

黄智超, 王丽丽. 改进 YOLOv11s 的农作物虫害标注与检测方法研究[J]. 智能计算机与应用, 2026, 16(2): 77-83. DOI: 10.20169/j. issn. 2095-2163. 25042401

改进 YOLOv11s 的农作物虫害标注与检测方法研究

黄智超^{1,2}, 王丽丽²

(1 吉林化工学院 信息与控制工程学院, 吉林 吉林 132022; 2 德州学院 计算机与信息学院, 山东 德州 253000)

摘要: 农作物虫害检测是现代农业中的核心技术之一, 对于提高农业生产效率和保护粮食安全具有重要意义。本文基于 IP102 数据集, 构建了一个更贴合实际应用的数据子集。在 YOLOv11s 目标检测框架基础上, 融合了双向特征金字塔网络 (BiFPN)、多尺度注意力聚合 (MSAA) 以及焦点交并比 (Focal IoU), 以提升模型的多尺度特征融合能力和小目标检测精度。此外, 本文提出了一种自动标注方法, 利用 T-Rex Label 结合 YOLOv11s 模型, 不仅节省了人工标注的时间, 还通过半自动标注流程显著提高了标注效率和数据质量, 从而有效解决了大规模数据集的标注问题。实验结果表明, 改进后的模型在平均精度、召回率上均优于 YOLOv5s, YOLOv5-EB, YOLOv5-GRNS, YOLOv8s, YOLOv10s, YOLOv11s 等算法, 为农作物虫害检测提供了一种准确、高效的解决方案。

关键词: 农作物虫害检测; 自动标注; YOLOv11s; BiFPN; MSAA; Focal IoU

中图分类号: TP319.41 **文献标志码:** A **文章编号:** 2095-2163(2026)02-0077-07

Study on enhanced annotation and detection techniques for crop pest identification using YOLOv11s

HUANG Zhichao^{1,2}, WANG Lili²

(1 School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 132022, Jilin, China;

2 School of Computer and Information, Dezhou University, Dezhou 253000, Shandong, China)

Abstract: Crop pest detection is one of the core technologies in modern agriculture and is of great significance for improving agricultural production efficiency and ensuring food security. The paper applies the IP102 dataset and constructs a data subset that is more suitable for practical applications. Based on the YOLOv11s object detection framework, it integrates Bidirectional Feature Pyramid Network (BiFPN), Multi-Scale Attention Aggregation (MSAA), and Focal IoU to enhance the model's multi-scale feature fusion capability and improve the accuracy of small object detection. In addition, this paper proposes an automatic annotation method. The method uses T-Rex Label in combination with the YOLOv11s model, which not only saves the time spent on manual labeling, but also significantly improves labeling efficiency and data quality through using the semi-automatic labeling process, effectively solving the problem of large-scale datasets annotation. The experimental results show that the improved model outperforms YOLOv5s, YOLOv5-EB, YOLOv5-GRNS, YOLOv8s, YOLOv10s, YOLOv11s, and other algorithms in terms of average precision and recall, providing an accurate and efficient solution for crop pest detection.

Key words: crop pest detection; automatic annotation; YOLOv11s; BiFPN; MSAA; Focal IoU

0 引言

随着全球农业生产的不断发展, 虫害问题成为制约农业生产的重要因素之一。虫害影响了作物的产量和质量, 从而导致经济损失。联合国粮食及农业组织在 2023 年发布的一份报告显示^[1], 全球每年约有四成农作物因虫害造成减产。传统人工检测方法难以满足农业生产的需求, 亟需更高效、精准的检

测手段来应对虫害问题。随着科技的不断进步, 机器学习技术也逐渐在各研究领域得到广泛应用, 这就为虫害检测提供了新的思路和方法。

机器学习主要依赖人工提取特征。例如, 李文勇等学者^[2]提出一种基于颜色和纹理等与形态无关的特征相结合并利用多类支持向量机分类器的多姿态害虫分类方法。翟治芬等学者^[3]运用最大类间方差阈值分割与多颜色分量组合技术, 实现作物

