

李芳, 危疆树, 王玉超, 等. 基于改进 YOLOv8n 模型的辣椒病害检测方法[J]. 江苏农业学报, 2025, 41(2): 323-334.
doi: 10.3969/j.issn.1000-4440.2025.02.013

基于改进 YOLOv8n 模型的辣椒病害检测方法

李芳¹, 危疆树¹, 王玉超², 张尧¹, 谢宇鑫¹

(1. 四川农业大学信息工程学院, 四川 雅安 625014; 2. 四川农业大学机电学院, 四川 雅安 625014)

摘要: 为了解决辣椒病害检测速度慢, 漏检率和误检率高的问题, 本研究以 YOLOv8n 为基线模型, 引入 Adown 下采样模块替代原模型骨干网络 (Backbone) 的卷积下采样层, 引入 SlimNeck 模块将原模型颈部网络中的卷积层和特征聚合模块 (C2f) 替换为混合卷积模块 (GSConv) 和跨阶段部分网络 (VoVGSCSP) 模块, 并利用辅助训练头 Aux Head (Auxiliary head) 融合原有的检测头, 构建改进的 YOLOv8n 模型 (YOLOv8n-ATA 模型)。最后利用辣椒炭疽病、褐斑病、脐腐病和细菌性叶斑病等 4 种病害影像数据集对改进后的模型性能进行分析。结果表明, 改进后模型的浮点计算量和模型大小比原 YOLOv8n 模型增加 19.5% 和 10.2%, 但模型对辣椒病害的识别精确率、平均精度均值 mAP_{50} 和 $mAP_{50:95}$ 分别提升 2.6 个百分点、2.9 个百分点和 2.9 个百分点, 同时每 1 s 传输帧数增加 15.1%。因此, 改进后的模型能够对辣椒病害进行有效识别, 较好实现模型识别准确度与效率的平衡。

关键词: 辣椒病害; YOLOv8n 模型; 目标检测; Adown 下采样模块; SlimNeck 模块; Aux Head 检测头
中图分类号: S24; TP391.4 **文献标识码:** A **文章编号:** 1000-4440(2025)02-0323-12

A chili disease detection method based on an improved YOLOv8n model

LI Fang¹, WEI Jiangshu¹, WANG Yuchao², ZHANG Yao¹, XIE Yuxin¹

(1. College of Information Engineering, Sichuan Agricultural University, Ya'an 625014, China; 2. College of Mechanical and Electrical Engineering, Sichuan Agricultural University, Ya'an 625014, China)

Abstract: To solve the problems of slow detection speed, high missed detection rate, and high false detection rate in chili disease detection, this study used YOLOv8n as the baseline model and constructed an improved YOLOv8n model (YOLOv8n-ATA model). The Adown downsampling module was introduced to replace the convolutional downsampling layer of the Backbone of the original model. The SlimNeck module was introduced to replace the convolutional layer and feature aggregation module (C2f) in the neck network of the original model with the hybrid convolution module (GSConv) and the cross-stage partial network (VoVGSCSP) module. Moreover, the auxiliary training head Aux Head (Auxiliary head) was used to fuse the original detection head. Finally, the performance of the improved model was evaluated using the image datasets of four chili diseases, such as anthracnose, brown spot, blossom-end rot and bacterial leaf spot. The results showed that the floating-point calculations and size of the improved model were 19.5% and 10.2% higher than those of the original YOLOv8n model. However, the identification accuracy, mAP_{50} and $mAP_{50:95}$ of the model for chili diseases increased by 2.6 percentage points, 2.9 percentage points and 2.9 percentage points, respectively. At the same time, the number of frames per second increased by 15.1%. Therefore, the improved model can effectively identify chili diseases and better achieve the balance between accuracy and efficiency of model recognition.

and efficiency of model recognition.

