

基于 YOLO v7 - ECA 模型的苹果幼果检测

宋怀波^{1,2} 马宝玲^{1,2} 尚钰莹^{1,2} 温毓晨^{1,2} 张姝瑾^{1,2}

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100;

2. 农业农村部农业物联网重点实验室, 陕西杨凌 712100)

摘要: 为实现自然环境下苹果幼果的快速准确检测,针对幼果期苹果果色与叶片颜色高度相似、体积微小、分布密集、识别难度大的问题,提出了一种融合高效通道注意力(Efficient channel attention, ECA)机制的改进 YOLO v7 模型(YOLO v7 - ECA)。在模型的3条重参数化路径中插入 ECA 机制,可在不降低通道维数的前提下实现相邻通道局部跨通道交互,有效强调苹果幼果重要信息、抑制冗余无用特征,提高模型效率。采集自然环境下苹果幼果图像 2 557 幅作为训练样本、547 幅作为验证样本、550 幅作为测试样本,输入模型进行训练测试。结果表明,YOLO v7 - ECA 网络模型准确率为 97.2%、召回率为 93.6%、平均精度均值(Mean average precision, mAP)为 98.2%、F1 值为 95.37%。与 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型相比,其 mAP 分别提高 15.5、4.6、1.6、1.8、3.0、1.8 个百分点,准确率分别提高 49.7、0.9、18.5、1.2、0.9、1.0 个百分点,F1 值分别提高 33.53、2.81、9.16、1.26、2.38、1.43 个百分点,召回率相较于 Faster R - CNN、SSD、YOLO v5、YOLO v6、YOLO v7 网络模型分别提高 5.0、4.5、1.3、3.7、1.8 个百分点;单幅图像检测时间为 28.9 ms,可实现苹果幼果的高效检测。针对幼果目标模糊、存在阴影和严重遮挡的情况,本研究采用 550 幅测试图像进行模型鲁棒性检验。在加噪模糊情况下,YOLO v7 - ECA 的 mAP 为 91.1%,F1 值为 89.8%,与 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型相比其 mAP 分别提高 26.3、21.0、5.4、8.0、11.5、8.9 个百分点,F1 值分别提高 27.19、7.08、8.50、4.20、3.94、4.67 个百分点;在阴影情况下,YOLO v7 - ECA 的 mAP 为 97.5%,F1 值为 95.36%,与 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型相比其 mAP 分别提高 14.8、8.8、2.1、2.4、5.4、2.5 个百分点,F1 值分别提高 21.51、2.60、10.49、1.53、3.23、2.56 个百分点;在严重遮挡情况下,YOLO v7 - ECA 的 mAP 为 98.6%,F1 值为 94.8%,与 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型相比其 mAP 分别提高 21.7、13.7、2.3、2.4、4.8、2.2 个百分点,F1 值分别提高 28.29、3.50、6.45、0.96、1.36、1.36 个百分点。该网络模型可在保证网络模型精度的同时拥有较快的检测速度,且对场景模糊、阴影和严重遮挡等影响具有较好的鲁棒性。该研究可为幼果实时检测系统提供有效借鉴。

关键词: 苹果幼果; 检测; YOLO v7; 高效通道注意力机制; 机器视觉

中图分类号: TP391.4

文献标识码: A

文章编号: 1000-1298(2023)06-0233-10

OSID:



Detection of Young Apple Fruits Based on YOLO v7 - ECA Model

SONG Huaibo^{1,2} MA Baoling^{1,2} SHANG Yuying^{1,2} WEN Yuchen^{1,2} ZHANG Shujin^{1,2}

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

2. Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China)

Abstract: In order to detect young apple fruits quickly and accurately in the natural environment, an improved YOLO v7 model (YOLO v7 - ECA) was proposed to solve the problems of high similarity, small size, dense distribution and difficult identification between young apple fruits and leaves. By inserting the ECA mechanism into the three reparameterized paths of the model, the local cross-channel interaction of adjacent channels could be carried out without reducing the channel dimension, which can effectively emphasize the important information of young apple fruits, suppress redundant and useless features, and improve the efficiency of the model. Totally 2 557 images of young apple fruits were collected as training samples, totally 547 images as validation samples, and 550 images as test samples in

收稿日期: 2022 - 10 - 29 修回日期: 2022 - 12 - 09

基金项目: 国家重点研发计划项目(2019YFD1002401)和国家自然科学基金项目(31701326)

作者简介: 宋怀波(1980—),男,教授,博士生导师,主要从事图像处理理论研究,E-mail: songyangfeifei@163.com

the natural environment, and input them into the model for training and testing. The YOLO v7 – ECA model was trained to have a precision of 97.2%, a recall rate of 93.6%, an mAP of 98.2%, and F1 value of 95.37%. Compared with the Faster R – CNN, SSD, Scaled – YOLO v4, YOLO v5, YOLO v6, YOLO v7 models, its mAP was increased by 15.5, 4.6, 1.6, 1.8, 3.0 and 1.8 percentage points, its precision was increased by 49.7, 0.9, 18.5, 1.2, 0.9 and 1.0 percentage points, its F1 value was increased by 33.53, 2.81, 9.16, 1.26, 2.38 and 1.43 percentage points, and its recall rate was increased by 5.0, 4.5, 1.3, 3.7 and 1.8 percentage points for Faster R – CNN, SSD, YOLO v5, YOLO v6 and YOLO v7 models, respectively; the image detection time was 28.9 ms, which could realize efficient detection of young apple fruits. Aiming at the fuzzy, shadowing and severe occlusion of young fruit targets, totally 550 test images were used to test the robustness of the model. The mAP of YOLO v7 – ECA was 91.1% and the F1 value was 89.8% under the condition of adding noise and fuzziness. Compared with the Faster R – CNN, SSD, Scaled – YOLO v4, YOLO v5, YOLO v6 and YOLO v7 models, its mAP was increased by 26.3, 21.0, 5.4, 8.0, 11.5 and 8.9 percentage points, and its F1 value was increased by 27.19, 7.08, 8.50, 4.20, 3.94 and 4.67 percentage points, respectively. The mAP of YOLO v7 – ECA was 97.5% and the F1 value was 95.36% in the shadow. Compared with the Faster R – CNN, SSD, Scaled – YOLO v4, YOLO v5, YOLO v6 and YOLO v7 models, its mAP was increased by 14.8, 8.8, 2.1, 2.4, 5.4 and 2.5 percentage points, and its F1 value was increased by 21.51, 2.60, 10.49, 1.53, 3.23 and 2.56 percentage points, respectively. The mAP of YOLO v7 – ECA was 98.6% and the F1 value was 94.8% under severe occlusion. Compared with that of the Faster R – CNN, SSD, Scaled – YOLO v4, YOLO v5, YOLO v6 and YOLO v7 models, its mAP was increased by 21.7, 13.7, 2.3, 2.4, 4.8 and 2.2 percentage points, and its F1 value was increased by 28.29, 3.50, 6.45, 0.96, 1.36 and 1.36 percentage points, respectively. Experiments showed that the proposed model was of high accuracy and speed, it was also robust to different interference situations such as blurred scene, shadow and severe occlusion. The research result can provide an effective reference for the detection system of apple young fruit.