作者简介: 黄智超(1999—), 男, 硕士研究生, 主要研究方向: 人工智能技术及应用。

通信作者: 王丽丽(1981—), 女, 博士, 教授, 主要研究方向: 农业信息化, 智能算法研究。Email: jsjxwll@126.com。

收稿日期: 2025-04-24

和土壤分离及病斑分割,再用分水岭分割法分离粘连棉叶,提取盲椿象危害区域的颜色、纹理和形状特征,最后结合朴素贝叶斯分类器,根据棉花盲椿象危害等级标准来识别危害等级。Ebrahimi 等学者^[4]采用差分核函数支持向量机方法进行草莓植株上的寄生虫分类和蓟马检测。这些机器学习的方法在性能上依赖于特征提取算法,且需要耗费较长的时间。

近年来,随着人工智能和计算机视觉技术的发展,深度学习在农作物虫害检测中展现出较好的性能^[5-6]。常见的算法有 Faster R-CNN^[7]和 Cascade R-CNN^[8]。Wang 等学者^[9]改进 Faster R-CNN 模型,采用 ResNet-101 网络提取出害虫的特征信息,并更改 anchor 的尺寸改进区域建议网络,更好地检测小目标尺寸害虫。刘凯旋等学者^[10]改进 Cascade R-CNN 模型对水稻害虫进行检测,利用特征金字塔网络融合多尺度特征图,并将 ROI Pooling 替换成 ROI Align,减少了水稻害虫信息的丢失,提高对水稻害虫的检测准确率。这些方法计算开销较大,难以满足农业生产环境中的实时检测需求,在复杂环境下的检测性能有待提高。

YOLO 系列算法^[11]因其检测速度快、实时性高,更适合农作物虫害检测场景。王梦妮等学者^[12]在 YOLOv5s 的基础上,将 SPPF 替换为 ASPP,来捕获更广泛的上下文信息,适应不同尺寸的虫害目标,引入 BiFPN^[13]和 CBAM,不仅提高特征融合的效率,还提高模型关注小目标特征的能力,但对于密集目标可能会引入冗余信息,导致特征表达不够紧凑。张荣华等学者^[14]提出在 YOLOv8 基础上,引入 GSConv 提高模型的感受野,部分 Conv 替换为 Ghost Convolution,采用 HorBlock 捕捉更长期的特征依赖关系,引入 BiFPN,使用 VoVGSCSP 模块提升微小目标检测,引入 CBAM 来强化田间虫害目标特征,使用 Wise-IoU 损失函数更多地关注普通质量样本,提高网络模型的泛化能力和整体性能。Guan 等学者^[15]在 YOLOv10 的基础上,引入 BiFPN 提高对不同尺度目标的检测性能,引入 SEAM 提高复杂环境下的特征提取能力,引入 GCNet 更准确地处理复杂信息,但对于不同种类的害虫,模型的分类精度存在差异。王昕等学者^[16]通过改进 CBAM 中的通道注意力机制和替换 Focus 操作,提高模型对小目标特征的提取能力,平均精度值达到了 87%。郭嘉璇等学者^[17]通过设计融入 GRN 注意力机制的编码器模块,提高对密集目标的识别精度,并利用形状交并比 (SIoU) 损失函数加快了模型收敛速度,其全类别平

均准确率在 IP8-CW 数据集上达到了 72.3%。这些方法在处理小目标和密集目标检测精度方面,仍有提升空间。

为了解决这些问题,本文提出一种基于改进 YOLOv11s 网络的农作物虫害检测方法,进一步提高检测的准确性和效率。具体改进如下:

(1) 引入 BiFPN 增强多尺度特征融合能力。BiFPN 的自顶向下和自底向上的双向特征融合路径,能有效提高特征表达能力和计算效率,对大小不一的虫害目标尤为适用,可增强模型对小目标和复杂背景的检测能力。

