doi:10.6041/j.issn.1000-1298.2024.03.024

基于改进 YOLO v5 的复杂环境下花椒簇识别与定位方法

黄　华　张　昊　胡晓林　聂兴毅
（兰州理工大学机电工程学院，兰州 730050）

摘要：花椒树产果量大，枝干纵横交错，树叶茂密，给花椒的自动化采摘带来了困难。因此，本文设计一种基于改进 YOLO v5 的复杂环境下花椒簇的快速识别与定位方法。通过在主干提取网络 CSPDarknet 的 CSPLayer 层和 Neck 的上采样之后增加高效通道注意力 ECA（Efficient channel attention）来简化 CSPLayer 层的计算量，提升了特征提取能力。同时在下采样层增加协同注意力机制 CA（Coordinate attention），减少下采样过程中信息的损失，强化特征空间信息，配合热力图（Grad－CAM）和点云深度图，来完成花椒簇的空间定位。测试结果表明，与原 YOLO v5 相比较，改进的网络将残差计算减少至 1 次，保证了模型轻量化，提升了效率。同帧数区间下，改进后的网络精度为 96.27%，对比 3 个同类特征提取网络 YOLO v5、YOLO v5－tiny、Faster R－CNN，改进后网络精确度 P 分别提升 5.37、3.35、15.37 个百分点，连株花椒簇的分离识别能力也有较大提升。实验结果表明，自然环境下系统平均识别率为 81.60%、漏检率为 18.39%，能够满足花椒簇识别要求，为移动端部署创造了条件。

关键词：花椒簇识别；农业机器人；深度学习；目标检测；YOLO v5
中图分类号：TP391　　**文献标识码**：A　　**文章编号**：1000-1298(2024)03-0243-09　　**OSID**：

Recognition and Localization Method for Pepper Clusters in Complex Environments Based on Improved YOLO v5

HUANG Hua　ZHANG Hao　HU Xiaolin　NIE Xingyi
(*School of Mechanical and Electrical Engineering*, *Lanzhou University of Technology*, *Lanzhou* 730050, *China*)

Abstract: Pepper trees yield is a substantial quantity of fruits, characterized by crisscrossed branches and dense foliage, resulting insignificant challenges for automated peppercorn picking. Therefore, a fast identification and localization method of pepper clusters in complex environment based on improved YOLO v5 was proposed. By adding efficient channel attention (ECA) after the CSPLayer of the backbone extraction network CSPDarknet and the upsampling layer of Neck to simplify the computation of the CSPLayer layer and improve the feature extraction capability. In the downsampling layer, coordinate attention (CA) was added to reduce the loss of information in the downsampling process, strengthen the spatial information of features, and cooperate with the heat map (Grad－CAM) and the depth map of the point cloud to complete the spatial localization of pepper clusters. The test results showed that the improved network over the original YOLO v5 reduced the residual computation to 1 time, which ensured the model was lightweight and the efficiency was improved. Under the same frame number interval, the accuracy of the improved network was 96.27%, comparing with three similar feature extraction networks YOLO v5, YOLO v5－tiny, and Faster R－CNN, the precision of the improved network was improved by 5.37 percentage points, 3.35 percentage points, and 15.37 percentage points, respectively, and the ability of separating and recognizing the pepper clusters of the successive plants was greatly improved. The experimental results showed that the average checking accuracy of the system in the natural environment was 81.60%, and the leakage rate was 18.39%, which can satisfy the pepper cluster recognition, and build the foundation for mobile deployment.

Key words: pepper cluster identification; agricultural robots; deep learning; target detection; YOLO v5

收稿日期：2023－08－09　修回日期：2023－09－13
基金项目：现代丝路寒旱农业发展资金项目（njyf2022－10）和国家自然科学基金项目（51965037）
作者简介：黄华（1978—），男，教授，博士生导师，主要从事智能农机装备和数字孪生等研究，E-mail：hh318872@126.com

0 引言

甘肃省作为中国最大的花椒种植省份,年产量超过 2.8×10^5 t[1-2],是本省重要的经济作物之一。但该地区的花椒种植区域多在山区,地形多样,并且花椒树枝干带刺,枝叶交错,给自动化采摘带来了极大的不便,因此花椒簇的识别与定位便成了采摘机器人要解决的首要问题。

在农作物识别领域,很多学者提出了基于机器视觉的检测方案。SENGUPTA 等[3]结合 Canny 边缘检测和 Hough 变换,实现了柑橘的复杂检测。HE 等[4]通过直方图和三元阈值分割法,定位主干枝条确定了采摘点。但传统的机器视觉检测受环境影响大,光照、遮挡均会影响识别效果,因此机器学习算法被用来提升识别的准确率。MAURYA 等[5]通过采用 SVM 分类器实现了蘑菇的分类。WANG 等[6]利用 3 种模态识别方法将农作物分为 3 类进行识别,精确度超过 98%。但机器学习算法仍有局限,对复杂环境下的特征识别能力较差,因此需要引入深度学习算法提升鲁棒性[7]。在各类深度学习网络中,YOLO 是最典型的整幅图像检测网络之一,相比 Faster R-CNN 的区域分类法,有更高的效率[8]。王卓等[9]通过改进 YOLO v4 网络,轻量化结构,大幅提升了网络计算能力。YOLO v5 网络相比 YOLO v4 有更优化的结构,并且响应速度更快[10-12]。ZHU 等[12]在 YOLO v5 网络中增加了 CBAM 模块,提升了模型对目标小尺寸特征的检测能力。

