

骆润玫, 殷惠莉, 刘伟康, 等. 基于 YOLOv5-C 的广佛手病虫害识别 [J]. 华南农业大学学报, 2023, 44(1): 151-160.
LUO Runmei, YIN Huili, LIU Weikang, et al. Identification of bergamot pests and diseases using YOLOv5-C algorithm[J]. Journal of South China Agricultural University, 2023, 44(1): 151-160.

基于 YOLOv5-C 的广佛手病虫害识别

骆润玫¹, 殷惠莉¹, 刘伟康¹, 胡 凯¹, 廖 飞¹, 刘泽乾¹, 曹亚芃¹, 李 强², 王卫星^{1,3}

(1 华南农业大学 电子工程学院, 广东 广州 510642; 2 广东省农业技术推广中心, 广东 广州 510520; 3 广东省农情信息监测工程技术中心, 广东 广州 510642)

摘要:【目的】为实现复杂背景下广佛手发病早期的病虫害快速精准识别, 提出一种基于 YOLOv5-C 的广佛手病虫害识别方法。【方法】使用 YOLOv5s 网络模型作为基础网络, 通过引入所提出的多尺度特征融合模块, 提高网络模型的特征提取与特征融合能力, 均衡提高每一类广佛手病虫害的识别准确率; 使用注意力机制模块提高网络模型对病虫害目标特征信息的关注度, 弱化复杂背景的干扰信息, 提高网络模型的识别准确率; 利用改进的 C3-SC 模块替换 PANet 结构中的 C3 模块, 在不影响网络模型识别性能的条件下减少网络模型的参数。【结果】基于 YOLOv5-C 的复杂背景下的广佛手病虫害识别, F1 分数为 90.95%, 平均精度均值为 93.06%, 网络模型大小为 14.1 Mb, 在 GPU 上每张图像平均检测时间为 0.01 s。与基础网络 YOLOv5s 相比, 平均精度均值提高了 2.45 个百分点, 7 个类别识别的平均准确率的标准差由 7.14 减少为 3.13, 变异系数由 7.88% 减少为 3.36%。平均精度均值比 RetinaNet、SSD、Efficientdet 和 YOLOv4 模型分别高 22.30、20.65、4.84 和 2.36 个百分点。【结论】该方法能快速准确地识别复杂背景下广佛手病虫害目标, 可为广佛手种植产业的智能化管理提供参考。

关键词: 广佛手; 病虫害识别; 目标检测; 轻量化; 多尺度特征融合; 注意力机制

中图分类号: S435.67; TP391

文献标志码: A

文章编号: 1001-411X(2023)01-0151-10

Identification of bergamot pests and diseases using YOLOv5-C algorithm

LUO Runmei¹, YIN Huili¹, LIU Weikang¹, HU Kai¹, LIAO Fei¹, LIU Zeqian¹, CAO Yapeng¹, LI Qiang², WANG Weixing^{1,3}

(1 School of Electronic Engineering, South China Agricultural University, Guangzhou 510642, China; 2 Guangdong Agricultural Technology Extension Center, Guangzhou 510520, China; 3 Guangdong Engineering Research Center for Agricultural Information Monitoring, Guangzhou 510642, China)

Abstract: 【Objective】 In order to achieve rapid and accurate identification of pests and diseases in the early disease stage of bergamot in complex background, we proposes a YOLOv5-C-based method for the identification of bergamot pests and diseases. 【Method】 The YOLOv5s network model was used as the base network. The multi-scale feature fusion module was proposed and introduced to improve the feature extraction and feature fusion capability of the network model, and to improve the recognition accuracies of different bergamot pests and diseases in a balanced manner. The attention mechanism module was used to upgrade the attention degree of the network model to the information of target features of pests and diseases, weaken the interference information of complex background, and raise the recognition accuracy of the network model. An improved C3-SC module was used to replace the C3 module in the PANet structure to decrease the number of

收稿日期: 2022-03-06 网络首发时间: 2022-06-09 11:52:18

首发网址: <https://kns.cnki.net/kcms/detail/44.1110.s.20220608.1605.002.html>

作者简介: 骆润玫, 硕士研究生, 主要从事深度学习在农业工程领域应用研究, E-mail: 1039273787@qq.com; 通信作者:

王卫星, 教授, 博士, 主要从事智慧农业与农业物联网研究, E-mail: weixing@scau.edu.cn

基金项目: 2021 年省级乡村振兴战略专项省级组织实施项目 (粤财农〔2021〕37 号)

parameters in the network model without lowering the network model recognition performance. 【Result】 *F1* score of 90.95% and mean average precision of 93.06% were achieved when identifying the bergamot pests and diseases under a complex background using the YOLOv5-C method. The size of network model was 14.1 Mb, and the average detection time was 0.01 s per image on the GPU. Comparing with the original YOLOv5s, the mean accuracy of YOLOv5-C increased by 2.45 percentage point, the standard deviation of the average precision for seven categories was reduced from 7.14 to 3.13, and the coefficient of variation decreased from 7.88% to 3.36%. Moreover, the mean average accuracy was 22.30, 20.65, 4.84 and 2.36 percentage point higher than those of RetinaNet, SSD, Efficientdet and YOLOv4, respectively. 【Conclusion】 The method can quickly and accurately identify bergamot pests and diseases under complex background, and provide a reference for intelligent management of bergamot cultivation industry.

Key words: Bergamot; Pest and disease identification; Object detection; Lightweight; Multi-scale feature fusion; Attention module

广佛手是首批受法律保护的 8 种岭南中药材之一,其根、茎、叶、花、果均可入药,具有疏肝理气、和胃止痛、燥湿化痰的功效^[1]。除药用外,广佛手提取出的精油,气味清香淡雅,具有舒缓压力、提神解乏的作用,被广泛应用于食品、药品、化妆品等行业^[2]。广佛手的种植规范要求严格,禁止滥用农药和化学药剂,因此,对广佛手早期的病虫害进行精准识别与防治,是提高广佛手产量与质量的关键。