(2) 引入 MSAA^[18]将多尺度特征提取与注意力机制相结合。MSAA 能有效增强对图像中不同尺寸目标的感知能力,通过生成不同尺度的注意力图对特征进行加权处理,突出关键信息,抑制背景干扰,从而提高检测精度。

(3) 采用 Focal IoU^[19]替换原损失函数以更好处理类别不平衡问题,减少背景干扰。Focal IoU 引入聚焦机制,增强对难例(如小物体和遮挡物)的学习,进一步提升模型在复杂场景下的精度和鲁棒性。

1 设计与方法

1.1 YOLOv11s 概述

YOLOv11s 的网络结构主要由主干网络、颈部网络、检测头组成。

主干网络负责提取图像的多尺度特征,包含卷积层、C3k2 模块、SPPF 模块、C2PSA 模块。卷积层结合高效的计算模块,在降低计算复杂度的同时,实现相同、甚至更优的特征提取效果。C3k2 模块以 2 个小卷积核(kernel size = 2)替代大卷积核,提升计算效率,减少模型参数量,增强特征表达能力。SPPF 模块通过多尺度池化操作提取全局特征,利于识别大目标或全局模式,增强了模型的整体识别能力。C2PSA 模块融合 CSP 和空间注意力机制,增强对图像特定区域的关注,提升小目标和复杂场景检测能力。图 1 呈现了 C3k2、C2PSA 和 SPPF 模块的组成单元、连接模式及特征交互机制。

颈部网络负责融合不同尺度特征并传递至预测头部,由 PANet 和 C3k2 模块组成。YOLOv11s 对 PANet 进行优化,使高低层特征间信息传递更高效,增强处理小目标和复杂场景的鲁棒性。通过多次上采样和拼接操作,整合不同层特征,C3k2 模块进一步提升特征融合效率和精度,C2PSA 模块增强空间注意力机制作用,聚焦关键区域。

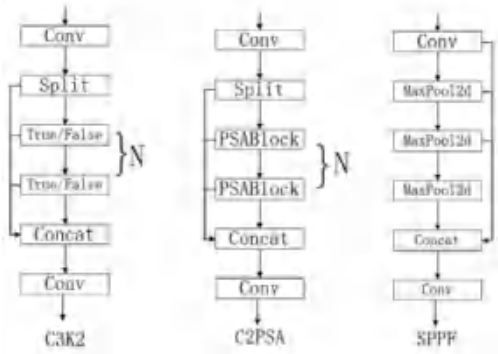


图 1 C3k2、C2PSA、SPPF 结构图

Fig. 1 Structural diagrams of the C3k2, C2PSA, and SPPF modules

检测头负责处理融合特征并生成检测结果。C3k2 模块进一步提取和处理多尺度特征,压缩增强特征图,为最终检测提供高质量输入。每条检测路径末端的卷积层输出目标框坐标、目标性得分及类别得分,检测层整合这些结果,输出完整目标检测信息。图 2 展示 YOLOv11s 的整体架构,包括主干网络、颈部网络和检测头的层次关系与连接方式。

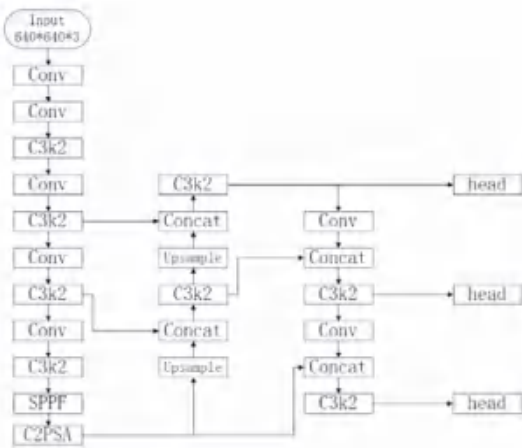


图 2 YOLOv11s 网络结构图

Fig. 2 Architecture of the YOLOv11s network

1.2 改进 YOLOv11s 模型结构

为解决农作物虫害检测中的目标数量多、背景复杂、小目标难以检测等问题,本文对 YOLOv11s 网络结构进行了优化,主要改进如下:引入 BiFPN;引入 MSAA;采用 Focal IoU 作为损失函数。改进后网络结构如图 3 所示。