上述方法虽然在模型结构、精度、速度上都有所改进,并且对于农作物的识别也给出了完整的可行性方案[13-14],但现阶段对于花椒簇的识别大多停留在机器视觉层面,与深度学习网络结合较少,仅通过图形算法应用在实际环境中,无法实现在机部署[15-16]。花椒果实相对于其他农作物果实较小,进行单体识别意义不大,如何在不同情况下精准识别花椒簇则是采摘机器人的关键。本文基于 YOLO v5 网络,针对花椒特殊的生长结构,在 Backbone 和 Neck 部分引入 ECA 和 CA 注意力机制,通过两种注意力机制的配合,有效简化残差边 CSPLayer 的结构,增强网络的特征提取能力,强化上采样特征的空间信息,结合 Gard-CAM 类梯度热力图和点云图[17-18],精确定位花椒簇三维位置。

1 材料与方法

1.1 花椒簇数据集

本文主要从复杂、多遮挡场景下花椒簇的定位与识别问题出发,以相对密集饱满的花椒簇为对象,如图 1 所示。



图 1 成熟期花椒簇

Fig. 1 Mature pepper cluster

本文使用 Inter 公司生产的 Real Sense d435i 型相机采集花椒成熟期内的多组图像(分辨率为 640 像素×640 像素)。试验田位于甘肃省陇南武都区花椒种植基地,该基地种植面积约为 133 hm^2,分别种植了知名品种“大红袍 1 号、3 号、特级”,选取其中长势最好的 1 hm^2 进行图像采集。上位机配合 Real Sense d435i 型相机为主要设备,自动模式调整焦距与曝光。考虑本算法后期的在机部署,图像拍摄过程中模拟实际的机械臂采摘视角,将相机布置于自动升降三脚架上,对高度为 1.5~2.0 m 的花椒簇沿竖直方向自下而上进行拍摄,相机水平无倾角。

花椒簇的识别难点主要在于花椒与传统果实不同,间隙小,颗粒生长方向不确定,一枝多簇。若无法对同一枝干上的“连株花椒簇”实现分离识别,就会导致花椒簇采摘点信息缺失,影响机械臂采摘,因此将数据集进行以下分类:两簇花椒根部生长位置相近且花椒簇相互重叠,则为“连株”,之后根据花椒簇前方是否有树叶遮挡将“连株”分为“连株无遮挡”、“连株有遮挡”;反之分类为“单株无遮挡”、“单株有遮挡”。每个类别 200 幅图像,原始图像共计 800 幅,图 2 为不同类别花椒簇图像示例。



(a) 单株有遮挡 (b) 单株无遮挡

(c) 连株有遮挡 (d) 连株无遮挡

图 2 花椒簇数据集种类

Fig. 2 Data-set types of pepper clusters

1.2 数据集增强和划分

本实验采用 4 种方法对原始数据集进行增强,分别为随机 90°转动、高斯噪声添加、随机镜像、饱

和度色相调整，以此来提高网络的识别准确率和实际使用中的鲁棒性，本文将 800 个图像编号设为随机数进行了 10 000 次触发，变换后共 10 000 幅，其中方法 1 共计 3 346 幅，方法 2 共计 1 605 幅，方法 3 共计 2 399 幅，方法 4 共计 2 650 幅。

图像增强后使用 LabelImg 仅标注“pepper”一个目标，构成“单目标”。并将 4 个不同类别的数据集整理为 VOC 格式，生成 XML 格式的声明文件，其中训练集、测试集、验证集比例为 7∶2∶1，格式和分类与 COCO 数据集相同，以此来实现“单类别目标”。

1.3　YOLO 模型改进

YOLO v5 网络作为经典的单目标检测方法之一，通过不断地采样预测物体的特征来实现目标检测[17-19]。考虑到本实验仅针对花椒这一目标进行识别和研究，为了提高本网络对于花椒识别的鲁棒性，加强对目标特征的识别，针对 YOLO v5 的主干网络 Backbone、颈部 Neck、头部 YOLO－Head 进行改进，引入注意力机制，如图 3 所示。

