

DOI: 10.13733/j.jcam.issn.2095-5553.2024.10.040

章璞, 乔波, 陈义明. 基于改进YOLOv8的多阶段草莓检测算法[J]. 中国农机化学报, 2024, 45(10): 274-280

Zhang Pu, Qiao Bo, Chen Yiming. A multi-stage strawberry detection algorithm based on improved YOLOv8 [J]. Journal of Chinese Agricultural Mechanization, 2024, 45(10): 274-280

# 基于改进YOLOv8的多阶段草莓检测算法\*

章璞, 乔波, 陈义明

(湖南农业大学信息与智能科学技术学院, 长沙市, 410128)

**摘要:**为实现温室复杂环境中草莓的快速、精确检测,提出一种基于改进YOLOv8的多阶段草莓检测算法。首先,针对温室环境下采集到的草莓数据集,使用LabelImg对数据集进行标注;其次,针对草莓体积小、环境复杂等问题,在主干网络中融入BiFormer动态注意力机制,实现更加灵活的计算分配和特征感知,使网络模型更加关注小目标检测,并提高其在复杂环境下的果实检测能力;最后,在Neck部分引入VanillaNet模块,以降低模型计算量,进一步提高对草莓的识别精度。试验结果表明,相较传统的YOLOv8,改进后的YOLOv8的平均精度均值 $mAP$ 提升4.6%,达到93.8%。改进后的YOLOv8具有更高的检测精度,同时在小目标检测方面表现出色,为后续的采摘机器人实时小目标检测提供支撑。

**关键词:**深度学习;草莓检测;YOLOv8;注意力机制;数据增强

**中图分类号:**TP391 **文献标识码:**A **文章编号:**2095-5553(2024)10-0274-07

## A multi-stage strawberry detection algorithm based on improved YOLOv8

Zhang Pu, Qiao Bo, Chen Yiming

(College of Information and Intelligence, Hunan Agricultural University, Changsha, 410128, China)

**Abstract:** In order to achieve fast and accurate detection of strawberries in complex greenhouse environments, a multi-stage strawberry detection algorithm based on an improved YOLOv8 was proposed. Firstly, the strawberry dataset, collected in greenhouse environments, was initially annotated by using LabelImg. Subsequently, in order to address issues such as the small size of strawberries and the complexity of the environment, a BiFormer dynamic attention mechanism was integrated into the backbone network. This integration allowed for more flexible computational allocation and feature perception, focusing the network model more on small object detection and enhancing its fruit detection capabilities in complex environments. Finally, a VanillaNet module was introduced in the Neck component to reduce the computational complexity of the model and further improve its strawberry recognition accuracy. Experimental results demonstrated that the improved YOLOv8, in comparison to the traditional YOLOv8, increased the  $mAP$  by 4.6%, reaching 93.8%. The improved YOLOv8 not only has higher detection accuracy, but also performs well in small target detection, which can provide support for the subsequent real-time small target detection of picking robots.

**Keywords:** deep learning; strawberry detection; YOLOv8; attention mechanism; data enhancement

## 0 引言

中国是全球最大的草莓生产国,产值达50亿美元,是第二大草莓生产国美国的3倍多<sup>[1]</sup>。草莓是中国农民增收致富的重要经济作物,草莓成熟度的识别

对于智能管理中的灌溉、施肥和环境控制等决策具有重要意义。由于草莓体积小、自然环境复杂等特点,使得草莓的检测比较困难。目前,草莓的识别主要依靠人工观测,工作效率低<sup>[2]</sup>。传统的人工采摘和生长观测对于小型植株和果实来说耗时且费力,可能导致高

收稿日期:2024年1月2日 修回日期:2024年3月21日

\* 基金项目:2022年度湖南省社会科学成果评审委员会课题(XSP22YBZ024)

第一作者:章璞,男,1999年生,湖南岳阳人,硕士研究生;研究方向为农业信息技术。E-mail: 389785727@qq.com

通讯作者:乔波,男,1981年生,湖南临澧人,博士,副教授;研究方向为自然语言处理、农业知识图谱与智慧农业。E-mail: qiaobo@hunau.edu.cn

