doi:10.6041/j.issn.1000-1298.2025.10.050

基于 YOLO v5 - MNv4S 与 RANSAC - GN 的柑橘姿态估计方法

李　丽[1,2]　张官明[1,2]　张云峰[1,2]　梁继元[1,2]　淳长品[3]

(1. 西南大学工程技术学院, 重庆 400715; 2. 重庆市丘陵山区农业装备重点实验室, 重庆 400715;
3. 西南大学柑橘研究所, 重庆 400700)

摘要: 为解决果园环境下的柑橘多变姿态估计问题,提高柑橘采收成功率,本文提出一种轻量化目标检测网络 YOLO v5 - MNv4S,与改进的随机采样一致性-高斯牛顿(Random sample consensus - Gauss Newton,RANSAC - GN)点云处理算法,根据柑橘生长及末端采摘流程定义柑橘姿态,构建了柑橘姿态实时估计系统。首先将 YOLO v5s 骨干网络优化为轻量化特征提取网络(MNv4 - Conv - S),大幅降低训练参数,使最终输出网络权重轻量化,减少计算量,提高识别效率;其次添加 CA 注意力机制及替换损失函数为 SIoU,解决了轻量化网络特征提取能力弱的问题。上述改进实现了 YOLO v5 - MNv4S 网络轻量化与优于 YOLO v5s 的检测能力。D435i 型相机采集图像后输入到 YOLO v5 - MNv4S 获取目标边界框,使用针孔模型输出柑橘区域点云,分割出柑橘表面点云,结合本文改进的 RANSAC - GN 点云算法拟合出准确稳定的柑橘参数,融合果蒂空间坐标,最终输出待采摘柑橘空间姿态结果。消融试验及网络对比试验结果表明,轻量化的 YOLO v5 - MNv4S 精度为 93.1%,参数量仅为 YOLO v5s 的 14.7%,对比 YOLO v5s - Ghost 以及 YOLO v7 、YOLO v8 等网络,其在参数大幅缩减的同时,识别精度最佳。柑橘定位与姿态识别试验结果表明,采用 RANSAC - GN 的柑橘参数拟合误差为(0.18,0.19,0.44) mm,姿态估计误差为 2.56°,姿态估计准确,真实果园环境下柑橘姿态估计结果与真实柑橘一致。该研究能满足果园环境下对柑橘姿态的识别,可为结构化柑橘果园机械采收设备提供技术支持。

关键词: 柑橘采摘; 目标检测; 姿态估计; 轻量化网络; 点云处理; MobileNetV4

中图分类号: TP242.3; S24　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)10-0558-09　　**OSID**:

Citrus Pose Estimation Based on YOLO v5 - MNv4S and RANSAC - GN

LI Li[1,2]　ZHANG Guanming[1,2]　ZHANG Yunfeng[1,2]　LIANG Jiyuan[1,2]　CHUN Changpin[3]

(1. *College of Engineering and Technology*, *Southwest University*, *Chongqing* 400715, *China*
2. *Chongqing Key Laboratory of Agricultural Equipment for Hilly and Mountainous Regions*, *Chongqing* 400715, *China*
3. *Citrus Research Institute*, *Southwest University*, *Chongqing* 400700, *China*)

Abstract: Aiming to address the challenge of estimating the variable pose of citrus in orchard environments and enhance the citrus harvesting success rate, a lightweight target detection network, YOLO v5 - MNv4S, along with an improved random sample consensus - Gauss Newton (RANSAC - GN) point cloud processing algorithm was introduced. The citrus pose was defined based on actual citrus growth and the final harvesting process, constructing a real-time citrus pose estimation system. The YOLO v5s backbone network was optimized to be a lightweight feature extraction network (MNv4 -Conv - S), which significantly reduced the training parameters so that the final output network weights were lightweight, reducing the amount of computation and improving the recognition efficiency. Additionally, the CA attention mechanism was incorporated, and the loss function was replaced with SIoU, addressing the weak feature extraction ability of the lightweight network. These improvements resulted in a lightweight YOLO v5 - MNv4S network with superior detection capabilities compared with YOLO v5s. After the D435i image acquisition, YOLO v5 - MNv4S was input to detect the target bounding box by

收稿日期: 2024 - 07 - 16　修回日期: 2024 - 08 - 20

基金项目: 重庆市杰出青年科学基金项目(2022NSCQ - JQX0030)、宜宾市双城协议保障科研经费科技项目(XNDX2022020015)、中央高校基本科研业务费项目(Swu - XDJH202302)和重庆市研究生科研创新项目(CYB23125)

作者简介: 李丽(1982—),女,教授,博士生导师,主要从事智能农机装备研究,E-mail: cqulily@163.com

using the pinhole model to output the citrus regional point cloud, segment the citrus surface point cloud, combined with the improved RANSAC－GN point cloud algorithm to fit accurate and stable citrus parameters, then fusion of the stem-end spatial coordinates, and ultimately the output of citrus spatial pose results to be harvested. Ablation experiments and network comparisons demonstrated that the lightweight YOLO v5－MNv4S achieved 93.1% accuracy, with only 14.7% of the parameters found in YOLO v5s. Compared with YOLO－Ghost, YOLO v7, YOLO v8, and other networks, it offered the best recognition accuracy with significantly reduced parameters. Experimental results for citrus localization and pose recognition showed that the citrus parameter fitting error using RANSAC－GN was (0.18, 0.19, 0.44) mm, and the pose estimation error was 2.56°. The pose estimation was accurate, and the estimation results of citrus pose in real orchard environments were consistent with real citrus. The research result can recognize citrus pose in orchard environments and provide technical support for structured citrus orchard mechanical harvesting equipment.