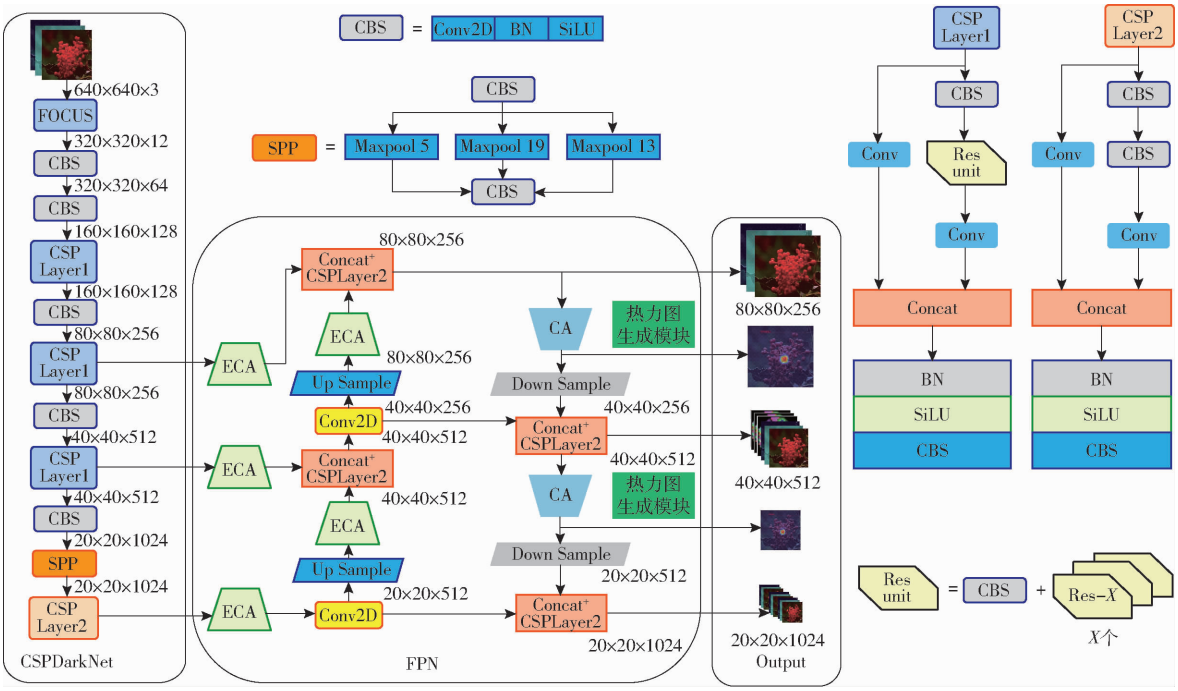


图 3　改进的 YOLO v5 结构图

Fig. 3　Improved YOLO v5 structure diagram

1.3.1　注意力机制模块

花椒簇为红色，其生长环境主体背景为绿色，虽然对比度较高，但是果实尺寸较小，且树叶茂密，环境复杂，为了更好地提升网络识别的准确率并简化网络，在 YOLO v5 网络中引入 2 个注意力机制，分别为高效通道注意力机制（Efficient channel attention，ECA）和协同注意力机制（Coordinate attention，CA），具体结构如图 4 所示。

与传统的 CBAM（Convolutional block attention module）模块通过特征通道和特征空间进行特征强化不同，ECA 在只增加少量参数的情况下提升了网络的特征提取能力。该模块通过全局平均池化 GPA（Global average pooling）计算通道维度的统计结果，其中通道统计值由自适应卷积内核 k_φ 和激活函数 $\sigma \in [0,1]$ 决定，卷积后的输出图与原图相乘，以此来达到强化高贡献值通道，弱化低贡献值通道的目的。其中 k_φ 计算公式为

$$k_\varphi = \left| \frac{\mathrm{lb}C_{\mathrm{in}} + 1}{2} \right|_{\mathrm{odd}} \tag{1}$$

式中　C_{in}——输入通道数

$|\cdot|_{\mathrm{odd}}$——最接近的奇数

CA 模块引入空间注意力机制的概念，与仅进行全局最大池化或者平均池化的模块不同，能够最大限度减少空间信息丢失，针对较密集的预测物有很好的效果。该模块将特征图从宽度和高度两个方向进行全局平均池化，将输入特征图（$C \times H \times W$）分为 Z_c^w（宽）和 Z_c^h（高）两部分，其中特征层 $Z_c^w(C \times H \times 1)$、$Z_c^h(C \times 1 \times W)$ 的计算公式为

$$\begin{cases} Z_c^w = \dfrac{1}{W} \displaystyle\sum_{0 \leqslant i < W} X_c(H,i) \\ Z_c^h = \dfrac{1}{H} \displaystyle\sum_{0 \leqslant j < H} X_c(j,W) \end{cases} \tag{2}$$

式中　H、W——高、宽方向像素数

i、j——当前像素数

c——图通道数目

为了强化空间特征，在得到 Z_c^w 和 Z_c^h 后；对 Z_c^h 的 x、y 方向进行均值池化，再进行卷积、标准化、激活函数计算，将结果进行堆叠后得到特征层 $1(C/r \times$

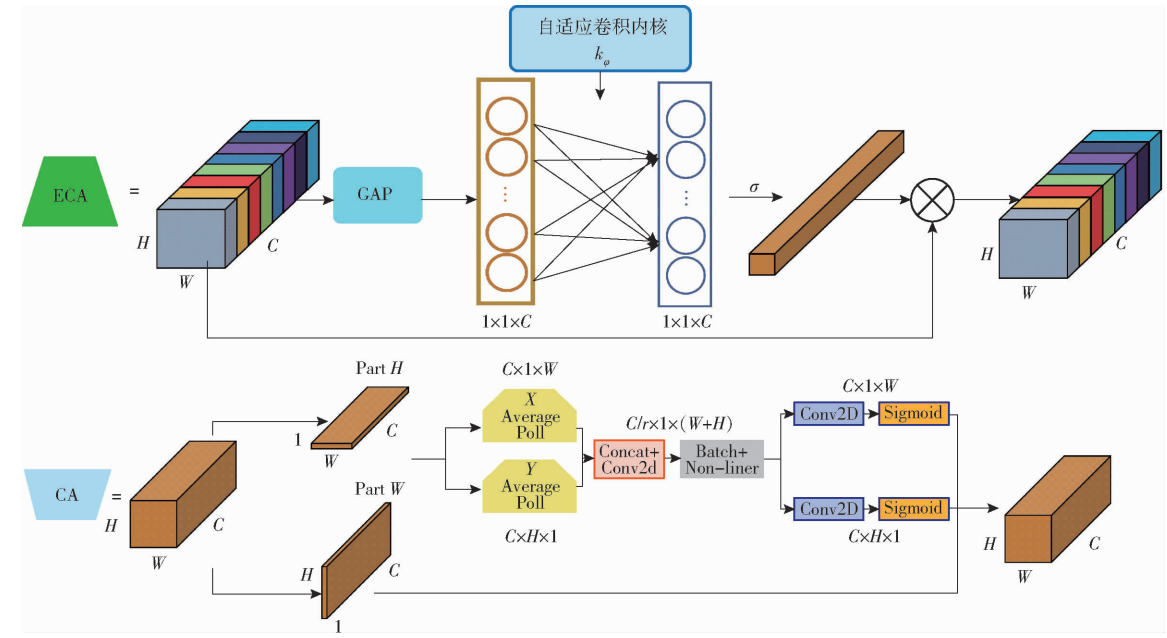


图 4 注意力机制模块

Fig. 4 Attention mechanism module

$1\times(W+H)$),又将特征层 1 分离强化特征层,得到增强的 $Z_c^h(C\times1\times W)$进行 1×1 卷积计算调整通道数,与 $Z_c^w(C\times H\times1)$两者相乘,最后输出强化后的特征。