额的人工成本,并且无法满足实时、快速的监测需求。

近年来,国内外在果实精确识别和成熟度分类方面的研究已取得显著进展。传统机器学习方法利用颜色、质地、形状等浅层特征,结合不同化学和物理性质的测量参考值,或结合光谱成像等数据来识别水果对象。Munera 等<sup>[3]</sup>使用 VIS-NIR 高光谱反射成像实现对油桃成熟度检测。Malik 等<sup>[4]</sup>基于改进的(HSV)色彩空间和增强的分水岭分割算法检测成熟番茄。马惠玲等<sup>[5]</sup>利用近红外高光谱图像技术结合 KNN 计算对苹果品种进行快速鉴别,取得了较高的检测精度。Indrabayu 等<sup>[6]</sup>等使用采用径向基函数(RBF)核函数的多类支持向量机(SVM)对草莓成熟度进行分类,准确率高达 85.64%。传统图像识别方法在精确度方面表现出色。然而,传统机器学习技术的局限性往往依赖于手工设计的特征提取方法,意味着在复杂背景下水果目标检测的特征表达能力会受到一定程度限制。因此,基于图像的水果检测技术需要进一步深入研究,以更好地适应复杂场景下的检测需求。

随着计算机视觉和深度学习技术的发展,尤其是在 2012 年 Krizhevsky 等<sup>[7]</sup>推出 AlexNet 之后,深度学习技术已经快速发展并广泛应用于各个领域。许多研究已经证明基于 CNN 的水果检测技术具有潜力。Appel 等<sup>[8]</sup>提出了一种基于改进 YOLO 的目标检测算法,即 CAM-YOLO 算法。该算法结合了 YOLOv5 目标识别和卷积注意力模块(CBAM),并应用非极大值抑制和距离交并集(DIoU)来增强对图像中重叠物体的识别。试验结果表明,CAM-YOLO 算法在重叠番茄和小番茄的检测效率上达到了 88.1%。Zhou 等<sup>[2]</sup>提出了一种基于 Faster R-CNN 的草莓检测方法,平均检测精度达到 86%。张小花等<sup>[9]</sup>提出了基于 EfficientDet-D1 的草莓快速检测及分类方法,该方法结合了 EfficientNet 网络的快速归一化特征加权融合特点,针对成熟和未成熟草莓的检测任务,平均精度达到了 96.71%,有较好的泛化性和鲁棒性。孙俊等<sup>[10]</sup>提出一种基于改进 YOLOv4-Tiny 的草莓检测模型,采用轻量级网络 GhostNet 作为特征提取网络,并在 GhostBottleneck 结构中嵌入卷积注意力模块,实现轻量化的同时将草莓的识别精度提升至 92.62%。

与传统的图像方法相比,卷积神经网络具有更强的特征提取能力、准确性以及鲁棒性和泛化能力<sup>[11]</sup>。然而,现有研究大多侧重于红熟期草莓的检测,对草莓多阶段的研究有限,且实际的草莓检测应用环境中,由于草莓体积小且密集分布,其外观特征与周围背景非常相似,导致检测困难。在检测过小的草莓目标方面,目前缺乏有效的算法来获得准确的结果。现有算法在检测过程中可能

会遇到误检、漏检等问题。为解决上述问题,设计一种基于改进 YOLOv8 的多阶段草莓检测算法,采用 BiFormer<sup>[12]</sup>动态注意力机制与轻量级结构 VanillaNet<sup>[13]</sup>相结合,从而提升对小目标的检测性能和效率。

## 1 材料与方法

### 1.1 数据集的采集与构建

本试验所采集的数据来自浙江衢州莲花基地,其采用温室高架栽培模式,种植场景如图 1 所示。草莓的品种为红颜,试验使用的图像采集设备为手机后置摄像头。试验通过采集动态视频数据,并使用 Python 对采集到的视频进行帧处理,共采集了温室环境中的草莓图像 463 张,为提高数据集的容量,使用 Python 脚本从网上随机爬取并针对性地筛选了小目标等影响下的草莓图像 345 张,总数共计 808 张原始图像。本试验使用的标记工具为 LabelImg 图像处理软件,并通过人工手动对草莓位置进行标记。将识别数据分为未成熟、半成熟和成熟草莓 3 种样本,按 8:1:1 的比例划分为训练集、验证集和测试集。



图 1 草莓种植场景

Fig. 1 Strawberry planting scene

Mosaic 方法是 YOLOv4 论文中提出的一种基于图像变换的数据增强方法。它通过一系列方法来拼接多个原始图像,从而创建一个合成图像。合成图像融合了多个图像的内容,包含了更多的信息,有助于训练模型更好地适应各种复杂场景。

