

基于 YOLO - SVP 的小尺寸蔬菜害虫检测模型研究

王春桃^{1,2} 谢伟斌¹ 肖德琴^{1,2}

(1. 华南农业大学数学与信息学院, 广州 510642; 2. 农业农村部华南热带智慧农业技术重点实验室, 广州 510642)

摘要: 有效的害虫监测对高品质蔬菜栽培至关重要。基于深度学习的害虫检测方法在大、中型害虫识别上表现出色,但其在小尺寸害虫检测方面仍面临挑战。为此,本文提出一种基于 YOLO 算法的小尺寸蔬菜害虫检测方法 (YOLO - SVP)。为强化关键小尺寸害虫特征并改进特征融合效果,提出了一种动态加权注意力 (DWA) 机制,并将其整合至 YOLO 11 的 C3k2 模块中,形成 C3k2 - DWA 模块。此外,为了在下采样过程中保留关键空间信息并减少小尺寸害虫特征损失,提出了一种空间到深度下采样 (SPD - Down) 模块。同时,为缓解小尺寸害虫检测中边界框回归的严重不足,引入归一化瓦瑟斯坦距离 (NWD) 损失函数。基于自建蔬菜害虫数据集进行了仿真试验,验证了所提 YOLO - SVP 方法的有效性。其 F1 值达 85.7%、mAP₅₀ 达 89.3%、mAP_{50,95} 达 54.9%;相较于 YOLO 11 基线模型,分别提高 4.5、3.8、4.3 个百分点。对于小尺寸害虫西花蓟马,其检测的 F1 值、mAP₅₀ 和 mAP_{50,95} 分别提升 6.3、8.5、5.0 个百分点。研究结果为适应精细农业中具有挑战性的小目标检测任务提供了一种深度学习架构改进范式,为蔬菜害虫有效监测提供重要支撑。

关键词: 小尺寸害虫检测; YOLO 11; 动态加权注意力; 空间到深度下采样

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2026)05-0364-09

OSID:



Detection Model for Small-sized Vegetable Pests Based on YOLO - SVP

WANG Chuntao^{1,2} XIE Weibin¹ XIAO Deqin^{1,2}

(1. College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China

2. Key Laboratory of Smart Agricultural Technology in Tropical South China, Ministry of Agriculture and Rural Affairs, Guangzhou 510642, China)

Abstract: Effective pest monitoring is crucial for high-quality vegetable cultivation. While deep learning-based pest detection methods excelling at detecting large- and medium-sized pests, they face challenges with small-sized pests. To address the problem, a you only look once (YOLO)-based small-sized vegetable pest detection method was presented, named YOLO - SVP. To emphasize crucial small-sized pest features and improve feature fusion, a dynamic weighting attention (DWA) mechanism was constructed and integrated into the C3k2 block of YOLO 11, yielding a new block denoted C3k2 - DWA. Additionally, to preserve critical spatial information during downsampling and reduce the loss of small pest features, a space-to-depth downsampling (SPD - Down) block was proposed. Besides, to alleviate the severe weakness of bounding box regression in the case of small pests, the normalized Wasserstein distance (NWD) loss function was introduced. Experimental simulation on a self-built vegetable pest dataset demonstrated the effectiveness of the proposed YOLO - SVP, which achieved 85.7% F1 score, 89.3% mAP₅₀, and 54.9% mAP_{50,95}, outperforming YOLO 11 by 4.5, 3.8, and 4.3 percentages points, respectively. For the *Frankliniella occidentalis* (small-sized pest), the detection performance improved the F1 score, mAP₅₀, and mAP_{50,95} by 6.3, 8.5, and 5.0 percentages points, respectively. This research provided a paradigm for adapting deep learning architectures to challenge small-sized object detection tasks in precision agriculture, which would provide important support for the effective monitoring of vegetable pests.

Key words: small-sized pest detection; YOLO 11; dynamic weighting attention; space-to-depth down sampling

收稿日期: 2025-06-25 修回日期: 2025-08-16

基金项目: 广东省现代农业产业技术体系创新团队项目(2024CXTD21)和国家自然科学基金项目(62172165)

作者简介: 王春桃(1979—),男,教授,博士生导师,主要从事农业人工智能和多媒体信息安全研究,E-mail: wangct@scau.edu.cn

0 引言

农业是国家生存与发展的基石。我国作为传统农业大国,长期受虫害问题困扰。虫害不仅会大幅降低作物产量,更会诱发多种作物病害,严重影响农产品质量,甚至对人类健康构成威胁。为实现从宏观的作物多样化布局^[1]到微观的特定功能植物配置^[2]等生态策略所追求的害虫控制效益,对害虫种群进行有效监测是首要前提。但目前我国大部分地区仍通过诸如田间人工巡查、灯光诱捕器和粘虫板等传统方法捕获害虫,作物保护专家随后对捕获的害虫进行种类识别和数量评估^[3]。然而,这种传统的害虫监测方法劳动成本高、效率低、易发生误检和漏检,在面对大规模农田或微小害虫时更为突出。因此,开发实时、有效的害虫监测方法至关重要。

随着信息技术的发展,基于目标检测的智能害虫识别方法日益涌现。深度学习技术的进步为害虫目标检测带来了创新性解决方案。基于深度学习的目标检测算法,如 Faster R-CNN^[4]、RetinaNet^[5]、SSD^[6]及 YOLO 系列^[7-15],能够自动从图像中提取高级特征,显著提升检测准确性和效率,被广泛用于害虫检测研究^[16-20]。尽管深度学习方法为传统方法的局限性引入了解决方案,但它们在检测小目标害虫时仍面临挑战。为此研究人员提出了多种基于目标检测算法的创新方法,包括改进网络结构、引入注意力机制等^[21-26]。