1.3.2 网络结构改进

参照图 3 主干提取网络,在未进行池化前能够提取丰富的细节特征,但随着卷积的计算和通道数的增加,感受野逐渐变小,抽象和降噪能力减弱,导致图像的特征提取能力减弱。为解决此问题,在 Backbone 中的 CSPLayer1 后加入 ECA 模块强化特征信息,使得 CSPLayer1 中残差单元的计算减少至 1 次。SPP(Spatial pyramid pooling)的引入能解决特征图尺寸变小时几何失真问题,简化了 CSPLayer2 的结构,将输入特征分为 2 部分,通过在后方加入 ECA 模块的方式,提升了小尺寸特征图的特征表现能力,减少了卷积计算次数,如图 3 所示。

原结构根据特征金字塔网络(Feature pyramid networks,FPN),在颈部(Neck)的 FPN 部分设计了两次上采样,采用内插值的方式扩大图像分辨率;为实现更好特征融合能力,在上采样后增加 ECA 模块,以此来融合不同感受野的关键信息,减少下采样造成的特征信息损失,并在残差模块后增加 CA 模块,强化图像空间位置信息,配合热力图,更精确地显示预测框的位置。

2 模型训练与分析

2.1 环境搭建

实验硬件平台为移动端计算机,处理器型号为 AMD R9-6000hx,8 核、16 线程,内存 16 GB,显卡型号为 Nvidia Geforce RTX 2060Ti,操作系统为 Ubuntu 18.04,深度学习框架为 Pytorch 1.2、Cudnn 7.6.6、Cuda 11,编程语言为 Python 3.7。

2.2 模型训练与评价指标

2.2.1 评价指标

为了更好地评估模型的准确率和泛化性,在此引入单类召回率 R、单类精确度 P、每类平均精度均值(mAP)、调和平均值 F1 值作为评价指标。

2.2.2 网络训练

将改进后的网络在 GPU 上进行训练,训练方法采用梯度下降法,模型为上文制作的 VOC 数据集,图像尺寸均为 640 像素×640 像素。训练次数为 120 次,标签平滑度为 5×10^{-3},训练学习率下降方式采用 step,最大学习率为 1×10^{-3},批量大小为 8。对主干网络进行冻结再解冻,通过这种方式提升模型训练效率,冻结部分迭代周期(Epoch)为 25[19-20]。

2.2.3 结果分析

为验证网络的改进效果,分别与 3 个主流网络 YOLO v5、YOLO v5-tiny、Faster R-CNN 在同一平台上进行训练并分析,为体现模型泛化性能,使用上文所述的验证集和测试集进行分析,图像共计 3 000 幅。

由表 1 可知,改进的 YOLO v5 网络最终召回率为 96.07%、精确度为 96.27%、F1 值为 95.95%。精确度相较于同类模型 YOLO v5-tiny、YOLO v5 分别提升 5.37、3.35 个百分点,召回率 R 分别提升 3.78、3.71 个百分点,F1 值提升 6.60、3.37 个百分点,相较于其他分类模型 Faster R-CNN 精确度、召

回率、F1 值分别提升 15.37、9.20、10.61 个百分点。改进的 YOLO v5 网络每类平均精度均值为 98.55%，相较于同类模型 YOLO v5 - tiny、YOLO v5 精度提升 3.68、3.31 个百分点，相较其他分类模型 Faster R - CNN 提升 9.58 个百分点。

表 1　花椒簇识别结果比较

Tab.1　Comparison of pepper cluster recognition results

识别模型	P/%	R/%	F1 值/%	mAP/%	检测速度/(f·s^{-1})
YOLO v5 - tiny	90.90	92.20	89.35	94.87	50.97
改进的 YOLO v5	96.27	96.07	95.95	98.55	40.87
YOLO v5	92.92	92.36	92.22	95.24	40.56
Faster R - CNN	80.90	86.87	85.34	88.97	55.32

出现以上结果的原因如下：

(1)改进后的模型在原基础上增加了 ECA 与 CA 注意力机制模块，并在 Neck 层增加了热力图输出模块，比原模型复杂，但由于残差计算次数的减少，体积小于 YOLO v5，因此精确度、召回率、检测速度、平均精度均值均有所提升。

(2)相较于 YOLO v5 - tiny，改进了 CSPLayer 层结构。整体网络在与 YOLO v5 - tiny 大小相近的情况下，显著增强了网络的精度，但由于上下采样过程中步骤的增加，帧数略低于 YOLO v5 - tiny。

(3)Faster R - CNN 与 YOLO 网络在结构和算法上有较大不同。采用 VGG16 网络作为 Backbone 使得 Faster R - CNN 的整体识别速度较快，但 Darknet53 相比 VGG 有更强的特征提取能力[9]。对比两者的综合能力，改进后的 YOLO v5，识别速度虽略慢于 Faster R - CNN，但整体性能有大幅提升，更符合移动端部署要求。