Key words: young apple fruit; detection; YOLO v7; efficient channel attention mechanisms; machine vision

0 引言

疏果前期苹果幼果目标识别对自动化疏果、农药用量控制、智能化表型监测以及果树新品种选育具有重要意义。幼果期苹果果色与叶片颜色相似、体积微小、分布密集,受光照影响,果实表面会出现高光、阴影等现象,果实叶片枝条相互重叠遮挡,导致果实识别难度增大^[1-3]。深度卷积神经网络相较于传统目标检测算法能够对高层次特征进行学习,为此类目标的有效检测提供了必要借鉴^[4-7]。

国内外众多研究已将深度卷积网络作为果实目标识别的主流算法,并在果实高效检测方面取得较好结果^[8-11]。为快速精准实现冬枣果实的识别,刘天真等^[12]在 YOLO v3 网络模型中融入注意力机制,可有效提高特征图的表现能力,相较于原始网络模型,各项指标有明显提高。龙燕等^[13]提出了一种改进的 FCOS 网络模型对自然环境下不同遮挡和密集程度的苹果目标进行识别,改进后的网络模型检测准确率提升 3.2 个百分点,mAP 提升 3.1 个百分点,试验表明,改进后的 FCOS 网络模型在苹果检测试验中具有更高的精度和鲁棒性能。何斌等^[14]提出了一种改进的 YOLO v5 网络模型进行夜间番茄

果实识别,夜间环境下,该网络模型对绿色番茄果实和红色番茄果实平均识别精度分别为 96.2% 和 97.6%,能够有效实现暗光下番茄果实的识别。为实现复杂环境下不同生长阶段果实的高效检测,赵辉等^[15]提出了一种改进的 YOLO v3 网络模型对不同成熟度的苹果果实进行识别,网络模型在测试集上的 mAP 为 96.3%,F1 值为 91.8%,在原有网络模型的基础上两项指标各自提高 3.8%。王立舒等^[16]在 YOLO v4 – Tiny 中融入卷积注意力模块进行不同通道间特征相关性的学习,有效降低了复杂环境对蓝莓目标的干扰。试验结果表明,该网络模型在蓝莓果实严重遮挡与逆光场景中,mAP 和准确率均较优,可同时满足识别精度与检测速度的需求。为实现复杂环境下苹果的高效准确检测,王卓等^[17]提出了一种轻量化 YOLO v4 – CA 网络模型,该网络模型平均检测精度为 92.23%,嵌入式平台上的检测速度为 15.11 f/s,与 SSD 和 Faster R – CNN 网络对比,其精度和速度均有明显提升。

针对疏果前期果实重叠遮挡严重且与背景颜色极为相似的难题,王丹丹等^[2]提出了 R – FCN 苹果幼果检测网络,用 28 662 幅苹果幼果图像进行网络模型训练、验证和测试,最终网络模型召回率为

85.7% ,准确率为 95.1% ,误识率为 4.9% ,单幅图像平均处理时间为 0.187 s ,基本能够实现近景色小目标识别。为进一步提高苹果幼果的检测准确率,宋怀波等^[1]将挤压激发块和非局部块两种注意力机制融入 YOLO v4 网络模型进行蔬果前期微小苹果果实的识别,采用 3 000 幅苹果幼果图像进行网络模型训练、验证和测试,最终该网络模型在 600 幅测试集上的平均精度达 96.9%。为改善苹果幼果在光照不均和严重遮挡等干扰条件下造成的果实漏检问题,JIANG 等^[18]提出了一种融合非局部注意力模块和卷积块注意力模块的 YOLO v4 网络模型,采用 3 000 幅低质量幼果图像进行训练测试,测试集在高光、阴影、模糊、严重遮挡情况下的平均精度分别为 98.0%、96.2%、97.0%、96.9% ,表明该方法可实现复杂干扰下苹果幼果的高效识别。

蔬果前期幼果识别相关研究在高精度检测方面虽已取得较好的结果,但仍然存在以下亟待解决的问题:①苹果幼果检测速度较慢,实时性检测难度大。②苹果幼果检测模型较大、内存开销高、难以完成应用端部署。③苹果幼果在模糊、含噪、存在阴影以及严重遮挡情况下仍然无法实现高精度检测。本文拟借鉴 YOLO v7 网络模型在速度和精度上的双重优势,并通过融合 ECA 的方式,进一步提升网络模型对幼果特征的捕获能力,以期在不明显增加网络深度和内存开销的情况下,实现苹果幼果的高效准确检测。

1 材料与方法

1.1 数据集

1.1.1 苹果幼果图像的采集与预处理

试验数据采集于西北农林科技大学园艺学院实验站,试验对象包括秦冠、秦脆、红富士、嘎啦等 300 多个品种,栽培方式为矮砧密植,行距 3 m。数据采集时间为 2020 年 5 月中旬至 6 月上旬,此时果径在 48 mm 以下,采集设备为 iPhone 8Plus 手机,采集时间段为 08:00—20:00,共采集 3 654 幅幼果图像,格式为 JPEG,分辨率为 3 024 像素 × 3 024 像素。如图 1 所示,拍摄方式为拍摄者垂直于地面站立,相距主干 0.4 ~ 1.5 m 手持手机设备在果树行间两侧分别进行拍摄。

为保证样本的多样性,分别采集清晨、正午、傍晚不同时间段;多云转晴、晴转多云、晴不同天气情况;顺光、逆光不同光照条件;无遮挡、轻微遮挡、严重遮挡不同遮挡程度下的幼果图像。图 2a ~ 2d 分别是在顺光、逆光、顺光强光直射、逆光强光直射下拍摄的幼果图像,由 4 幅图像对比可知,顺光状态下