然而,针对小尺寸害虫检测的研究仍存在不足。本研究数据集涵盖华南地区五大蔬菜害虫,其中,烟粉虱与西花蓟马成虫体长显著微小,导致多数目标检测方法在本数据集上的害虫检测精度受到严重限制。单阶段目标检测算法(如 YOLO 系列)因其结构简单、速度快,已成为害虫检测的主流方法,但其在小尺寸目标检测性中表现欠佳。在不影响大中型目标害虫检测性能的基础上为进一步提高小尺寸目标害虫检测性能,本文基于 YOLO 11 模型进行小尺寸害虫检测优化,提出一个基于动态加权注意力与空间到深度下采样的小尺寸蔬菜害虫检测模型 YOLO-SVP。为验证 YOLO-SVP 模型的有效性,设计并开展了系统的试验评估,主要包括:在自建 VegpestSet 数据集上对改进部分进行消融试验;将 YOLO-SVP 与当前主流 YOLO 系列模型以及专门针对小目标或蔬菜害虫检测的先进模型进行性能对比。

1 数据集构建

1.1 数据采集

采用本团队自建的害虫数据集^[27]。该数据集

于2017年6—8月在广东省广州市南沙区东升农场,通过划分20个独立监测单元并部署20套智能监测设备采集数据。采用尺寸为35 cm×25 cm的黄色粘虫板吸引并捕获害虫,粘虫板每6~7 d更换1次。使用一台1 600万像素的相机拍摄被粘虫板捕获的迁飞性蔬菜害虫。控制盒被设定于每天10:00在自然光下拍摄1幅图像,并自动上传到云服务器。

1.2 数据预处理与标注

在筛选阶段,从收集到的蔬菜害虫图像中手动排除了与害虫无关的图像后,保留396幅3 120像素×4 208像素的高分辨率JPEG图像。为了满足YOLO 11模型输入要求,图像被裁剪为640像素×640像素子图像。本研究聚焦于5种类型蔬菜害虫:小菜蛾、烟粉虱、瓜实蝇、黄曲条跳甲和西花蓟马。表1为其外观和特征详情。

表1 VegpestSet数据集害虫特征和外观

Tab. 1 Characteristics and appearance display of pests in VegpestSet

害虫种类	害虫成虫特征	维基百科图像	数据集图像
小菜蛾	体长约6 mm 灰黑色或赭褐色		
烟粉虱	体长约0.9 mm 体淡黄色,翅白色		
瓜实蝇	体长7~9 mm 黄褐色		
黄曲条跳甲	体长1.8~2.4 mm 黑色		
西花蓟马	体长1~1.7 mm 淡黄色或褐色		

使用LabelImg工具在这些图像中手动标注5类害虫,其他害虫种类和无关背景被视为背景类别,未进行标注。标注信息保存为XML文件,并与图像一起转换为YOLO格式数据集,命名为VegpestSet。图1为部分图像处理过程。

1.3 数据增强

为了增强模型泛化性,对上述制作的数据集应

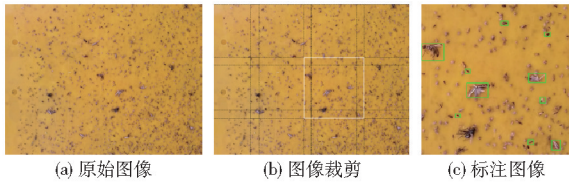


图1 图像处理过程

Fig. 1 Image processing procedure

用多种数据增强技术,模拟图像在尺度、方向、光照和局部遮挡等方面的自然变化,以有效扩展训练数据多样性。

图2为数据增强方法处理结果,通过数据增强方法产生了9 171幅新图像,使数据集图像总量增至23 891幅。将增强后的图像集按比例8:1:1随机选取划分为训练集、验证集和测试集,分别形成19 111、2 389、2 391幅图像。数据集统计信息如表2所示。

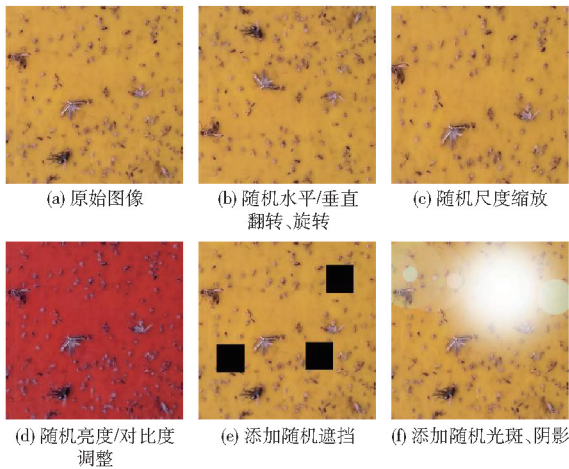


图2 数据增强方式

Fig. 2 Data augmentation methods

表2 VegpestSet数据集害虫标注数量统计

Tab. 2 Statistics of annotated vegetable pests in VegpestSet

害虫类别	训练集	验证集	测试集	类别总数
小菜蛾	9 091	1 110	1 155	11 356
烟粉虱	19 596	2 420	2 406	24 422
瓜实蝇	3 990	541	511	5 042
黄曲采跳甲	22 282	2 892	2 680	27 854
西花蓟马	9 347	1 175	1 234	11 756
总计	64 306	8 138	7 986	80 430

2 YOLO-SVP模型

2.1 YOLO 11目标检测模型

YOLO 11作为YOLO系列的模型,兼具高效性与精确性优势,特别适合实时快速推理任务^[15]。该模型由4部分组成:输入、骨干网络、颈部网络和检测头。将预处理后图像输入骨干网络进行特征提取,后经颈部网络特征融合,生成的特征图像分