随着农业机械化与信息化的快速发展,农作物病虫害识别的方法由传统的人工识别方法,发展为基于深度学习技术的识别方法^[3-5]。Yuan 等^[6]将迁移学习与 2 种深度学习结构相结合,对 8 种作物病害图像进行分类,准确率为 95.93%。Chen 等^[7]利用 VGGNet 增强模型,对水稻病虫害进行分类,准确率为 92.00%。Lü等^[8]基于 AlexNet 网络,从不同角度验证了特征增强算法能有效提高网络识别能力,所提出的方法能准确识别自然环境中采集的玉米病害图像。陆健强等^[9]提出一种基于 Mixup 算法和卷积神经网络的柑橘黄龙病果实识别模型,最终模型的准确率达到 94.29%。Hassan 等^[10]提出了一种用于植物病虫害识别的新型深度学习模型,并在 3 个公开的植物病虫害数据集上进行验证,均取得了良好的识别效果。上述基于深度学习的农作物病虫害叶片分类研究,为深度学习技术应用于农作物病虫害识别提供了参考。

随着深度学习的快速发展,基于深度学习的目标检测算法被广泛应用于农作物病虫害识别,该方法不仅能快速识别病虫害的类别,还能精确定位病斑、害虫在图像中的位置,进一步促进了精准农业的发展^[11]。常用的目标检测网络有 Faster RCNN^[12]、SSD^[13]、YOLO 系列^[14-17]等,可用于识别

农作物病虫害^[12-21]。其中, YOLO 网络模型是一种 one-stage 的目标检测算法,基于回归分析可直接获得目标分类结果与目标定位信息,具有结构简单、计算量小、运算速度快等特点,被广泛应用于农作物病虫害识别研究^[21-23]。Mamdouh 等^[21]提出了改进的 YOLOv4 模型,利用数据增强和加入负样本的方法提高橄榄果蝇的检测精度,最终的识别准确率为 96.68%。张博等^[22]提出了一种基于改进 YOLOv3 的农作物害虫识别算法,平均识别准确率为 88.07%。胡嘉沛等^[23]提出了一种改进 YOLO4-tiny 模型的柑橘木虱识别方法,平均识别精度为 96.16%。YOLOv5s 是 YOLO 系列网络的最新模型,具有检测精度高、推理速度快等特点,能有效识别小目标或被遮挡目标。

自然环境下拍摄的广佛手早期病虫害图像背景复杂,部分病斑与枝干和光斑相似,且病斑、病虫害体积较小,易造成病虫害目标的误检与漏检。为提高广佛手病虫害的识别准确率,结合 SqueezeNet 网络^[24]以及 YOLO 系列网络的优点,提出一种轻量级目标检测模型 YOLOv5-C,以期为实现田间快速精准监测广佛手早期病虫害提供参考。

1 材料与方法

1.1 试验数据采集

试验数据采集于广东省肇庆乐城镇广佛手园,采集时间为 2021 年 8 月和 10 月。使用手持式 Nikon 单反相机 D3100 和高清手机,对自然环境下的广佛手幼苗盆栽和成熟植株叶片上早期的病斑、病虫害进行拍摄,拍摄时相机距离病斑叶片约 0.5 m,图像分辨率为 4 608×3 072 和 3 000×3 000。为真实还原田间复杂环境,本研究的试验数据采集天气包

括晴天、阴天和雨后,采集时间为 08:30—12:00 和 14:00—18:00,拍摄角度包括东、南、西、北 4 个方向。在自然环境下采集的广佛手早期病虫害图像中,病虫害类别包括褐斑病、溃疡病、煤烟病、附生绿球藻、潜叶蛾、双钩巢粉虱和玉带凤蝶,共 7 类,最终数据集共 5969 张原始图像。

表 1 广佛手病虫害数据集基本信息
Table 1 Basic information of bergamot pest and disease dataset

类别 Category	原始图像 Original image	褐斑病病斑 Labeled brown spot	溃疡病病斑 Labeled canker	煤烟病病斑 Labeled dark mildew	附生绿球藻 Labeled chlorococcum	潜叶蛾 Leaf miner	玉带凤蝶 <i>Papilio polytes</i>	双钩巢粉虱 Mealworm
训练集 Training set	4 292	608	1 150	1 654	638	1 463	1 258	740
验证集 Validation set	1 076	145	272	438	155	348	301	200
测试集 Test set	601	91	155	214	90	200	200	91
合计 Total	5 969	844	1 577	2 306	883	2 011	1 759	1 031

1.3 构建广佛手病虫害识别网络

针对在自然环境下,广佛手早期病斑和幼虫体积较小、田间拍摄的图像背景复杂、病变叶片褶皱导致部分病斑特征被破坏等问题,本研究提出了能快速、准确地检测田间广佛手病虫害的 YOLOv5-C 目标检测模型。根据广佛手病虫害数据集的实际情况,参照 SqueezeNet 网络的模型压缩思想对网络模型结构进行压缩,降低模型的复杂度。同时,参照不同的多尺度特征融合思想^[25-27],添加改进的多尺度特征融合模块,进一步提升网络模型的识别准确率。

1.3.1 YOLOv5 网络模型 YOLOv5 网络模型检测精度高、推理速度快,能够有效识别小目标和被遮挡的目标。该网络包含 4 种基础架构,根据网络特定位置的特征提取不同数量的模块和卷积核,将网络分为 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x,对应的网络参数量依次为 7.08×10^6 、 2.11×10^7 、 4.67×10^7 、 8.73×10^7 ,模型大小依次为 14.4、42.5、93.8、175.2 Mb。随着网络模型参数的增加,基础网络模型的复杂度随之增加,虽然网络模型的准确率有所提升,但需要消耗更多的计算资源及硬件资源,网络模型的实时性也随之降低。为实现广佛手病虫害的快速精准识别,选择实时性最优、权重文件最小、最适合部署至嵌入式设备的 YOLOv5s 网络作为基础网络。YOLOv5s 的主干网络为 Darknet-53,颈网络为 PANet 路径聚合网络^[27],主干网络由 Focus 模块、卷积单元、C3 模块(含有 3 个卷积模块的 CSP 模块^[28])以及 SPP 模块^[29]组成。本研究使用 YOLOv5s 作为基础网络,网络结构如图 1 所示。

1.2 构建广佛手病虫害数据集

使用 labelImg 工具对 5 969 张原始图像中的广佛手病虫害目标进行标注,并将构建的广佛手发病早期的病虫害数据集,按照 72%、18% 和 10% 的比例随机划分为训练集(4 292 张)、验证集(1 076 张)和测试集(601 张)。数据集的基本信息如表 1 所示。