Key words: citrus picking; object detection; pose estimation; lightweight network; point cloud processing; MobileNetV4

0　引言

我国柑橘栽培面积位居世界首位，成为世界第一大柑橘生产国[1]，根据《中国统计年鉴》，2022 年我国柑橘产量为 $6.003\,9\times10^7$ t。在人口老龄化与柑橘人工采摘需要大量劳动力的冲突背景下，柑橘自动收获技术亟待取得突破。智慧农业科技是农业领域又一次新技术革命[2]，国家有关部门出台了一系列政策文件鼓励因地制宜发展柑橘产业，凸显了智能化柑橘收获技术的研发对柑橘产业的重要性[3]。

自然环境下的柑橘生长姿态是无规律的，人工采摘柑橘时会首先确定果梗方位，再进行剪切，实现柑橘自动化采摘必须克服姿态多变所带来的挑战。在真实果园环境下，准确估计柑橘的姿态并采用适当的末端采收姿势是成功采收柑橘的前提，在移动末端执行器之前获取柑橘的姿态可提高收获成功率[4]。为了实现复杂环境下水果自动化采收，需为采摘机器人实现稳定且准确的姿态识别系统[5]。

随着神经网络的发展，深度学习广泛应用于水果信息获取。文献[6－8]研究聚焦于目标检测与网络轻量化，目前农业信息化研究大量聚焦于目标检测任务[9－11]，后续定位工作的研究较少。文献[12－13]研究目标检测与定位，但部分水果如柑橘、番茄等在姿态未知情况下，采摘成功率偏低。文献[14－16]研究使用深度相机的 RGB 信息，由于输入会受到实际环境中各种噪声的影响，姿态估计准确性有待提高。文献[17－18]研究使用了 RGB 和深度信息，但由于深度相机的深度信息不可靠[19]和实际环境的复杂性，姿势估计的稳定性有待提高。综上所述，目前水果采摘视觉技术主要包含识别、定位、姿态估计等方面。受限于普通商用 RGB－D 相机传感器的性能与自然环境特性，深度信息质量不稳定，RGB 图像质量易受环境影响。利用普通商用深度相机实现果园环境下的柑橘姿态估计对于柑橘自动化采摘研究至关重要。

本文提出一种从实际柑橘生长及末端采摘流程出发，并全面利用 RGB－D 信息识别、定位并估计柑橘姿态的方法。采用本文提出的 YOLO v5－MNv4S 目标检测网络与 RANSAC－GN 点云处理算法实现树上柑橘姿态的实时估计。以期提高柑橘采摘机器人在实际果园中的柑橘采摘成功率，减少柑橘损失。

1　图像数据获取与预处理

1.1　柑橘与果蒂图像采集

本文研究对象为现代化柑橘园种植的柑橘，数据集为使用 Intel D435i 型相机采集于重庆市北碚区西南大学柑橘研究所与九龙坡区“橙里城外”庄园的芸香科柑橘属脐橙类图像。考虑到真实果园的复杂性，相机置于柑橘 0.4～1.5 m 处并选择不同角度，不同时间采集图像 1 000 幅。数据集按照比例 8∶1∶1划分训练集、验证集与测试集，目标数据分布见表 1。使用 Labelimg 标注软件标注柑橘与果蒂，其中果蒂为柑橘与果梗连接处，如图 1d～1f 所示。

表 1　标注目标的数量分布

Tab.1　Distribution of number of labeled targets

数据集	柑橘	果蒂	总计
训练集	2 364	754	3 118
验证集	351	116	467
测试集	319	89	408
总计	3 034	959	3 993

1.2　图像数据预处理

为提高网络鲁棒性，防止模型训练中过拟合，对训练数据集使用多种数据增强技术来处理，包括亮度调节、Cutout、剪切、镜像、添加噪声。增强后的图像如图 2 所示。

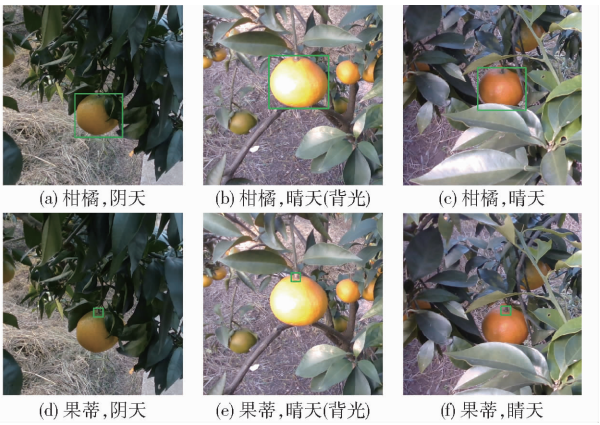


图 1 不同条件下标注的图像数据集

Fig. 1 Labeled image dataset under different conditions

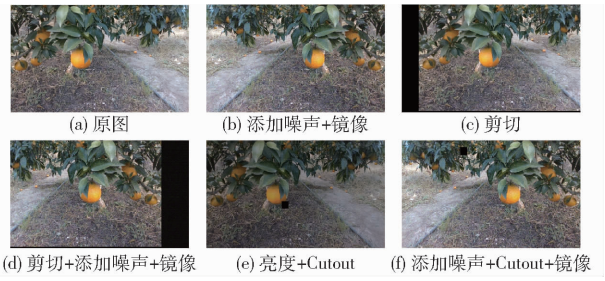


图 2 数据增强后的图像示例

Fig. 2 Image under data enhancement

2 复杂环境下的柑橘姿态估计算法

2.1 原始 YOLO v5s 网络

当前目标检测的深度学习算法可分为两类:One-stage、Two-stage。YOLO 系列采用 One - stage 结构[20],v5 系列平衡检测精度与检测速度,权重文件较小,较新的 YOLO 架构如 v8,其参数量多、模型大,检测精度相比原始 v5 提升微弱。本文选择改进原始 YOLO v5 网络完成自然环境下的柑橘检测任务。YOLO v5s 目标检测网络由 3 部分组成:骨干网络提取目标特征,将原始的高分辨输入图像转换为多层低分辨率特征图;颈部网络对骨干网络提取的浅层次特征进行不同尺度特征融合;检测网络对颈部网络输出的不同尺度特征图进行回归预测,得出图像的检测框及目标类别等信息[21]。