别由负责大、中、小3种尺度目标检测的3个专用检测头并行处理。这些检测头输出各目标的边界框坐标与类别置信度,经后处理算法获得最终检测结果。

尽管YOLO 11在公开数据集上检测性能表现卓越,但将其应用于VegpestSet数据集时效果不佳,主要存在问题:①数据集中害虫普遍微小,图像中占比区域极小,限制模型可提取的特征信息量。②图像包含复杂背景干扰,导致模型易受无关细节干扰。③常规下采样操作导致关键特征丢失,进一步削弱小尺寸害虫检测能力。

2.2 小尺寸蔬菜害虫检测模型YOLO-SVP

鉴于YOLO 11在目标检测速度与精度平衡方面的优异表现,选用其作为基准模型。为了使YOLO 11可检测小尺寸害虫,提出改进模型YOLO-SVP,旨在确保大中型蔬菜害虫检测性能不降低的情况下,提升模型小尺寸蔬菜害虫检测能力。

为增强对小目标的特征提取能力,提出C3k2-DWA模块,通过将动态加权注意力机制嵌入原C3k2模块,优化小尺寸害虫特征提取。同时,采用SPD-Down模块替换YOLO 11中的下采样卷积层,以减少信息损失并保留更多空间信息。此外,使用NWD损失替代CIoU损失,以更好地适应小目标检测。改进后网络架构如图3所示。

2.2.1 C3k2-DWA模块

本研究提出改进的特征提取模块C3k2-DWA,以解决原始C3k2模块在小目标检测中难以有效聚焦小尺寸害虫特征且无法充分融合全局与局部特征的局限。受XIE等^[28]提出的多尺度融合模块启发,设计动态加权注意力机制(DWA)并集成至C3k2-DWA模块,其与C3k2模块结构对比如图4所示。

如图4a所示,C3k2模块首先通过CBS层(卷积-批归一化-SiLU激活函数)处理输入特征,经通道拆分操作分为两个分支。第1分支通过串联C3k模块提取空间与语义特征,第2分支则采用残差连接提供全局特征,保持网络轻量化。两分支特征经拼接后形成最终输出。该架构虽能高效处理多尺度特征,但其特征融合方式较为简单,缺乏有效的整合机制,导致关键特征未能得到充分强调。

图4b为在基础结构上引入动态加权注意力(DWA)机制的C3k2-DWA模块。输入特征经CBS层处理后拆分为双分支,第1分支通过C3k模块提取局部细节,第2分支保留全局特征并通过残差连接维持效率。核心改进在于对双分支分别施加动态加权注意力处理:每个分支通过全局平均池化、 1×1 卷积调整通道特征分布,接着利用Sigmoid函数生

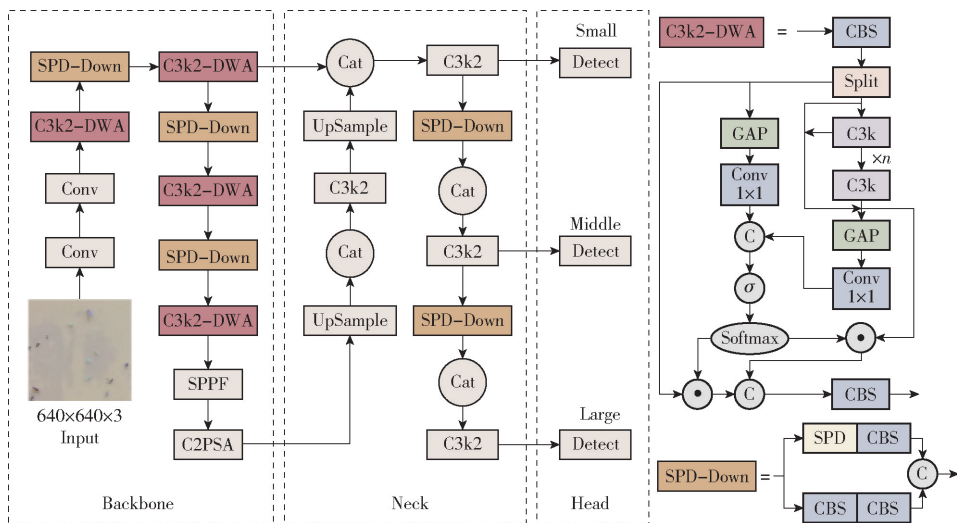
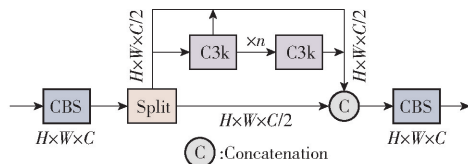
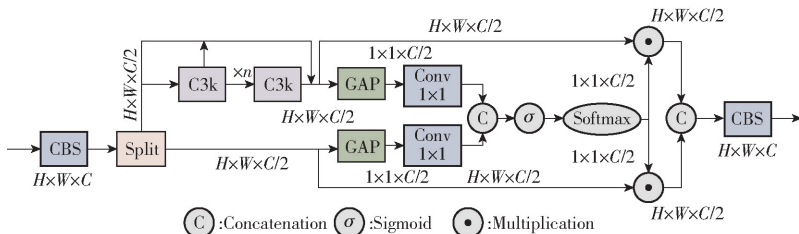


图 3 YOLO-SVP 模型网络结构图

Fig. 3 YOLO-SVP network architecture



(a) C3k2 模块结构图



(b) C3k2-DWA 模块结构图

图 4 C3k2 和改进后 C3k2-DWA 模块结构图比较

Fig. 4 Comparison of C3k2 block and improved C3k2-DWA block

成通道注意力权重并经 Softmax 函数精细化,并将权重矩阵与对应特征图逐元素相乘,选择性强化小目标检测关键区域。最终,增强后的特征拼接后经 CBS 层输出。C3k2-DWA 模块通过动态加权注意力机制增强对关键区域的关注,显著提升模型对小目标特征的敏感性,优化了 VegpestSet 数据集中微小害虫的检测性能。为了保持网络计算效率,仅在骨干网络中将 C3k2 模块替换为 C3k2-DWA 模块,以在精度和效率之间达到最佳平衡。

