

DOI: 10.13718/j.cnki.xdzk.2025.02.015

王元昊, 娄欢欢, 罗红品, 等. 基于改进 YOLOv8 算法对被遮挡柑橘的识别与定位优化 [J]. 西南大学学报(自然科学版), 2025, 47(2): 171-183.

基于改进 YOLOv8 算法对被遮挡柑橘的识别与定位优化

王元昊, 娄欢欢, 罗红品, 付兴兰, 李光林

西南大学 工程技术学院, 重庆 400715

摘要: 针对果园环境中柑橘果实相互重叠和被枝叶遮挡, 导致机器视觉识别柑橘果实与定位目标柑橘空间位置难度较大的问题, 提出了一种基于 YOLOv8-SAM 的改进算法。通过增加 BAM(Bottleneck Attention Module)注意力机制提高模型对被遮挡柑橘的识别准确率, 运用 SAM(Segment Anything Model)算法对被遮挡柑橘轮廓形状进行识别, 并运用边缘检测法结合双目立体相机三维稠密深度点云得到被遮挡柑橘有效轮廓边, 使用最小二乘法拟合出被遮挡柑橘的完整轮廓以确定目标柑橘果实更精确的空间坐标位置。试验结果表明: 该算法可以准确识别并分离目标柑橘果实, 同时更精确地定位柑橘果实空间坐标。改进的 YOLOv8-SAM 算法在果园环境中对被遮挡柑橘果实的识别平均精度达到 91.1%, 对被遮挡柑橘形心空间坐标的平均定位误差相比传统定位方法减少了 16.22 mm, 平均果径误差降低了 7.99%, 可为柑橘采摘机器人对重叠与被遮挡果实的准确识别提供参考。

关键词: 柑橘采摘; 机器视觉; 空间定位; 轮廓重建; 遮挡果

实; 图像处理

中图分类号: TP391.41

文献标志码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1673-9868(2025)02-0171-13

Optimization of Identification and Localization of Occluded Citrus Based on Improved YOLOv8 Algorithm

WANG Yuanhao, LOU Huanhuan, LUO Hongpin,
FU Xinglan, LI Guanglin

College of Engineering and Technology, Southwest University, Chongqing 400715, China

Abstract: In response to the challenges for machine vision to identify citrus fruit and locate the spatial posi-

收稿日期: 2024-01-24

基金项目: 国家自然科学基金项目(31971782); 重庆市科委产业化重点专项(cstc2018jszx-cyzdX0051)。

作者简介: 王元昊, 硕士研究生, 主要从事智能农业装备研究。

通信作者: 李光林, 博士, 教授, 博士研究生导师。

tion of target citrus in orchards due to overlapping fruit and occlusion by branches and leaves, a modified algorithm based on YOLOv8-SAM was proposed. The model's accuracy in identifying occluded citrus fruit was improved by adding BAM (Bottleneck Attention Module) attention mechanism. The contour shape of occluded citrus fruit was identified using SAM (Segment Anything Model) algorithm, and effective contour edges were obtained by combining edge detection with a binocular camera's 3D dense point cloud. The complete contour of the occluded citrus fruit was fitted using least squares to determine the more precise spatial coordinate position of the target citrus fruit. The experimental results show that the algorithm can accurately identify and separate the target citrus fruit, and more precisely locate the spatial coordinate of the citrus fruit. The average identification accuracy of the modified YOLOv8-SAM algorithm for occluded citrus fruit in the orchard environment is 91.1%, and the average spatial coordinate positioning error of the citrus fruit's center compared to traditional positioning methods is reduced by 16.22 mm, and the average fruit diameter error is reduced by 7.99%. This algorithm can provide reference for accurate identification of overlapping and occluded citrus fruit by citrus harvesting robots.

Key words: citrus picking; machine vision; spatial positioning; contour reconstruction; occluded fruit; image processing

我国是世界第一大柑橘生产国, 约占世界柑橘产量的三分之一^[1]。据国家统计局和联合国粮食及农业组织统计, 2020 年我国柑橘种植面积超过 $2.8 \times 10^6 \text{ hm}^2$, 产量超过 $5 \times 10^7 \text{ t}$, 分布于全国 19 个省(自治区、直辖市)^[2]。目前, 柑橘果园仍采用人工采摘的方式收获成熟柑橘果实, 采摘过程耗时长, 劳动强度大, 人工成本高, 如何实现自动化采摘代替人工采摘作业成为柑橘果园管理中亟待解决的问题^[3]。

近年来, 随着机器视觉和人工智能技术的迅猛发展, 新的图像处理与分析方法层出不穷^[4]。然而在自然环境下, 柑橘多个果实相互遮挡、重叠, 以及果实部分受枝叶遮挡的现象非常严重, 给目标果实的检测识别与空间定位带来了很大的困难^[5]。图像分割技术在农业生产方面的应用吸引了众多学者的关注^[6]。刘振宇等^[7]提出了一种凝聚层次聚类算法用于检测被遮挡果实, 将圆心点集作为改进型凝聚层次聚类分类器的输入, 并对输出结果进行圆拟合实现遮挡果实的识别, 对未被遮挡的果实识别率高达 93.53%, 但对被大面积遮挡的果实识别效果较差, 仅为 82.71%。李立君等^[8]基于改进的凸壳理论对被遮挡油茶果进行定位, 对被遮挡油茶果的正确识别率达到了 93.21%, 平均分割误差和平均重合度分别为 5.53% 和 93.43%。李硕等^[9]提出了基于 YCbCr 颜色空间模型和 Otsu 阈值分割技术, 应用边缘检测法和凸包算法提取被遮挡果实的轮廓特征来定位目标果实, 果实平均检出率达到 90.48%。吴德刚等^[10]提出了一种利用 K-means 聚类方法分离被遮挡的苹果目标与背景, 采用 Canny 算法和凸壳理论相融合的方法获得苹果被遮挡部分真实轮廓, 运用三点定圆理论实现了被遮挡苹果整体目标的边缘定位, 平均重叠系数为 88.921%。通过对上述文献的分析可以看出, 现有算法多对被遮挡果实的二维图像进行识别, 虽具有较高的准确度, 但未结合果实的三维空间信息, 无法获得被遮挡果实的空间坐标, 难以对果实采摘作业的空间标定提供实质性的帮助。