Key words: chili diseases; YOLOv8n model; target detection; Adown downsampling module; SlimNeck module; Aux Head

收稿日期: 2024-06-18

基金项目: 四川省科技厅关键技术攻关项目(22ZDYF0095)

作者简介: 李芳(2001-), 女, 四川巴中人, 硕士研究生, 研究方向为机器视觉和目标检测。(E-mail) 18116718920@163.com

com

通讯作者: 危疆树, (E-mail) weijiangshu66@163.com

辣椒 (*Capsicum annuum* L) 是茄科辣椒属一年

生或有限多年生草本或灌木植物,具有适应性强、风味多样、营养丰富等特点,备受消费者喜爱^[1]。辣椒是中国重要的蔬菜产品,2022 年辣椒种植面积达 $2.23 \times 10^6 \text{ hm}^2$,总产量约 $6.4 \times 10^7 \text{ t}$,产值超过 2.7×10^{11} 元^[2]。辣椒种植过程中,病害是导致辣椒产量和品质下降的主要原因之一,及时精准地诊断辣椒病害不但可以提升辣椒品质与产量,同时还能降低环境负担、增强生产效能^[3]。

随着信息技术的发展,人工智能和图像识别技术在现代农业中得到了广泛应用。目前基于卷积神经网络的深度学习技术越来越受到关注。基于深度学习的目标检测算法主要分为单阶段检测算法与双阶段检测算法两类。单阶段检测算法常以某个神经网络模型直接预测图像中的目标位置和类别,常用的模型有 YOLO 系列^[4]、SSD^[5]、RetinaNet^[6]等,具有快速、简洁及适应性强的优势;双阶段检测算法一般分为两个独立的阶段进行检测,首先提取检测目标所在区域,再进行目标分类和边界框回归。经典的双阶段检测算法有 R-CNN^[7]、Faster R-CNN^[8]、SPP-Net^[9]等,双阶段检测算法具有较高的识别准确度,但识别效率相对较低。

在基于深度学习的作物病害识别方面已有较多研究。储鑫等^[10]在 YOLOv4 模型的基础上,利用 K 均值聚类算法调整先验框维度,将宽度因子为 0.25 的 MobileNetV1 替代 CSPDarknet53 网络来提取特征进行模型优化。优化后的模型参数量比原模型减少 80%,对番茄叶片病害识别的平均精度均值 (mAP) 达 98.76%。施杰等^[11]通过集成 SimAm 注意力机制和引入 EIOU 损失函数对 YOLOv5s 模型进行优化,优化后的模型对玉米病虫害的识别平均精度提高至 88.8%。刘诗怡等^[12]基于改进的 YOLOv7 模型,强化了模型对黄瓜叶片微小病斑的识别性能。周维等^[13]利用轻量化的 GhostNet 网络替换 YOLOv4 的主干网络,构建了 YOLOv4-GhostNet 模型,构建的模型对水稻病虫害的识别精度和效率均得到提升。邹玮等^[14]基于 YOLOv2 模型对辣椒叶片上的蚜虫进行定位和识别,结果显示,该模型对辣椒叶片上的蚜虫识别平均准确率达到 96.49%。尚俊平等^[15]在 MobileNetV3 的基础上,构建了 MobileLU 模型,并用于辣椒病害的识别,结果发现该模型的识别准确率达 98.2%。李西兴等^[16]引入注意力机制对 MaxViT 模型进行了改进,改进后的模型对

辣椒病毒病、根腐病的识别能力显著提升,识别准确率达到 98.10%。姚飞^[17]优化了 Faster R-CNN 模型,并将其用于辣椒炭疽病、病毒病、白粉病、青枯病和疫病等病害的识别,有效降低原模型的漏检率,提高模型的检测效率。丁文宽^[18]基于卷积神经网络和机器视觉开发了自然环境中辣椒果实的识别与定位,为智能采摘机器人的研发提供基础。王启锟等^[19]在 YOLOv5 模型中引入注意力机制,有效降低自然环境中辣椒果实识别中的漏检或错检问题,模型的识别平均准确率达 93.1%。Pratap 等^[20]利用改进的 EffecientNetB4 模型对辣椒卷叶病、叶斑病等进行识别和分类,取得了较好的效果。Kanaparthi 等^[21]利用深度学习技术比较了不同 SqueezeNet 层数和训练参数对模型识别准确率的影响,实现了辣椒叶片病害的早期检测。