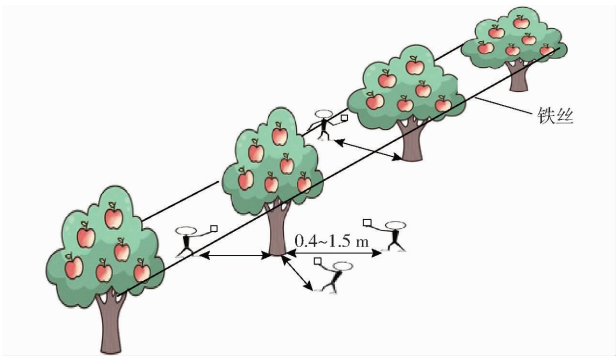


图 1 采集方式
Fig. 1 Acquisition method



图 2 不同光照和遮挡情况

Fig. 2 Different lighting and occlusion conditions

拍摄的图像,依据颜色和形状特征易于区分目标;逆光状态下拍摄的图像,整体偏暗,在幼果果色与背景颜色相似的情况下,易将幼果果实目标漏检错检;顺光强光直射图像存在过曝,造成幼果目标部分泛白;逆光强光直射图像由于图像欠曝,造成幼果目标整体偏黑,欠曝过曝情况下所成图像整体加强幼果果实目标的识别难度。图 2e ~ 2h 分别展示了幼果果实目标在自然场景下被相邻果实、叶片、枝条、铁丝等物体遮挡的现象。图 2i、2j 为不同遮挡程度下的图像,本研究设定苹果幼果表面被遮挡未超过 30% 属于轻微遮挡,被遮挡超过 30% 属于严重遮挡。

为增强算法鲁棒性与泛化性,随机选取 600 幅苹果幼果图像,随机均分为 2 组,为 2 组图像分别添加 1% 的椒盐噪声和 5% 的高斯噪声,对添加噪声后的图像各自再次随机均分为 2 组,任选其中 1 组对其进行高斯滤波,最终得到 300 幅加噪图像和 300 幅模糊图像,分别模拟算法在迁移应用时遇到的镜头不清、对焦不准以及特殊天气等情况。图 3a 分别为添加 1% 椒盐噪声以及进行高斯滤波的图像,图 3b 分别为添加 5% 高斯噪声以及进行高斯滤波的图像。

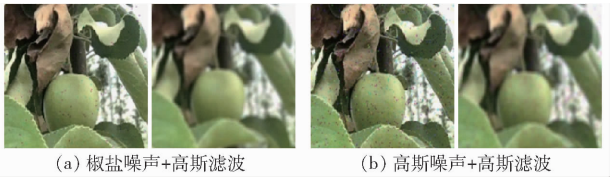


图 3 不同加噪和模糊情况

Fig.3 Different noise and blur conditions

1.1.2 数据集制备

通过加噪滤波预处理后,3 654 幅苹果幼果图像的具体信息如表 1 所示。

表 1 数据集拍摄信息
Tab.1 Dataset shooting information

影响因素	图像类别	图像数量/幅
天气情况	晴转多云	顺光 556
		逆光 644
	多云转晴	顺光 988
		逆光 1 012
	晴	顺光 190
		逆光 264
遮挡程度	未遮挡	665
	轻微遮挡	774
	严重遮挡	2 215
平滑程度	光滑	3 054
	模糊加噪	600
光照情况	亮光	1 734
	阴影	1 920

利用 LabelImg 软件进行标注,框选苹果幼果果实,标注标签为“apple”,每幅图像生成对应.xml 文件。将标注完毕的数据集按照训练集、验证集、测试集比例 7:1.5:1.5 进行划分,最终获得 2 557 幅训练集图像、547 幅验证集图像、550 幅测试集图像。如表 2 所示,将测试集图像按照所设标准进行分组,为网络模型进行分组测试做准备,检验网络模型在各种干扰情况下的有效性。

表 2 测试集图像分组信息
Tab.2 Test set image grouping information

影响因素	图像类别	图像数量/幅
遮挡程度	未遮挡	60
	轻微遮挡	60
	严重遮挡	60
平滑程度	光滑	75
	模糊加噪	75
光照情况	亮光	110
	阴影	110

1.2 方法

深度卷积网络因其丰富的特征表现能力已广泛应用于各类图像分类^[19]、目标检测^[20-23]、语义分割^[24]等视觉任务中。为避免苹果幼果产生冗余信

息造成检测效果不佳等情况,本研究分别引入挤压激发(Squeeze-and-excitation, SE)注意力^[25]、卷积块注意力(Convolutional block attention module, CBAM)^[26]、协调注意力(Coordinate attention, CA)^[27]、ECA 机制^[28]对 YOLO v7 目标检测算法进行了优化试验。通过对比发现,SE 注意力机制与 CBAM 机制依赖通道相关性进行特征校准,采用两层全连接层捕获所有通道间的相互依赖关系进行显式建模,不可避免地造成参数和计算量的增加以及维度的缩减;而 CA 机制更关注于强调位置信息捕获目标结构,针对复杂生长环境下,生长趋势不可控且与背景颜色极其相近不易区分的苹果幼果目标而言,特征强调作用较弱。因此本研究利用 ECA 机制进行 YOLO v7 网络模型的改进与优化工作。

1.2.1 ECA 机制

ECA 机制结构如图 4 所示,图中 H 为输入图像的高度, W 为输入图像的宽度, C 为特征图通道数,其为避免 SE 注意力机制维度缩减,采用全局平均池化对每个通道中的特征图进行压缩操作后,自适应选择卷积核大小为 K 的一维卷积替换两层全连接层来决定跨通道信息交流的覆盖范围,最后与未压缩的原始特征图相乘,完成特征图的重新校准,实现网络选择性强调重要特征、抑制无用特征。一维卷积核大小 K 与通道数 C 成正比映射关系,满足

$$K = \varphi(C) = \left\lceil \frac{\ln C}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}}$$
 (1)

式中,映射参数 $\gamma = 2$ 和 $b = 1$ 为按照文献[28]取的经验值。该网络模型不降维的局部跨信道交互策略在只增加少量参数,显著降低网络模型复杂度的同时,获得明显的性能增益,是一种极轻的通道注意力机制。

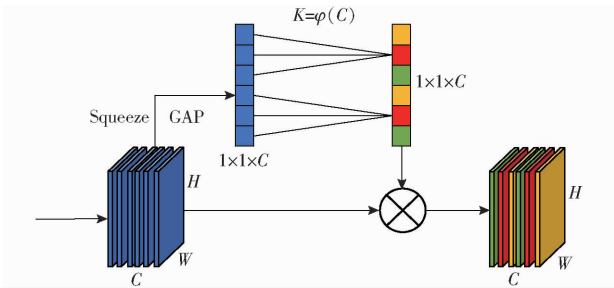


图 4 ECA 结构示意图

Fig.4 Schematic of ECA structure

ECA 机制采用一维卷积,能够有效避免全连接层降维带来的副作用,并且其自适应卷积核大小 K 通过交互相邻通道信息可以有效捕获苹果幼果中易被忽略漏检的小目标,通过跨通道式的信息交互,可在不明显增加内存开销和网络深度的情况下,避免苹果幼果图像中非目标冗余特征的学习,同时有效