因此, 为解决柑橘果实受枝叶与果实遮挡严重, 难以对目标果实进行准确空间定位的问题, 本文提出了一种使用双目立体相机三维稠密深度点云并基于 YOLOv8-SAM 算法对目标柑橘进行识别与空间定位的方法, 同时运用边缘检测法提取被遮挡果实的轮廓特征, 对其进行椭圆拟合补完, 测算果实直径并推断果实形心空间位置, 实现对被遮挡柑橘果实的准确分割与精确空间定位, 为柑橘采摘的空间标定提供有效帮助。

1 材料与方法

1.1 数据集构建

本次试验使用的图像数据为果园实地拍摄所得, 拍摄地点为西南大学柑桔研究所, 拍摄设备使用的是 Canon 60D 和手机, 图片保存格式为 JPG, 分辨率为 $4\,032\times3\,072$ 。图 1 为部分采集到的图像数据。

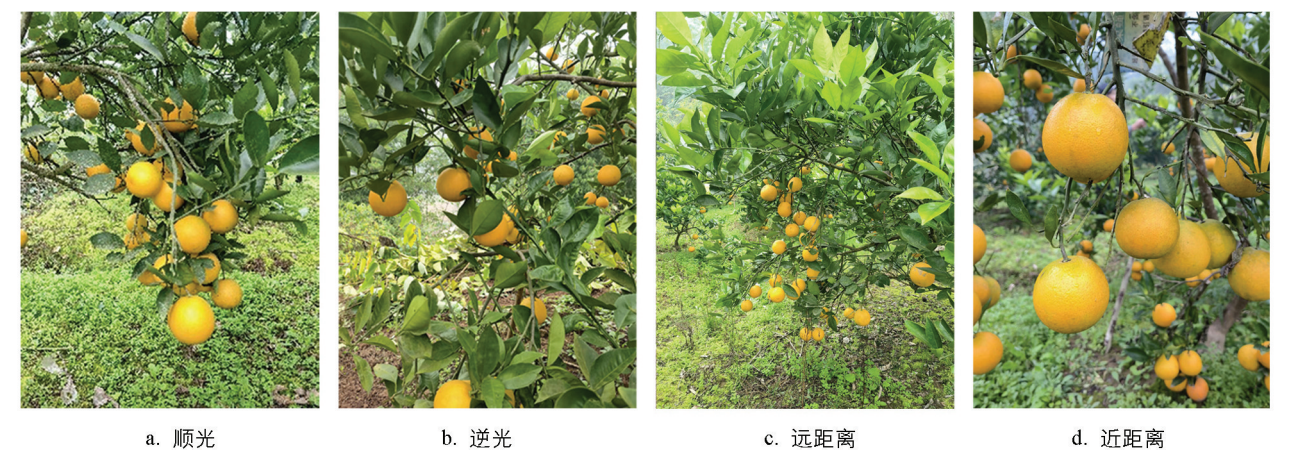


图 1 柑橘图像数据

为了对真实采摘环境中的目标柑橘果实进行检测识别, 采集了 1 796 张不同角度、光照、距离、大小的待采摘柑橘图片, 使用 Labellmg 工具对原始图像进行标注, 并随机划分为训练集 1 200 张、验证集 300 张、测试集 296 张。同时, 为了提高模型的鲁棒性和泛化能力, 防止过拟合, 在模型训练前对数据集进行数据增强。增强方式包括翻转、等比例缩放、随机平移、模糊、调节亮度等。数据增强后得到 4 382 张待采摘柑橘图片, 其中训练集 2 921 张、验证集 730 张、测试集 731 张。

1.2 试验环境与训练策略

本文训练使用的硬件环境为 CPU Intel(R) Xeon(R) Gold 6242R CPU @ 3.10 GHz, GPU Tesla V100-PCIE 32 GB $\times$ 2, 内存 126 GB, 软件环境为 Ubuntu 20.04.1 操作系统, Python 3.8, PyTorch 1.10, Torchvision 0.11, CUDA 11.2, cuDNN 8.2.0。

识别试验的计算机硬件环境为 CPU Intel(R) Core(TM) i7-10875H @ 2.30 GHz, GPU GeForce RTX 2060, 内存 8 GB $\times$ 2, 软件环境为 Win10 操作系统, Python 3.8, PyTorch 1.10, Torchvision 0.11, CUDA 11.3, cuDNN 8.2.0。

经多次调整参数测试后, 最终确定训练时选用的模型参数如表 1 所示。

表 1 模型相关参数

参数	数值	参数	数值
图像尺寸	640 $\times$ 640	最大学习率	0.01
优化器	Adam	动量	0.9
Batch Size	32	Epoch	200

1.3 评价指标

识别的精度是对柑橘果实目标检测的重要评价指标, 因此本文选择准确率 P (Precision)、召回率 R (Recall)、平均精度 A_p (Average Precision)、平均精度均值 m_{AP} (Mean Average Precision) 作为目标检测模型的评价指标。其中 P 可以反映模型的查准率, R 可以反映模型的查全率, A_p 可以反映单个类别的平均精度, m_{AP} 可以反映所有类别的平均精度均值^[11]。具体计算方法如下:

$$P = \frac{T_p}{(T_p + F_p)} \times 100\% \tag{1}$$

$$R = \frac{T_p}{(T_p + F_N)} \times 100\% \quad (2)$$

$$A_p = \int_0^1 P \cdot (R) dR \quad (3)$$

$$m_{AP} = \frac{1}{N} \sum_{i=1}^N A_{P_i} \quad (4)$$

其中: T_p 表示模型预测结果为正样本, 实际结果也为正样本, 即正确识别出目标果实的数量; F_p 表示模型预测结果为正样本, 实际结果为负样本, 即误报的负样本数量; F_N 表示模型预测结果为负样本, 实际结果为正样本, 即错误识别的样本数量。