虽然利用深度学习技术进行辣椒病害检测已取得较多成果,但上述研究主要针对单一背景的病害,并没有考虑自然环境的复杂性,此外,现有的辣椒病害数据集还不够丰富,难以覆盖众多的辣椒病害,这限制了现有模型在生产中的应用。针对这些问题,本研究以 YOLOv8n 为基线模型,提出了一种改进的辣椒病害检测方法,旨在提高自然环境背景下辣椒病害的检测精度,提高模型效率,为模型在移动设备中的应用提供技术支持。

1 材料与方法

1.1 数据集

本研究选取炭疽病、褐斑病、脐腐病和细菌性叶斑病等 4 种辣椒生产中常见的病害进行识别。病害示例如图 1 所示。原始数据集由 2022 年科大讯飞辣椒病虫害图像识别挑战赛中使用的部分病害图片和课题组在巴中市辣椒种植地拍摄的病害图片构成,共 1 025 张影像。为解决病害样本分布不均衡的问题,通过旋转与缩放技术对原始数据进行增强,增强后的数据集共有 1 937 张影像。使用 LabelImg 软件对所有数据集中的病害类型进行标注,并按照 7:2:1 的比例划分为训练集、验证集和测试集,分别包含 1 355 张影像、387 张影像和 195 张影像。

1.2 辣椒病害检测模型

1.2.1 YOLOv8 模型 YOLOv8 模型是 YOLO 系列中性能较好的单阶段检测模型,相比于 Faster R-CNN 等双阶段目标检测模型具有更快的检测速度和更高的

准确率^[22]。YOLOv8 模型包括 4 个主要部分:接收图像的输入部分,负责图像特征学习和提取的主干网络(Backbone),进行池化和特征融合的颈部(Neck)网络,检测类别和检测目标回归框^[23]及输出的检测头(Head)。YOLOv8 模型有 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 和 YOLOv8x 等版本,它们通过调整网络架构的深度、宽度和参数量等来满足不同的应用需求。随着模型规模的增加,网络的复杂度、计算

量和推理时间都逐渐增加,同时精度也得到提升。YOLOv8n 是最小且最快的版本,适用于资源受限的实时检测任务;YOLOv8s 则适合中等精度需求的场景;YOLOv8m 在精度和速度之间提供良好的平衡,适合复杂应用;YOLOv8l 适用于需要较高精度的任务,如复杂图像分析;而 YOLOv8x 则提供最佳精度,适合高精度要求的任务,如医学影像分析等。