1.2.1 BiFPN

BiFPN 通过双向特征融合路径和加权特征融合模块,有效提升了特征金字塔的表达能力和计算效率。BiFPN 的双向特征融合路径包括自顶向下和自底向上的过程。在自顶向下的路径中,对不同尺度的特征进行上采样,并将上采样后的特征与低层特

征进行融合,有助于保留图像的细节信息;在自底向上的路径中,反向传递信息,进一步融合低层和高层特征,使得模型能够充分利用高层的语义信息,从而增强对多尺度目标的识别能力。BiFPN 在每一层的特征图融合过程中采用加权求和的方法,而非传统 FPN 中的简单加法或连接操作。通过为每个特征层分配一个权重,模型能够更加灵活地融合不同尺度的特征,避免信息丢失或特征不均衡的问题。这种加权特征融合模块能够根据特征的重要性动态调整融合方式,进一步提升了模型对不同尺寸物体的识别能力。

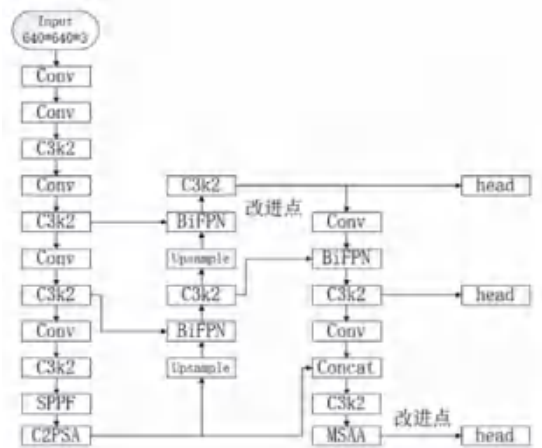


图 3 改进模型结构图

Fig. 3 Structural diagram of the enhanced YOLOv11s - based detection model

在农作物虫害检测任务中,引入 BiFPN 更有效地处理小目标和复杂背景。在检测较小的虫害目标时,BiFPN 能够通过自顶向下的路径保留足够的细节信息,同时通过自底向上的路径充分利用高层的语义信息,从而更准确地识别出小目标。此后,加权特征融合模块能够根据图像的不同区域动态调整特征融合的权重,使得模型在复杂背景中能够更好地聚焦于虫害目标,提高检测的准确性和鲁棒性。BiFPN 结构如图 4 所示。

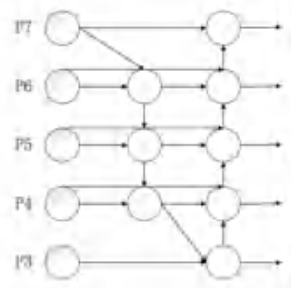


图 4 BiFPN 结构图

Fig. 4 Architecture of the BiFPN module

而实现高效的半自动标注流程。T-Rex Label 标注结果如图 7 所示, T-Rex Label 工具所生成的半自动标注结果在标注效果上与 labellmg 的人工标注相近, 但在标注效率方面具有优势。然而, 在对图像进行目标计数时, T-Rex Label 并不能保证每张图像中的目标都被完全识别, 仍存在一定的人工补充标注工作量。因此, 本文结合 YOLOv11s 模型, 设计并实现了一种标注效率更高的半自动标注工具。

