检测平均速度、召回率与全类平均精度三项指标均有明显增强,提高采摘机器人在工作过程中的精度与速度。

3) 在实验室进行模拟环境中机器人的采摘试验,验证改进模型能够应用于采摘机器人并准确实现采摘功能的技术可实现性,突破设施环境中复杂情况下的部分条件制约,满足番茄采摘机器人的采摘识别需求。

参 考 文 献

- [1] 贾玉琴. 甘肃省设施农业和蔬菜产业的发展现状、潜力与对策分析[J]. 中国瓜菜, 2023, 36(9): 144-150.
Jia Yuqin. Development status, potential and countermeasure analysis of facility agriculture and vegetable industry in Gansu [J]. China Cucurbits and Vegetables, 2023, 36(9): 144-150.
- [2] 赵敬, 杨化伟, 刘光辉, 等. 我国果蔬采摘机器人技术发展现状及展望[J]. 农业装备与车辆工程, 2023, 61(7): 23-28.
Zhao Jing, Yang Huawei, Liu Guanghui, et al. Development status and prospects of fruit and vegetable picking robots technology in China [J]. Agricultural Equipment & Vehicle Engineering, 2023, 61(7): 23-28.
- [3] 戴军. 机器视觉技术在瓜菜检测应用中的研究进展[J]. 中国瓜菜, 2023, 36(11): 1-9.
Dai Jun. Research progress of machine vision technology in the detection of cucurbits and vegetables [J]. China Cucurbits and Vegetables, 2023, 36(11): 1-9.
- [4] 冯俊惠, 李志伟, 戎有丽, 等. 基于改进 Hough 圆变换算法的成熟番茄果实识别[J]. 中国农机化学报, 2021, 42(4): 190-196.
Feng Junhui, Li Zhiwei, Rong Youli, et al. Identification of mature tomatoes based on an algorithm of modified circular Hough transform [J]. Journal of Chinese Agricultural Mechanization, 2021, 42(4): 190-196.
- [5] 金超杞, 梁喜凤, 章艳, 等. 基于 Matlab 的番茄果实串检测提取[J]. 江苏农业科学, 2018, 46(8): 209-211.
Jin Chaoqi, Liang Xifeng, Zhang Yan, et al. Tomato fruit string detection and extraction based on Matlab [J]. Jiangsu Agricultural Science, 2018, 46(8): 209-211.
- [6] 梁晓婷, 庞琦, 杨一, 等. 基于 YOLOv4 模型剪枝的番茄缺陷在线检测[J]. 农业工程学报, 2022, 38(6): 283-292.
Liang Xiaoting, Pang Qi, Yang Yi, et al. Online detection of tomato defects based on YOLOv4 model pruning [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(6): 283-292.
- [7] Bhagya C, Shyna A. An overview of deep learning based object detection techniques [C]. 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT). IEEE, 2019: 1-6.
- [8] Jiao L, Zhang F, Liu F, et al. A survey of deep learning-based object detection [J]. IEEE access, 2019, 7: 128837-128868.
- [9] Hary C, Mandala S. Object detection analysis study in images based on deep learning algorithm [C]. 2022 International Conference on Data Science and Its Applications (ICoDSA). IEEE, 2022: 226-231.
- [10] Zhou Mingzhe. Research advanced in deep learning object detection [C]. 2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), Dalian, China, 2022: 1318-1322.
- [11] 吕志远, 张付杰, 魏晓明, 等. 采用组合增强的 YOLOX-ViT 协同识别温室番茄花果[J]. 农业工程学报, 2023, 39(4): 124-134.
Lü Zhiyuan, Zhang Fujie, Wei Xiaoming, et al. Synergistic recognition of tomato flowers and fruits in greenhouse using combination enhancement of YOLOX-ViT [J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(4): 124-134.
- [12] 李天华, 孙萌, 丁小明, 等. 基于 YOLO v4+HSV 的成熟期番茄识别方法[J]. 农业工程学报, 2021, 37(21): 183-190.
Li Tianhua, Sun Meng, Ding Xiaoming, et al. Tomato recognition method at the ripening stage based on YOLO v4 and HSV [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(21): 183-190.
- [13] 杨坚, 钱振, 张燕军, 等. 采用改进 YOLOv4-tiny 的复杂环境下番茄实时识别[J]. 农业工程学报, 2022, 38(9): 215-221.
Yang Jian, Qian Zhen, Zhang Yanjun, et al. Real-time recognition of tomatoes in complex environments based on improved YOLOv4-tiny [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(9): 215-221.
- [14] 何斌, 张亦博, 龚健林, 等. 基于改进 YOLO v5 的夜间温室番茄果实快速识别[J]. 农业机械学报, 2022, 53(5): 201-208.
He Bin, Zhang Yibo, Gong Jianlin, et al. Fast recognition of tomato fruit in greenhouse at night based on improved YOLO v5 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 201-208.
- [15] 朱智惟. 番茄采摘机器人果实识别与定位技术研究[D]. 合肥: 合肥工业大学, 2022.
Zhu Zhiwei. Tomato picking robot fruit recognition and positioning technology research [D]. Hefei: Hefei University of Technology, 2022.
- [16] 王海楠, 弋景刚, 张秀花. 番茄采摘机器人识别与定位技