2.2 改进的 YOLO v5 - MNv4S 网络结构

本文目标识别网络以 YOLO v5s 作为基础网络,为后续能将本文提出的姿态估计算法部署于算力不高的计算平台,本文将 MobileNetV4[22] 提出的轻量化网络 MNv4 - Conv - S 替换 YOLO v5 骨干网络用于提取目标特征信息,替换后如图 3 中骨干网络所示。在修改后的特征提取骨干网络中加入 CA (Coordinate attention)[23] 注意力机制,以提高神经网络对图像中柑橘与果蒂目标特征的关注度,其次损失函数从 CIoU 变更为 SIoU[24],经过以上修改的网络被命名为 YOLO v5 - MNv4S,其整体结构如图 3 所示。

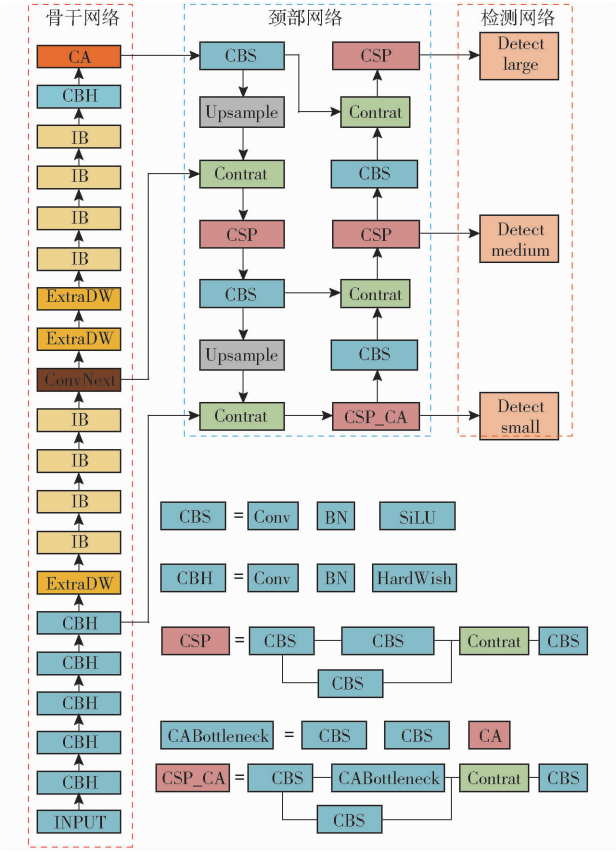


图 3 YOLO v5 - MNv4S 网络结构图

Fig. 3 YOLO v5 - MNv4S network structure diagram

2.2.1 轻量化特征提取网络

MNv4 - Conv - S 是由谷歌提出的 MobileNetV4 系列中的一款轻量级神经网络结构,适用于低性能移动设备和嵌入式设备。MobileNetV4 主要由 UIB 与 Mobile MQA 注意力机制组成。UIB (Universal inverted bottleneck)是一种统一灵活的结构,合并了倒置瓶颈(IB)[25]、ConvNext[26]、前馈网络(FFN)[27]和一种新的额外深度卷积(Extra DepthWise)变体,结构如图 4 所示。使用原始 UIB 网络结构中的 Extra DW 与 IB 模块来搭建本文所述 MNv4 - Conv - S 特征提取骨干网络。

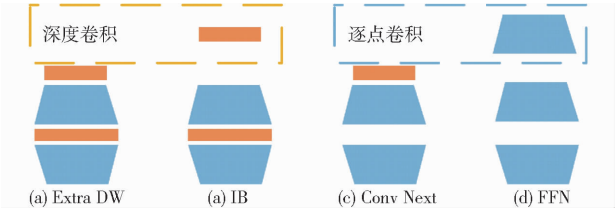


图 4 UIB 块的 4 种结构

Fig. 4 Four structures of UIB blocks

2.2.2 添加注意力机制

本研究所检测目标在实际环境下存在特殊形状

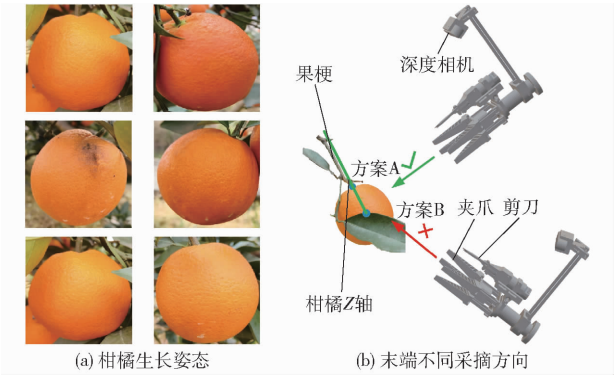
特征及特殊空间位置特征，为此在 YOLO v5 的检测网络中加入 CA 注意力机制可有效提高训练模型检测速度、网络特征提取能力和检测精度。CA 注意力不仅考虑了通道信息，还考虑了方向相关的位置信息，足够灵活和轻量，能简单地插入到轻量级网络 MobileNetV4 中。

2.2.3　修改损失函数

YOLO 的损失函数由 3 部分组成，预测框损失、置信度损失和分类损失，修改原始 YOLO v5 预测框损失函数为 SIoU。SIoU 考虑期望回归之间向量的角度，它可以使预测框快速漂移到最近的轴，随后则只需要回归一个坐标，这有效地减少了自由度的总数，可以提高训练速度和推理准确性。