兼顾重要特征。将 ECA 机制融入 YOLO v7 网络模型，可有效增强苹果幼果的检测性能。

1.2.2 YOLO v7 - ECA 网络模型构建

YOLO v7 网络模型结构主要分为 Input、

Backbone 和 Head 3 部分^[29]，其网络模型结构如图 5 所示。

Backbone 部分引入高效层聚合网络 (Efficient layer aggregation networks, ELAN) 模块，其由两个分

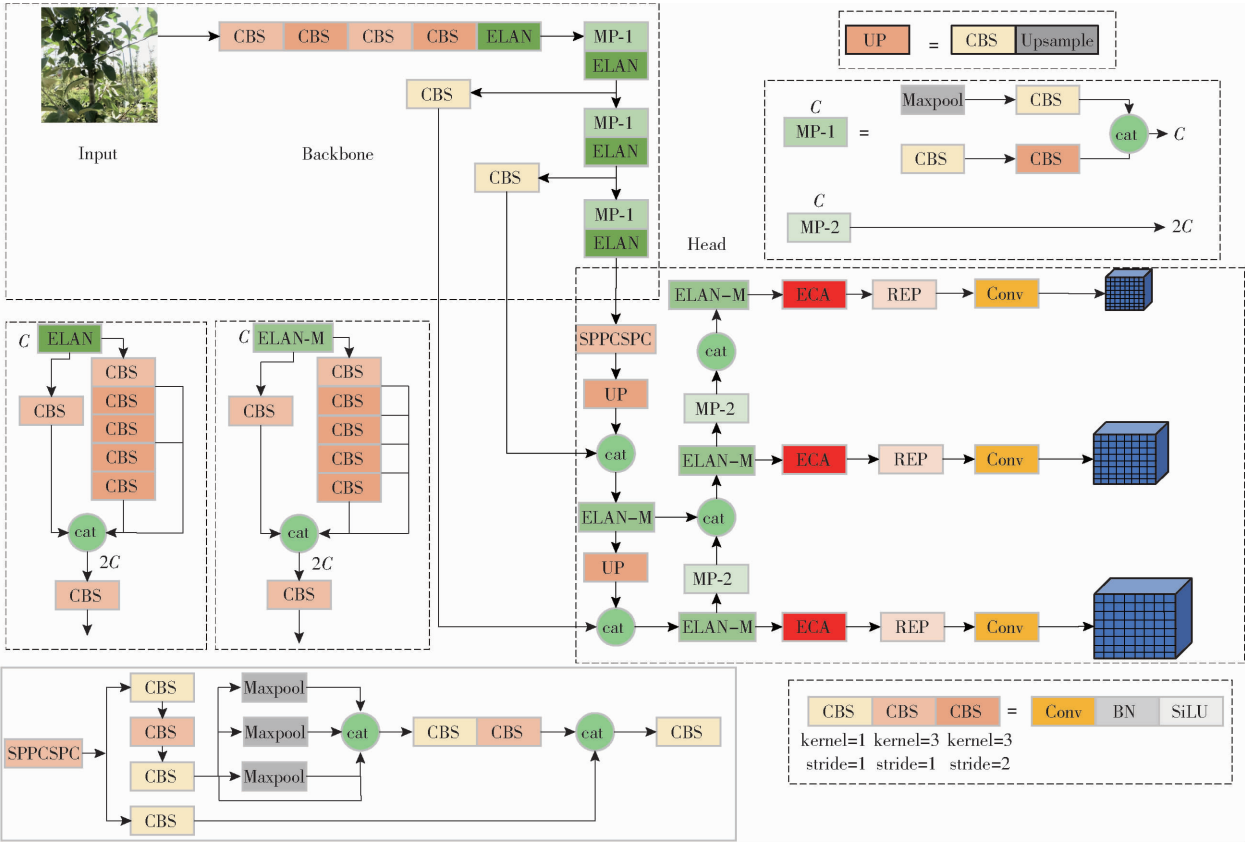


图 5 YOLO v7 - ECA 网络结构示意图

Fig. 5 Schematic of YOLO v7 - ECA network structure

支构成，通过控制最短和最长的梯度路径，同时增强网络的特征学习能力和鲁棒性能；MP 模块将分别使用最大池化和卷积块进行下采样的结果进行叠加，在减少计算量的同时增大感受野，向后有效传递全局信息。Head 部分主要由 SPPCSPC 模块、MP 模块、ELAN - M 模块以及重参数化 (Re-parameterized, REP) 模块构成；SPPCSPC 模块中空间金字塔池化层采用 $1 \times 1, 5 \times 5, 9 \times 9, 13 \times 13$ 共 4 个不同尺度最大池化适应不同分辨率图，用以区分不同尺寸的目标。网络坐标损失采用 CIoU_Loss，其同时考虑到重叠面积、中心点距离和长宽比 3 种几何参数使预测框更加符合真实框；标签分配策略采用 YOLO v5 的跨网格搜索以及 YOLOX 的匹配策略。

如图 5 所示，为在不明显增加内存开销和网络深度的情况下对苹果幼果特征进行有效学习，本研究在 YOLO v7 网络模型的 3 条重参数化路径中插入极轻的 ECA 机制，在不降低通道维数的情况下，进行相邻通道局部跨通道交互，该方式既可通过强调苹果幼果重要信息提高网络模型识别效果，又能抑制冗余无用特征提高网络模型效率。

1.3 试验平台

试验在 64 位 Windows 10 操作系统下运行，处理器为 Intel (R) Xeon (R) CPU E5 - 1620 v4 @ 3.50 GHz，内存为 32 GB，显示适配器为 NVIDIA GeForce RTX 2080Ti，使用 PyTorch 框架、PyCharm 编辑器、Python 编程语言进行网络模型训练。

1.4 评价指标

为验证网络模型有效性，采用准确率 (Precision, P)、召回率 (Recall, R)、F1 值、平均精度均值 (mAP)、模型权值文件大小、检测时间 6 个指标进行评价。

2 结果与分析

2.1 试验结果

本试验共训练 500 轮次，每 25 轮次保存一个权重文件，mAP、 R 、 P 在前 100 轮次较为振荡，之后持续上升并趋于饱和。并采用相同的数据集对 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 以及 YOLO v7 - ECA 等 7 个网络模型进行训练测试，训练结果如表 3 所示。从网络模型

检测精度角度, YOLO v7 - ECA 网络模型 mAP 高于 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型, 其值分别提高 15.5、4.6、1.6、1.8、3.0、1.8 个百分点; P 分别提高 49.7、0.9、18.5、1.2、0.9、1.0 个百分点; R 相较于 Faster R - CNN、SSD、YOLO v5、YOLO v6、YOLO v7 网络模型分别提高 5.0、4.5、1.3、3.7、1.8 个百分点。