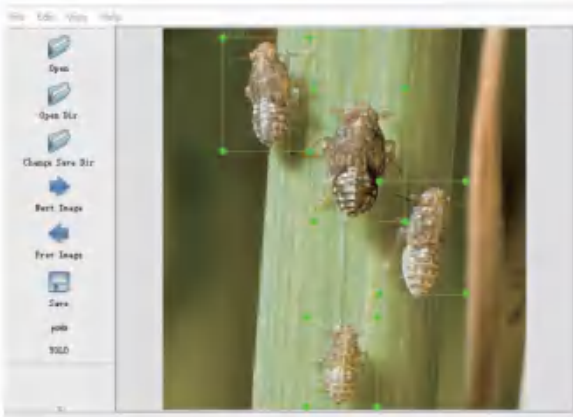


图 6 labellmg 标注图

Fig. 6 Annotation example using labellmg

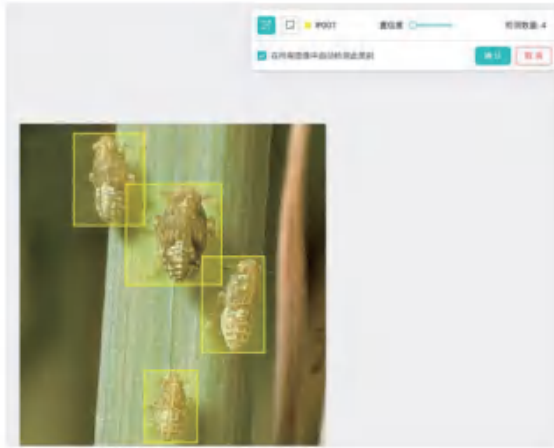


图 7 T-Rex Label 标注图

Fig. 7 Illustration of the annotation process with T-Rex Label

(2) 自动标注扩展。将已有的标注数据作为训练集, 用于训练 YOLOv11s 模型。利用该模型对剩余未标注的数据进行预测, 自动生成标注结果。自动标注结果如图 8 所示。由图 8 可知, 模型生成的标注效果与人工标注相近, 但整体效率更高。除了生成标注图外, 系统还会同步生成对应的标签文件。用户可通过查看标注结果图, 筛选出需要修改的样本, 并借助 labellmg 工具进行修正。通过这种方式, 人工工作量被大幅减少, 仅需进行少量审核与校正, 就可有效提升标注数据的质量与准确性。

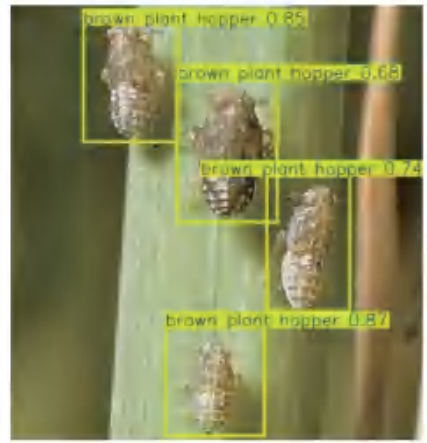


图 8 自动标注图

Fig. 8 Results of automatic annotation

2.3 实验环境及评价指标

操作系统为 Linux, CPU 型号为 Intel(R) Core (TM) i7-11700F @ 2.50 GHz, GPU 型号为 NVIDIA GeForce RTX 3080 Ti, 运行内存为 64 GB, 8 TB 机械硬盘, 编程语言为 Python3. 10. 12, 深度学习框架为 Pytorch2. 5. 1, GPU 加速库为 CUDA 12. 4、CUDNN 9. 1. 0。训练过程中的模型参数设置: 分辨率为 640×640, 批量大小为 16, 迭代轮数为 100, 优化器为 SGD, 初始学习率为 0. 01。