2.3　柑橘姿态定义

基于本课题组所研制末端执行器，柑橘采摘流程如图 5 所示。在方案 A 中夹爪中心首先移动到柑橘中心，控制夹爪电机夹紧柑橘，此时末端上的剪刀动作剪断柑橘果梗。



(a) 柑橘生长姿态　　(b) 末端不同采摘方向

图 5　柑橘生长姿态与末端采摘方向

Fig. 5　Citrus growth pose and direction of end-effector picking

剪断果梗需要获取图 5 中多变的果梗方位，并调整末端姿态。脐橙果梗位于柑橘中心与果蒂连接延长线附近，故定义柑橘质心与果蒂的方向为柑橘姿态 Z 轴。

2.4　三维坐标转换

选用 Intel D435i 型相机可以方便获取 RGB－D 信息。提前标定获取相机内参等参数，是建立二维图像到三维坐标准确对应关系的前提[28]。

使用 Dynamic Calibrator 工具箱及具有固定尺寸的动态校准标定板进行标定，获取内部参数，D435i 型相机标定结果如表 2 所示。

获得准确的内部参数后，利用针孔相机模型可计算二维坐标对应的三维空间坐标[29]。点 $\boldsymbol{C}$ 三维坐标计算式为

$$\boldsymbol{C} = \boldsymbol{K}D[u \quad v \quad 1]^{\mathrm{T}} \tag{1}$$

式中　$\boldsymbol{K}$——相机内参矩阵

D——计算像素点 P'对应的三维空间点 C 与相机平面的距离

$[u \quad v \quad 1]^{\mathrm{T}}$——点 $\boldsymbol{C}$ 像素坐标的齐次表示形式

表 2　D435i 型相机内部参数标定结果

Tab. 2　D435i camera intrinsics calibration results

参数	相机标定结果
相机内部参数	$\begin{bmatrix} 909.0874 & 0 & 640.9918 \\ 0 & 907.8764 & 360.3285 \\ 0 & 0 & 1.0 \end{bmatrix}$
旋转矩阵	$\begin{bmatrix} 1.0 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}$
投影矩阵	$\begin{bmatrix} 909.0874 & 0 & 640.9918 \\ 0 & 0 & 907.8764 \\ 360.3285 & 0 & 0 \\ 0 & 1.0 & 0 \end{bmatrix}$

2.5　基于 RANSAC－GN 的柑橘姿态精确估计

2.5.1　柑橘姿态精确估计流程

柑橘姿态的估计依赖于柑橘中心点的准确计算，本文提出了一种基于随机采样一致性-高斯牛顿（Random sample consensus－Gauss Newton，RANSAC－GN）的柑橘中心点估计方法。该算法应用于目标识别算法之后，获取柑橘目标点云即可计算准确的柑橘中心点，结合果蒂坐标输出柑橘姿态。柑橘姿态估计流程如图 6 所示。

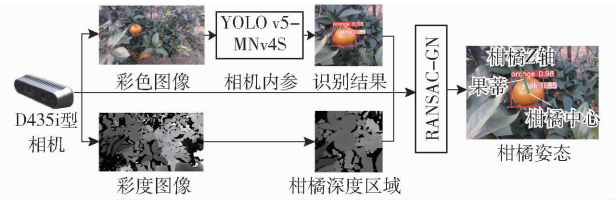


图 6　柑橘姿态估计流程图

Fig. 6　Citrus pose estimation process

RANSAC－GN 基于随机采样一致性进行改进。RANSAC 采用迭代的方式从 1 组包含离群或者错误数据的被观测数据中估算出数学模型的参数，相较于最小二乘法，可以剔除掉不合格数据[30]。D435i 型深度相机输出的深度信息包含无效深度带，以及部分深度误差，为此提出 RANSAC－GN 来减少计算误差。

2.5.2　RANSAC－GN 算法流程

RANSAC－GN 在原始 RANSAC 算法之后增加后处理流程，使用高斯牛顿优化方法来迭代减少误差，流程如下：

（1）由于柑橘类似于球体，从球体方程定义优化目标函数

$$f(\boldsymbol{X}_k)=\sum_{i=0}^{n}\left[(P_i^x-x_k)^2+(P_i^y-y_k)^2+(P_i^z-z_k)^2-r_k^2\right] \tag{2}$$

其中 $\boldsymbol{X}_k=[x_k \quad y_k \quad z_k \quad r_k]$

式中 n——点云中点的数量

x_k、y_k、z_k、r_k——待优化柑橘参数

P_i^x、P_i^y、P_i^z——柑橘表面点云中各点三维坐标

$\boldsymbol{X}_k$——柑橘参数

(2) RANSAC 给出初值向量 $\boldsymbol{X}_k$($k=1,2,\cdots,n$), $\boldsymbol{X}_0=(x_0,y_0,z_0,r_0)$。

(3)对与第 k 次迭代,计算 $f(\boldsymbol{X}_k)$ 雅可比矩阵 $\boldsymbol{J}_k$,计算代价 f_k,其等于当前 $f(\boldsymbol{X}_k)$ 值。

(4)求解增量方程

$$\boldsymbol{H}\Delta\boldsymbol{x}=\boldsymbol{g} \tag{3}$$

其中

$$\boldsymbol{H}=\boldsymbol{J}_k^{\mathrm{T}}\boldsymbol{J}_k \tag{4}$$

$$\boldsymbol{g}=-\boldsymbol{J}_k^{\mathrm{T}}f_k \tag{5}$$

(5)如果 Δx 小于设定误差,停止迭代,否则 $\boldsymbol{X}_k=\boldsymbol{X}_{k-1}+\Delta\boldsymbol{x}$,返回步骤(3)循环。

3 试验结果与分析

3.1 试验环境

图像采集设备为 Intel Realsense D435i 型相机,采集图像的分辨率设置为 1 280 像素 × 720 像素。在 Linux 服务器(Intel Xeon Gold 6348 CPU @ 2.6 GHz;503 GB 内存; NVIDIA A800 80GB X2),Ubuntu 22 操作系统下搭建了 PyTorch 深度学习框架,利用 Python 语言编写程序代码来训练与评估 YOLO v5 - MNv4S。

3.2 模型性能评估指标

为验证模型性能的指标,本文使用精确率(P)、召回率(R)和平均精度均值(mAP)3 个常用指标评估模型。

3.3 模型性能分析

3.3.1 注意力对比试验

为验证不同注意力机制对改进模型的影响,本文将 SE、CBAM、ECA、CA、SimAM 等 5 种注意力机制按照图 4 所示网络结构,仅替换其中注意力模块,损失函数使用原始 CIoU。观察表 3 可知,基于神经科学提出 SimAM 注意力机制结果优于 SE、CBAM、ECA,这与文献[31]所给出试验结果一致。相较于其余 4 种注意力机制,CA 注意力机制关注了柑橘及其果蒂之间的特殊空间及形状关系,其试验各项结果最佳。综上所述,本文选择将 CA 注意力机制插入轻量化网络中。