1.3.2 多尺度特征融合模块 广佛手早期病虫害的病斑、病虫体积较小,为有效提高广佛手病虫害目标的识别准确率,本研究针对广佛手病虫害数据集的特点,提出了一种多尺度特征融合模块(Multi-scale feature fusion module, MSFFM),以提高网络模型特征提取、特征融合的能力。该模块在骨干网络中选定 3 个特征图,包括顶层特征图 U_1 、中层特征图 U_2 和足够低层特征图 U_3 。其中,顶层特征图 U_1 通过上采样操作使得特征图尺寸、通道数与中层特征图保持一致,得到特征图 U_1^2 ;低层特征图通过下采样操作同样将特征图尺寸、通道数与中层特征图保持一致,得到特征图 U_3^2 。将特征图 U_1^2 、 U_2 、 U_3^2 进行融合,获得信息更丰富的特征图 U_0 ,将融合后的特征图 U_0 与 YOLOv5s 的颈网络 PANet 路径聚合网络相结合使用,进一步提高网络模型的识别准确率。图 2 为 MSFFM 的多尺度特征融合示意图。

1.3.3 注意力机制模块 针对广佛手病虫害目标体积较小、形状多变且部分病斑和病虫与光斑、背景枝干相似等问题,本研究在网络模型中添加了注意力机制模块(Convolutional Block Attention Module, CBAM)^[30],该模块由通道注意力模块和空间注意力模块组成,输入特征图依次经过通道注意力模块和空间注意力模块,获得通道维度和空间维度的注意力图,将输入特征图与获得的注意力图相乘以实现特征图的自适应优化,使得网络模型更加关注于目标的特征信息,弱化图像中的背景信息,提高网络模型的识别性能。

其中,通道注意力模块通过探索不同通道间特征图的关系,选择目标特征的关键信息并提高这部

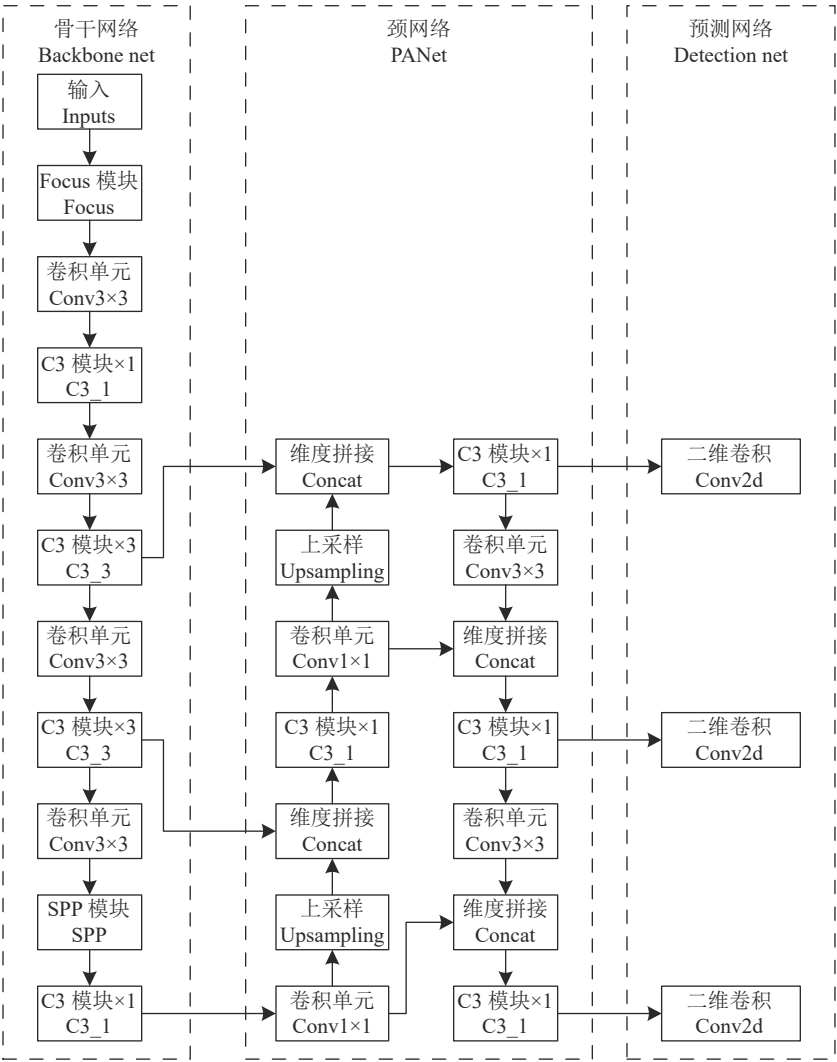


图 1 YOLOv5s 网络模型结构

Fig. 1 Network structure of YOLOv5s

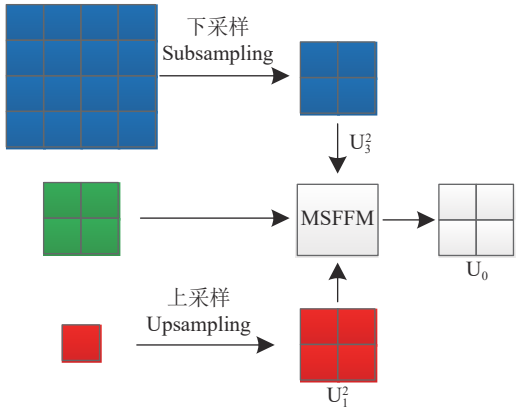


图 2 MSFFM 的多尺度特征融合示意图

Fig. 2 Schematic diagram of multi-scale feature fusion of MSFFM

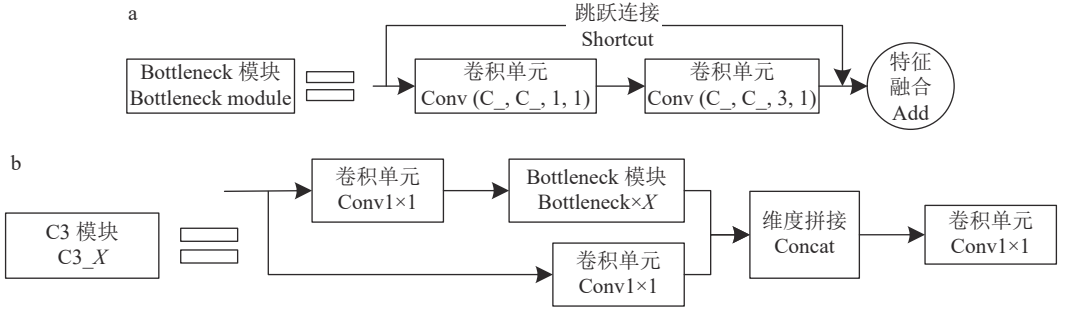
分信息的表达能力,使得网络模型更加关注目标相关的特征信息,更有效地提取目标特征。