为对比“单类别目标”数据集对网络性能的影响，在相同网络与训练参数的情况下，使用 LabelImg 对数据集内的 4 类花椒图像分别标注为“pepper1”、“pepper2”、“pepper3”、“pepper4”，训练后得到多类别目标的 $P-R$ 曲线如图 5b 所示，对比图 5a，单类别目标 $P-R$ 曲线相对稳定，改进的 YOLO v5 所生成的 $P-R$ 曲线基本位于其余 3 条曲线上方，且平均精度($P-R$ 曲线与坐标轴所围面积)最大。图 5b 中 4 个网络召回率均出现了明显下降，导致平均精度值减小。

出现以上结果的原因如下：多类别目标识别导致负样本增加，使得回归值最终未能“归一”，产生过拟合，增加了下采样时的特征信息，降低了网络学习效率。

综上所述，由于在机部署对图像识别精度要求更高而不是更准的图像分类能力，在数据集中进行分类处理可以强化每类特征的占比，提高网络单目标识别能力，而进行单类别目标识别可以更准确地检测目标物，相比多类别目标识别有更高的在机部署优势。

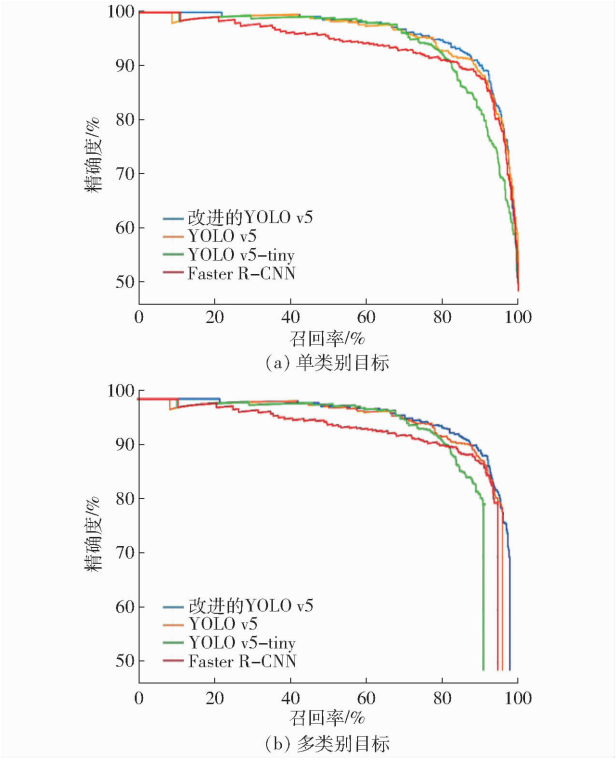


图 5　$P-R$ 曲线

Fig.5　$P-R$ curves

3　实验

3.1　可视化分析

为直观地显示改进后网络的识别能力，对识别结果进行可视化展示。将 4 个网络在测试集上进行实验对比，测试集图像共计 1 000 幅，从中选取 1 幅检测结果，作为验证图展示，该图像包涵了上文提到的 4 类情况，并且花椒簇生长密集，受树叶大面积遮挡，环境光线复杂，能够模拟采摘过程中相机所拍摄的实际情况，符合“单类别目标”的识别条件。

实验平台、训练方式、参数均相同，对比改进的 YOLO v5 与其余 3 个网络识别能力，结果如图 6 所示。

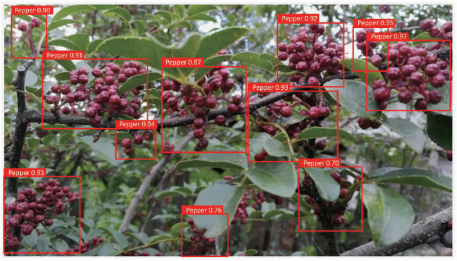
首先将验证集分为Ⅰ、Ⅱ、Ⅲ 3 个区域，如图 6a 所示，通过 4 个指标将其余网络同改进的 YOLO v5 网络进行对比。指标 1 为 3 个区域内花椒簇的有效识别数 T_p；为了方便后续花椒自动化采摘，连株花椒簇能否成功分离并识别为独立植株，是自动采摘关键，以此设计指标 2，区域Ⅰ、Ⅱ内的连株分离率 D，即两区域识别枝数与两区域总簇数的比值；指标 3，全图识别率 P_1；指标 4 为区域Ⅱ遮挡部分置信度加权平均值 A_C。



(a) 图像划分示意图



(b) YOLO v5



(c) 改进的YOLO v5



(d) YOLO v5-tiny



(e) Faster R-CNN

图 6 可视化分析验证图

Fig. 6 Visual analysis validation charts

由图 6 和表 2 可得,YOLO v5、改进的 YOLO v5、YOLO v5 - tiny、Faster R - CNN 对于复杂情况下花椒簇识别置信度区间分别为 0.77 ~ 0.90、0.89 ~ 0.95、0.64 ~ 0.86、0.58 ~ 0.73,并且在改进的 YOLO v5 网络中,置信度高于 0.9 的检测结果约占全部结果的 77.78%,数值跨度相对较小,平稳度高,均优于其余 3 类网络。

表 2 图像检测结果

Tab. 2 Detection results of image

指标	网络			
	YOLO v5	改进的 YOLO v5	YOLO v5 - tiny	Faster R - CNN
T_p	5 654	6 938	4 578	2 936
D/%	63.47	85.91	69.54	20.57
P_1/%	85.32	89.88	79.52	67.83
A_C	0.81	0.87	0.75	0.54

如图 6 所示,区域Ⅰ、Ⅱ中共有 6 株花椒簇,在增加注意力机制模块后,由于上下采样过程中对于特征的保留程度有所增加,相同感受野的情况下改进后的网络特征更加明显,分离率更高。区域Ⅰ、Ⅱ下半部分环境光线复杂,有大面积遮挡,改进的 YOLO v5 检测数为 3 个,较 YOLO v5 与 YOLO v5 - tiny 分别多 1 个和 2 个,并且能够根据枝干所在位置绘制预测框,精简的网络结构保证了较高的检测速度,进一步提升了检测性能。