表 3 注意力对比试验结果

Tab. 3 Attentional comparison experiment results

%

注意力机制	精确率	召回率	平均精度均值
SE	82.1	64.3	73.7
CBAM	84.0	65.9	75.5
ECA	87.9	67.8	73.7
SimAM	91.6	63.0	76.5
CA	94.0	68.2	80.1

3.3.2 损失函数对比试验

将 4 种损失函数 GIoU、DIoU、EIoU、SIoU 分别替换 YOLO v5 原有损失函数 CIoU,网络结构见图 4,注意力机制为 CA。观察表 4,GIoU 在预测框与真实框为包含情况时退化为 IoU,难以区分目标中心相对位置,其各项指标均为最低。DIoU 最小化目标中心距离,精度取得大幅提升,CIoU 在 DIoU 上添加纵横比尺度损失,各项指标均有所提升。EIoU 在 CIoU 基础上将预测框纵横比分开计算损失,召回率提升较大。SIoU 综合考虑上述各项损失,有效地提升召回率与平均精度均值,试验中 SIoU 取得最高召回率与平均精度均值,精确率比最高低 0.9 个百分点。综合评价模型 3 项指标,选取 SIoU 作为损失函数。

表 4 损失函数对比试验结果

Tab. 4 Loss function comparison experiment results

%

损失函数	精确率	召回率	平均精度均值
GIoU	77.4	66.5	73.7
DIoU	93.8	63.2	74.6
EIoU	92.5	75.1	80.6
CIoU	94.0	68.2	80.1
SIoU	93.1	80.8	82.9

3.3.3 消融试验对比结果分析

为验证改进方法的有效性,采用消融试验进行验证。使用相同的数据集、测试平台及模型参数进行对比试验。YOLO v5 的 3 种改进方法分别为替换骨干网络为 MNv4 - Conv - S、添加 CA 注意力机制和修改损失函数为 SIoU,试验方法如下:①在原始 YOLO v5 算法中分别加入 3 种改进方式,分别观察效果。②加入全部的改进模块,将 YOLO v5 - MNv4S 与原 YOLO v5 进行比对。从表 5 可以看出,轻量化网络 MNv4 - Conv - S 大幅降低了参数量,CA 注意力机制与 SIoU 损失函数提高了总体网络的性能。所提出改进网络平均精度均值仅降低 3.6 个百分点,精确率提高 0.8 个百分点,召回率提高 5.7 个百分点,模型内存占用量仅为原始模型的 17.4%,

参数量为原始模型的 14.7%。说明上述改进在更少的参数量实现优于原始 YOLO v5 检测性能，足够轻量，便于部署于低性能计算平台。

3.3.4　不同网络对比结果分析

为验证 YOLO v5－MNv4S 对柑橘以及果蒂的检测效果，使用 YOLO v5s、YOLO v3、YOLO v7 和 YOLO v8s，及轻量网络 YOLO v5s－Ghost、YOLO v5s－MobileNetv3 等模型在相同环境下对柑橘数据进行训练和测试。观察表 6，本文提出的改进模型与其他检测模型相比，精确率均有所提高，较改进之前的 YOLO v5s 网络精度提升 0.8 个百分点，同时对比综合性能最好的 YOLO v7，精度提升 0.7 百分点。同时在添加多种改进因素的情况下大幅降低了参数数量，对比原始 YOLO v5s，参数量下降 85.2%，对比综合性能最佳的 YOLO v7，参数量仅为其 2.8%。在轻量级网络中对比性能最佳的 YOLO v5s－Ghost，在精度提升 0.7 个百分点下，参数量下降 79.6%。综上所述，本研究所提出的 YOLO v5－MNv4S 是以上 7 种网络模型最轻量化的。对比轻量级网络，其精确率与召回率最佳，与大参数量网络 YOLO v7 对比，有更高的精度，YOLO v5－MNv4S 在最轻量化的前提下，实现了最高精度与较高的召回率，综合性能最优，故本文选取 YOLO v5－MNv4S 进行目标检测任务。

表 5　消融试验对比结果

Tab. 5　Comparative results of ablation experiments

MNv4－Conv－S	SIoU	CA	精确率/%	召回率/%	平均精度均值/%	模型内存占用量/MB	参数量
×	×	×	92.3	75.1	86.5	14.4	7.02×10^6
√	×	×	88.9	70.8	76.9	2.5	1.04×10^6
√	√	×	84.6	74.5	80.5	2.5	1.04×10^6
√	×	√	94.0	68.2	80.1	2.5	1.04×10^6
√	√	√	93.1	80.8	82.9	2.5	1.04×10^6

注：“√”表示使用此改进方法；“×”表示不使用此改进方法。

表 6　与主流模型的对比试验结果

Tab. 6　Comparative test results with mainstream models

算法模型	精确率/%	召回率/%	平均精度均值/%	模型内存占用量/MB	参数量
YOLO v5s	92.3	75.1	86.5	14.4	7.02×10^6
YOLO v3	93.1	69.6	83.4	123.6	6.15×10^6
YOLO v7	92.4	83.3	90.2	74.8	3.64×10^6
YOLO v8s	85.0	73.7	85.1	22.5	1.13×10^6
YOLO v5s－Ghost	92.4	77.6	86.6	10.6	5.09×10^6
YOLO v5s－MobileNetv3	87.1	74.0	82.3	10.5	5.02×10^6
YOLO v5－MNv4S	93.1	80.8	82.9	2.5	1.04×10^6

3.3.5　实际柑橘目标识别试验结果分析

为验证本文提出的改进网络能应用于柑橘果园进行实际柑橘、果蒂目标识别，YOLO v5－MNv4S 网络模型与原始 YOLO v5s 网络的识别结果如图 7、8 所示。图 7 为柑橘果蒂识别对比，可以看出原始 YOLO v5 出现了边缘误检，而本文提出的 YOLO v5－MNv4S 没有出现漏检情况。图 8 为多柑橘目标检测对比情况，在图 8 中原始 YOLO v5 在右下边缘未检测到柑橘，YOLO v5－MNv4S 能成功检测。综上，YOLO v5－MNv4S 在参数量远低于原始 YOLO v5 情况下，综合对比性能优于原始 YOLO v5。