文中算法的主要功能是为被遮挡的柑橘果实进行更准确的识别与空间定位, 得到目标果实的形心空间坐标与果实直径, 为柑橘采摘机器人末端执行器的运动目标点提供参考。

利用目标果实拟合轮廓与人工标注果实轮廓区域内部像素的面积重合度作为算法拟合的评价指标^[12], 计算公式为:

$$C = \frac{|Q \cap Q_i|}{Q} \times 100\% \quad (5)$$

式中: C 为重合度; Q 为人工标注果实轮廓区域内像素个数; Q_i 为拟合轮廓内像素个数。

在实际试验过程中, 很难准确获取目标果实的实际形心空间坐标, 不适合以形心空间位置误差来评价算法的定位精度。因此, 仅将被遮挡柑橘果实的推断空间位置与其未被遮挡状态下推断的空间位置进行对比, 来评价算法的定位效果^[13]。

算法定位误差的计算公式为:

$$\Delta = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \quad (6)$$

式中: Δ 为算法定位误差; x_i, y_i, z_i 为未被遮挡状态下的柑橘果实推断三维空间坐标; x_j, y_j, z_j 为同一柑橘果实被遮挡状态下的算法推断三维空间坐标。

果径误差的计算公式为:

$$E = \frac{|D - D_i|}{D} \times 100\% \quad (7)$$

式中: E 为果径推断误差; D 为目标果实真实直径; D_i 为推断果实直径。

2 算法

2.1 改进的 YOLO 目标检测与分割算法

由于传统的图像处理方法对柑橘果实识别与定位的误差较大, 且难以分离相互重叠的柑橘果实, 本文提出一种结合改进的 YOLOv8 与 SAM 算法的目标检测与分割模型, 较传统方法提高了识别准确率, 具有较强的鲁棒性, 可适应果园中的复杂环境变化。

YOLOv8 是目前较前沿的一阶段目标检测算法, 能在保证较高检测精度的同时保持较快的检测速度^[14]。本文将 YOLOv8 网络作为目标检测模型来构建空间定位系统, 图 2 为 YOLOv8 网络模型的结构示意图。

YOLOv8 网络的检测模型主要分为 4 个模块: 输入模块、骨干网络模块、检测头模块和输出模块。输入模块负责把待检测的柑橘图片输入到网络模型中, 通过 Mosaic 数据增强、自适应锚框计算与自适应图像缩放等方法, 对输入的柑橘图像进行预处理, 并将原始图像统一缩放到同一标准尺寸; 骨干网络模块负责提取图片特征, 其中 C2f 模块在保证模型轻量化的同时能获得更加丰富的梯度流信息; 检测头模块采用 FPN(Feature Pyramid Networks)自顶向下和 PAN(Pixel Aggregation Network)自底向上的双塔结构传递

特征, 能进一步提高模型特征提取能力; 输出模块把提取的特征经过非极大抑制得到图片中检测目标的置信度与坐标^[15]。

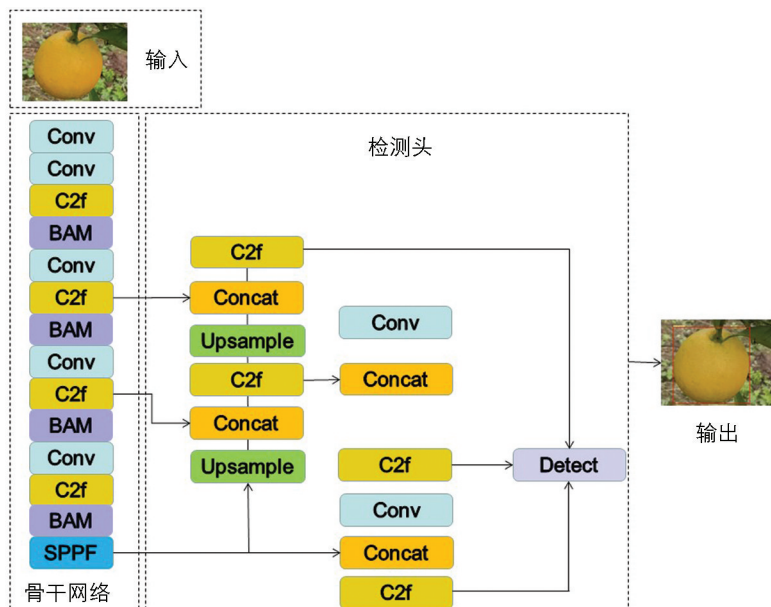


图 2 YOLOv8 网络模型结构

由于柑橘果实相互重叠极易被识别为同一果实, 且 YOLOv8 网络模型中的骨干网络模块隐含层容易忽视被遮挡过大的果实特征信息, 而 BAM 注意力机制模块能对这部分特征信息进行注意力重构, 形成一种分层的注意力机制, 可以有效地抑制背景特征, 使模型更加聚焦于前景特征, 从而增强高级语义, 实现更高的性能和更准确的识别效果^[16]。BAM 模块结构如图 3 所示。