2.2.2 SPD-Down 模块

在卷积神经网络中,传统的最大池化、平均池化和步长卷积等下采样操作常导致特征信息丢失,尤其在深度网络中,重复下采样会加剧小目标检测困难。为解决这一问题,提出下采样模块 SPD-Down。

SPD-Down 模块采用双并行分支架构实现创新性下采样。第 1 分支引入 SUNKARA 等^[29]提出的空间到深度层,该层通过特征重组机制转换输入

数据,后接 1×1 卷积调整通道维度;第 2 分支采用 1×1 卷积结合步长为 2 的 3×3 卷积执行传统下采样操作。两分支输出特征经通道拼接形成最终下采样特征图,如图 5(注:scale 值设置为 2)所示。空间到深度层变换过程包含 3 步核心操作:将输入特征图按缩放因子 $scale$ 分割为网格状子区域,随后将空间维度特征重新排列至通道维度,再沿通道轴合并重组。这种独特转换使尺寸为 (S, S, C) 的输入特征图转变为 $(S/scale, S/scale, C \times scale^2)$ 的输出形式,在有效压缩空间分辨率的同时,将原始空间信息完整保留至深度通道。通过将 YOLO 11 网络中所有传统下采样卷积替换为 SPD-Down 模块,将能显著缓解传统下采样技术导致的小目标特征丢失问题。此外,这种空间到深度的转换机制避免了大卷积核计算,在完整保留小目标空间信息同时优化了计算效率。

2.2.3 NWD 损失函数

YOLO 11 中的定位损失采用 ZHENG 等^[30]提

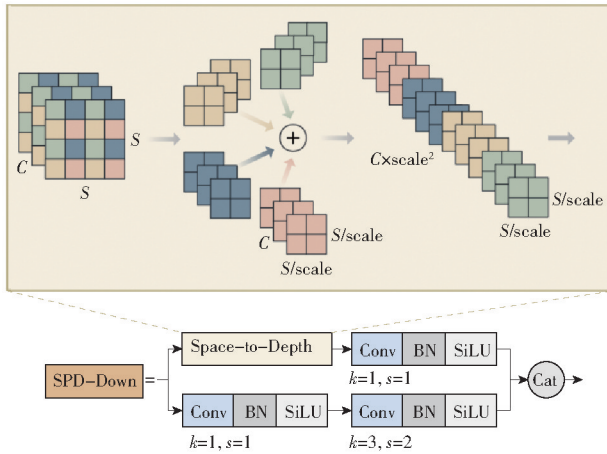


图5 SPD-Down 模块结构图

Fig. 5 Structural diagram of SPD-Down block

出的 CIoU 损失函数。该函数通过综合考量边界框的重叠区域、中心点距离及宽高比,为不同形状目标提供鲁棒的定位评估。当预测框与真实框无交集或存在完全包含关系时,CIoU 损失无法生成有效梯度以优化网络,在小目标检测中尤为突出——预测框的轻微像素偏移可能导致无交集状态,且微小目标更易被预测框完全覆盖,引发漏检现象。

为突破此局限,引入基于最优传输理论瓦瑟斯坦距离的 NWD 损失函数^[31]替代原 CIoU 损失。在 YOLO 11 网络中,损失函数 L 主要由定位损失 (L_{loc})、分类损失 (L_{cls}) 和置信度损失 (L_{conf}) 组成。

将 YOLO 11 损失函数中的定位损失项 (L_{loc}) 替换为 NWD 损失函数,该替换策略针对小目标检测任务具有显著优越性。与传统基于 IoU 的度量相比,NWD 损失通过将边界框建模为高斯分布并引入瓦瑟斯坦距离衡量分布差异,实现了对小目标位置变化的精确捕捉。

表3 YOLO 11 和 YOLO-SVP 对比试验结果

Tab. 3 Comparison of experimental results between YOLO 11 and YOLO-SVP

模型	类别	P	R	F1 值	mAP_{50}	$mAP_{50,95}$
YOLO 11	所有类别	83.7	78.9	81.2	85.5	50.6
	小菜蛾 (<i>Plutella xylostella</i>)	89.0	88.4	88.7	94.0	60.7
	烟粉虱 (<i>Bemisia tabaci</i>)	81.4	72.4	76.6	81.3	41.7
	瓜实蝇 (<i>Bactrocera cucurbitae</i>)	84.2	82.3	83.2	89.7	62.4
	黄曲条跳甲 (<i>Phyllotreta striolata</i>)	88.1	91.5	89.8	94.0	53.1
	西花蓟马 (<i>Frankliniella occidentalis</i>)	76.1	59.9	67.0	68.4	35.0
YOLO-SVP (本方法)	所有类别	87.7	83.7	85.7	89.3	54.9
	小菜蛾 (<i>Plutella xylostella</i>)	92.0	92.5	92.2	95.7	64.2
	烟粉虱 (<i>Bemisia tabaci</i>)	86.2	80.1	83.0	87.2	46.9
	瓜实蝇 (<i>Bactrocera cucurbitae</i>)	88.1	86.2	87.1	91.3	67.5
	黄曲条跳甲 (<i>Phyllotreta striolata</i>)	91.0	93.0	92.0	95.4	55.9
	西花蓟马 (<i>Frankliniella occidentalis</i>)	81.1	66.9	73.3	76.9	40.0

2.3 试验环境与参数设置

试验均在配备 Ubuntu 20.04.1 操作系统的服务器上完成,硬件配置包括 Intel® Core™ i9-10900K 的处理器及 NVIDIA GeForce RTX 3080 图像处理器。服务器搭配 CUDA 11.6 驱动。软件环境基于 Python 3.9 和 Pytorch 1.13.1 框架。模型训练参数中批次大小设为 16,训练周期设为 200,初始学习率设为 0.01。