1.3.4 C3 模块的改进 由于广佛手病虫害的类别相对较少,在不影响网络识别准确率的前提下,可

适当对 YOLOv5s 网络模型结构进行压缩剪枝。其中, YOLOv5s 中的 C3 模块由 3 个卷积模块和 1 个 Bottleneck 模块构成, C3_X 表示 C3 模块中的 Bottleneck 模块数量为 X 个。Bottleneck 模块和 C3 模块的结构分别如图 3a、3b 所示。

本研究借鉴 SqueezeNet 网络的模型压缩思想,对 C3 模块进行改进,提出了 C3-S 模块。C3-S 模块中的 Bottleneck-S 瓶颈模块,由 2 个 1×1 卷积模块和 1 个 3×3 卷积模块构成,首先通过 1×1 卷积模块 Conv1 将输入特征图的通道数 C_{in} 压缩至输入的四分之一,即 $C_{in}/4$;将压缩后的特征图分别传入 1×1 卷积模块 Conv2 和 3×3 卷积模块 Conv3,得到两个通道数都为 $C_o/2$ 的特征图;将 2 个特征图进行维度拼接,所获得的便是 Bottleneck-S 瓶颈模块的输出特征图。Bottleneck-S 模块和 C3-S 模块的结构分别如图 4a、4b 所示。

在卷积神经网络模型中,无偏置情况下,每个卷积层的参数数量计算如下:



卷积单元 Conv 模块括号内的参数依次表示输入通道数、输出通道数、卷积核大小和步长

The number of input channels, output channels, the size of the convolution kernel and the step size are indicated by the parameters in parentheses of Conv module

图 3 Bottleneck 瓶颈模块 (a) 和 C3 模块 (b) 的结构示意图

Fig. 3 Structural diagram of the Bottleneck module (a) and C3 module (b)

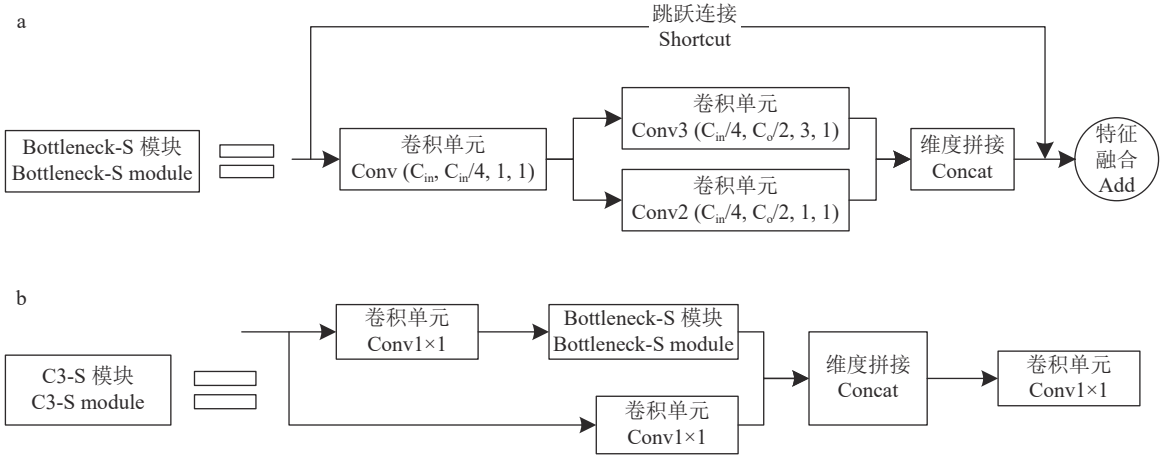


图 4 Bottleneck-S 瓶颈模块 (a) 和 C3-S 模块 (b) 的结构示意图

Fig. 4 Structural diagram of Bottleneck-S module (a) and C3-S module (b)

$$\text{params} = C_o \times (k_w \times k_h \times C_{in}), \quad (1)$$

式中, params 表示卷积层的参数数量, C_o 表示输出通道数, k_w 表示卷积核宽, k_h 表示卷积核高, C_{in} 表示输入通道数。

原始 C3 模块中的 Bottleneck 瓶颈模块, 由 1 个 1×1 卷积模块和 1 个 3×3 卷积模块构成, 2 个卷积模块的输入和输出通道数量相同, 均为 C , 则 Bottleneck 瓶颈模块的 2 个卷积层所需的参数数量 P_1 如式 (2) 所示:

$$P_1 = C \times (1 \times 1 \times C) + C \times (3 \times 3 \times C) = 10C^2, \quad (2)$$

式中, P_1 表示所需的参数数量, C 表示 2 个卷积层的输入和输出通道数。

使用 C3-S 模块替换 C3 模块, 在网络模型相同位置的 Bottleneck-S 瓶颈模块中, 3 个卷积层所需的参数数量为 P_2 , 计算如式 (3) 所示:

$$P_2 = \frac{C}{4} \times (1 \times 1 \times C) + \frac{C}{2} \times \left(1 \times 1 \times \frac{C}{4}\right) + \frac{C}{2} \times \left(3 \times 3 \times \frac{C}{4}\right) = \frac{3}{2}C^2, \quad (3)$$

式中, P_2 表示所需的参数数量, C 表示 Bottleneck-S 瓶颈模块的输入和输出通道数量。

虽然在卷积模块中还包含了批量归一化操作, 需要少量的参数, 但是在改进的 YOLOv5-C 模型中, 使用 C3-S 模块替换 C3 模块的部分, 卷积层的参数远远多于批量归一化操作所需的参数, 故直接使用卷积层的参数近似代表 2 个瓶颈模块所需的参数。根据式 (2) 和式 (3) 所得的参数数量 P_1 和 P_2 , 可知 Bottleneck-S 瓶颈模块的参数数量约为 Bottleneck 瓶颈模块的参数数量的 15%。