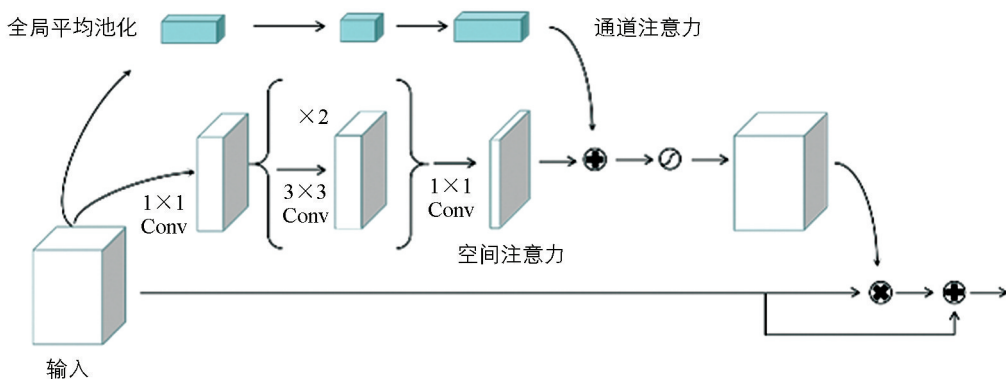


图 3 BAM 模块结构图

在骨干网络模块中增加 BAM 注意力机制模块, 置于 C2f 模块之后, BAM 模块会通过 2 个分离的路径 Channel 和 Spatial, 得到 1 个 Attention Map, 可以很好地和 YOLOv8 目标检测算法进行融合, 使提取到的高层特征更加丰富, 更加适用于柑橘果实的检测。

通过引入混合注意力模块, 共同结合通道注意力和空间注意力, 可以有效检测目标轮廓特征, 从而得到更好的对重叠果实和被遮挡果实的检测效果。其计算方法为:

$$\mathbf{M}(\mathbf{F}) = \sigma(\mathbf{M}_c(\mathbf{F}) + \mathbf{M}_s(\mathbf{F})) \quad (8)$$

式中: σ 为 Sigmoid 函数; $\mathbf{M}_c(\mathbf{F})$ 为通道注意力, $\mathbf{M}_c(\mathbf{F}) \in \mathbf{R}^C$; $\mathbf{M}_s(\mathbf{F})$ 为空间注意力, $\mathbf{M}_s(\mathbf{F}) \in \mathbf{R}^{H \times W}$ 。

$$\mathbf{M}_c(\mathbf{F}) = \text{BN}(\mathbf{W}_1(\mathbf{W}_0 \text{AvgPool}(\mathbf{F}) + \mathbf{b}_0) + \mathbf{b}_1) \quad (9)$$

式中: BN 为 Batch Normalization, 即批归一化; $\mathbf{W}_0 \in \mathbf{R}^{C/r \times C}$; $\mathbf{b}_0 \in \mathbf{R}^{C/r}$; $\mathbf{W}_1 \in \mathbf{R}^{C \times C/r}$; $\mathbf{b}_1 \in \mathbf{R}^C$ 。

$$\mathbf{M}_s(\mathbf{F}) = \text{BN}(f_3^{1 \times 1}(f_2^{2 \times 2}(f_1^{3 \times 3}(f_0^{1 \times 1}(\mathbf{F})))))) \quad (10)$$

式中: Z_c 为相机坐标系下的深度值; $\mathbf{K}$ 为相机的内参矩阵; $\mathbf{R}$ 为外参旋转矩阵; $\mathbf{t}$ 为平移矩阵。

对双目立体相机进行标定, 得到标定结果的相机内部参数矩阵 $\mathbf{K}$ 为:

$$\mathbf{K} = \begin{bmatrix} 1\ 060.680\ 1 & 0 & 952.720\ 0 \\ 0 & 1\ 060.040\ 0 & 524.030\ 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (12)$$

则彩色镜头焦距为 $(f_x, f_y) = (1\ 060.680\ 1, 1\ 060.040\ 0)$, 主光点为 $(c_x, c_y) = (952.720\ 0, 524.030\ 0)$ 。

2.3 常见的空间定位方法

利用改进的 YOLOv8 模型可检测出柑橘果实所在的图像区域, 得到目标果实所在图像区域的长方形框坐标 $(x_{\min}, y_{\min})$ 和 $(x_{\max}, y_{\max})$ 。可推断得到果实中心点坐标为 $((x_{\min} + x_{\max})/2, (y_{\min} + y_{\max})/2)$, 代入双目立体相机算法中分别计算目标果实左右边距离, 即 $((x_{\min} + x_{\max})/2, y_{\min})$ 与 $((x_{\min} + x_{\max})/2, y_{\max})$ 的三维坐标, 从而得到目标柑橘果实的直径特征^[21]。同时, 将目标果实的中心点坐标代入深度图像中, 得到该点的深度值信息, 加上目标果实半径即可推断出目标果实的空间三维坐标。这也是常见的用于推断未被遮挡柑橘果实空间坐标的方法^[22]。

但对于被遮挡的目标柑橘果实, 通常有 2 点原因会对目标果实空间三维坐标造成较大的影响: ① 识别框仅能框选出未被遮挡部分的目标柑橘, 无法对被遮挡部分进行预测, 因此可能对柑橘果径的测算造成较大偏差; ② 识别框的中心点空间三维坐标并不一定为目标柑橘果实的坐标, 也有可能为前置遮挡物的空间坐标, 以此来测算被遮挡柑橘果实的坐标可能会有较大偏差^[23]。

随机选取未被遮挡与被遮挡柑橘果实各 20 组, 测算其果实直径, 实际值采用游标卡尺手动测量, 测量相对误差值如图 7 所示。

可以看出, 在对未被遮挡柑橘目标进行果径识别时, 其相对误差值较小, 仅在 1% 左右; 而在对被遮挡柑橘目标进行果径识别时, 其误差值较大, 在 0.6% 到 15.6% 不等, 波动明显。这主要是受遮挡面积和遮挡部位变化的影响, 因此需要一种算法能够尽量规避遮挡面积和遮挡部位带来的误差影响。

2.4 优化后的空间定位算法

柑橘果实可以近似为一个椭圆形, 因此可将被遮挡的柑橘果实问题抽象为对被遮挡的椭圆形进行补完。当得到被遮挡果实的有效轮廓边时, 可用椭圆拟合算法求得完整的椭圆^[24]。