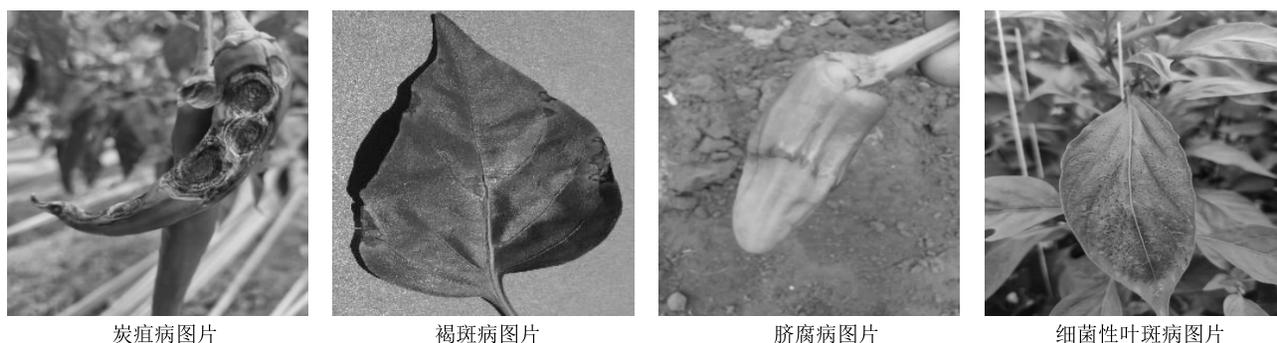


图 1 数据集示例图片

Fig.1 Example images from the dataset

为提高自然环境中辣椒病害检测准确率和效率,本研究在 YOLOv8n 模型的基础上,通过一系列的改进措施构建辣椒病害识别模型 YOLOv8n-ATA,模型的整体结构如图 2 所示。改进措施主要有 3 个方面:首先将 Backbone 中原有的卷积(Conv)下采样层替换为 Adown 下采样层,以降低模型的计算量和复杂度,其次引入轻量块 SlimNeck,将颈部网络中的卷积层和特征聚合块(C2f)替换为混合卷积块(GSConv)和跨阶段部分网络(VoVGSCSP)模块,保留原有颈部网络连接通道,减少参数堆叠,以提升模型检测精度和加速模型的推理;最后用辅助训练头 AUXHead(Auxiliary head)融合原有的检测头,更好地融合辣椒病害深层和浅层特征,并综合主检测头和辅助检测头的预测结果,实现更全面、更准确的目标检测。

1.2.2 Adown 下采样模块 检测目标较小时,常规的下采样操作可能会导致检测目标信息丢失或者模糊,从而影响模型的检测性能。为提升模型捕捉目标特征的能力,减少计算量,本研究引入 Adown 下采样模块^[24],替换 YOLOv8n 主干网络(Backbone)的 3 个卷积(Conv)下采样层,Adown 下采样模块的结构如图 3 所示。

Adown 下采样模块可减少特征图的空间维度,

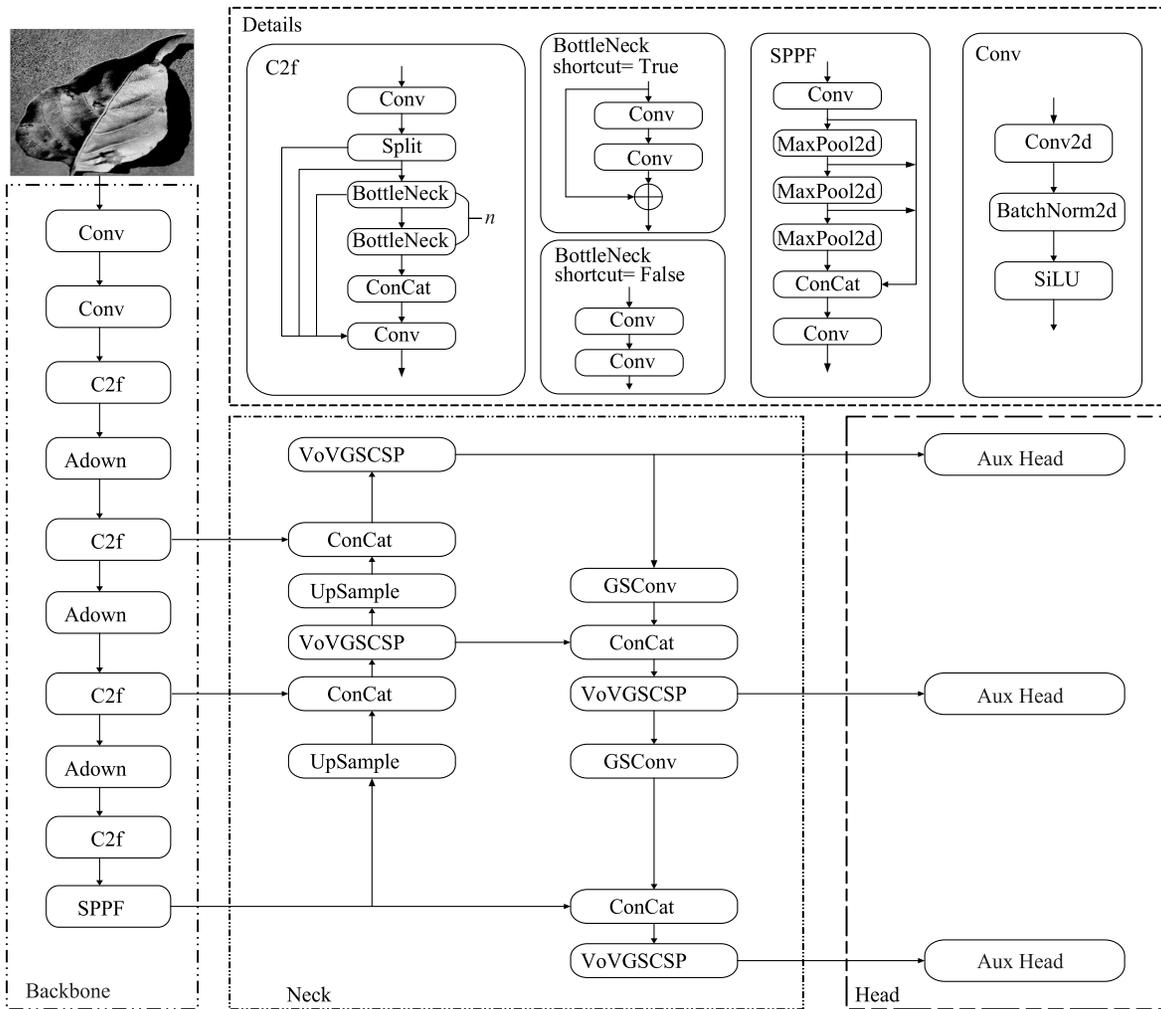
能更好地检测不同尺度的病害信息,特别适合较小检测目标的特征提取和识别。Adown 下采样模块先对输入特征进行平均池化操作,再将结果分为两个部分,分别进行卷积和最大池化处理,这样一方面可加强参数的共享,另一方面可减少模型的参数量,最后将处理后的特征图拼接起来作为下采样处理后的结果返回,这样能保留较多的输入特征,有效减少信息的损失。