为了增强模型的泛化性能,本文在训练过程中应用 Mosaic 技术。该技术通过随机选取 4 张图片,并对它们进行随机裁剪、翻转、缩放和色域变化,然后将它们拼接在一起,作为输入传入模型中。具体的图像处理方法和参数如表 1 所示。

表 1 Mosaic 数据增强

Tab. 1 Mosaic data augmentation

方法	参数
图像 HSV 色调增强	0.015
图像 HSV 饱和度增强	0.7
图像 HSV 明度增强	0.4
图像平移	0.1
图像缩放	0.5
图像左右翻转	0.5
马赛克技术	1.0

## 1.2 方法

### 1.2.1 YOLOv8网络模型

YOLOv8由Ultralytics于2023年1月发布<sup>[14]</sup>。基于前一系列YOLO模型的成功,YOLOv8引入了新的功能和改进,这使其在目标检测方面表现出更高的效率,同时仍然保持高精度。YOLOv8的架构如图2所示,主要可分为骨干特征提取网络(Backbone)、特征融合网络(Neck)和检测头(Head)三部分。YOLOv8包含五种模型,分别是YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l和YOLOv8x。其中YOLOv8s在检测性能和模型大小之间取得了最佳平衡。因此,本文采用一阶段检测算法YOLOv8s作为基线模型。

YOLOv8使用CSPDarknet53作为骨干网络,它是YOLOv7中使用的Darknet53网络的改进版本<sup>[15]</sup>。相较于YOLOv5,YOLOv8采用了更轻量化的c2f模块,取代了YOLOv5中的C3模块,这一改进在保持模型轻量化的同时,提供了更丰富的梯度流信息,有助于提高检测性能。在骨干网络末端,则仍使用最流行的空间金字塔快速池化(SPPF)模块,SPPF是SPP<sup>[16]</sup>的变体,它是一种更高效且边缘设备友好的信息融合模块。在特征融合网络部分,YOLOv8使用PAN-FPN结构进行特征融合,与YOLOv5和YOLOv7模型相比,YOLOv8在PAN结构中去除了上采样后的卷积运算,以实现轻量化,并保持原有性能。在最后的检测头部分(Head),采用目前主流的解耦头结构(Decoupled Head)。与YOLOv6和YOLOX类似。综上所述,YOLOv8引入了诸多改进,这使得它在目标检测任务中具有较高的性能

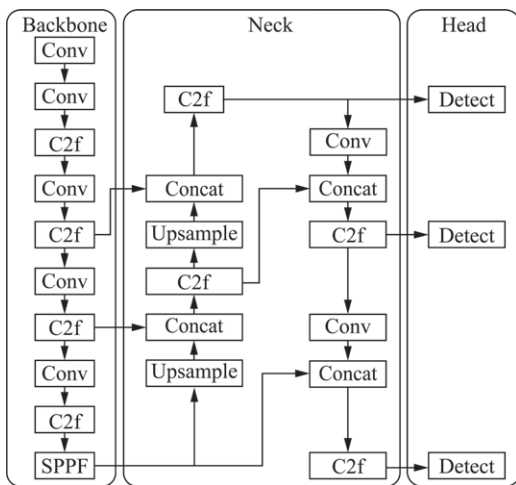


图2 YOLOv8网络结构

Fig. 2 YOLOv8 network structure

### 1.2.2 YOLOv8算法的改进

YOLOv8在众多方面表现出色,然而在复杂场景中对小目标的识别仍存在不足之处。为了更好地应对

草莓检测任务的需求,对原始的YOLOv8结构进行了优化,改进了对小目标的检测效果,并进一步提高了整体的检测精度。改进YOLOv8的多阶段草莓检测算法结构如图3所示。

1) 为了使模型能够更好地关注草莓的细节特征,在YOLOv8骨干网络部分,引入CVPR2023的最新成果BiFormer注意力机制,替换原有的C2f模块,从而提升模型对小目标的检测性能。

2) 为了进一步提高对草莓的识别精度,在Neck部分,引入了VanillaNet模块,以降低模型计算量,使得模型能够在提高性能的同时减少计算资源的消耗,从而提高模型的效率和实用性。