本文采用 Canny 检测算子对分割后的柑橘目标进行轮廓提取^[25]。

平面上任意位置的一个椭圆, 其中心坐标为 (x_0, y_0) , 半长轴为 a , 半短轴为 b , 长轴偏角为 θ , 椭圆的一般方程为:

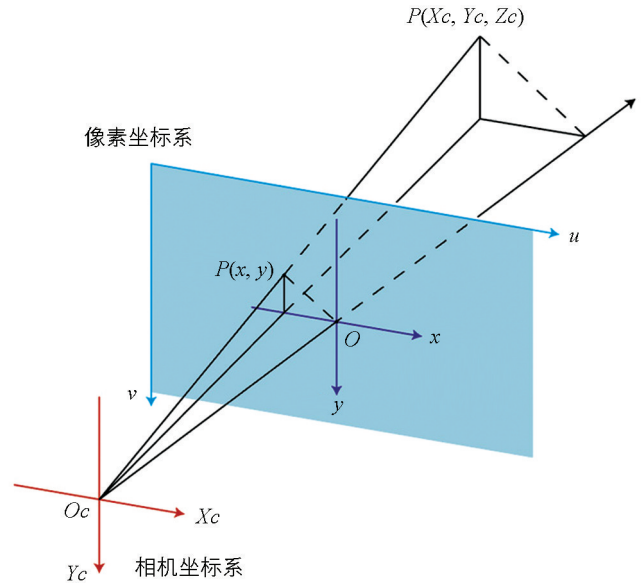


图 6 各坐标系转化示意图

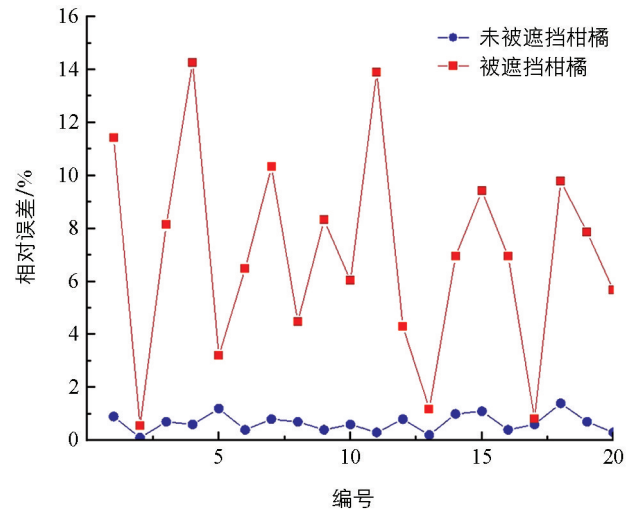


图 7 柑橘果实直径测量相对误差

$$x^2 + Axy + By^2 + Cx + Dy + E = 0$$

(13)

$$\begin{cases} x_0 = \frac{2BC - AD}{A^2 - 4B} \\ y_0 = \frac{2D - AD}{A^2 - 4B} \\ a = \frac{\sqrt{2(ACD - BC^2 - D^2 + 4BE - A^2E)}}{\sqrt{(A^2 - 4B)(B - \sqrt{A^2 + (1 - B)^2} + 1)}} \\ b = \frac{\sqrt{2(ACD - BC^2 - D^2 + 4BE - A^2E)}}{\sqrt{(A^2 - 4B)(B + \sqrt{A^2 + (1 - B)^2} + 1)}} \\ \theta = \arctan\left(\sqrt{\frac{a^2 - b^2B}{a^2B - b^2}}\right) \end{cases}$$

(14)

其中： A 、 B 、 C 、 D 、 E 分别为椭圆参数。
现采集到多个测量点坐标 (x_i, y_i) ，所拟合的目标函数为：

$$F(A, B, C, D, E) = \sum_{i=1}^n (x_i^2A + x_iy_iB + x_iC + y_iD + E)^2$$

(15)

为了使 F 最小，需使 F 的各项偏导为 0，即：

$$\frac{\partial F}{\partial A} = \frac{\partial F}{\partial B} = \frac{\partial F}{\partial C} = \frac{\partial F}{\partial D} = \frac{\partial F}{\partial E} = 0$$

(16)

求解线性方程组即可解出拟合的椭圆参数^[26]。
当得到被遮挡柑橘果实的轮廓时，基于双目立体相机的点云图像，可以获得该轮廓上点的深度信息。经多次试验尝试，取该轮廓上深度值最大的 30 个点的像素坐标，以其定位被遮挡柑橘轮廓的有效边，采用最小二乘法进行椭圆拟合，使推断的实际柑橘果实轮廓最接近真实柑橘果实轮廓。

考虑到实际柑橘果实轮廓为类圆形，在实际定位过程中可能由于双目立体相机点云坐标的误差产生一定的偏差，在进行椭圆拟合时，重复选取 50 次，若椭圆的轴比(椭圆长轴与短轴的比值)为 0.8~1.2，且椭圆的焦距不大于或小于 50 次定位焦距平均值的 50%，则视该次拟合为一次有效的拟合，计算所有有效拟合椭圆的均值，获得最终拟合效果，从而确保柑橘果实目标定位的精度。

拟合后的效果如图 8 所示，可以看出，该算法拟合后的果实轮廓接近真实的柑橘果实轮廓。

柑橘采摘机器人在工作过程中常以目标柑橘的形心作为末端执行器目标点的参考坐标。得到目标柑橘果实推算轮廓后，结合轮廓有效边的空间坐标，可推得被遮挡果实的拟合直径，从而求出果实中心点空间坐标。