为进一步提高网络模型的特征提取能力, 将 CBAM 模块与 C3-S 模块相结合, 构成 C3-SC 模块, 结构如图 5 所示。

1.3.5 YOLOv5-C 广佛手病虫害识别网络 以 YOLOv5s 网络模型为基础网络, 使用所提出的 C3-SC 模块替换颈网络 PANet 中的 C3 模块, 有效减少网络模型参数的同时, 增强网络模型对广佛手病虫害目标的特征提取能力、弱化干扰背景特征; 在 PANet 结构中添加 MSFFM, 提高网络模型的特征融合能力, 进一步提高网络模型在复杂自然环境下

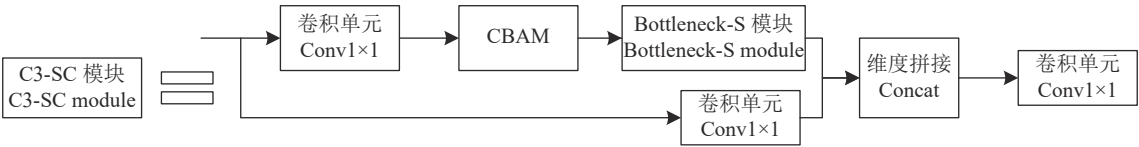


图 5 C3-SC 模块结构示意图

Fig. 5 Structural diagram of C3-SC module

对广佛手病虫害目标的识别准确率。YOLOv5-C 网络模型结构参数如表 2 所示。

表 2 YOLOv5-C 网络模型结构参数¹⁾

Table 2 Structural parameters of YOLOv5-C network model

序列号 Serial No.	当前层的输入来源 Input source of current layer	网络层 Network layer	输入尺寸 Input size	通道数 No. of channels
0	-1	Focus	640×640×3	32
1	-1	Conv3×3	320×320×32	64
2	-1	C3_1	160×160×64	64
3	-1	Conv3×3	160×160×64	128
4	-1	C3_3	80×80×128	128
5	-1	Conv3×3	80×80×128	256
6	-1	C3_3	40×40×256	256
7	-1	Conv3×3	40×40×256	512
8	-1	SPP	20×20×512	512
9	-1	C3_1	20×20×512	512
10	[-1, 6, 4]	MSFFM		256
11	-1	C3-SC	40×40×256	256
12	9	Conv1×1	20×20×512	256
13	-1	Upsample	20×20×256	
14	[-1, 6, 11]	Concat		
15	-1	C3-SC	40×40×768	256
16	-1	Conv1×1	40×40×256	128
17	-1	Upsample	40×40×128	
18	[-1, 4, 11]	Concat1		
19	-1	C3-SC	80×80×512	128
20	-1	Conv3×3	80×80×128	128
21	[-1, 16]	Concat		
22	-1	C3-SC	40×40×256	256
23	-1	Conv3×3	40×40×256	256
24	[-1, 12]	Concat		
25	-1	C3-SC	20×20×512	512
26	[19, 22, 25]	Detect		

1)序列号10为多尺度特征融合模块,序列号11、15、19、22、25为C3-SC模块,序列号18为多尺度维度拼接模块;C3_X表示C3模块中串联的瓶颈模块为X个

1) Serial number 10 is the proposed multi-scale feature fusion module, serial number 11, 15, 19, 22, 25 are the proposed C3-SC modules, serial number 18 is the multi-scale dimension module; C3_X represents the number of bottleneck modules in series of C3 module is X

1.4 试验参数设置与评价指标

1.4.1 试验平台与模型参数设置 在 Windows 10 操作系统下,基于 Python3.8.1 和 Pytorch1.8.0 搭建的深度学习框架进行网络模型的训练与测试。试验所采用的 CPU 为 Intel Xeon Glod 6 240@2.60 GHz, 192 Gb RAM, GPU 为 NVIDIA GeForce GTX 3090Ti×2, 使用 CUDA 11.1 版本并行计算框架配合 CuDNN 8.0.5 版本的深度神经网络加速库。YOLOv5-C 模型训练时输入图像大小为 640×640, 采用 Adam 优化器,训练批次为 16,训练迭代次数为 300。在网络模型的训练阶段,设置初始学习率为 0.001,使用余弦退火衰减策略更新学习率,目标类别数为 7。

1.4.2 网络模型评价指标 本研究使用准确率 (Precision)、召回率 (Recall)、F1 分数、平均准确率 (Average precision, AP)、平均精度均值 (Mean of average precision, mAP)、模型大小和检测速度作为网络模型的评价指标。其中,准确率、召回率、F1 分数及平均准确率的计算如式 (4)~(7) 所示。

Precision = TP / (TP + FP) × 100%, (4)

Recall = TP / (TP + FN) × 100%, (5)

F1 = (2 × Precision × Recall) / (Precision + Recall) × 100%, (6)

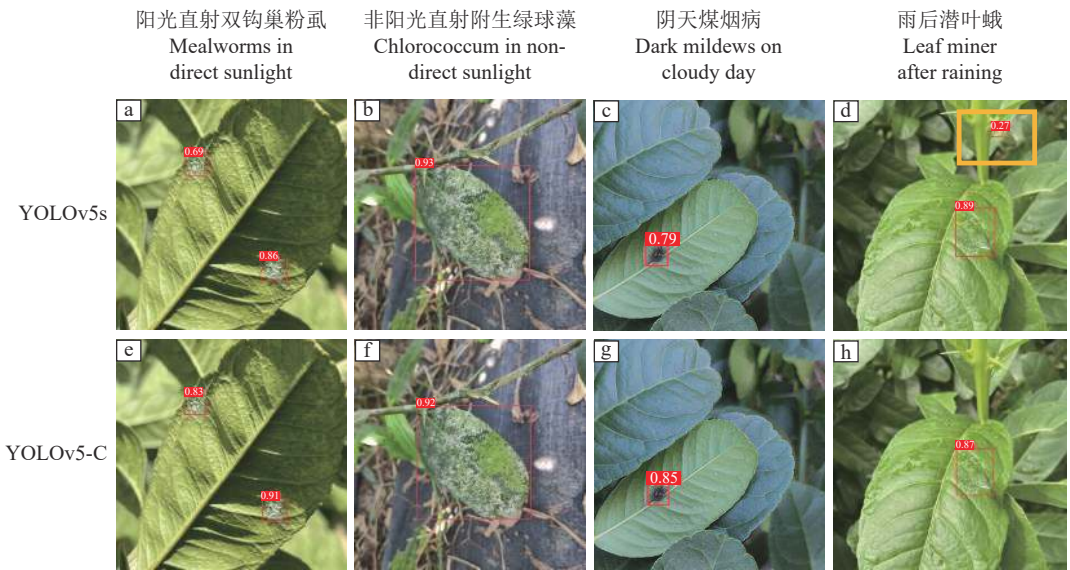
AP = ∫₀¹ Precision(Recall) dRecall × 100%, (7)