选用准确率、召回率、平均精度均值 (mAP) 作为评估模型性能的指标。

2.4 不同模型检测结果对比

为了验证改进模型的有效性, 本文选用 YOLOv5s、YOLOv8s、YOLOv10s、YOLOv11s、YOLOv5-EB^[16] 以及 YOLOv5-GRNS^[17], 与改进模型在同一个训练集下做对比。指标为准确率、召回率、平均精度均值, 不同模型的实验结果见表 1。表 1 中数据显示, 改进后的模型准确率、召回率以及平均精度均值分别为 85. 9%、85. 8%、89. 2%。相比于 YOLOv5s、YOLOv5-EB、YOLOv11s, 模型准确率提升了 1. 2 个百分点、1. 8 个百分点、1. 3 个百分点, 召回率提升了 1. 6 个百分点、9. 7 个百分点、1. 4 个百分点, 平均精度均值提升了 1. 6 个百分点、5. 7 个百分点、0. 8 个百分点。与 YOLOv8s、YOLOv10s、YOLOv5-GRNS 比, 虽然准确率分别下降 1. 0 个百分点、1. 2 个百分点、0. 6 个百分点, 但召回率上升 4. 8 个百分点、5. 2 个百分点、3. 5 个百分点, 平均精度均值上升 0. 6 个百分点、1. 1 个百分点、0. 5 个百分点。

从模型的整体性能上看, 改进后的模型在农作物虫害检测方面具有更高的召回率和平均精度均值优势, 能够对害虫进行更为精确的识别。具体分析

如下:

(1)BiFPN 的贡献。BiFPN 通过双向特征融合,提升了模型对多尺度目标的检测能力。

(2)MSAA 的贡献。MSAA 引入多尺度注意力机制,使模型在处理不同大小的虫害时自动强调重要特征,减少背景噪声的干扰,从而提高了检测精度。

(3)Focal IoU 的贡献。Focal IoU 通过聚焦难以识别的虫害,降低了背景对训练的影响,使模型更专注于处理小物体、遮挡物和难以识别的病虫草害。

表 1 不同模型的实验结果

Table 1 Experimental results comparison of different models %

模型	准确率	召回率	平均精度均值
YOLOv5s	84.7	84.2	87.6
YOLOv8s	86.9	81.0	88.6
YOLOv10s	87.1	80.6	88.1
YOLOv11s	84.6	84.4	88.4
YOLOv5-EB ^[16]	84.1	76.1	83.5
YOLOv5-GRNS ^[17]	86.4	82.3	88.7
改进模型	85.9	85.8	89.2

2.5 消融实验

为了验证改进模块对模型性能贡献,本文设计了4组消融实验,分别在YOLOv11s基础上依次引入BiFPN、MSAA和Focal IoU。通过平均精度均值(mAP)对各组模型的性能进行评估,实验结果见表2。

表 2 消融实验结果

Table 2 Results of ablation experiment

模型	平均精度均值/%
YOLOv11s	88.4
YOLOv11s+BiFPN	89.0
YOLOv11s+BiFPN+MSAA	89.2
YOLOv11s+BiFPN+MSAA+Focal IoU	89.4

通过表2可以看出,YOLOv11s, YOLOv11s+BiFPN, YOLOv11s+BiFPN+MSAA, YOLOv11s+BiFPN+MSAA+Focal IoU依次引入的模块,使平均精度均值依次提升了0.6、0.2、0.2个百分点。

3 结束语

(1)在数据标注方面,本文提出了一种结合T-Rex Label工具和YOLOv11s模型的半自动标注方法。该方法通过初始人工标注、自动标注扩展的流程,大幅减少了人工标注时间,并通过人工审核确保了标注的高准确性。这种标注方案提高了标注效率,为大规模数据集的标注提供了高效解决方案。

(2)本文针对农作物虫害检测中存在的小目标

检测困难、背景复杂等问题,提出了一种基于改进YOLOv11s的检测方法。在检测方法上,通过在YOLOv11s网络结构中引入BiFPN、MSAA和Focal IoU,提升了模型对多尺度目标的检测能力和复杂场景下的鲁棒性。实验结果表明,改进后的模型平均精度均值(mAP)达到89.2%,相比原YOLOv11s提升了0.8个百分点,且在召回率和平均精度上均优于YOLOv5s、YOLOv8s和YOLOv10s、YOLOv5-EB、YOLOv5-GRNS等模型,验证了其在农作物虫害检测中的高效性和准确性。