图 8 拟合效果图

3 结果与分析

3.1 改进后的网络模型效果

3.1.1 注意力对比试验

为了验证 BAM 注意力机制的有效性，本文将 BAM 注意力模块与其他热门注意力模块进行对比试

验, 将 BAM 注意力模块与未添加注意力模块、BoTNet(Bottleneck Transformers)注意力模块、SimAM (Simple Attention Module)注意力模块进行对比, 如表 2 所示。可以发现, 引入 BAM 注意力模块后的检测精度更高。

表 2 不同注意力模块对比试验

注意力模块	$P/\%$	$R/\%$	$m_{AP}/\%$
无	86.5	85.9	89.3
BoTNet	87.0	85.7	89.4
SimAM	87.4	86.2	89.6
BAM	88.4	87.4	91.1

注: P 、 R 、 m_{AP} 分别为准确率、召回率和平均精度均值。下同。

3.1.2 不同模型对比试验

为了验证本文改进 YOLOv8 网络的有效性, 选取目前热门的深度学习目标检测算法作为对比, 使用同一数据集分别对 Faster RCNN、YOLOv5、YOLOv8 和改进的 YOLOv8 网络进行训练和测试。如表 3 所示, 使用引入 BAM 注意力模块后的改进 YOLOv8 网络模型相比其他网络模型, 具有最高的平均精度均值。

表 3 不同网络模型的检测性能对比

模块	$P/\%$	$R/\%$	$m_{AP}/\%$
Faster RCNN	83.6	82.6	84.6
YOLOv5	85.4	84.3	88.9
YOLOv8	86.5	85.9	89.3
改进 YOLOv8	88.4	87.4	91.1

3.2 检测效果对比

3.2.1 果实识别效果

选取特定环境下的柑橘果实图片, 分别使用 YOLOv8 网络(图 9a)与改进后的 YOLOv8 网络(图 9b)进行检测。

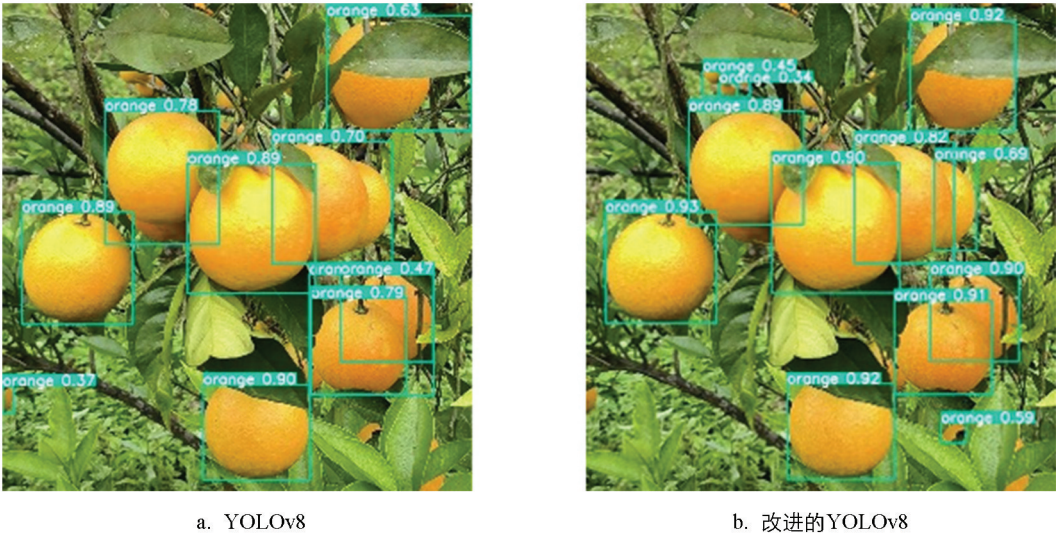


图 9 原始 YOLOv8 网络与改进 YOLOv8 网络模型效果对比

由图 9 可知, 原始的 YOLOv8 网络模型在果实重叠与被枝叶遮挡的情况下, 出现一定的漏检和识别偏

差的情况，通常出现在果实重叠面积和被枝叶遮挡面积过大时；而改进的 YOLOv8 网络模型对这种情况有更好的识别效果，漏检目标更少且平均置信度更高。综合分析可知，改进的 YOLOv8 网络模型对该环境下的柑橘果实检测具有良好的泛化性。

3.2.2 拟合效果

为了验证本文提出的改进方法的有效性和实用性，选取 120 幅被遮挡柑橘果实图像进行试验。按照目标柑橘果实被遮挡情况的区分，可将测试图像分为 3 类各 40 幅，如图 10 所示：a 类为如图 10a 所示的多个果实同时出现但未被遮挡；b 类为如图 10b 所示的果实部分区域被枝叶遮挡导致目标轮廓不完整；c 类为如图 10c 所示的果实之间相互重叠遮挡。