- 术研究进展[J]. 中国农机化学报, 2020, 41(5): 188-196.
- Wang Hainan, Yi Jinggang, Zhang Xiuhua, et al. Research progress on recognition and localization technology of tomato picking robot [J]. Journal of Chinese Agricultural Mechanization, 2020, 41(5): 188-196.
- [17] Han W, Hao W, Sun J, et al. Tomatoes maturity detection approach based on YOLOv5 and attention mechanisms [C]. 2022 IEEE 4th International Conference on Civil Aviation Safety and Information Technology (ICCSIT). IEEE, 2022: 1363-1371.
- [18] 代国威, 樊景超, 胡林. 采用天气增强与八度卷积改进 YOLOv5 的番茄检测模型构建[J]. 山东农业科学, 2022, 54(11): 138-149.
- Dai Guowei, Fan Jingchao, Hu Lin. Improved tomato detection model construction based on YOLOv5 using weather augmentation and octave convolution [J]. Shandong Agricultural Sciences, 2022, 54(11): 138-149.
- [19] 郎松, 曹选, 张艳微, 等. 融合改进 YOLOv5 算法的图像全站仪全自动测量方法[J]. 仪器仪表学报, 2022, 43(5): 120-127.
- Lang Song, Cao Xuan, Zhang Yanwei, et al. Fully automated measurement method of image total station based on the improved YOLOv5 algorithm [J]. Chinese Journal of Scientific Instrument, 2022, 43(5): 120-127.
- [20] 彭继慎, 孙礼鑫, 王凯, 等. 基于模型压缩的 ED-YOLO 电力巡检无人机避障目标检测算法[J]. 仪器仪表学报, 2021, 42(10): 161-170.
- Peng Jishen, Sun Lixin, Wang Kai, et al. ED-YOLO power inspection UAV obstacle avoidance target detection algorithm based on model compression [J]. Chinese Journal of Scientific Instrument, 2021, 42(10): 161-170.
- [21] Hao K, Chen G, Zhao L, et al. An insulator defect detection model in aerial images based on multiscale feature pyramid network [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-12.
- [22] Khalfaoui A, Badri A, Mourabit I E L. Comparative study of YOLOv3 and YOLOv5's performances for real-time person detection [C]. 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET). IEEE, 2022: 1-5.
- [23] Li B, Hou Y, Che W. Data augmentation approaches in natural language processing: A survey [J]. Ai Open, 2022, 3: 71-90.
- [24] Zhang Y F, Ren W, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146-157.
- [25] Zheng Z, Wang P, Ren D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [J]. IEEE Transactions on Cybernetics, 2021, 52(8): 8574-8586.
- [26] Loss G Z S I U. More powerful learning for bounding box regression [J]. arXiv Preprint arXiv: 2205.12740, 2022.
- [27] Rezatofighi H, Tsoi N, Gwak J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.
- [28] Tong Z, Chen Y, Xu Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism. arXiv 2023 [J]. arXiv preprint arXiv: 2301.10051.