未来工作将进一步优化模型性能,探索更广泛的数据集和应用场景,以推动农作物虫害检测技术的持续发展。

参考文献

- [1] 联合国粮食及农业组织.《粮食与农业状况2023》[R]. 罗马:联合国粮食及农业组织,2023.
- [2] 李文勇,李明,陈梅香,等.基于机器视觉的作物多姿态害虫特征提取与分类方法[J].农业工程学报,2014,30(14):154-162.
- [3] 翟治芬,徐哲,周新群,等.基于朴素贝叶斯分类器的棉花盲椿象危害等级识别[J].农业工程学报,2015,31(1):204-211.
- [4] EBRAHIMI M A, KHOSHTAGHAZA M H, MINAEI S, et al. Vision-based pest detection based on SVM classification method[J]. Computers and Electronics in Agriculture, 2017, 137: 52-58.
- [5] 刘浏.基于深度学习的农作物害虫检测方法研究与应用[D].合肥:中国科学技术大学,2020.
- [6] 彭琳希.基于深度学习的农作物害虫检测方法研究与应用[J].信息与电脑(理论版),2021,33(19):76-79.
- [7] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [8] CAI Z, VASCONCELOS N. Cascade R-CNN: High quality object detection and instance segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(5):1483-1498.
- [9] WANG Tiewei, ZHAO Longgang, LI Baohua, et al. Recognition and counting of typical apple pests based on deep learning[J]. Ecological Informatics, 2022, 68: 101556.
- [10] 刘凯旋,黄操军,李亚鹏,等.一种基于级联R-CNN的水稻害虫检测算法[J].黑龙江八一农垦大学学报,2021,33(5):106-111.
- [11] JEGHAM N, KOH C Y, ABDELATTI M, et al. Evaluating the evolution of YOLO (You Only Look Once) models: A comprehensive benchmark study of YOLO11 and its predecessors[J]. arXiv preprint arXiv,2411.00201, 2024.
- [12] 王梦妮,顾寄南,王化佳,等.基于改进YOLOv5s模型的茶叶嫩芽识别方法[J].农业工程学报,2023,39(12):150-157.
- [13] TAN Mingxing, PANG Ruoming, LE Q V. EfficientDet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 10781-10790.
- [14] 张荣华,白雪,樊江川.复杂场景下害虫目标检测算法:YOLOv8-Extend[J].智慧农业,2024,6(2):49-61.

- [15] GUAN Sitong, LIN Yiming, LIN Guoyu, et al. Real-time detection and counting of wheat spikes based on improved YOLOv10[J]. *Agronomy*, 2024, 14(9): 1936.
- [16] 王昕, 董琴, 杨国宇. 基于优化 CBAM 改进 YOLOv5 的农作物病虫害识别[J]. *计算机系统应用*, 2023, 32(7): 261-268.
- [17] 郭嘉璇, 王蓉芳, 南江华, 等. 融入全局相应归一化注意力机制的 YOLOv5 农作物害虫识别模型[J]. *农业工程学报*, 2024, 40(8): 159-170.
- [18] LIU Mushui, DAN Jun, LU Ziqian, et al. CM-UNet: Hybrid CNN - Mamba UNet for remote sensing image semantic segmentation[J]. *arXiv preprint arXiv*, 2405. 10530, 2024.
- [19] ZHANG Yifan, REN Weiqiang, ZHANG Zhang, et al. Focal and efficient IoU loss for accurate bounding box regression [J]. *Neurocomputing*, 2022, 506: 146-157.
- [20] WU Xiaoping, ZHAN Chi, LAI Yukun, et al. Ip102: A large-scale benchmark dataset for insect pest recognition [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2019: 8787-8796.
- [21] MENG SUWAN K, PALACIO J C R, RYO M. ChatGPT and general-purpose AI count fruits in pictures surprisingly well[J]. *arXiv preprint arXiv*, 2404. 08515, 2024.