1.2.3 SlimNeck YOLOv8n 模型中 Neck 网络通过特征融合技术将不同尺度的特征图进行拼接,恢复因下采样而丢失的空间分辨率的同时,还能保留高层的语义信息。Neck 中的 C2f 模块在特征融合中起重要作用,但其也增加了模型的复杂度,导致训练和推理时间增加。因此,为更好地提升模型性能,本研究将 Neck 中的 Conv 和 C2f 模块替换为 SlimNeck^[25]的混合卷积模块(GSConv)和跨阶段部分网络(VoVGSCSP)模块。

GSConv 是由标准卷积(SC)、深度可分离卷积(DSC)和混洗(Shuffle)操作共同组成的混合卷积。通过 Shuffle 操作,将 SC 生成的信息传递到 DSC 的各个部分,可最大限度地减少 DSC 缺陷对模型的影响并有效利用 DSC 的优势,GSConv 结构如图 4 所示。首先输入一个普通卷积进行下采样,然后使用

深度卷积(DWConv),并将两个 Conv 的结果拼接起来,最后采用重塑、转置等方法对输出张量进行 Shuffle 操作,实现对通道维度的分割和重组。这种分组卷积的方式可以提高模型对不同区域的感知能力,同时提高特征多样性,有效增强目标检测的性能。基于混合卷积模块(GSConv),结合一次性聚合

方法构建跨阶段部分网络模块(VoVGSCSP),如图 5 所示。使用 VoVGSCSP 模块来替换原模型中的 C2f 模块,能够较大程度地减轻计算负担,强化模型提取病害特征的能力,有利于模型较好地捕捉目标关键信息。



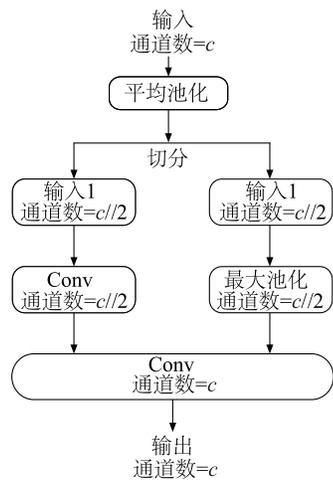
Backbone:主干网络;Neck:颈部网络;Head:检测头(预测网络);Conv:卷积层;C2f:特征聚合模块;Adown:下采样卷积模块;SPPF:快速空间金字塔池化;VoVGSCSP:跨阶段部分网络模块;UpSample:上采样;Concat:连接操作;GSConv:混合卷积模块;Aux Head:Auxiliary head 检测头;Split:切分操作;BottleNeck:瓶颈模块;MaxPool2d:最大池化;BatchNorm2d:归一化;SiLU:激活函数。