式中, TP 表示被正确检测到的病虫害目标数量, FP 表示被网络错误认为是病虫害目标的数量, FN 表示漏检的病虫害目标数量。

2 结果与分析

2.1 YOLOv5-C 与 YOLOv5s 检测结果对比

为了直观地体现改进的 YOLOv5-C 网络模型的性能,分别使用 YOLOv5s 和 YOLOv5-C 网络模型对测试集图像进行检测,图 6 对比展示了这 2 种网络模型对不同天气拍摄的不同病虫害图像的检测结果,图 7 对比展示了 2 种网络模型对与枝干或光斑相似的病斑的检测效果。

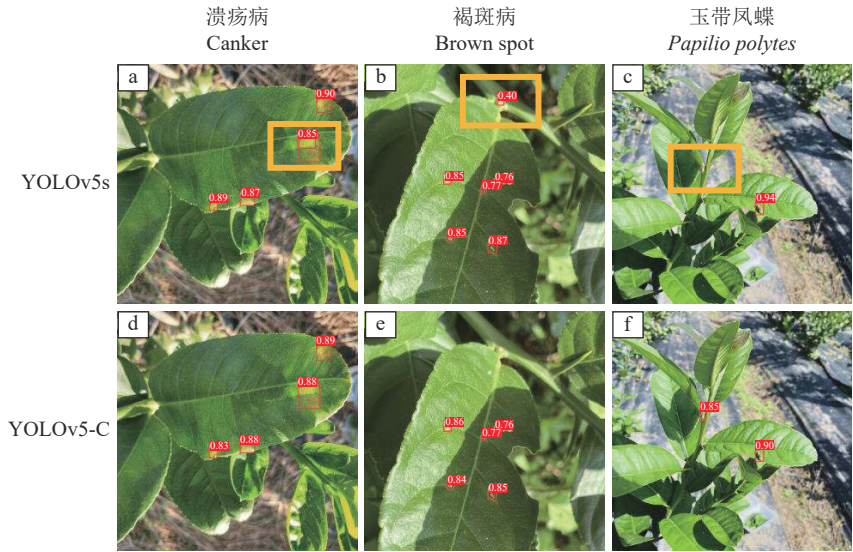


图中数据为置信度; 图中红色框为网络模型的检测结果, 黄色框为人工标注的被网络模型漏检或误检的目标

The data in figures represent confidence levels; The red box in the figure is the detection result of network model, the yellow box is the artificially labeled target that is missed or misdetected by the network model

图 6 不同天气情况下 YOLOv5-C 和 YOLOv5s 检测结果

Fig. 6 Detection results of YOLOv5-C and YOLOv5s under different weather conditions



图中数据为置信度; 图中红色框为网络模型的检测结果, 黄色框为人工标注的被网络模型漏检或误检的目标

The data in figures represent confidence levels; The red box in the figure is the detection result of network model, the yellow box is the artificially labeled target that is missed or misdetected by the network model

图 7 与枝干或光斑相似的病虫害目标的检测效果

Fig. 7 The detection effect of pests and diseases similar to branches or spots

如图 6 所示, 2 种网络模型均能准确识别阳光直射和非阳光直射的广佛手病虫害图像。通过对比图 6a、6c、6e 和 6g 可知, YOLOv5-C 网络模型识别的目标置信度更高。雨后拍摄的潜叶蛾图像中, 叶片和枝干均有水珠, 增加了目标检测的难度, 通过对比图 6d 和 6h 可知, YOLOv5s 网络模型存在误检的病虫害目标, 即图 6d 中黄色框内的红色目标框, 误将水珠识别为潜叶蛾病虫, 而 YOLOv5-C 网络模型能准确识别病虫害目标, 抗干扰能力更强。

部分广佛手病虫害的病斑与枝干或光斑相似, 直接使用 YOLOv5s 网络模型进行检测易造成病虫害目标的误检或漏检, YOLOv5s 网络模型的检测结果如图 7 中的 a~c 所示, 其中, 图 7a 中存在重复框选同一目标的情况, 图 7b 中误将枝干上的光斑检测为褐斑病目标, 图 7c 中漏检了枝干处的玉带凤蝶目标。YOLOv5-C 网络模型能有效抑制环境干扰, 准确识别与枝干或光斑相似的广佛手病虫害目标, 检测结果如图 7d~7f 所示。由图 6、7 可知,

YOLOv5-C 网络模型具有更优的抗干扰能力,能更准确地识别复杂背景下广佛手病虫害目标。

2.2 消融试验

为验证所提出的各种改进策略的有效性,本研究进行了一系列消融试验,在同一平台同一试验框架,使用相同的训练参数对网络模型进行训练,在测试集中分别验证了所提出的改进模块对网络模型性能的影响,以及多个模块共同作用对网络模型性能的改进效果,结果如表 3 所示。

表 3 各个模块消融试验

Table 3 Ablation experiments for each module

网络 模型 Model	准确 率/% Precision	召回 率/% Recall	F1 分数/% F1 score	精度 均值/% Mean average precision	模型大 小/Mb Model size
YOLOv5s	90.82	86.50	88.61	90.61	14.4
YOLOv5s+MSFFM	92.31	86.43	89.27	91.93	15.8
YOLOv5s+C3-S	92.62	86.41	89.41	91.20	12.7
YOLOv5s+C3-SC	93.41	88.06	90.66	92.32	12.7
YOLOv5-C	93.00	88.99	90.95	93.06	14.1