2.4 评价指标

为评估所提出的蔬菜害虫检测模型性能,采用精确率 (P)、召回率 (R)、F1 值、平均精度均值 (mAP)、浮点运算量 (FLOPs) 和模型参数量来评估模型的准确性和复杂度。在评估 mAP 时, mAP_{50} 使用交并比 (IoU) 阈值为 0.5 进行计算,对所有类别的 AP 值取平均。 $mAP_{50,95}$ 代表了 IoU 阈值从 0.5 到 0.95、以 0.05 为步长的范围内取平均值。浮点运算量用来评估模型复杂度,模型参数量指模型中可训练参数的总数量,反映了模型的大小和内存需求。

3 试验结果与分析

3.1 YOLO-SVP 模型的性能

为验证 YOLO-SVP 模型的有效性,利用 1.2 节构建的 VegpestSet 数据集及 2.3 节给出的试验设置,进行改进前后模型仿真试验对比,结果如表 3 所示。由表 3 可知,YOLO-SVP 模型在精确率 (87.7%)、召回率 (83.7%)、F1 值 (85.7%)、 mAP_{50} (89.3%) 和 $mAP_{50,95}$ (54.9%) 等指标上均优于基准模型 YOLO 11。与 YOLO 11 相比,YOLO-SVP 模型的精确率和召回率分别提升 4.0、4.8 个百分点,F1 值提高 4.5 个百分点, mAP_{50} 和 $mAP_{50,95}$ 分别提升 3.8、4.3 个百分点,显示出全面的性能提升。

试验涉及的5类蔬菜害虫中,小菜蛾、瓜实蝇和黄曲条跳甲属大中型害虫,烟粉虱和西花蓟马则属小型害虫。针对小型害虫,YOLO-SVP所有指标均取得显著进步:烟粉虱与西花蓟马F1值分别提升6.4、6.3个百分点,mAP₅₀分别提高5.9、8.5个百分点,mAP_{50,95}分别增长5.2、5.0个百分点,验证了该模型在小目标检测场景的有效性。大中型害虫的检测结果也呈现不同程度的提升:小菜蛾的检测性能在F1值与mAP_{50,95}方面都提升3.5个百分点,mAP₅₀提升1.7个百分点;黄曲条跳甲在F1值、mAP₅₀与mAP_{50,95}上分别增长2.2、1.4、2.8个百分点。瓜实蝇检测性能也显著改善,F1值提升3.9个百分点,mAP₅₀和mAP_{50,95}分别增长1.6、5.1个百分点。

总体而言,YOLO-SVP模型不仅提升了小目标害虫检测性能,还优化了大中型目标害虫检测能力。这主要源于三大改进:C3k2-DWA模块通过动态注意力机制增强了模型对中小型害虫的聚焦能力;SPD-Down模块保留了更多空间信息,避免了小目标特征的丢失;NWD损失函数优化了小目标的损失计算,降低了误检和漏检率,显著提高了检测精度。

3.2 与不同检测模型的性能比较

为全面评估本文所提的YOLO-SVP性能,将其与主流具有较优性能的YOLO v5、YOLO v8、YOLO v10等YOLO系列模型进行对比。表4展示试验仿真结果。

表4 与YOLO模型的对比试验结果

Tab.4 Comparative experiment results with YOLO models

模型	F1值/mAP ₅₀ /mAP _{50,95}			浮点运算量	参数量
	%	%	%		
YOLO v5	80.5	81.2	43.5	1.35×10^{11}	5.32×10^7
YOLO v8	78.5	82.9	45.6	1.66×10^{11}	4.37×10^7
YOLO v10	80.9	84.4	49.0	1.28×10^{11}	2.59×10^7
YOLO 11	81.2	85.5	50.6	8.73×10^{10}	2.53×10^7
YOLO-SVP	85.7	89.3	54.9	8.16×10^{10}	2.54×10^7

由表4可知,在所有比较方法中,YOLO-SVP在F1值、mAP₅₀和mAP_{50,95}指标上均为最优性能。相较于其他模型,其F1值最大提升7.2个百分点、最小提升4.5个百分点、平均提升5.4个百分点;mAP₅₀相应提升8.1、3.8、5.8个百分点;mAP_{50,95}最大、最小、平均增益达11.4、4.3、7.7个百分点。这些优势源于C3k2-DWA模块增强的特征提取能力、NWD损失函数优化的目标检测精度。

在计算效率方面,YOLO-SVP以 8.16×10^{10} 的浮点运算量在对比模型中表现出最高效率,且参数量

保持在 2.54×10^7 的轻量级水平。相比参数规模相近的YOLO v10,YOLO-SVP实现更高检测精度的同时降低了 4.64×10^{10} 浮点运算量;相较于基准模型YOLO 11,参数量仅增加 1.0×10^5 而浮点运算量显著降低 5.7×10^9 ,同时全面提升精度指标。这归功于SPD-Down模块避免大卷积核计算开销的设计。

为进一步评估本文所提的YOLO-SVP在小目标检测上的性能,将其与先进的小目标检测算法FBRT-YOLO^[32]和IDD-YOLO^[33]及蔬菜害虫检测模型YOLO-SIP^[27]进行对比。表5为试验仿真结果。

表5 与小目标检测模型和蔬菜害虫检测模型的对比试验结果

Tab.5 Comparative experiment results with small object detection models and vegetable pest detection models

模型	F1值/mAP ₅₀ /mAP _{50,95}			浮点运算量	参数量
	%	%	%		
YOLO-SIP	77.8	78.2	37.0	1.13×10^{11}	5.36×10^7
IDD-YOLO	78.7	82.5	44.5	1.64×10^{10}	9.83×10^6
FBRT-YOLO	81.1	84.8	49.8	1.20×10^{11}	1.49×10^7
YOLO-SVP	85.7	89.3	54.9	8.16×10^{10}	2.54×10^7