图 10 自然场景下柑橘果实不同遮挡情况测试图像

利用本文算法对图像进行处理，中间过程效果如图 11 所示。

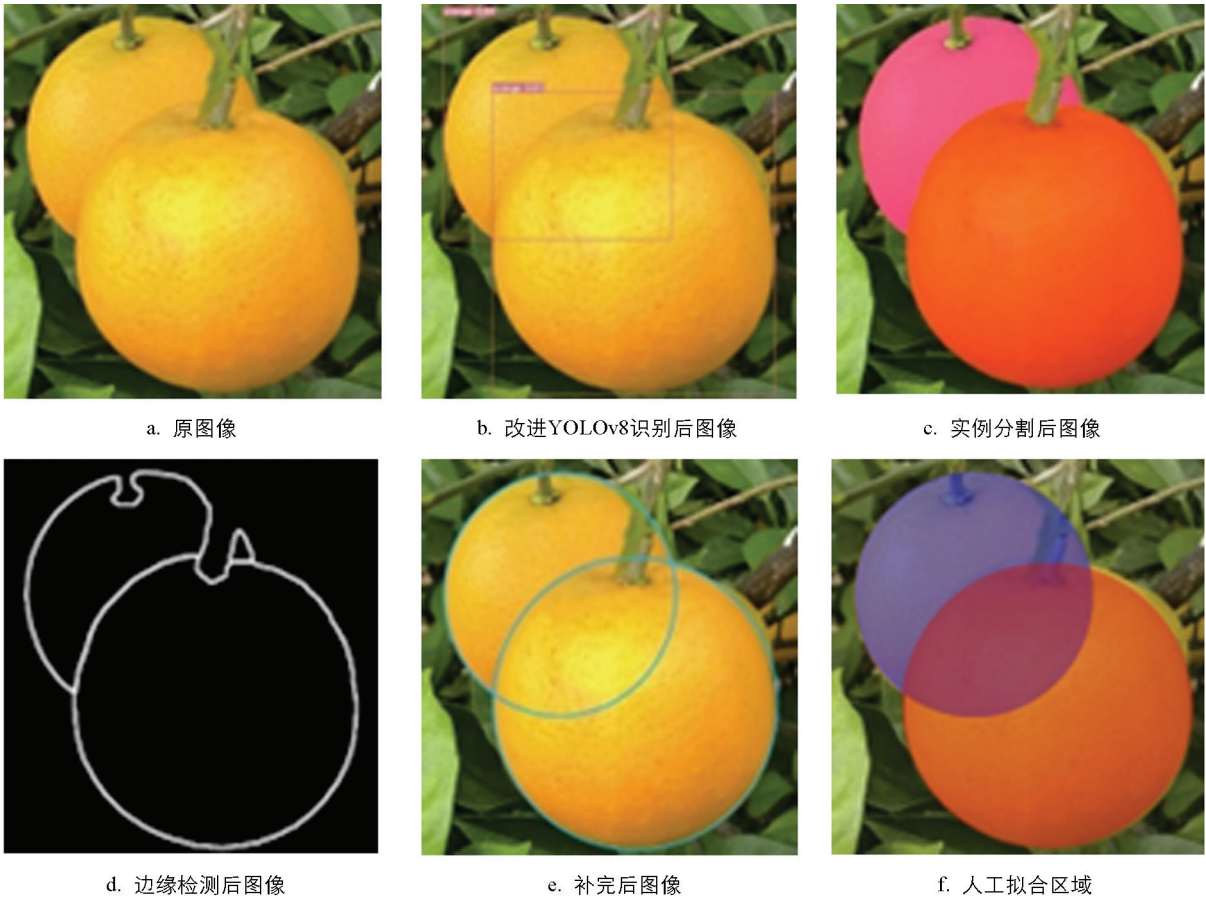


图 11 自然场景下被遮挡柑橘果实定位过程及结果

分别用凸包算法^[23]与本文改进算法对目标果实与实际果实像素面积重合度进行对比验证, 对于 a 类的果实, 果实像素重合度较高, 重合度均值分别达到 92.61% 与 93.04%, 2 种方法检测效果相近, 均能达到采摘精度要求; 对于 b、c 类果实, 凸包算法对于果实轮廓缺失严重的情况难以做到有效识别, 重合度均值为 63.85%, 且波动范围较大, 而本文算法得到的重合度均值为 86.73%, 且波动范围较小。其中, 当两果实重叠程度超过 78% 时, 算法容易将其识别为同一果实, 导致识别失败。

3.2.3 定位检测效果

分别用未改进的算法(如 2.3 节所述)与本文改进算法计算 3 类目标柑橘果实形心空间坐标位置与果实直径, 得到如表 4 所示的定位误差与果径误差对比数据。

表 4 算法定位误差与果径误差比较

类型	未改进算法		改进算法	
	定位误差/mm	果径误差/%	定位误差/mm	果径误差/%
a 类	1.35	0.97	1.31	0.94
b 类	16.49	7.75	2.36	1.01
c 类	21.63	10.48	3.31	1.25
平均	13.16	6.40	2.33	1.07

由表 4 可知, 本文算法对未被遮挡的分离果实(a 类)的定位与果径识别没有较大的提升, 但在被枝叶遮挡与果实重叠的情况下, 本文算法有着明显的提升效果, 平均定位误差减小了 16.22 mm, 平均果径误差降低了 7.99%, 证明了本文算法能够较为准确地进行被遮挡柑橘果实的定位检测与果径识别。

为直观验证改进后的 YOLOv8-SAM 算法的检测效果, 对同一果园环境进行对比测试。图 12 显示了在果园环境中算法改进前后运行对比的效果。



图 12 果园环境中算法改进前后运行结果

在图 12 所示的环境中, 共有 12 颗目标柑橘果实, 改进前的算法检测到 9 颗柑橘果实, 漏检 3 颗, 且对于大面积重叠果实与枝叶遮挡的情况有明显的形心位置判断偏差; 而改进后的算法, 仅漏检 1 颗, 相比改进前, 漏检数目更少且对目标果实的形心位置和果径判断更为准确。综合分析可知, 改进后的 YOLOv8-SAM 算法对被遮挡柑橘的目标识别和定位效果更好, 更适用于果园采摘环境。

4 结论

针对柑橘果实相互重叠与被枝叶遮挡导致的对机器视觉定位目标柑橘空间位置造成误差的问题, 本文

提出了一种基于 YOLOv8-SAM 算法对被遮挡柑橘的位置形状进行识别,并运用边缘检测法结合双目立体相机三维稠密深度点云得到被遮挡柑橘的有效轮廓边,使用最小二乘法拟合出被遮挡柑橘的完整轮廓以确定目标柑橘更精确的空间坐标位置。主要结论如下:

1) 提出一种改进的 YOLOv8 网络结构,在骨干网络模块中加入 BAM 注意力机制模块,使其对重叠果实与被枝叶遮挡果实有更好的识别效果。改进网络在果园环境下的柑橘果实检测中达到了 88.4% 的准确率、87.4% 的召回率以及 91.1% 的平均精度均值,相较于原始的 YOLOv8 网络模型,改进的 YOLOv8 网络在果实重叠与被枝叶遮挡环境下的识别准确率提高了 1.9 个百分点,平均精度均值提升了 1.8 个百分点。