从表 3 可以看出,添加了 MSFFM 模块的网络模型,相比于原始的 YOLOv5s 网络模型,准确率提高了 1.49 个百分点,平均精度均值提高了 1.32 个百分点,说明 MSFFM 模块能提高网络模型的特征提取和特征融合能力,丰富所提取的特征信息,提高网络模型的识别准确率。使用 C3-S 模块替换 YOLOv5s 原始网络中 PANet 结构中的 C3 模块,使网络模型减小了 1.7 Mb,小幅度提高了网络检测效果,进一步说明了 C3-S 模块能在不影响网络模型识别性能的基础上,使网络模型轻量化。C3-SC 模块是在 C3-S 模块中添加 CBAM 注意力机制模块,使用 C3-SC 模块替换原始 YOLOv5s 网络模型 PANet 结构中的 C3 模块,相较于原始 YOLOv5s 网络模型,准确率提高了 2.56 个百分点,召回率提高了 1.56 个百分点,说明 C3-SC 模块能提高检测目标的特征提取能力,抑制背景干扰,提高网络模型的识别准确率;F1 分数提高了 2.05 个百分点,说明 C3-SC 模块能提高网络模型的综合性能,使网络模型更稳健;由于 CBAM 模块的参数少,所以网络模型的大小仍保持为 12.7 Mb。

试验结果表明,本研究所提出的每个改进策略均能提高网络模型的性能。相比原始的 YOLOv5s 网络模型,YOLOv5-C 网络模型具有更好的识别性能,mAP 提高了 2.45 个百分点,F1 分数提高了

2.34 个百分点,网络模型减小了 0.3 Mb。

2.3 消融试验中各个类别的平均准确率对比

为进一步分析所提出的网络模型对各个病虫害类别识别准确率的影响,消融试验中广佛手潜叶蛾虫害、双钩巢粉虱虫害、玉带凤蝶虫害、褐斑病、附生绿球藻病、溃疡病和煤烟病的 AP 如图 8 所示。

观察图 8 可知,原始的 YOLOv5s 网络模型,呈现出各个病虫害类别的 AP 差异较大,所提出的 MSFFM 模块和 CBAM 模块能改善这一情况。经计算,原始 YOLOv5s 网络模型各个类别 AP 的标准差为 7.14,变异系数为 7.88%;所提出的 YOLOv5-C 网络模型各个类别 AP 的标准差为 3.13,变异系数为 3.36%,各个病虫害类别的 AP 更均衡。

所提出的 YOLOv5-C 网络模型,在提高网络模型识别准确率的同时,综合考虑了各个病虫害类别的识别效果,该网络模型的平均精度均值为 93.06%。相比于原始的 YOLOv5s 网络模型,潜叶蛾病虫害和煤烟病的平均准确率分别提高了 14.17 和 2.71 个百分点。所提出的 YOLOv5-C 网络模型各个类别的 AP 情况,更符合广佛手病虫害识别的实际应用需求。

2.4 不同检测模型性能对比

将本文所提出的算法与 RetinaNet、SSD、Efficientdet-D1、YOLOv4、主干网络为 MobilenetV1 的 YOLOv4、YOLOv4-tiny 和 YOLOv5s 目标检测网络进行对比,不同网络的检测结果如表 4 所示。由表 4 数据可知,YOLOv5-C 算法获得了最高的召回率、F1 分数、平均精度均值,其中平均精度均值比 RetinaNet、Efficientdet-D1、SSD、YOLOv4 的分别高 22.30、20.65、4.84 和 2.36 个百分点,其中,所提出的 YOLOv5-C 网络模型大小仅为 YOLOv4 网络模型的 5.77%。相比于轻量级结构 YOLOv4-mobilenet 和 YOLOV4-tiny, YOLOv5-C 的平均精度均值分别高了 7.43 和 24.06 个百分点。本研究所构建的广佛手病虫害数据集,是由自然环境下拍摄的发病早期的病虫害目标构成,图像背景复杂,病斑、虫体的体积较小,且形态多变,常用的目标检测网络易产生漏检、误检,检测精度不高。

相较于上述常用的目标检测算法,本研究所提出的 YOLOv5-C 网络模型具有更优的识别性能,其中,F1 分数为 90.95%,平均精度均值为 93.61%,网络模型大小仅为 14.1Mb,每张图像的平均检测时间仅为 0.01s, YOLOv5-C 网络模型更适合完成在自然环境下对广佛手病虫害的识别任务。

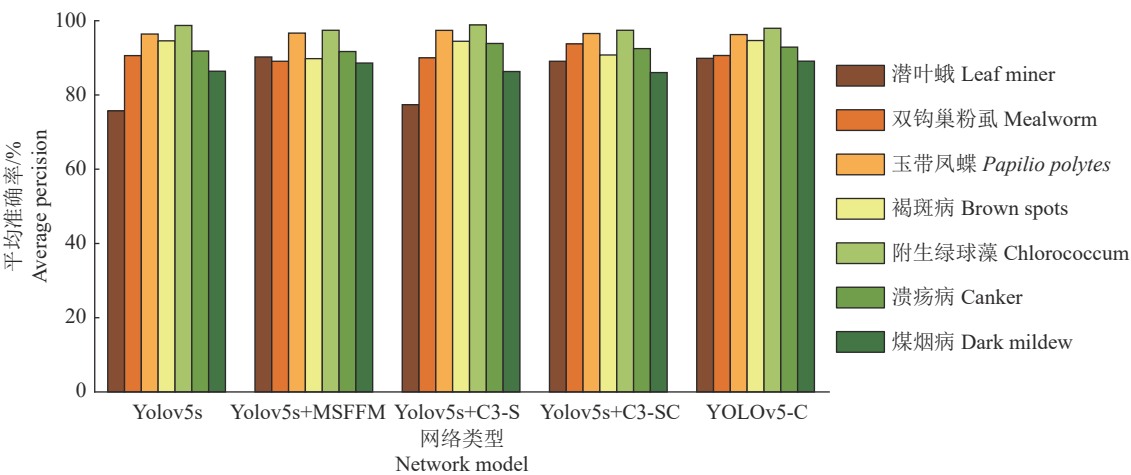


图 8 消融试验中各个类别识别的平均准确率

Fig. 8 Average accuracy of each category in ablation experiments

表 4 不同网络的检测结果对比

Table 4 The detection results of different network

网络模型	骨干网络	准确率/%	召回率/%	F1分数/%	精度均值/%	模型大小/Mb	检测时间/s
Network model	Backbone network	Precision	Recall	F1 score	Mean average precision	Model size	Test time
RetinaNet	ResNet-50	89.13	63.49	74.16	70.76	139.4	0.037
SSD	VGG16	91.59	76.71	83.49	88.22	93.7	0.018
Efficientdet-D1	EfficientNet	93.85	63.77	75.94	72.41	25.7	0.084
YOLOv4	CSPDarknet53	90.55	87.28	88.88	90.70	244.5	0.042
YOLOv4-mobilenet	MobileNetV1	85.61	80.60	83.03	85.63	51.2	0.026
YOLOV4-tiny	CSPDarknet53-tiny	90.50	63.50	74.63	69.00	22.5	0.012
YOLOv5s	CSPDarknet53	90.82	86.50	88.61	90.61	14.4	0.010
YOLOv5-C	CSPDarknet53	93.00	88.99	90.95	93.06	14.1	0.010