(a) 原始YOLO v5识别柑橘、果蒂　(b) YOLO-MNv4S识别柑橘、果蒂

图 7　柑橘、果蒂识别结果对比

Fig. 7　Comparison of citrus and stem-end detection results



(a) 原始YOLO v5识别　(b) YOLOv5-MNv4S识别

图 8　多柑橘目标检测结果对比

Fig. 8　Comparison results for multiple citrus targets

3.4　基于 RANSAC－GN 柑橘姿态估计试验结果分析

本文在 Ubuntu 18.04 系统（AMD 7840HS CPU@3.8 GHz;32 GB 内存），基于 ROS1 框架，使用C++开发了柑橘姿态估计算法。

为验证柑橘参数计算的准确性，试验使用双目相机（D435i 型）与三脚架、激光雷达（robosense helios 16）、游标卡尺等仪器进行三维坐标测量，试验场景如图 9 所示。

试验步骤为：

（1）将激光雷达固定安装在三脚架上，雷达坐标系 X 轴正对柑橘，调整三脚架来设定激光雷达与柑橘的距离，启动 RSView 软件并设定激光雷达与深度相机的坐标变换参数，获取柑橘在双目相机下的准确坐标。

（2）取待检测柑橘高、宽之和的 1/4，作为柑橘半径值，统计柑橘真值，见表 7 第 4 列。



图 9 柑橘参数误差试验场景

Fig. 9 Citrus parameter error experiment scenarios

(3)将深度相机安装至三脚架上,启动柑橘参数计算程序,输出 RANSAC、RANSAC－GN 2 种算法计算柑橘质心坐标与半径。

(4)不同距离下选取 2 个柑橘,每个柑橘获取 2 次计算质心坐标与半径,汇总测量结果见表 7。

统计试验结果并绘制 *XYZ* 轴误差折线图,如图 10 所示,RANSAC－GN 所计算误差小,结果更稳定。改进 RANSAC－GN 算法 *X* 轴最大误差 0.37 mm,平均误差 0.18 mm; *Y* 轴最大误差 0.72 mm,平均误差 0.19 mm; *Z* 轴最大误差 0.91 mm,平均误差0.44 mm,柑橘半径最大误差 1.64 mm,平均误差0.58 mm。RANSAC－GN 参数估计结果优于原始 RANSAC。依据文献[31]抓取容许误差 *X* 轴 20 mm, *Y* 轴 15 mm, *Z* 轴 15 mm, RANSAC－GN 参数估计结果完全满足使用要求。

表 7 柑橘空间参数估计结果

Tab. 7 Results of spatial parameter estimation for citrus mm

柑橘编号	柑橘宽	柑橘高	真实值	RANSAC	RANSAC－GN
1	78.04	80.51	7.31,19.43,300.07,39.76	(7.55,17.96,299.27,38.34) (7.33,19.52,304.60,42.80)	(7.18,19.18,300.16,39.32) (7.22,19.27,300.42,39.56)
2	65.04	78.60	7.54,－2.66,294.87,35.91	(6.46,－3.18,296.35,35.61) (7.10,－2.62,292.27,32.12)	(7.64,－2.45,295.87,35.24) (7.31,－2.26,294.94,34.39)
3	75.74	80.14	12.11,－2.98,399.93,38.97	(10.19,－4.31,392.91,32.83) (12.43,－2.48,396.12,36.16)	(12.27,－3.07,399.73,39.26) (12.35,－3.15,399.02,38.62)
4	75.93	80.57	－5.02,23.54,399.37,39.13	(－4.66,23.43,402.58,41.39) (－4.88,23.11,398.13,38.38)	(－5.06,23.62,399.79,39.22) (－5.36,24.26,399.62,38.90)
5	72.37	81.33	0.35,－31.02,499.63,38.43	(1.66,－31.02,503.37,42.12) (－0.02,－31.41,496.14,36.01)	(0.58,－31.08,499.27,38.76) (0.21,－30.96,500.57,40.07)
6	80.14	79.39	－13.12,－9.68,499.86,39.88	(－12.72,－9.58,499.16,38.86) (－13.04,－9.97,494.91,35.06)	(－13.24,－9.72,499.79,38.99) (－13.49,－9.76,500.48,39.55)

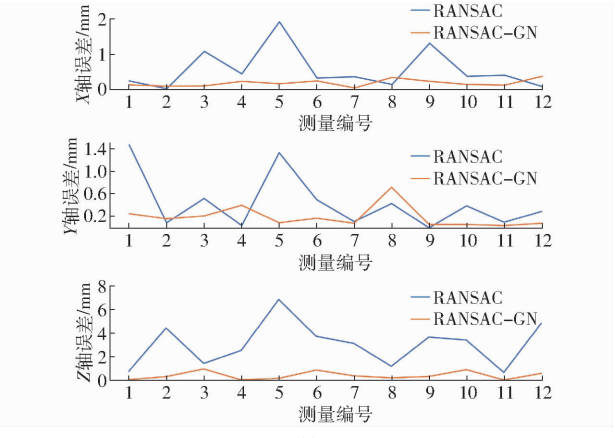


图 10 *XYZ* 轴误差折线图

Fig. 10 Line graph of *XYZ* axis error

为准确量化估计姿态误差,将柑橘置于可调整角度平台进行试验,对不同角度下的同一柑橘姿态估计 100 次,统计其平均误差,结果如表 8 所示,试验场景如图 11 所示。表中 RANSAC－GN 所计算姿态误差均小于原始RANSAC 算法,说明改进之后的

表 8 不同角度下的位姿估计平均误差

Tab. 8 Mean error of position estimation at different angles (°)