图 2 YOLOv8n-ATA 模型结构图

Fig.2 Structure diagram of YOLOv8n-ATA model

1.2.4 Aux Head 检测头 自然环境中,辣椒病害的病斑尺寸及空间位置往往存在较大差异,这给病害特征的提取和目标检测带来了极大的挑战。此外,训练初期主检测头可能会受到梯度消失或梯度爆炸的影响而导致学习困难。针对这些问题,本研究在 YOLOv8n 的检测头部分增加 Aux 辅助训练头^[26],

构成 Aux Head 检测头(图 6)。辅助检测头可以提供额外的梯度反馈,能加速模型的收敛。Aux 辅助训练头仅在训练时使用,而在模型推理过程中仍使用普通检测头。Aux Head 有助于网络中间层学习到更多的信息,从而帮助模型训练,增强模型检测不同尺寸目标的性能。



Conv: 卷积层。

图 3 Adown 下采样模块结构图

Fig.3 Structure diagram of the Adown downsampling module

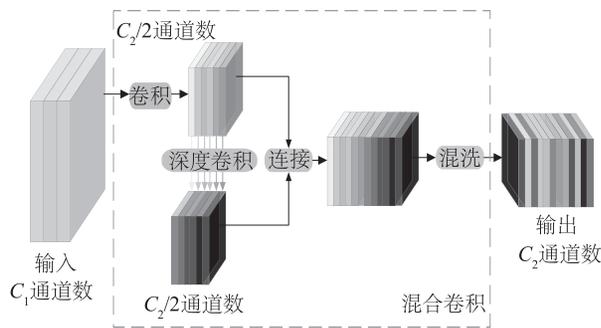


图 4 GSConv 结构图

Fig.4 Structure diagram of the GSConv module

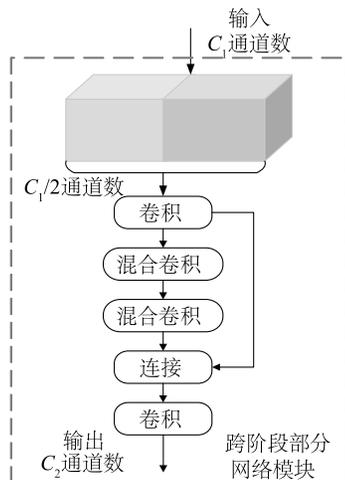
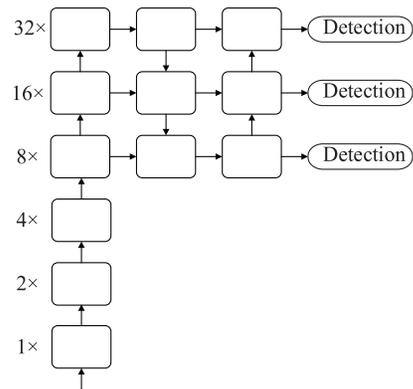


图 5 VoVGSCSP 模块结构图

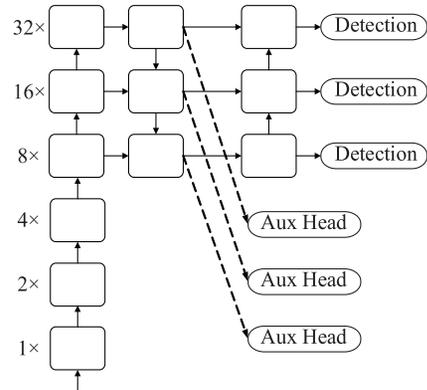
Fig.5 Structure diagram of the VoVGSCSP module