3 结论

本研究提出了 1 种基于 YOLOv5-C 目标检测网络的复杂背景下广佛手病虫害识别方法。为均衡提高每一类广佛手病虫害的识别准确率,提出了一种多尺度特征融合模块 MSFFM,该模块与 CBAM 注意力机制模块共同作用,使得每一类广佛手病虫害的平均识别准确率均高于 89.11%。在保障每一个病虫害类别都有较高识别准确率的前提下,综合提高网络模型的识别准确率。

本研究为减少网络模型的参数,提出了一种改进的 C3-SC 模块,在不影响网络模型识别性能的前提下,将该模块与 MSFFM、CBAM 模型相结合,构建了轻量级目标检测网络 YOLOv5-C,网络模型大小为 14.1Mb。

本研究构建了自然环境下的广佛手病虫害数据集用于网络模型的训练与测试,试验结果表明,所提出的 YOLOv5-C 网络模型的检测准确率为 93%,召回率为 88.99%,F1 分数为 90.95%,平均精

度均值为 93.06,在 GPU 上每张图像的平均检测时间为 0.01 s,本研究所提出的网络模型的各项性能均高于常用的目标检测算法,具有较高的应用价值,可为后续自然场景下农作物病虫害的精准识别研究提供参考。

本研究仍存在一定的局限性,广佛手病虫害的数据采集与标注需要耗费大量时间和精力,今后将进一步研究病虫害数据集的智能构建,促进智慧农业的发展。

参考文献:

[1] 曹瑞,郝二伟,杜正彩,等. 广佛手化学成分、药理作用的研究进展及质量标志物的预测分析[J/OL]. [2022-02-11]. <https://kns.cnki.net/kcms/detail/21.1546.r.20220210.1726.076.html>.

[2] 张思获,杨海燕,曾俊,等. 佛手的研究进展[J]. 中华中医药杂志, 2018, 33(8): 3510-3514.

[3] 翟肇裕,曹益飞,徐焕良,等. 农作物病虫害识别关键技术综述[J]. 农业机械学报, 2021, 52(7): 1-18.

[4] LI L, ZHANG S, WANG B. Plant disease detection and classification by deep learning: A review[J]. IEEE Ac-

- cess, 2021, 9: 56683-56698.
- [5] 兰玉彬, 王天伟, 陈盛德, 等. 农业人工智能技术: 现代农业科技的翅膀[J]. *华南农业大学学报*, 2020, 41(6): 1-13.
- [6] YUAN Y, FANG S, CHEN L. Crop disease image classification based on transfer learning with DCNNs[C]// Pattern Recognition and Computer Vision. Cham: Springer, 2018: 457-468.
- [7] CHEN J, CHEN J, ZHANG D, et al. Using deep transfer learning for image-based plant disease identification[J]. *Computers and Electronics in Agriculture*, 2020, 173: 105393.
- [8] LÜ M, ZHOU G, HE M, et al. Maize leaf disease identification based on feature enhancement and dms-robust alexnet[J]. *IEEE Access*, 2020, 8: 57952-57966.
- [9] 陆健强, 林佳翰, 黄仲强, 等. 基于 Mixup 算法和卷积神经网络的柑橘黄龙病果实识别研究[J]. *华南农业大学学报*, 2021, 42(3): 94-101.
- [10] HASSAN S M, MAJI A K. Plant disease identification using a novel convolutional neural network[J]. *IEEE Access*, 2022, 10: 5390-5401.
- [11] ZOU Z, SHI Z, GUO Y, et al. Object detection in 20 years: A survey[EB/OL]. 2019, arXiv: 1905.05055. <https://arxiv.org/abs/1905.05055>.
- [12] 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: 91-99.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [14] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 779-788.
- [15] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 7263-7271.
- [16] REDMON J, FARHADI A. Yolov3: An incremental improvement[EB/OL]. 2018, arXiv: 1804.02767. <http://arxiv.org/abs/1804.02767>.
- [17] BOCHKOVSKIY A, WANG C, LIAO H M. Yolov4: Optimal speed and accuracy of object detection[EB/OL]. 2020, arXiv: 2004.10934. <https://arxiv.org/abs/2004.10934>.
- [18] ZHOU G, ZHANG W, CHEN A, et al. Rapid detection of rice disease based on FCM-KM and faster R-CNN fusion[J]. *IEEE Access*, 2019, 7: 143190-143206.
- [19] SUN J, YANG Y, HE X, et al. Northern maize leaf blight detection under complex field environment based on deep learning[J]. *IEEE Access*, 2020, 8: 33679-33688.
- [20] LÜ 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.
- [21] MAMDOUH N, KHATTAB A. YOLO-based deep learning framework for olive fruit fly detection and counting[J]. *IEEE Access*, 2021, 9: 84252-84262.
- [22] 张博, 张苗辉, 陈运忠. 基于空间金字塔池化和深度卷积神经网络的作物害虫识别[J]. *农业工程学报*, 2019, 35(19): 209-215.
- [23] 胡嘉沛, 李震, 黄河清, 等. 采用改进 YOLOv4-Tiny 模型的柑橘木虱识别[J]. *农业工程学报*, 2021, 37(17): 197-203.
- [24] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50× fewer parameters and <0.5 MB model size[EB/OL]. 2016, arXiv: 1602.07360. <https://arxiv.org/abs/1602.07360>.
- [25] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with Convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Boston: IEEE, 2015: 1-9.
- [26] LIN T, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2117-2125.
- [27] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 8759-8768.
- [28] WANG C, LIAO H M, WU Y, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C/OL]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. IEEE Computer Society, 2020: 390-391.
- [29] 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.
- [30] WOO S, PARK J, LEE J, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). Cham: Springer, 2018: 3-19.