组别	RANSAC	RANSAC－GN
0°	2.03	1.08
5°	3.15	1.48
10°	5.72	2.56

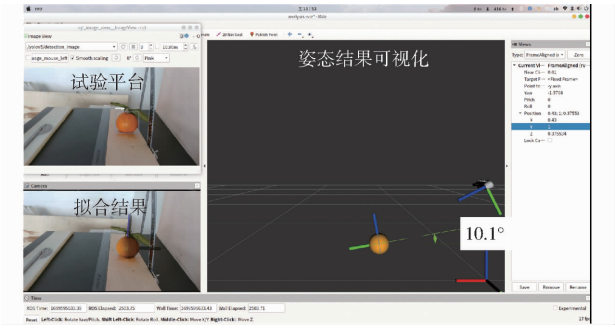


图 11 可调角度平台柑橘姿态估计试验场景

Fig. 11 Adjustable angle platform citruspose estimation experiment

RANSAC－GN 可有效地收敛角度误差范围，位姿误差仅为 2.56°。

最后将本文所提出姿态估计算法部署在柑橘收获机器人平台，于西南大学柑橘研究所柑橘园内进行姿态估计试验，姿态估计结果输出在 Rviz 软件中，如图 12 所示。真实果园下的柑橘果梗与拟合果梗基本重合一致，能应用于柑橘收获机器人进行实地收获。

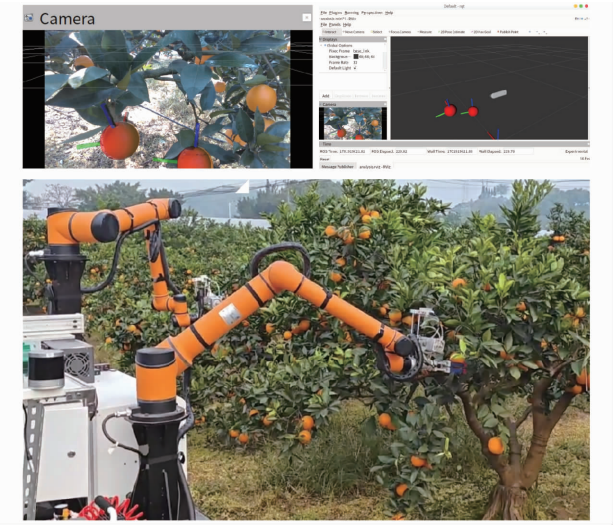


图 12　果园柑橘姿态估计试验

Fig. 12　Orchard citrus pose estimation experiment

4　结论

（1）本文基于 YOLO v5－MNv4S 检测模型与 RANSAC－GN 点云处理算法实现了真实柑橘果园环境下的柑橘姿态识别。

（2）采用最新的 MobileNetV4 轻量级网络 MNv4－Conv－S 替换 YOLO v5 的主干网络，添加 CA 注意力机制以及修改为 SIoU 损失函数，模型相比 YOLO v5 参数量降低了 85.3%，模型内存占用量减小了 82.6%，识别精确率提升 0.8 个百分点，召回率提升 5.7 个百分点，对比综合性能最佳的 YOLO v7 精准度提升 0.7 个百分点。YOLO v5－MNv4S 检测模型使用极低的参数量拥有最佳的精度，更适宜低算力平台部署。

（3）提出 RANSAC－GN 改进点云处理算法并在复杂果园环境下进行柑橘位置估计试验。试验结果显示，其定位误差在 500 mm 距离内为（0.18，0.19，0.58）mm，RANSAC－GN 相较于原始 RANSAC 算法极大收敛了误差范围。

（4）使用 C＋＋实现了柑橘姿态估计系统，柑橘位姿可视化输出在 Rviz 中，分别于实验室、柑橘果园条件下进行柑橘姿态估计试验。使用 RANSAC－GN 柑橘位姿估计算法后输出的位姿在实验室条件下位姿误差仅为 2.56°，实际果园环境下估计柑橘姿态与实际柑橘姿态结果吻合。可为复杂果园环境下的柑橘采收设备研究提供技术支持。

参　考　文　献

[1] 聂扬眉. 乡村振兴下我国柑橘产业发展情况——基于价值链模型的实证[J]. 中南民族大学学报(自然科学版)，2023，42(2)：283－288.
NIE Yangmei. China's citrus industry under rural revitalization: empirical evidence based on the value chain model[J]. Journal of South-Central University for Nationalities(Natural Science Edition)，2023，42(2)：283－288. (in Chinese)

[2] 曹冰雪，李鸿飞，赵春江，等. 智慧农业科技创新引领农业新质生产力发展路径[J]. 智慧农业(中英文)，2024，6(4)：116－127.
CAO Bingxue，LI Hongfei，ZHAO Chunjiang，et al. The path of smart agricultural technology innovation leading development of agricultural new quality productivity[J]. Smart Agriculture，2024，6(4)：116－127. (in Chinese)

[3] 何宇. 四川柑橘产业供给侧结构性改革绩效评价及推进路径研究[D]. 成都：四川农业大学，2024.
HE Yu. Performance evaluation and of supply side structural reform of citrus industry in Sichuan research on promotion path [D]. Chengdu：Sichuan Agricultural University，2024. (in Chinese)

[4] LUO L，YIN W，NING Z，et al. In-field pose estimation of grape clusters with combined point cloud segmentation and geometric analysis[J]. Computers and Electronics in Agriculture，2022，200：107197.

[5] 朱立学，赖颖杰，张世昂，等. 基于改进 U－Net 的火龙果采摘图像分割和姿态估计方法[J]. 农业机械学报，2023，54(11)：180－188.
ZHU Lixue，LAI Yingjie，ZHANG Shiang，et al. Image segmentation and pose estimation method for pitaya picking robot based on enhanced U－Net[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(11)：180－188. (in Chinese)

[6] 赵方，左官芳，顾思睿，等. 基于改进 YOLO v5s 的温室番茄检测模型轻量化研究[J]. 江苏农业科学，2024，52(8)：200－209.
ZHAO Fang，ZUO Guanfang，GU Sirui，et al. Lightweight research of greenhouse tomato detection model based on improved YOLO v5s[J]. Jiangsu Agricultural Sciences，2024，52(8)：200－209. (in Chinese)

[7] 郑文轩，杨瑛. 基于频域数据增强与轻量化 YOLO v7 模型的成熟期香梨目标检测方法[J]. 农业机械学报，2024，55(5)：244－253.
ZHENG Wenxuan，YANG Ying. Mature stage pear detection method based on frequency domain data augmentation and lightweight YOLO v7 model[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(5)：244－253. (in Chinese)