相较于目前先进的小目标检测算法,所提出的YOLO-SVP在F1值、mAP₅₀和mAP_{50,95}指标上都最优,较FBRT-YOLO分别提升4.6、4.5、5.1个百分点,较IDD-YOLO分别提升7.0、6.8、10.4个百分点。而相较于同类较优的蔬菜害虫检测模型YOLO-SIP,YOLO-SVP的F1值提升7.9个百分点,mAP₅₀提升11.1个百分点,mAP_{50,95}提升17.9个百分点。这主要因为FBRT-YOLO、IDD-YOLO和YOLO-SIP分别基于特征提取能力受限的YOLO v8、YOLO v5和YOLO v4改进,而YOLO-SVP依托新型YOLO 11架构,并融合提出的C3k2-DWA、SPD-Down和NWD损失函数3项改进,在精度与效率上实现全面提升。

综上所述,YOLO-SVP凭借C3k2-DWA模块的特征优化、SPD-Down模块的高效采样以及NWD损失函数的检测校准,在蔬菜害虫检测任务中实现了精度与效率的协同突破,为实际农业应用提供了最优解决方案。

3.3 消融试验

为评估数据增强技术的有效性,通过消融试验比较了未经增强的原始数据集和应用基于1.3节设置的复合数据增强后的数据集训练的YOLO-SVP模型。结果表明,复合数据增强显著优化了模型性能:精确率和召回率分别提高了4.3、1.5个百分点,F1值增加2.9个百分点,mAP₅₀提高3.0个百分点,

mAP_{50,95}提升6.4个百分点。该增强策略有效缓解了单一农场数据分布的局限性,显著提升了模型的泛化能力。

为了评估这些改进的有效性,进行相应的消融试验,结果如表6所示。基准YOLO 11的F1值为81.2%,mAP₅₀和mAP_{50,95}分别为85.5%和50.6%。单独引入C3k2-DWA模块后,F1值提升至83.2%,mAP₅₀升至86.8%。添加SPD-Down模块后,精确率、召回率和F1值分别提升0.4、3.4、2.0

个百分点,且计算效率显著优化(浮点运算量减少 6.3×10^9 ,参数量减少约 3.9×10^6),mAP₅₀和mAP_{50,95}分别提升1.8、4.0个百分点。引入NWD损失函数后,召回率达到最优的83.8%,显著提升4.9个百分点,同时mAP₅₀和mAP_{50,95}分别提升2.9、2.6个百分点。3项改进的协同效应最终使YOLO-SVP达到综合性能最优:F1值达到85.7%,mAP₅₀和mAP_{50,95}分别为89.3%和54.9%,并保持较低浮点运算量(8.16×10^{10})。

表6 消融试验结果

Tab. 6 Results of ablation experiments

C3k2-DWA	SPD-Down	NWD	P/%	R/%	F1值/%	mAP ₅₀ /%	mAP _{50,95} /%	浮点运算量	参数量
			83.7	78.9	81.2	85.5	50.6	8.73×10^{10}	2.53×10^7
✓			83.3	83.1	83.2	86.8	53.3	8.75×10^{10}	2.70×10^7
	✓		84.1	82.3	83.2	87.3	54.6	8.10×10^{10}	2.14×10^7
		✓	85.2	83.8	84.5	88.4	53.2	8.73×10^{10}	2.53×10^7
✓		✓	86.8	83.5	85.1	88.7	53.9	8.75×10^{10}	2.70×10^7
✓	✓	✓	87.7	83.7	85.7	89.3	54.9	8.16×10^{10}	2.54×10^7

注:“✓”表示添加此模块。

为直观对比模型改进效果,绘制了YOLO 11和YOLO-SVP的mAP曲线及训练损失曲线,如图6所示。图6a显示,YOLO-SVP在mAP₅₀和mAP_{50,95}指标上显著超越了基线模型,且其mAP曲线在训练过程中稳定上升,验证了其更优的检测性能。图6b展示了不同改进方案的训练损失值变化,所有改进

方案均获得比基线模型更优的损失收敛效果和更低的损失终值。值得注意的是,在NWD损失函数引入后损失曲线呈现显著改善,相比原始模型曲线更平滑、波动更小,证实该损失函数特别适用于小目标检测任务。通过融合C3k2-DWA模块、SPD-Down模块和NWD损失函数,YOLO-SVP实现了最佳的mAP曲线形态和损失下降趋势。这些改进方案显著增强了基线模型的检测能力,在目标识别性能与模型优化效率层面取得双重突破。

3.4 可视化结果比较

图7(注:第1列为真实框,第2列和第3列为YOLO 11和YOLO-SVP检测结果,第4列和第5列为对应的热图;在图像中,白色箭头表示未检测,红色箭头表示错误检测)通过对比YOLO 11与YOLO-SVP的检测结果及其对应的梯度加权类激活映射(Grad-CAM)热力图,直观揭示了YOLO-SVP模型在目标定位与特征聚焦能力上的显著提升。在检测性能方面,YOLO 11表现出明显的局限性,包括对小型害虫目标的显著漏检(如白色箭头所示)以及对非目标区域的误检(如红色箭头所示)。相比之下,YOLO-SVP虽然也存在个别误检,但其对小型、密集目标的捕捉能力显著增强,整体检测结果可靠性更高。

热力图的对比进一步揭示了模型的差异。如第4列所示,YOLO 11生成的热力图呈现出高度分散的激活模式,激活区域(以暖色调表示)范围广且边界模糊,未能有效聚焦于目标害虫,且常常出现在与