3.2 系统漏检及误差分析

在利用本文所设计的网络进行花椒簇的检测时,仍会出现图 7 所示的漏检和误检的情况,由图 7 可看出部分花椒簇被漏检(图 7a、7b 中红圈),部分连珠花椒簇被误检为单株(图 7c 中红圈),以及置信度较低的情况(图 7d 中红框)。原因如下:

(1)由于 YOLO 网络对于形状较为规则且聚集的对象检测效果较好,但自然环境下的花椒簇生长较为松散,如图 7b 中红圈所示,当遮挡或重叠现象较为严重时,位于前方的花椒簇会遮挡住后方花椒簇的大部分特征导致漏检。同样,当树叶或其他杂质遮挡住花椒簇的绝大部分特征时,如图 7a 所示,也会造成漏检的情况。



(a) 遮挡导致漏检 (b) 生长松散造成漏检

(c) 特征相似造成误检 (d) 光照影响导致置信度低

图 7 漏检误检示例

Fig. 7 Example diagrams of missed and false detection

(2)受光照影响的花椒簇图像如图 7d 所示,当光照不均匀或处于树叶阴影中时,会极大地影响花椒簇的对比度,导致特征识别和提取产生困难,因其与环境背景较为相似,网络在识别时无法分离特征,

但通过较为相似的外形，仍成功地定位了花椒簇的位置，整体置信度较低。

（3）当花椒生长时，会出现两株较为茂盛的花椒簇重叠的情况（图 7c），由于本网络在下采样过程中是基于花椒簇上花椒粒的形状进行的特征强化，根部特征不明显，此时模型在处理此情况时，则无法完成“连株”花椒簇的分离，造成误检。

3.3　花椒簇定位方法

3.3.1　基于 Grad－CAM 的花椒质点系定位

先验框的选取位置和预测框的绘制位置均由网络训练后得出，预测过程不透明对结果的解释有限。多数研究者通过在预测框内任意取点的方式来获得目标物的深度信息[21－22]。但考虑到花椒簇与普通果实不同，单株产量高、中心空洞大、果实位置乱，如果仅通过上述方式，很容易获得空洞，得到错误的深度信息，难以确定花椒簇的位置，无法实现精确的采摘动作。

在 CA 注意力模块的强化下，计算每幅下采样完成后特征图的平均梯度来获得一组权重 l_k^c，计算公式为

$$l_k^c = \frac{1}{Z}\sum_i\sum_j \frac{\partial y^c}{\partial A_{ij}^k} \tag{3}$$

式中　l_k^c——第 k 个特征图对类别 c 的权重

Z——特征图数量

y^c——对应类别 c 的得分

A_{ij}^k——第 k 个特征图中坐标为(i,j)的像素

加权求和后即可得出 Grad－CAM 类梯度热力图。热力图能够直观显示区域关注度高低，特征越明显，对应位置先验框绘制越密集，则热度越高。取热力值为 1 的部分作为花椒簇的等效质点系，能有效避免由于树叶遮挡造成的质点系混乱，如图 8 所示。

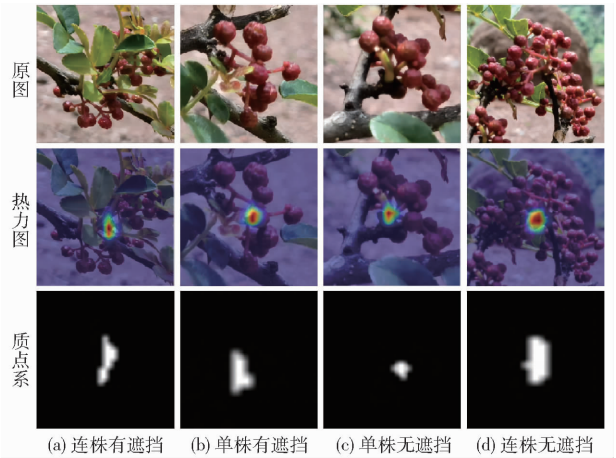


图 8　各种情况下的等效质点系

Fig. 8　Equivalent particle systems in various situations

3.3.2　花椒簇采摘点二维定位

在得到的等效坐标系内求出质心坐标(x_i,y_i)，取质点系内像素和为质量 m_i，$\boldsymbol{T}_c$ 为质点到质心的矢径，由平面质点系质心计算方法可得等效质点系的质心，将其视为花椒簇采摘点的二维坐标(x,y)[16]。

$$\boldsymbol{T}_c = \frac{\sum_{i=1}^{N} m_i\boldsymbol{r}_i}{\sum_{i=1}^{N} m_i} = \frac{\sum_{i=1}^{N} m_i\boldsymbol{r}_i}{M} \tag{4}$$

$$\begin{cases} x = \dfrac{\sum_i m_i x_i}{M} \\ y = \dfrac{\sum_i m_i y_i}{M} \end{cases} \tag{5}$$

式中　M——所有质点的质量和

$\boldsymbol{r}_i$——质点到质心的矢径

N——质点个数

3.3.3　花椒簇采摘点三维定位

在得到花椒簇的二维采摘点坐标后，需借助相机获取深度信息，依靠 Real Sense d435i 型相机的一对红外立体传感器（Stereo IR Pair），产生左右 2 幅灰度图像，配合红外激光发射器（IR Projector）对所采集的 RGB 图像生成点云图像，在得到点云信息后对二维采摘点坐标(x,y)进行测距获得深度信息，测距原理与双目相机相同，利用左右图像的视差求解距离[23]。得到深度信息后，即可确定花椒簇采摘点的三维坐标(x,y,z)，具体流程如图 9 所示。同时对平面点云图测距能够有效避免花椒簇结构空洞造成的深度信息不准确。