表 3 网络模型性能对比

Tab.3 Network model performance comparison

网络模型	P /%	R /%	mAP/%	F1 值/%	权值文件大小/MB	检测时间/ms
Faster R - CNN(voc_ resnet. pth)	47.5	88.6	82.7	61.84	108.0	135.8
SSD(voc_ vgg. pth)	96.3	89.1	93.6	92.56	90.6	124.4
Scaled - YOLO v4(yolov4 - csp. weights)	78.7	95.3	96.6	86.21	200.0	53.2
YOLO v5(yolov5l. pt)	96.0	92.3	96.4	94.11	88.5	35.3
YOLO v6(yolov6s. pt)	96.3	89.9	95.2	92.99	36.2	12.7
YOLO v7(yolov7. pt)	96.2	91.8	96.4	93.94	71.3	28.9
YOLO v7 - ECA	97.2	93.6	98.2	95.37	71.3	28.9

YOLO v7 - ECA 网络模型 F1 值为 95.37%, 优于 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型, 其值分别提高 33.53、2.81、9.16、1.26、2.38、1.43 个百分点, 网络模型整体较为稳定。相较于 Scaled - YOLO v4 网络模型通过堆叠模型大小, 提高模型精度; YOLO v6 网络模型牺牲模型精度, 降低模型大小; YOLO v7 - ECA 网络模型在保持苹果幼果高精度检测的同时又能满足实时检测的要求。

2.2 注意力机制性能验证

为验证不同注意力机制对 YOLO v7 网络模型的优化效果, 本试验同时构建了 SE、CBAM、CA、ECA 网络模型进行苹果幼果目标检测性能分析, 如表 4 所示。可见, YOLO v7 - ECA 网络模型相较于其他 4 个网络模型, 权值文件大小无明显增加; 相较于 YOLO v7、YOLO v7 - SE、YOLO v7 - CBAM、YOLO v7 - CA F1 值分别提高 1.43、1.17、0.02、0.36 个百分点, 网络模型稳定性较强; 相较于 YOLO v7 mAP、 R 、 P 分别提高 1.8、1.8、1.0 个百分点, 有效降低了网络模型的错检率和漏检率。试验

表 4 注意力机制性能对比

Tab.4 Performance comparison of attention mechanisms

网络模型	P /%	R /%	mAP/%	F1 值/%	权值文件大小/MB
YOLO v7	96.2	91.8	96.4	93.94	71.3
YOLO v7 - SE	97.4	91.2	97.2	94.20	71.4
YOLO v7 - CBAM	96.0	94.7	98.2	95.35	71.4
YOLO v7 - CA	97.0	93.1	98.2	95.01	71.4
YOLO v7 - ECA	97.2	93.6	98.2	95.37	71.3

YOLO v7 - ECA 网络模型整体漏检率和错检率较低, 基本能够达到苹果幼果准确检测的要求。从网络模型检测速度角度, YOLO v7 - ECA 的最佳权值文件大小为 71.3 MB, 单幅图像检测时间为 28.9 ms, 与 YOLO v7 相比网络模型大小和检测速度基本未发生变化, 能够有效实现苹果幼果的实时检测和网络模型的移植部署。

表明 ECA 机制能够在不明显增加权值的情况下, 有效提高网络模型检测性能。

2.3 讨论

为测试网络模型在不同干扰情况下苹果幼果的检测效果, 本试验将测试集数据按照不同平滑程度、光照条件、遮挡情况分为 3 组, 并分别输入上述 7 个网络模型中对 3 种情况下的网络模型性能进行测试。

2.3.1 不同平滑程度检测对比

分别选取 75 幅光滑和模糊加噪苹果幼果图像检验上述 7 个网络模型在目标识别设备出现脏污、近距离离聚焦不清、摄像机和镜头不匹配、光圈发生变化以及特殊天气造成苹果幼果图像模糊不清问题的检测效果。试验结果如表 5 所示, 在图像光滑情况下 YOLO v7 - ECA 的 P 为 94.5%, R 为 94.1%, mAP 为 97.3%, 相较于其余 6 个网络模型能够实现苹果幼果高精度检测。其中 Faster R - CNN 网络模型 P 为 54.1%, R 为 79.6%, mAP 为 71.5%, 基本无法完成苹果幼果高精度检测任务; 在图像模糊加噪情况下 YOLO v7 - ECA 的 mAP 为 91.1%, 网络模型依旧能够准确检测苹果幼果目标, 相较于 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型 mAP 分别提高 26.3、21.0、5.4、8.0、11.5、8.9 个百分点, P 分别提高 42.7、0.4、14.1、1.5、0.7、1.3 个百分点, R 分别提高 10.1、11.2、3.9、6.0、6.1、6.9 个百分点, F1 值分别提高 27.19、7.08、8.50、4.20、3.94、4.67 个百分点, 有效提高了网络模型召回率, 降低了苹果幼果在模糊噪声干扰下的漏检率。

表 5 不同平滑程度检测结果

Tab.5 Detection results of different smoothness		degrees		%		
网络模型	类别	P	R	mAP	F1 值	
Faster R - CNN	光滑	54. 1	79. 6	71. 5	64. 42	
	模糊加噪	54. 7	73. 2	64. 8	62. 61	
SSD	光滑	97. 2	89. 1	88. 9	92. 97	
	模糊加噪	97. 0	72. 1	70. 1	82. 72	
Scaled - YOLO v4	光滑	78. 5	95. 8	96. 3	86. 29	
	模糊加噪	83. 3	79. 4	85. 7	81. 30	
YOLO v5	光滑	97. 1	90. 4	95. 4	93. 63	
	模糊加噪	95. 9	77. 3	83. 1	85. 60	
YOLO v6	光滑	98. 1	88. 7	93. 1	93. 16	
	模糊加噪	96. 7	77. 2	79. 6	85. 86	
YOLO v7	光滑	97. 7	90. 3	95. 6	93. 85	
	模糊加噪	96. 1	76. 4	82. 2	85. 13	
YOLO v7 - ECA	光滑	94. 5	94. 1	97. 3	94. 30	
	模糊加噪	97. 4	83. 3	91. 1	89. 80	