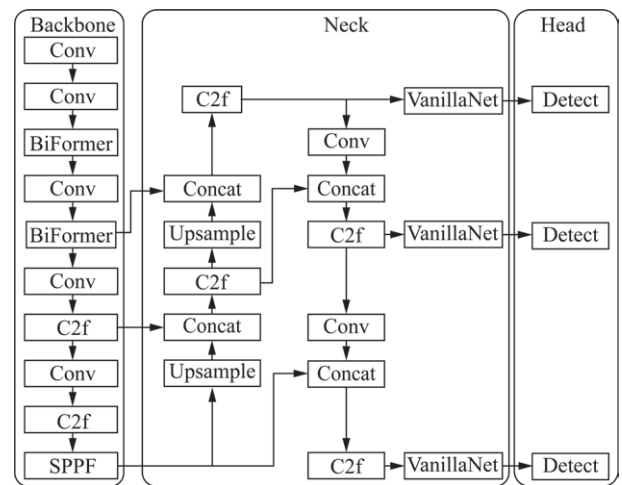


图3 改进YOLOv8网络结构

Fig. 3 Improved YOLOv8 network structure

如图3所示,改进后的YOLOv8引入了BiFormer注意力机制和VanillaNet模块,相比于其他常用的目标检测算法(如YOLOv5、YOLOv7、SSD<sup>[17]</sup>和Faster R-CNN<sup>[18]</sup>),在整体精度上更高,并且得益于BiFormer注意力机制对于小目标草莓的检测性能也更好。在GFLOPs方面,改进后的YOLOv8其计算复杂度远低于YOLOv7、SSD和Faster R-CNN,然而,尽管引入VanillaNet模块成功降低了BiFormer注意力机制带来的计算量问题,但改进后的YOLOv8仍然引入了一定的计算量。

### 1.2.3 BiFormer注意力机制

BiFormer基于Transformer<sup>[19]</sup>模型架构设计,在多个任务上表现优异,包括图像分类、目标检测、实例分割和语义分割等。从BiFormer在COCO2017数据集上的试验结果来看,BiFormer注意力机制在处理小目标和密集检测方面尤为突出。草莓生长环境复杂,检测过程中常出现漏检和误检问题。这是因为草莓植株和果实较小且密集分布,导致对其进行准确检测具有挑战性。为了解决这些问题,引入了BiFormer机制,



以便检测模型过滤掉大部分与目标无关的信息,更好地聚焦于待检测目标。与传统注意力相比,BiFormer大幅减少了计算量和内存需求,同时又保留了全局建模能力,实现性能和效率之间的平衡。BiFormer注意力机制的结构如图4所示。

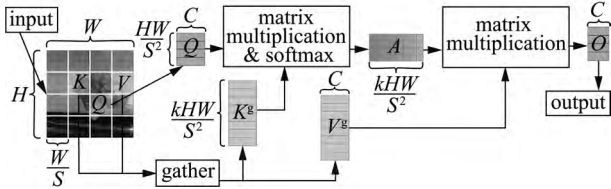


图 4 BiFormer 注意力机制工作流程

Fig. 4 BiFormer attention mechanism workflow

从图4可以看出,BiFormer首先划分输入特征  $X \in R^{H \times W \times C}$  为  $S \times S$  个不同区域,即  $X^r \in R^{S^2 \times \frac{HW}{S^2} \times C}$ ,以此确保每个区域都包含一定数量的特征信息。划分后,利用线性映射,通过权值矩阵  $W^q$ 、 $W^k$  和  $W^v$  获得  $Q, K, V \in R^{S^2 \times \frac{HW}{S^2} \times C}$ ,计算如式(1)所示。

$$Q = X^r W^q, K = X^r W^k, V = X^r W^v \quad (1)$$

其中,  $Q, K, V$  分别为查询向量、键向量和值向量。

随后,对于每个划分区域,计算查询向量  $Q$  和键向量  $K$  的平均值。即  $Q^r, K^r \in R^{S^2 \times C}$ 。并以此构建邻接矩阵  $A^r$ ,计算区域间语义相关性。

$$A^r = Q^r (K^r)^T \quad (2)$$

不同于传统注意力机制,BiFormer通过索引矩阵只选择与查询区域最相关的前  $k$  个区域,以此避免冗余计算。

$$I^r = \text{topkIndex}(A^r) \quad (3)$$

由于与查询区域最相关的区域分散在整个特征图当中,直接计算会导致内存利用率较低,因此BiFormer收集分散的键向量和值向量,并在收集的键值对中应用细粒度的 token-to-token 注意力机制,即  $K^g, V^g \in R^{S^2 \times \frac{HW}{S^2} \times C}$ 。

$$K^g = \text{Gather}(K, I^r), V^g = \text{Gather}(V, I^r) \quad (4)$$

最终得到BiFormer注意力机制式如式(5)所示。

$$O = \text{Attention}(Q, K^g, V^g) + \text{LCE}(V) \quad (5)$$

其中,  $\text{LCE}$  (Local Context Enhancement)<sup>[20]</sup> 是基于深度卷积的局部上下文增强函数。

传统的注意力机制中,会对所有的键值对进行全局计算,这导致了计算复杂度较高的问题。然而,在BiFormer中,引入了双层路由注意力机制,来优化这个问题。双层路由注意力机制的关键思想是只关注与查询区域最相关的前  $k$  个区域。通过选择与查询最相关的  $k$  个区域,可以避免在最不相关的区域进行冗余计算,从而大幅减少计算量。这种做法充分利用了稀