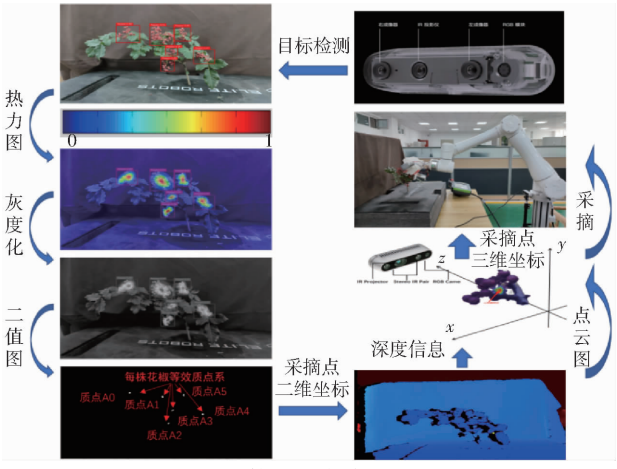


图 9　花椒簇三维定位方法流程图

Fig. 9　Diagram of 3D positioning method for pepper clusters

3.3.4　识别与定位实验验证

为验证网络的实际应用效果，本文在实验室内搭建花椒树模型，进行自动化固定位置模拟识别与定位，如图 10 所示。采摘机器人由 ELIBOT 六轴协作机器人、Real Sense d435i 型相机、自主设计采集器、上位机组成。参与实验的对象共 5 组，每组 20

枝,每个枝干上含 5 ~ 7 簇花椒,理论采摘点应与花椒簇数目相同,实验记录每组花椒簇总数目 N、识别花椒簇数目 X、末端执行器采摘数目 S、单枝平均总耗时。



图 10 采摘点空间距离示意图

Fig. 10 Diagram of spatial distance of picking point

记录每组采集器几何中心点 A 与花椒簇采摘点 B 的空间直线距离 D,来验证定位精度,根据程序设计当点 A 运动至花椒簇采摘点的三维坐标点 $B(x,y,z)$时,通过关节 4 的转动来模拟采摘动作,所以点 A 与采摘点 $B(x,y,z)$的空间直线距离 D 将直接决定能否顺利采摘。在实际测量时,将点 B 近似为花椒簇形态学中心,通过游标卡尺测量两点距离,得到实际定位距离 D。

模拟采摘结果如表 3 所示,面对复杂的情况,整体系统依然能保证较高的检测水准。5 组平均采摘率为 89.19%,配合自主设计的采集器以固定角度执行采摘,平均单株收获时长为 14.9 s。当 D 小于 20 mm 时刀头均能采集,所以三维定位和采集器设计均符合设计要求。第 1 组受光照和树叶遮挡较小,漏检率最低,识别率最高,第 3 组由于花椒簇生长较为集中,形态饱满,采集过程较为顺利,采摘率最高。5 组实验平均识别率为 81.60%,平均漏检率为 18.39%。

出现以上结果的原因如下:

表 3 定位结果与实验验证结果

Tab. 3 Positioning and experience results

组号	花椒簇总数目	识别数	采摘数	识别率/%	采摘率/%	漏检率/%	平均距离/mm
1	124	107	95	86.37	88.78	13.63	19.52
2	136	113	97	83.05	85.85	16.95	17.44
3	147	121	112	82.34	92.56	17.66	15.79
4	139	105	93	75.51	88.57	24.49	18.35
5	114	92	83	80.76	90.21	19.24	16.77

(1)识别率在实际实验中有所下降,是由于计算方法与实验条件所致,识别率的计算过程中未统计假正样本,原因是实验为视频流的检测,与验证集图像不同,视频帧数高,无法逐帧暂停,并且视角处于动态变化中,假正样本数量显示不明显,无法统计。

(2)机械臂与相机采用“眼在手外”的安装方式,相机距离检测物体较远,并且机械臂会对位置较高的目标物产生遮挡,导致漏检。

4 结论

(1)针对 YOLO v5 原网络,在 Backbone、Neck 部分引入注意力机制模块 ECA 和 CA,增强了上采样中网络的特征提取能力和泛化性,将残差单元的计算次数简化至一次,强化了下采样中的空间位置信息,配合热力图,更清晰地显示网络预测框的绘制流程和花椒簇结构信息。经对比,该网络在各方面均优于其他网络,能够为自动化采摘系统提供理论支持。

(2)花椒簇与常见的果实不同,为了实现连株花椒簇的分离识别定位,根据热力图提取了花椒簇的等效质点,处理后得到花椒簇的二维点,配合相机得出的点云图,精确定位了花椒簇的三维位置。

(3)提出的改进 YOLO v5 网络,经实际实验验证,对花椒簇这一特征进行了强化识别,对连株花椒实现了分离识别,能满足花椒簇的定位与识别。但在不同光照条件下或大面积遮挡情况下,由于相机部署位置较远,仍然存在部分漏检和误检的情况,后续将继续改进方案,强化检测效果。

参 考 文 献

[1] 齐锐丽. 基于机器视觉花椒目标识别与定位技术研究[D]. 汉中:陕西理工大学, 2020.