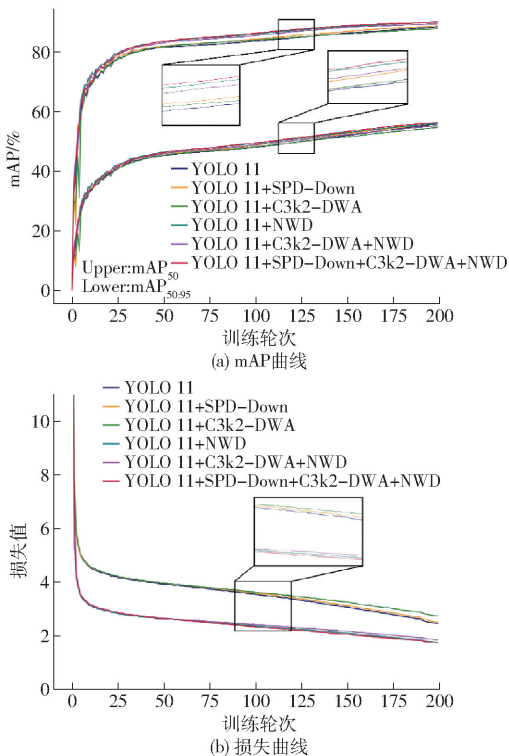


图6 YOLO 11 和其改进后模型的性能比较

Fig. 6 Performance comparison of YOLO 11 model and its improved models

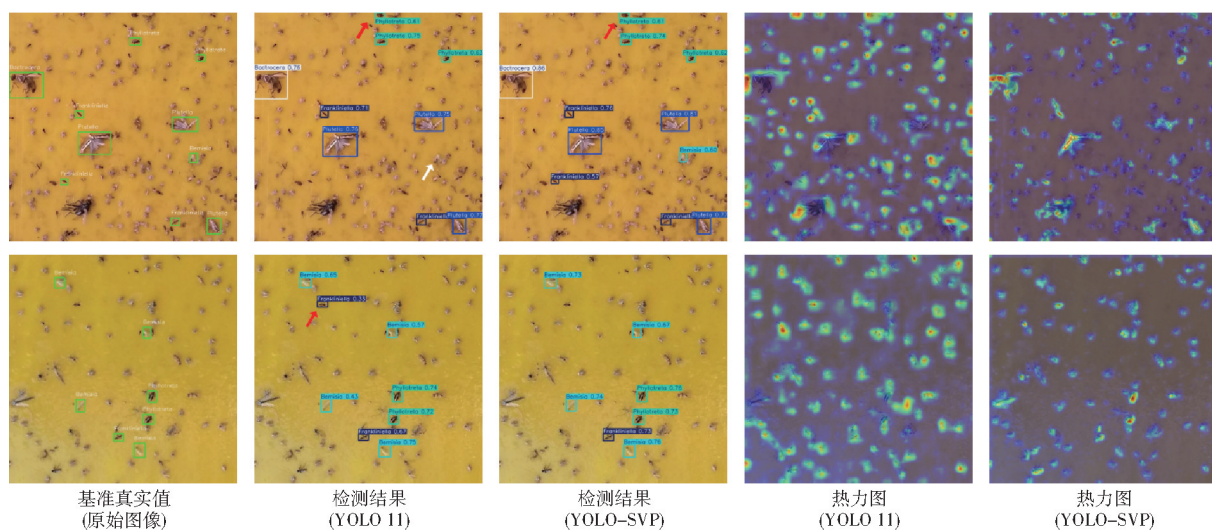


图7 YOLO 11 和 YOLO-SVP 检测结果和热图可视化

Fig. 7 Visualization of detection results and heatmaps for YOLO 11 and YOLO-SVP

害虫无关的背景纹理或结构上,这种分散的注意力模式导致了误检和漏检问题,尤其是对小型目标的漏检。相比之下,第5列的YOLO-SVP生成的热力图表现出更集中且精准的激活模式,高激活区域紧密聚焦于实际害虫的位置,特别是易被遗漏的小型目标,形成清晰的热点。同时,背景和非目标区域的激活响应被有效抑制(以冷色调表现),显著减少了模型对无关信息的敏感性,提升了对小目标的识别能力并降低了误检率。

4 结论

(1) 提出一种基于动态加权注意力与空间到深度下采样的小尺寸蔬菜害虫检测模型YOLO-SVP,特别针对华南地区5类蔬菜害虫中的小型目标进行优化。

(2) 在骨干网络中引入动态加权注意力机制(C3k2-DWA模块),增强对小目标害虫关键特征

的聚焦与融合能力。通过空间深度转换层替代传统卷积下采样(SPD-Down下采样模块),减少小目标特征丢失。采用归一化瓦瑟斯坦距离度量(NWD损失函数),精准捕捉小型害虫位置变化。

(3) 在自制VegpestSet数据集上试验结果表明:YOLO-SVP以F1值85.7%、 mAP_{50} 89.3%、 $mAP_{50,95}$ 54.9%全面超越基线模型YOLO 11及其他先进模型。尤其对西花蓟马等小型害虫检测提升显著:F1值提升6.3个百分点, mAP_{50} 提高8.5个百分点, $mAP_{50,95}$ 提高5.0个百分点。相较当前先进的小目标检测方法FBRT-YOLO、IDD-YOLO与同类最优的蔬菜害虫检测模型YOLO-SIP,YOLO-SVP在F1值、 mAP_{50} 和 $mAP_{50,95}$ 指标上都最优,较FBRT-YOLO分别提升4.6、4.5、5.1个百分点,较IDD-YOLO分别提升7.0、6.8、10.4个百分点,较YOLO-SIP分别提升7.9、11.1、17.9个百分点。

参 考 文 献

- [1] JAWORSKI C C, THOMINE E, RUSCH A, et al. Crop diversification to promote arthropod pest management: a review[J]. Agriculture Communications, 2023, 1(1): 100004.
- [2] WANG J, LI S, YUAN W, et al. Ecological wisdom: functional plant *Cosmos bipinnata* enhances fitness traits of predator *Harmonia axyridis* strengthening biological control in greenhouse production[J]. Agriculture Communications, 2025, 3(3): 100097.
- [3] 王春桃, 梁炜健, 郭庆文, 等. 农业害虫智能视觉检测研究综述[J]. 中国农机化学报, 2023, 44(7): 207-213. WANG Chuntao, LIANG Weijian, GUO Qingwen, et al. Review on computer-vision-based detection of agricultural pests[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(7): 207-213. (in Chinese)
- [4] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.
- [5] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [6] LIU W, ANGUOLOV D, ERHAN D, et al. SSD: single shot multiBox detector[C]//Computer Vision-ECCV 2016, 2016: 21-37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [8] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517-6525.
- [9] REDMON J, FARHADI A. Yolov3: an incremental improvement[R]. arXiv Preprint, arXiv:1804.02767, 2018.