疏性,使得计算更加高效。另外,BiFormer注意力机制还针对GPU进行了优化,仅使用适用于GPU的密集矩阵乘法运算。这种优化进一步提高了计算效率,同时减少了内存占用。

### 1.2.4 VanillaNet 模块

VanillaNet是一种基于极简主义设计的神经网络模型,由华为诺亚和悉尼大学的研究者共同提出。该模型旨在应对复杂神经网络部署所面临的挑战,它以极简主义的设计理念为基础,舍弃了过多的深度、shortcut以及self-attention等复杂的操作,简化了网络结构,提供了在资源有限环境下高效的解决方案,并在计算机视觉任务中取得了令人满意的结果。因此,本文在Neck部分引入VanillaNet模块来提高模型的精度,并减少由BiFormer注意力机制引入的模型计算量。VanillaNet其框架如图5所示,包括主干、主体和全连接层三部分。其借鉴AlexNet和VGG等传统深度网络的结构设计,输入特征的大小被向下采样,而通道在每个阶段被加倍。以6层的网络结构为例,主干部分采用了步长为4的  $4 \times 4 \times 3 \times C$  卷积进行特征变换;对于主体部分的每个阶段,采用MaxPool进行特征下采样以减小尺寸和特征图,且为了保留特征图主要信息的同时,实现最小的计算成本,采用一个  $1 \times 1$  卷积进行特征处理,并且在每个  $1 \times 1$  卷积层之后应用激活函数。此外,在每一层之后还添加了BN<sup>[21]</sup>来简化网络的训练过程。需要注意的是,第5层卷积层不增加通道数,并使用平均池化层。对于全连接层部分,用于进行分类处理。

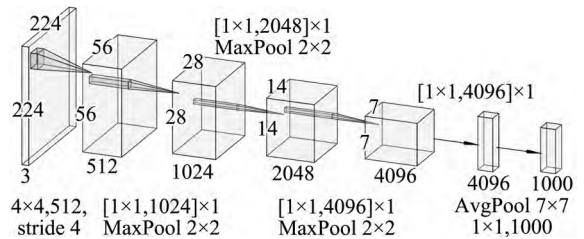


图 5 VanillaNet 网络结构

Fig. 5 VanillaNet network framework

## 2 试验结果与分析

### 2.1 试验环境以及参数设置

本文试验均在同一硬件环境下运行,具体包括:CPU为Intel(R)Xeon(R)Silver 4214 R CPU@2.40 GHz, GPU为NVIDIA GeForce RTX3080 Ti,显存为12 GB。试验所采用的操作系统环境为Linux Ubuntu,深度学习框架为PyTorch 2.0.0, Python版本为3.8, Cuda版本为11.8。

试验中使用的输入图像大小为640像素  $\times$  640像素,

优化器为 Adam, 权重衰减率为 0.000 5, 初始学习率为 0.001, 批尺寸为 16, 训练 200 个轮次 (Epoch)。此外, 为了增强模型泛化性, 本文还采用了 Mosaic 数据增强方法。模型训练结束后, 使用保存的权重在测试集上对模型进行性能评估。

## 2.2 评价指标

本次试验的模型性能评估指标包括  $mAP@0.5$  和  $mAP@0.5:0.95$  两项, 同时针对模型改进引入的计算量问题, 引入  $GFLOPs$  参数作为模型计算效率指标。 $GFLOPs$  越低, 表明模型的计算效率就越高。平均精度均值  $mAP$  是目标检测中最常用的评价指标之一, 它反映了模型的整体精度。 $mAP$  越大, 则表示模型的整体精度越高。

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$R = \frac{TP}{TP + FN} \quad (7)$$

$$AP = \int_0^1 P(R) dR \quad (8)$$

$$mAP = \sum_{i=1}^K AP_i / K \quad (9)$$

式中:  $P$ ——精度;

$R$ ——召回率;