[8] 王卓，王健，王枭雄，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 real-time apple detection method based on improved YOLO v4 [J]. Transactions of the Chinese Society for Agricultural Machinery，2022，53(8)：294－302. (in Chinese)

[9] ZHANG Y, LI L, CHUN C, et al. Multi-scale feature adaptive fusion model for real-time detection in complex citrus orchard environments[J]. Computers and Electronics in Agriculture, 2024, 219: 108836.
[10] LAWAL O M. Real-time cucurbit fruit detection in greenhouse using improved YOLO series algorithm[J]. Precision Agriculture, 2024, 25: 347 - 359.
[11] XU L, WANG Y, SHI X, et al. Real-time and accurate detection of citrus in complex scenes based on HPL - YOLOv4[J]. Computers and Electronics in Agriculture, 2023, 205: 107590.
[12] 郭辉, 陈海洋, 高国民, 等. 基于 YOLO v5m 的红花花冠目标检测与空间定位方法[J]. 农业机械学报, 2023, 54(7): 272 - 281.
GUO Hui, CHEN Haiyang, GAO Guomin, et al. Safflowercorolla object detection and spatial positioning methods based on YOLO v5m[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(7): 272 - 281. (in Chinese)
[13] 李丽, 梁继元, 张云峰, 等. 基于改进 YOLO v5 的复杂环境下柑橘目标精准检测与定位方法[J]. 农业机械学报, 2024, 55(8): 280 - 290.
LI Li, LIANG Jiyuan, ZHANG Yunfeng, et al. Accurate detection and localization method of citrus targets in complex environments based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 280 - 290. (in Chinese)
[14] KIM J, PYO H, JANG I, et al. Tomato harvesting robotic system based on Deep - ToMaToS: deep learning network using transformation loss for 6D pose estimation of maturity classified tomatoes with side-stem[J]. Computers and Electronics in Agriculture, 2022, 201: 107300.
[15] SUN Q, ZHONG M, CHAI X, et al. Citrus pose estimation from an RGB image for automated harvesting[J]. Computers and Electronics in Agriculture, 2023, 211: 108022.
[16] ZHANG F, GAO J, SONG C, et al. TPMv2: an end-to-end tomato pose method based on 3D key points detection[J]. Computers and Electronics in Agriculture, 2023, 210: 107878.
[17] LIN G, TANG Y, ZOU X, et al. Guava detection and pose estimation using a low - cost RGB - D sensor in the field[J]. Sensors (Basel, Switzerland), 2019, 19(2): 428.
[18] LI H, ZHU Q, HUANG M, et al. Pose estimation of sweet pepper through symmetry axis detection[J]. Sensors (Basel, Switzerland), 2018, 18(9): 3083.
[19] 乔欣, 葛晨阳, 邓鹏超, 等. ToF 相机的有效深度数据提取与校正算法研究[J]. 智能科学与技术学报, 2020, 2(1): 72 - 79.
QIAO Xin, GE Chenyang, DENG Pengchao, et al. Research on valid depth data extraction and correction for ToF camera[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(1): 72 - 79. (in Chinese)
[20] 翟亚红, 王杰, 徐龙艳, 等. 基于改进 YOLO v5n 的舍养绵羊行为识别方法[J]. 农业机械学报, 2024, 55(4): 231 - 240.
ZHAI Yahong, WANG Jie, XU Longyan, et al. Behavior recognition of domesticated sheep based on improved YOLO v5n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 231 - 240. (in Chinese)
[21] 李丽, 卢世博, 任浩, 等. 基于改进 YOLO v5 的复杂环境下桑树枝干识别定位方法[J]. 农业机械学报, 2024, 55(2): 249 - 257.
LI Li, LU Shibo, REN Hao, et al. Mulberry branch identification and location method based on improved YOLO v5 in complex environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 249 - 257. (in Chinese)
[22] QIN D, LEICHNER C, DELAKIS M, et al. MobileNetV4—universal models for the mobile ecosystem[R]. Ithaca: Cornell University Library, arXiv.org, 2024.
[23] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2021: 13708 - 13717.
[24] GEVORGYAN Z. SIoU Loss: more powerful learning for bounding box regression[R]. Ithaca: Cornell University Library, arXiv.org, 2022.
[25] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 4510 - 4520.
[26] LIU Z, MAO H, WU C Y, et al. A ConvNet for the 2020s[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2022: 11966 - 11976.
[27] BEBIS G, GEORGIOPOULOS M. Feed-forward neural networks[J]. IEEE Potentials, IEEE, 1994, 13(4): 27 - 31.
[28] 张昌松, 李伟. 基于 YOLO v5 的直播油菜激光间苗系统设计与试验[J]. 农业机械学报, 2024, 55(4): 40 - 52.
ZHANG Changsong, LI Wei. Research and experiment of direct seeding rape laser seedling system based on YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 40 - 52. (in Chinese)
[29] 朱立成, 韩振浩, 赵博, 等. 基于 VINS - MONO 和改进 YOLO v4 - Tiny 的果园自主寻筐方法[J]. 农业机械学报, 2023, 54(8): 97 - 109.
ZHU Licheng, HAN Zhenhao, ZHAO Bo, et al. Autonomous basket searching method for orchards transporter based on VINS-MONO and improved YOLO v4 - Tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(8): 97 - 109. (in Chinese)
[30] 李明煌, 苏力德, 张永, 等. 基于改进 YOLOv8n - pose 和三维点云分析的蒙古马体尺自动测量方法[J]. 智慧农业(中英文), 2024, 6(4): 91 - 102.
LI Minghuang, SU Lide, ZHANG Yong, et al. Automatic measurement of mongolian horse body based on improved YOLOv8n-pose and point cloud analysis[J]. Smart Agriculture, 2024, 6(4): 91 - 102. (in Chinese)
[31] YANG L, ZHANG R-Y, LI L, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks[C]// Proceedings of the 38th International Conference on Machine Learning. PMLR, 2021: 11863 - 11874.