- [10] BOCHKOVSKIY A, WANG C, LIAO H M. Yolov4: optimal speed and accuracy of object detection[J]. arXiv Preprint, arXiv:2004.10934, 2020.
- [11] LI C, LI L, JIANG H, et al. YOLOv6: a single-stage object detection framework for industrial applications[R]. arXiv Preprint, arXiv:2209.02976, 2022.
- [12] WANG C, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2013 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464 – 7475.
- [13] WANG C Y, YE H I, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information[C]//arXiv Preprint, arXiv:2402.13616, 2024.
- [14] WANG A, CHEN H, LIU L, et al. YOLOv10: real-time end-to-end object detection[C]//arXiv Preprint, arXiv:2405.14458, 2024.
- [15] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[C]//arXiv Preprint, arXiv:2410.17725, 2024.
- [16] 肖德琴, 黄一桂, 张远琴, 等. 基于改进 Faster R – CNN 的田间黄板害虫检测算法[J]. 农业机械学报, 2021, 52(6): 242 – 251.
XIAO Deqin, HUANG Yigui, ZHANG Yuanqin, et al. Pest detection algorithm of yellow plate in field based on improved Faster R – CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(6): 242 – 251. (in Chinese)
- [17] 王泰华, 郭亚州, 张家乐, 等. 基于改进 YOLO v5s 的水稻害虫识别研究[J]. 农业机械学报, 2024, 55(11): 39 – 48.
WANG Taihua, GUO Yazhou, ZHANG Jiale, et al. Rice pest identification based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 39 – 48. (in Chinese)
- [18] 赵辉, 黄鏢, 王红君, 等. 基于改进 YOLO v7 的农田复杂环境下害虫识别算法研究[J]. 农业机械学报, 2023, 54(10): 246 – 254.
ZHAO Hui, HUANG Biao, WANG Hongjun, et al. Pest identification method in complex farmland environment based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 246 – 254. (in Chinese)
- [19] 彭红星, 徐慧明, 高宗梅, 等. 基于改进 YOLOF 模型的田间农作物害虫检测方法[J]. 农业机械学报, 2023, 54(4): 285 – 294, 303.
PENG Hongxing, XU Huiming, GAO Zongmei, et al. Insect pest detection of field crops based on improved YOLOF model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(4): 285 – 294, 303. (in Chinese)
- [20] 梁勇, 邱荣洲, 李志鹏, 等. 基于 YOLO v5 和多源数据集的水稻主要害虫识别方法[J]. 农业机械学报, 2022, 53(7): 250 – 258.
LIANG Yong, QIU Rongzhou, LI Zhipeng, et al. Identification method of major rice pests based on YOLO v5 and multi-source datasets[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 250 – 258. (in Chinese)
- [21] LV Z, JIN H, ZHEN T, et al. Small object recognition algorithm of grain pests based on SSD feature fusion[J]. IEEE Access, 2021, 9: 43202 – 43213.
- [22] TIAN Y, WANG S, LI E, et al. MD – YOLO: multi-scale dense YOLO for small target pest detection[J]. Computers and Electronics in Agriculture, 2023, 213: 108233.
- [23] WEN C, CHEN H, MA Z, et al. Pest – YOLO: a model for large-scale multi-class dense and tiny pest detection and counting [J]. Frontiers in Plant Science, 2022, 13: 973985.
- [24] WANG R, JIAO L, XIE G, et al. S – RPN: sampling-balanced region proposal network for small crop pest detection[J]. Computers and Electronics in Agriculture, 2021, 187: 106290.
- [25] DONG S, TENG Y, JIAO L, et al. ESA – Net: an efficient scale-aware network for small crop pest detection[J]. Expert Systems with Applications, 2024, 236: 121308.
- [26] 邓姗姗, 黄慧, 马燕. 基于改进 Faster R – CNN 的小目标检测算法[J]. 计算机工程与科学, 2023, 45(5): 869 – 877
DENG Shanshan, HUANG Hui, MA Yan. A small object detection algorithm based on improved Faster R – CNN [J]. Computer Engineering & Science, 2023, 45(5): 869 – 877. (in Chinese)
- [27] GUO Q, WANG C, XIAO D, et al. Automatic monitoring of flying vegetable insect pests using an RGB camera and YOLO – SIP detector[J]. Precision Agriculture, 2023, 24(2): 436 – 457.
- [28] XIE L, LI C, WANG Z, et al. Shisrnet: super-resolution and classification network for low-resolution breast cancer histopathology image[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2023: 23 – 32.
- [29] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer, 2022: 443 – 459.
- [30] ZHENG Z, WANG P, LIU W, et al. Distance – IoU loss: faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993 – 13000.
- [31] WANG J, XU C, YANG W, et al. A normalized Gaussian Wasserstein distance for tiny object detection[J]. arXiv Preprint, arXiv:2110.13389, 2021.
- [32] XIAO Y, XU T, XIN Y, et al. FBRT – YOLO: faster and better for real-time aerial image detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(8): 8673 – 8681.
- [33] LU Y, LI D, LI D, et al. A lightweight insulator defect detection model based on drone images[J]. Drones, 2024, 8(9): 431.