$TP$ ——真正例;

$FN$ ——假反例;

$FP$ ——假正例;

$AP$ ——不同召回率下准确率平均值;

$K$ ——总类别数。

## 2.3 消融试验

为验证本文所提出的改进方法的有效性和准确性, 进行消融试验, 如表 2 所示。以 YOLOv8s 作为基线模型, 逐步融入 BiFormer 注意力机制和 VanillaNet 模块, 并逐一对比分析改进点。从表 2 可知, 使用 BiFormer 替换 YOLOv8 骨干网络 (Backbone) 当中的 C2f 模块后,  $mAP@0.5$  提高了 3.2%,  $mAP@0.5:0.95$  提高了 4.7%, 证明了引入 BiFormer 注意力机制改善

YOLOv8 对小目标检测性能不足的可行性。此外, 在 Neck 部分引入 VanillaNet 模块, 相比于基线模型,  $mAP@0.5$  提高了 2.4%,  $mAP@0.5:0.95$  提高了 5.2%。并且, 因为其极简的设计,  $GFLOPs$  相比于基线模型降低了 16%。在融入 BiFormer 的基础上, 引入 VanillaNet 模块后, 不仅  $mAP@0.5$  达到了 93.8%, 提高了 1.4%, 而且缓解了融入注意力机制带来的计算量问题, 使模型  $GFLOPs$  降低了 14%, 进一步证明了引入 VanillaNet 减少计算量、提高精度的可行性。

表 2 消融对比试验

Tab. 2 Ablation comparative experiment

YOLOv8	BiFormer	VanillaNet	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	$GFLOPs$
✓			89.2	66.5	28.4
	✓		92.4	71.2	35.5
		✓	91.6	71.7	23.6
✓	✓	✓	93.8	71.5	30.4

## 2.4 对比试验

将改进后的 YOLOv8 算法与其他常用的目标检测算法 (包括 YOLOv5、YOLOv7、SSD 和 Faster R-CNN) 进行比较, 以进一步验证改进算法的有效性和其在检测性能方面的出众表现。试验采用了相同的数据集划分, 并在相同的试验平台上进行。

如表 3 所示, 改进后的 YOLOv8 相比基线模型,  $mAP@0.5$  提高了 4.6%,  $mAP@0.5:0.95$  提高了 5%。相比于其他的单阶段算法 YOLOv7、YOLOv5 和 SSD,  $mAP@0.5$  也分别提高了 16.4%、3.2% 和 5%。此外, 与两阶段算法 Faster R-CNN 相比, 改进后的 YOLOv8 在  $mAP@0.5$  上提高了 5.8%。在  $GFLOPs$  方面, 改进后的 YOLOv8 比起 YOLOv7、SSD 和 Faster R-CNN, 具有更低的计算复杂度。相比 YOLOv5 和基线模型, 改进后的 YOLOv8 虽然引入了一定的计算量, 但能够提供更可靠和精确的检测结果。相比当前主流算法, 本文提出的算法显然有更出色的检测性能。

表 3 对比试验

Tab. 3 Contrast experiment

模型	$mAP@0.5/\%$			$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	$GFLOPs$
	未成熟	半成熟	成熟草莓			
YOLOv8	89.1	83.9	94.8	89.2	66.5	28.4
YOLOv7	75.3	75.8	81	77.4	45	103.2
YOLOv5	89.2	87.8	95	90.6	69.2	15.8
Faster R-CNN	78.5	91	94.6	88	—	370.2
SSD	86.3	87.5	92.6	88.8	—	62.7
本文算法	91.9	94	95.4	93.8	71.5	30.4

## 2.5 检测效果对比

为了更好地观察本文提出的算法的检测效果,从数据集中选取部分图片进行测试。如图 6 所示,展示了测试集人工标注图像和改进前后的检测结果。可以看出,基线模型 YOLOv8 对于小目标草莓的检测效果存在一定不足,存在漏检和误检现象。改进后的算法拥有更出色的检测性能,得益于 BiFormer 注意力机制,改进后的算法对小目标的检测效果得到了有效提升,能在复杂背景的干扰下准确的识别小目标草莓。