图 6 为模糊加噪情况下苹果幼果原始标注图像以及 7 种不同网络模型检测结果图像。按照图 6a

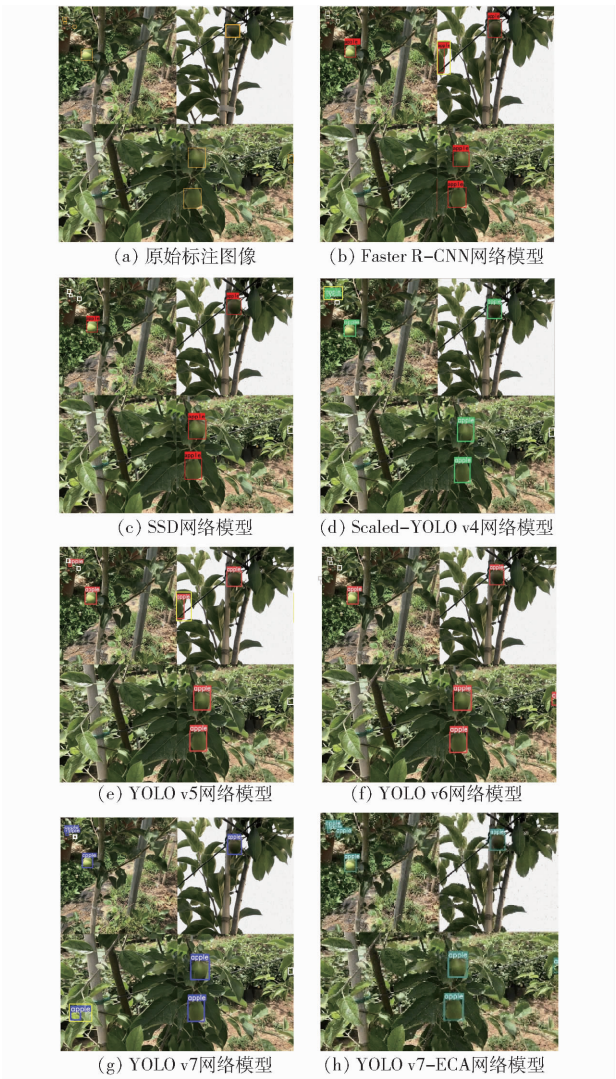


图 6 不同平滑程度检测结果对比

Fig. 6 Comparison of detection results of different smoothness levels

所示原始标注图像,将检测图像中漏检苹果幼果目标用白色框标出,错检苹果幼果目标用黄色框标出。由图 6b ~ 6g 可得, Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 分别漏检 4、4、2、3、3、2 个幼果果实目标; Faster R - CNN、Scaled - YOLO v4、YOLO v5、YOLO v7 各错检 1 个果实目标。图 6h 为 YOLO v7 - ECA 模糊加噪情况下的检测结果图像,由图可知 YOLO v7 - ECA 网络模型无漏检错检的情况。试验表明, YOLO v7 - ECA 网络模型在苹果幼果图像模糊加噪情况下能够完成高精度目标检测任务。

2.3.2 不同光照条件检测对比

为实现全方位苹果幼果果实检测,数据集拍摄从顺光、逆光不同拍摄角度进行,根据光线射入角度不同图像呈现亮光状态和阴影状态。本试验各选取 110 幅亮光和阴影图像输入 7 个网络模型对不同光照条件下的苹果幼果进行测试,试验结果如表 6 所示。

表 6 不同光照条件检测结果

Tab.6 Detection results under different lighting		conditions		%		
网络模型	类别	P	R	mAP	F1 值	
Faster R - CNN	亮光	56. 6	89. 1	81. 3	69. 23	
	阴影	64. 1	87. 1	82. 7	73. 85	
SSD	亮光	97. 4	91. 8	91. 7	94. 52	
	阴影	97. 1	88. 8	88. 7	92. 76	
Scaled - YOLO v4	亮光	81. 9	97. 0	97. 7	88. 81	
	阴影	77. 7	93. 5	95. 4	84. 87	
YOLO v5	亮光	97. 1	94. 8	97. 3	95. 94	
	阴影	97. 3	90. 6	95. 1	93. 83	
YOLO v6	亮光	98. 5	92. 1	95. 2	95. 19	
	阴影	97. 4	87. 4	92. 1	92. 13	
YOLO v7	亮光	98. 0	92. 4	97. 5	95. 12	
	阴影	95. 1	90. 6	95. 0	92. 80	
YOLO v7 - ECA	亮光	97. 8	96. 7	99. 0	97. 25	
	阴影	97. 3	93. 5	97. 5	95. 36	

亮光状态下 YOLO v7 - ECA 的 P 为 97.8% , R 为 96.7% , mAP 为 99% , 整体检测精度优于其余 6 个网络模型,其中 Faster R - CNN 网络模型 P 为 56.6% , R 为 89.1% , mAP 为 81.3% , P 与 mAP 较低,完成苹果幼果高精度检测存在困难;阴影状态下 YOLO v7 - ECA 的 mAP 为 97.5% , 其值相较于 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型分别提高 14.8、8.8、2.1、2.4、5.4、2.5 个百分点, $F1$ 值分别提高 21.51、2.60、10.49、1.53、3.23、2.56 个百分点,有效解决了背景偏暗情况下幼果不易检测的难题。

图 7 为阴影状态下苹果幼果原始标注图像以及 7 种不同网络模型检测图像。由图 7b ~ 7g 可得, Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 分别漏检 5、5、4、4、4、5 个幼果果实目标, YOLO v5 错检 1 个幼果果实目标。由图 7h 可得, YOLO v7 - ECA 网络模型在整体图像偏暗时无漏检错检的情况。试验表明, YOLO v7 - ECA 网络模型在苹果幼果图像整体偏暗时依旧能够实现目标果实的准确检测。