2) 提出一种对被遮挡柑橘果实轮廓的重建方法,基于图像分割与深度图像数据对被遮挡柑橘果实有效轮廓边进行补完,并获得果实形心空间坐标,利用该方法获得的重建轮廓与实际果实轮廓重合度高。对未被遮挡或被轻度遮挡的果实,重合度均值达到 93.04%;对被枝叶遮挡或果实重叠的情况,重合度均值达到 86.73%。

3) 使用本文算法模型对目标柑橘果实的空间位置与果径进行识别时,对形心位置定位的平均误差减少了 16.22 mm,平均果径误差降低了 7.99%,证明了本文提出的改进算法能够较为准确地识别被遮挡柑橘果实目标并对其进行定位与果径测量,验证了本文算法用于自动化柑橘采摘的可行性,为柑橘果实及相似水果在自然环境下的识别和定位提供了参考,为水果采摘机器人的研制提供了支撑。

参考文献:

- [1] 潘一凡,刘永忠,黄钰轩,等.我国柑橘栽培管理技术现状调查和发展思考[J].华中农业大学学报,2023,42(4): 140-149.
- [2] 吕佳,李帅军,曾梦瑶,等.基于半监督 SPM-YOLOv5 的套袋柑橘检测算法[J].农业工程学报,2022,38(18): 204-211.
- [3] 黄彤斌,黄河清,李震,等.基于 YOLOv5 改进模型的柑橘果实识别方法[J].华中农业大学学报,2022,41(4): 170-177.
- [4] 李欣,王玉德.基于颜色模型和阈值分割的有遮挡的柑橘果实识别算法[J].计算技术与自动化,2022,41(2): 136-140.
- [5] 宋怀波,尚钰莹,何东健.果实目标深度学习识别技术研究进展[J].农业机械学报,2023,54(1): 1-19.
- [6] 刘好,刘洒,杨长辉,等.基于双目立体视觉的重叠柑橘空间定位[J].中国农业科技导报,2020,22(9): 104-112.
- [7] 刘振宇,丁宇祺.自然环境中被遮挡果实的识别方法研究[J].计算机应用研究,2020,37(S2): 333-335,339.
- [8] 李立君,阳涵疆.基于改进凸壳理论的遮挡油茶果定位检测算法[J].农业机械学报,2016,47(12): 285-292,346.
- [9] 李颀,强华.基于双目视觉与深度学习的番茄本体特征检测系统[J].南方农业学报,2020,51(1): 237-244.
- [10] 吴德刚,赵利平,陈乾辉.遮挡苹果目标的三点定位算法研究[J].黄河科技学院学报,2022,24(8): 40-45.
- [11] 刘忠意,魏登峰,李萌,等.基于改进 YOLOv5 的橙子果实识别方法[J].江苏农业科学,2023,51(19): 173-181.
- [12] 周宏平,金寿祥,周磊,等.基于多模态图像的自然环境下油茶果识别[J].农业工程学报,2023,39(10): 175-182.
- [13] 林耀海,吕钟亮,杨长才,等.自然场景图像中的重叠蜜柚识别及试验[J].农业工程学报,2021,37(24): 158-167.
- [14] 王小荣,许燕,周建平,等.基于改进 YOLOv7 的复杂环境下红花采摘识别[J].农业工程学报,2023,39(6): 169-176.
- [15] VARGHESE R, SAMBATH M. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Ro-

bustness [C] //2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), April 18-19, 2024, Chennai, India. IEEE, 2024: 1-6.

[16] QI J T, LIU X N, LIU K, et al. An Improved YOLOv5 Model Based on Visual Attention Mechanism: Application to Recognition of Tomato Virus Disease [J]. Computers and Electronics in Agriculture, 2022, 194: 106780.

[17] 张俊宁, 毕泽洋, 闫英, 等. 基于注意力机制与改进 YOLO 的温室番茄快速识别 [J]. 农业机械学报, 2023, 54(5): 236-243.

[18] KIRILLOV A, MINTUN E, RAVI N, et al. Segment Anything [C] //2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 3992-4003.

[19] SILWAL A, DAVIDSON J R, KARKEE M, et al. Design, Integration, and Field Evaluation of a Robotic Apple Harvester [J]. Journal of Field Robotics, 2017, 34(6): 1140-1159.

[20] LIN G C, TANG Y C, ZOU X J, et al. Color-, Depth-, and Shape-Based 3D Fruit Detection [J]. Precision Agriculture, 2020, 21(1): 1-17.

[21] WU G, LI B, ZHU Q B, et al. Using Color and 3D Geometry Features to Segment Fruit Point Cloud and Improve Fruit Recognition Accuracy [J]. Computers and Electronics in Agriculture, 2020, 174: 105475.

[22] 陈志健, 伍德林, 刘路, 等. 复杂背景下油茶果采收机重叠果实定位方法研究 [J]. 安徽农业大学学报, 2021, 48(5): 842-848.

[23] 刘德儿, 朱磊, 冀炜臻, 等. 基于 RGB-D 相机的脐橙实时识别定位与分级方法 [J]. 农业工程学报, 2022, 38(14): 154-165.

[24] WANG Y W, CHEN Y F, WANG D F. Recognition of Multi-Modal Fusion Images with Irregular Interference [J]. PeerJ Computer Science, 2022, 8: e1018.

[25] 张勤, 庞月生, 李彬. 基于实例分割的番茄串视觉定位与采摘姿态估算方法 [J]. 农业机械学报, 2023, 54(10): 205-215.

[26] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C] //2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. IEEE, 2014: 580-587.

责任编辑 柳剑

崔玉洁