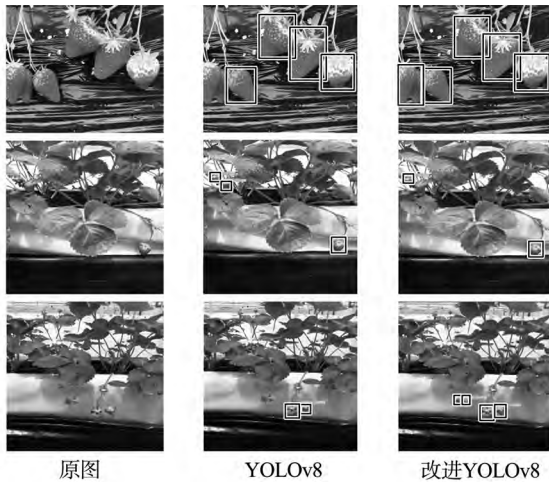


图 6 检测结果对比

Fig. 6 Comparison of test results

## 3 结论

1) 针对温室环境下采集的草莓数据集,使用 LabelImg 对数据集进行标注;针对 YOLOv8 在小目标检测方面存在的不足,在 YOLOv8 骨干网络(Backbone)部分融入 BiFormer 注意力机制,来替换原骨干部分中的 C2f 模块,使其对于草莓检测中的小目标更加敏感,并有效降低漏检和误检情况的发生;最后,在 Neck 部分,引入 VanillaNet 模块,进一步提高算法的整体检测精度,并降低融入 BiFormer 注意力机制所带来的计算量问题。

2) 通过消融试验来看,在 YOLOv8 s 的基础上分别引入 BiFormer 与 VanillaNet,  $mAP@0.5$  分别提高 3.2% 和 2.4%,在融入 BiFormer 的基础上,再引入 VanillaNet 模块,  $mAP@0.5$  提高 1.4%,并且使模型 GFLOPs 降低 14%,验证本文提出的改进方案的有效性。

3) 对比试验表明,相比 YOLOv8,改进后的算法在  $mAP@0.5$  和  $mAP@0.5:0.95$  上分别提升 4.6% 和 5%。相比其他的检测算法 YOLOv7、YOLOv5、SSD 和 Faster R-CNN,  $mAP@0.5$  也分别提高 16.4%、3.2%、5% 和 5.8%。本文提出的算法具备更加出色的检测性能,并且能够有效提升对小目标的检测效果。

4) 改进后的 YOLOv8 算法在整体精度上表现更

好,并且对于小目标草莓的检测性能也更出色,这是与其他常用目标检测算法(例如 YOLOv5、YOLOv7、SSD 和 Faster R-CNN)相比的优势。然而,改进后的 YOLOv8 具有一定的计算量。因此,在后续的工作中,应继续降低算法的计算复杂度,使其更适用于嵌入式设备,且减少计算资源的需求。此外,还应进一步增加更多具有复杂场景的草莓图像,以丰富训练数据集,从而进一步提升算法的泛化性和实用性。

## 参 考 文 献

- [1] Hernández-Martínez N R, Blanchard C, Wells D, et al. Current state and future perspectives of commercial strawberry production: A review [J]. Scientia Horticulturae, 2023, 312: 111893.
- [2] Zhou C, Hu J, Yue J, et al. A novel greenhouse-based system for the detection and plumpness assessment of strawberry using an improved deep learning technique [J]. Frontiers in Plant Science, 2020, 11: 559.
- [3] Munera S, Amigo M J, Blasco J, et al. Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging [J]. Journal of Food Engineering, 2017, 214: 29-39.
- [4] Malik H M, Zhang T, Li H, et al. Mature tomato fruit detection algorithm based on improved HSV and watershed algorithm [J]. IFAC Papersonline, 2018, 51 (17): 431-436.
- [5] 马惠玲,王若琳,蔡骋,等. 基于高光谱成像的苹果品种快速鉴别[J]. 农业机械学报, 2017, 48(4): 305-312.  
Ma Huiling, Wang Ruolin, Cai Cheng, et al. Rapid identification of apple varieties based on hyperspectral imaging [J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(4): 305-312.
- [6] Indrabayu I, Arifin N, Areni I S. Strawberry ripeness classification system based on skin tone color using multi-class support vector machine [C]. 2019 International Conference on Information and Communications Technology (ICOIACT), 2019: 191-195.
- [7] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25.
- [8] Appe S N, Arulselvi G, Balaji G N. CAM-YOLO: tomato detection and classification based on improved YOLOv5 using combining attention mechanism [J]. PeerJ Computer Science, 2023, 9: e1463.
- [9] 张小花,李浩林,李善军,等. 基于 EfficientDet-D1 的草莓快速检测及分类[J]. 华中农业大学学报, 2022, 41 (6): 262-269.