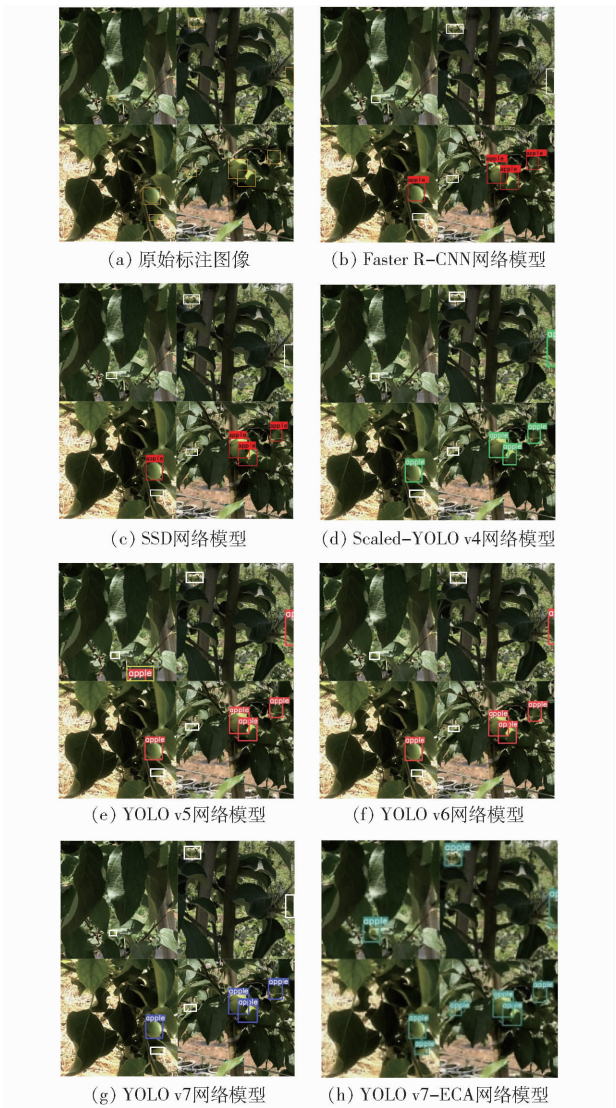


图 7 不同光照条件检测结果对比

Fig. 7 Comparison of detection results of different lighting conditions

2.3.3 不同遮挡情况检测对比

为验证不同遮挡下的幼果检测效果,本试验分别选取 60 幅无遮挡、轻微遮挡、严重遮挡图像进行测试,结果如表 7 所示。由表 7 可得,在无遮挡状态下本网络模型的 P 为 98.3%, R 为 98.3%, mAP 为 99.6%, 轻微遮挡状态下本网络模型的 P 为 100%, R 为 99.3%, mAP 为 99.7%, 整体检测精度优于其余 6 个网络模型, SSD、Scaled - YOLO v4、YOLO v5、

YOLO v6、YOLO v7 网络模型在苹果幼果目标无遮挡以及轻微遮挡时基本能够实现高精度检测, 其中 Faster R - CNN 网络模型在苹果幼果目标无遮挡以及轻微遮挡时, 牺牲 P , 提高 R , 网络模型稳定性较差, 难以完成苹果幼果高精度检测; 严重遮挡情况下本网络模型的 mAP 为 98.6%, 相较于 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网络模型分别提高 21.7、13.7、2.3、2.4、4.8、2.2 个百分点, R 分别提高 10.0、9.6、0.8、1.9、4.2、3.4 个百分点, $F1$ 值分别提高 28.29、3.50、6.45、0.96、1.36、1.36 个百分点, 能够在各种遮挡干扰下准确识别出苹果幼果目标。

表 7 不同遮挡情况检测结果

Tab. 7 Detection results of different occlusion

conditions					%
网络模型	类别	P	R	mAP	$F1$ 值
Faster R - CNN	无遮挡	58.6	97.3	94.8	73.15
	轻微遮挡	59.7	96.0	90.8	73.62
	严重遮挡	54.8	84.6	76.9	66.51
SSD	无遮挡	97.3	97.3	97.3	97.30
	轻微遮挡	98.6	98.0	97.9	98.30
	严重遮挡	98.6	85.0	84.9	91.30
Scaled - YOLO v4	无遮挡	78.7	99.1	99.4	87.73
	轻微遮挡	84.7	100	99.4	91.72
	严重遮挡	83.5	93.8	96.3	88.35
YOLO v5	无遮挡	99.9	98.7	99.5	99.30
	轻微遮挡	98.9	98.3	99.5	98.60
	严重遮挡	95.0	92.7	96.2	93.84
YOLO v6	无遮挡	100	96.5	99.4	98.22
	轻微遮挡	98.0	99.3	99.4	98.65
	严重遮挡	96.7	90.4	93.8	93.44
YOLO v7	无遮挡	99.9	96.5	99.4	98.17
	轻微遮挡	98.7	98.0	99.4	98.35
	严重遮挡	95.8	91.2	96.4	93.44
YOLO v7 - ECA	无遮挡	98.3	98.3	99.6	98.30
	轻微遮挡	100	99.3	99.7	99.65
	严重遮挡	95.0	94.6	98.6	94.80

图 8 为严重遮挡状态下苹果幼果原始标注图像及其检测结果。由图 8b ~ 8g 可得, Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 分别漏检 4、3、4、5、4、4 个幼果果实目标, Faster R - CNN 和 Scaled - YOLO v4 各错检 1 个幼果果实目标。由图 8h 可得 YOLO v7 - ECA 网络模型在幼果果实受枝条和叶片遮挡严重的情况下依旧能够准确识别幼果, 无果实目标的漏检错检。试验表明, YOLO v7 - ECA 网络模型在苹果幼果遮挡严重时依旧能够实现目标果实的准确检测。

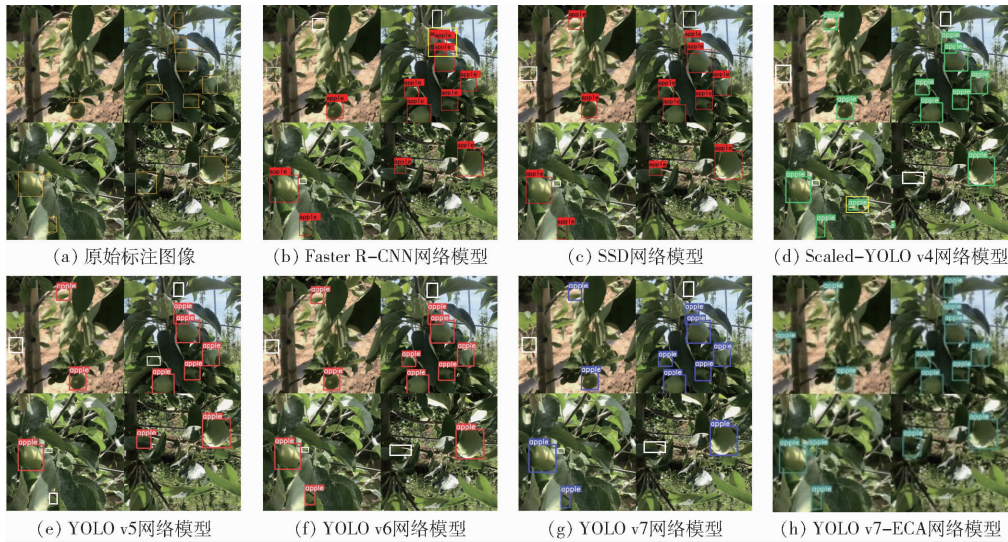


图 8 不同遮挡情况检测结果对比

Fig. 8 Comparison of detection results of different occlusion conditions

3 结论

(1) 在 YOLO v7 网络模型中插入极轻的 ECA 机制构建了 YOLO v7 - ECA 网络模型进行苹果幼果的检测,将采集到的 3 654 幅幼果图像输入网络模型进行训练,获得最优网络模型 P 为 97.2%、 R 为 93.6%、 mAP 为 98.2%、 $F1$ 值为 95.37%,与 Faster R - CNN、SSD、Scaled - YOLO v4、YOLO v5、YOLO v6、YOLO v7 网路模型相比,其 mAP 分别提高 15.5、4.6、1.6、1.8、3.0、1.8 个百分点, P 分别提高 49.7、0.9、18.5、1.2、0.9、1.0 个百分点,能够准确识别苹