[2] 王丹丹,宋怀波,何东健. 苹果采摘机器人视觉系统研究进展[J]. 农业工程学报,2017,33(10):59 - 69.
WANG Dandan, SONG Huaibo, HE Dongjian. Research advance on vision system of apple picking robot[J]. Transactions of the CSAE, 2017, 33(10): 59 - 69. (in Chinese)

[3] SENGUPTA S, LEE W S. Identification and determination of the number of immature green citrus fruit in a canopy under different ambient light conditions[J]. Biosystems Engineering, 2018, 117(1): 51 - 61.

[4] HE Z L, XIONG J T, LIN R, et al. A method of green litchi recognition in natural environment based on improved LDA

classifier[J]. Computers & Electronics in Agriculture, 2017, 140(3): 159 – 167.

[5] MAURYA P, SINCH N P. Mushroom classification using feature-based machine learning approach[C]// Proceedings of 3rd International Conference on Computer Vision and Image Processing,2020: 197 – 206.

[6] WANG Y, DU J, ZHANG H, et al. Mushroom toxicity recognition based on multi-grained cascade forest[J]. Scientific Programming, 2020, 8(3): 8849011.

[7] 熊俊涛, 刘振, 汤林越,等. 自然环境下绿色柑橘视觉检测技术研究[J]. 农业机械学报, 2018, 49(4):45 – 52.
XIONG Juntao, LIU Zhen, TANG Linyue, et al. Visual detection technology of green citrus under natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(4): 45 – 52. (in Chinese)

[8] FU L, FENG Y, WU J, et al. Fast and accurate detection of kiwifruit in orchard using improved YOLOv3-tiny model[J] Precision Agriculture, 2021, 22(3):754 – 776.

[9] 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)

[10] 刘成良,贡亮,苑进,等.农业机器人关键技术研究现状与发展趋势[J].农业机械学报,2022,53(7):1 – 22,55.
LIU Chengliang,GONG Liang,YUAN Jin, et al. Current status and development trends of agricultural robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 1 – 22,55. (in Chinese)

[11] 高新阳, 魏晟, 温志庆,等. 改进 YOLO v5 轻量级网络的柑橘检测方法[J].计算机工程与应用, 2023, 59(11): 212 – 221.
GAO Xinyang, WEI Sheng, WEN Zhiqing, et al. Citrus detection method based on improved YOLO v5 lightweight network [J]. Computer Engineering and Applications, 2023, 59(11): 212 – 221. (in Chinese)

[12] ZHU X, LYU S, WANG X, et al. TPH – YOLOv5: improved YOLO v5 based on transformer prediction head for object detection on crone-captured scenarios[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE, 2021: 2778 – 2788.

[13] 杨萍,郭志成. 花椒采摘机器人视觉识别与定位求解[J]. 河北农业大学学报, 2020, 43(3): 121 – 129.
YANG Ping,GUO Zhicheng. Vision recognition and location solution of *Zanthoxylum bungeanum* picking robot[J]. Journal of Hebei Agricultural University, 2020, 43(3): 121 – 129. (in Chinese)

[14] 余小东,杨孟辑,张海清,等. 基于迁移学习的农作物病虫害检测方法研究与应用[J]. 农业机械学报, 2020, 51(10): 252 – 258.
YU Xiaodong, YANG Mengji, ZHANG Haiqing, et al. Research and application of crop diseases detection method based on transfer learning[J]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(10):252 – 258. (in Chinese)

[15] WANG Qilong, WU Banggu, ZHU Pengfei,et al. ECA – Net: efficient channel attention for deep convolutional neural networks [C]// IEEE 2020 IEEE/CVF Conferenceon Computer Vision and Pattern, 2020: 11531 – 11539.

[16] 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. (in Chinese)

[17] 赵迪, 叶盛波, 周斌. 基于 Grad – CAM 的探地雷达公路地下目标检测算法[J]. 电子测量技术, 2020, 43(10):113 – 118.
ZHAO Di, YE Shengbo, ZHOU Bin. Ground penetrating radar anomaly detection based on convolution Grad – CAM[J]. Electronic Measurement Technology, 2020, 43(10): 113 – 118. (in Chinese)

[18] 许景辉,邵明烨,王一琛,等. 基于迁移学习的卷积神经网络玉米病害图像识别[J]. 农业机械学报,2020,51(2):230 – 236,253.
XU Jinghui,SHAO Mingye,WANG Yichen, et al. Recognition of corn leaf spot and rust based on transfer learning with convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 230 – 236, 253. (in Chinese)

[19] 龙满生,欧阳春娟,刘欢,等. 基于卷积神经网络与迁移学习的油茶病害图像识别[J]. 农业工程学报,2018,34(18):194 – 201.
LONG Mansheng, OUYANG Chunjuan, LIU Huan, et al. Image recognition of *Camellia oleifera* diseases based on convolutional neural network & transfer learning[J]. Transactions of the CSAE, 2018, 34(18):194 – 201. (in Chinese)

[20] WANG H,ZHU Y,GREEN B,et al. Axial-deep lab: stand-alone axial-attention for panoptic segmentation[C]// European Conference on Computer Vision, 2020: 108 – 126.

[21] HOU Q, ZHOU D,FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2021:13713 – 13722.

[22] SHRIVASTAVA V K,PRADHANM K. Rice plant disease classification using color features: a machine leaning paradigm[J]. Journal of Plant Pathology, 2021, 103(1): 17 – 26.

[23] 项荣,应义斌,蒋焕煜,等. 基于双目立体视觉的番茄定位[J]. 农业工程学报,2012,28(5):161 – 167.
XIANG Rong, YING Yibin, JIANG Huanyu, et al. Localization of tomatoes based on binocular stereo vision[J]. Transactions of the CSAE, 2012, 28(5): 161 – 167. (in Chinese)