- Zhang Xiaohua, Li Haolin, Li Shanjun, et al. Rapid detection and classification of strawberries based on EfficientDet-D1 [J]. Journal of Huazhong Agricultural University, 2022, 41(6): 262–269.
- [10] 孙俊, 陈义德, 周鑫, 等. 快速精准识别棚内草莓的改进 YOLOv4-Tiny 模型[J]. 农业工程学报, 2022, 38(18): 195–203.
- Sun Jun, Chen Yide, Zhou Xin, et al. Fast and accurate recognition of the strawberries in greenhouse based on improved YOLOv4-Tiny model [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(18): 195–203.
- [11] Bharman P, Saad A S, Khan S, et al. Deep learning in agriculture: A survey [J]. Asian Journal of Research in Computer Science, 2022, 13: 28–47.
- [12] Zhu L, Wang X, Ke Z, et al. BiFormer: Vision transformer with bi-level routing attention [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 10323–10333.
- [13] Chen H, Wang Y, Guo J, et al. VanillaNet: The power of minimalism in deep learning [J]. Advances in Neural Information Processing Systems, 2024, 36.
- [14] Terven J, Córdova-Esparza D M, Romero-González J A. A comprehensive review of yolo architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS [J]. Machine Learning and Knowledge Extraction, 2023, 5(4): 1680–1716.
- [15] Soylu E, Soylu T. A performance comparison of YOLOv8 models for traffic sign detection in the robotaxi-full scale autonomous vehicle competition [J]. Multimedia Tools and Applications, 2023, 83(8): 25005–25035.
- [16] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904–1916.
- [17] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector [C]. Computer Vision—ECCV 2016: 14th European Conference, 2016: 21–37.
- [18] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. Advances in Neural Information Processing Systems, 2015, 28.
- [19] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30.
- [20] Ren S, Zhou D, He S, et al. Shunted self-attention via multi-scale token aggregation [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 10853–10862.
- [21] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]. International Conference on Machine Learning. Pmlr, 2015: 448–456.

(上接第 273 页)

- [13] 万祖毅, 田永中, 刘旭东, 等. 基于轻小型无人机影像的柑橘树株数提取[J]. 湖北农机化, 2020(7): 140–144.
- [14] 胡馨月, 倪海明, 戚大伟. 基于无人机影像的树木株数提取[J]. 森林工程, 2021, 37(1): 6–12.
- [15] Mukhtar H, Khan M Z, Khan M, et al. Wheat plant counting using UAV images based on semi-supervised semantic segmentation [C]. 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), 2021.
- [16] Luthfan N H, Tomoya W, Tsutomu M, et al. Machine learning techniques to predict soybean plant density using UAV and satellite-based remote sensing [J]. Remote Sensing, 2021, 13: 2548.
- [17] 付虹雨, 崔国贤, 崔丹丹, 等. 基于无人机遥感影像的亚麻株数识别[J]. 中国麻业科学, 2020, 42(6): 249–256.
- [18] Reza M N, Na I S, Lee K H. Automatic counting of rice plant numbers after transplanting using low altitude UAV images [J]. International Journal of Contents, 2017, 13.
- [19] 牛亚晓, 张立元, 韩文霆, 等. 基于无人机遥感与植被指数的冬小麦覆盖度提取方法[J]. 农业机械学报, 2018, 49(4): 212–221.
- Niu Yaxiao, Zhang Liyuan, Han Wenting, et al. Fractional vegetation cover extraction method of winter wheat based on UAV remote sensing and vegetation index [J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(4): 212–221
- [20] 应兴亮, 黄菲, 余启源. 基于无人机遥感植被指数的松材线虫病枯死木自动提取研究[J]. 广东蚕业, 2023, 57(1): 57–61.
- [21] 敖登, 杨佳慧, 丁维婷, 等. 54种植被指数研究进展综述[J]. 安徽农业科学, 2023, 51(1): 13–21, 28.
- Ao Deng, Yang Jiahui, Ding Weiting, et al. Review of 54 vegetation indices [J]. Journal of Anhui Agricultural Sciences, 2023, 51(1): 13–21, 28.
- [22] 韩青松, 贾振红, 杨杰, 等. 基于改进的Otsu算法的遥感图像阈值分割[J]. 激光杂志, 2010, 31(6): 33–34.
- Han Qingsong, Jia Zhenhong, Yang Jie, et al. Remote sensing image thresholding segmentation based on the modified Otsu algorithm [J]. Laser Journal, 2010, 31(6): 33–34.