果幼果目标。

(2) 由不同干扰情况下的对比试验可知,YOLO v7 - ECA 在苹果幼果图像模糊加噪、阴影以及严重遮挡情况下依旧能够实现果实目标的准确检测。模糊加噪情况下 mAP 为 91.9%,阴影情况下 mAP 为 97.5%,严重遮挡情况下 mAP 为 98.6%。

(3) YOLO v7 - ECA 网络模型最佳权值文件大小为 71.3 MB,单幅图像检测时间为 28.9 ms,该网络模型在保持较高精度的同时又能拥有较快的检测速度,其网络模型内存消耗低,可进一步实现网络模型的迁移部署。

参 考 文 献

[1] 宋怀波,江梅,王云飞,等. 融合卷积神经网络与视觉注意机制的苹果幼果高效检测方法[J]. 农业工程学报, 2021, 37(9): 297 - 303.
SONG Huaibo, JIANG Mei, WANG Yunfei, et al. Efficient detection method for young apples based on the fusion of convolutional neural network and visual attention mechanism[J]. Transactions of the CSAE, 2021, 37(9): 297 - 303. (in Chinese)

[2] 王丹丹,何东健. 基于 R - FCN 深度卷积神经网络的机器人疏果前苹果目标的识别[J]. 农业工程学报,2019,35(3):156 - 163.
WANG Dandan, HE Dongjian. Recognition of apple targets before fruits thinning by robot based on R - FCN deep convolution neural network[J]. Transactions of the CSAE, 2019, 35(3): 156 - 163. (in Chinese)

[3] 潘云飞,周艳,何磊,等. 果园管理工作中疏花疏果的研究进展[J]. 中国农机化学报,2021,42(11):198 - 204.
PAN Yunfei, ZHOU Yan, HE Lei, et al. Research progress of flower and fruit thinning in orchard management[J]. Journal of Chinese Agricultural Mechanization, 2021, 42(11): 198 - 204. (in Chinese)

[4] JIA W, TIAN Y, LUO R, et al. Detection and segmentation of overlapped fruits based on optimized Mask R - CNN application in apple harvesting robot[J]. Computers and Electronics in Agriculture, 2020, 172: 105380.

[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.
SONG Huaibo, SHANG Yuying, HE Dongjian. Review on deep learning technology for fruit target recognition[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 1 - 19. (in Chinese)

[6] WU D, LV S, JIANG M, et al. Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments[J]. Computers and Electronics in Agriculture, 2020, 178: 105742.

[7] 尚钰莹,张倩如,宋怀波. 基于 YOLOv5s 的深度学习在自然场景苹果花朵检测中的应用[J]. 农业工程学报, 2022, 38(9): 222 - 229.
SHANG Yuying, ZHANG Qianru, SONG Huaibo. Application of deep learning using YOLOv5s to apple flower detection in natural scenes[J]. Transactions of the CSAE, 2022, 38(9): 222 - 229. (in Chinese)

- [8] ZHANG W, WANG J, LIU Y, et al. Deep-learning-based in-field citrus fruit detection and tracking[J]. Horticulture Research, 2022, 9: uhac003.
- [9] WANG Z, JIN L, WANG S, et al. Apple stem/calyx real-time recognition using YOLO-v5 algorithm for fruit automatic loading system[J]. Postharvest Biology and Technology, 2022, 185: 111808.
- [10] 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.
- [11] 宋怀波,王亚男,王云飞,等.基于YOLO v5s的自然场景油茶果识别方法[J].农业机械学报,2022,53(7):234-242.
SONG Huaibo, WANG Ya'nan, WANG Yunfei, et al. *Camellia oleifera* fruit detection in natural scene based on YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 234-242. (in Chinese)
- [12] 刘天真,滕桂法,苑迎春,等.基于改进YOLO v3的自然场景下冬枣果实识别方法[J].农业机械学报,2021,52(5):17-25.
LIU Tianzhen, TENG Guifa, YUAN Yingchun, et al. Winter jujube fruit recognition method based on improved YOLO v3 under natural scene[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 17-25. (in Chinese)
- [13] 龙燕,李南南,高研,等.基于改进FCOS网络的自然环境下苹果检测[J].农业工程学报,2021,37(12):307-313.
LONG Yan, LI Nannan, GAO Yan, et al. Apple fruit detection under natural condition using improved FCOS network[J]. Transactions of the CSAE, 2021, 37(12): 307-313. (in Chinese)
- [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. (in Chinese)
- [15] 赵辉,乔艳军,王红君,等.基于改进YOLOv3的果园复杂环境下苹果果实识别[J].农业工程学报,2021,37(16):127-135.
ZHAO Hui, QIAO Yanjun, WANG Hongjun, et al. Apple fruit recognition in complex orchard environment based on improved YOLOv3[J]. Transactions of the CSAE, 2021, 37(16): 127-135. (in Chinese)
- [16] 王立舒,秦铭霞,雷洁雅,等.基于改进YOLOv4-Tiny的蓝莓成熟度识别方法[J].农业工程学报,2021,37(18):170-178.
WANG Lishu, QIN Mingxia, LEI Jieya, et al. Blueberry maturity recognition method based on improved YOLOv4-Tiny[J]. Transactions of the CSAE, 2021, 37(18): 170-178. (in Chinese)
- [17] 王卓,王健,王泉雄,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)
- [18] JIANG M, SONG L, WANG Y, et al. Fusion of the YOLOv4 network model and visual attention mechanism to detect low-quality young apples in a complex environment[J]. Precision Agriculture, 2022, 23(2): 559-577.
- [19] SU F, LU Q, LUO R Z. Review of image classification based on deep learning[J]. Telecommunication Science, 2019, 35(11): 58-74.
- [20] ZAIDI S S A, ANSARI M S, ASLAM A, et al. A survey of modern deep learning based object detection models[J]. Digital Signal Processing, 2022, 126: 103514.
- [21] TONG K, WU Y, ZHOU F. Recent advances in small object detection based on deep learning: a review[J]. Image and Vision Computing, 2020, 97: 103910.
- [22] XIAO Y, TIAN Z, YU J, et al. A review of object detection based on deep learning[J]. Multimedia Tools and Applications, 2020, 79(33): 23729-23791.
- [23] WU X, SAHOO D, HOI S C H. Recent advances in deep learning for object detection[J]. Neurocomputing, 2020, 396: 39-64.
- [24] MO Y, WU Y, YANG X, et al. Review the state-of-the-art technologies of semantic segmentation based on deep learning[J]. Neurocomputing, 2022, 493: 626-646.
- [25] HU J, SHEN L, SUN G, et al. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [26] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [27] 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.
- [28] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 11531-11539.
- [29] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. arXiv e-prints, arXiv:2207.02696,2022.