1.3 模型对辣椒病害的识别性能比较

首先比较了 YOLOv8n、YOLOv8s、YOLOv8m、



(a)标准检测头



(b)改进后的Aux Head检测头

Detection: YOLOv8 原始标准检测头; Aux Head: Aux 辅助训练头。

图 6 标准检测头与 Aux Head

Fig.6 Standard detection head and Aux Head

YOLOv8l 和 YOLOv8x 模型对辣椒炭疽病的识别性能;然后比较了 YOLOv8n 和 YOLOv8n-ATA 模型对辣椒病害识别性能的差异,并开展模型改进方案的消融试验(表 1),明确不同改进模块对辣椒病害识别性能的影响,最后将改进后的 YOLOv8n-ATA 模型与其他主流目标检测模型 YOLOv5s、YOLOv7-Tiny、YOLOv7-W6、YOLOv9-Tiny、SSD 和 RetinaNet 进行比较,以验证改进后的模型在辣椒病害识别中的优势。

1.4 系统开发环境及模型评价指标

1.4.1 模型系统开发环境及参数设置 本研究中改进的 YOLOv8n 模型在 Windows10 专业版操作系统下开发,开发环境为 Pycharm,所用语言为 Python3.8.18, CPU 型号为 Intel (R) Xeon (R) Gold 6230R CPU @ 2.10GHz,运行内存为 256 GB, GPU 型号为 NVIDIA GeForce RTX 4090,显存为 24 GB,并使用 CUDA 对 GPU 进行加速。基于深度学习框

架 Torch 1.9.1 进行模型训练,具体参数设置如下:图片尺寸 640×640 像素,使用随机梯度下降算法(SGD)优化器,共迭代训练 200 次,学习率 0.008,动量(Momentum)0.937,批量大小(BatchSize)8。在训练过程中使用马赛克(Mosaic)数据增强,不使用混合提升(MixUp)数据增强。为了使模型更准确地学习,不使用预训练权重,模型从头开始训练。

表 1 消融试验方案设计

Table 1 Design of ablation experiment plan

方案	基线模型	Adown	SlimNeck	Aux Head
0	√	-	-	-
1	√	√	-	-
2	√	-	√	-
3	√	-	-	√
4	√	√	√	-
5	√	√	-	√
6	√	-	√	√
7	√	√	√	√

√表示方案中采用;-表示方案中不采用。

1.4.2 模型评价指标 研究中为表达模型对辣椒病害的检测性能及检测目标在图像中的具体位置和置信度,采用混淆矩阵(Confusion matrix)和热力图(Heatmap)进行模型对辣椒病害识别性能的评估。具体的指标包括精确率(Precision)、召回率(Recall)、交并比(Intersection over Union, IoU)阈值为 0.5 时的平均精度均值(mAP_{50})、IoU 阈值为 0.5 到 0.95 范围的平均精度均值($mAP_{50:95}$)、浮点运算量、模型大小、参数量和每 1 s 传输帧数(FPS),计算方法如下:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

表 2 不同版本 YOLOv8 模型对炭疽病害的识别效果

Table 2 The recognition performance of different versions of the YOLOv8 model on anthrax diseases

版本	精确率 (%)	召回率 (%)	平均精度均值 mAP_{50} (%)	平均精度均值 $mAP_{50:95}$ (%)	浮点运算量 (G)	模型大小 (M)	参数量 ($\times 10^6$)	传输帧数 (帧, 1 s)
YOLOv8n	95.2	89.5	93.2	54.9	8.2	5.9	3.0	31.2
YOLOv8s	95.7	89.5	92.2	54.8	28.6	21.4	11.1	14.5
YOLOv8m	95.9	90.4	92.1	56.1	79.1	49.6	25.9	3.9
YOLOv8l	94.9	89.1	92.3	55.0	165.4	83.6	43.6	0.9
YOLOv8x	95.0	90.4	93.4	56.4	258.1	130.0	68.2	0.3

$$AP = \int_0^1 Precision(Recall) dRecall \quad (3)$$

$$mAP = \frac{1}{C} \sum_{i=1}^C AP \quad (4)$$

$$FPS = \frac{1000}{P+I+PS} \quad (5)$$

式中, C 为病害类别数量,研究中取值为 4; TP 表示预测为正样本实际也为正样本的数量; FP 表示预测为正样本但实际为负样本的数量; FN 表示预测为负样本但实际为正样本的数量; P 为数据从原始输入到模型输入之间的处理时间; I 为模型对图像进行前向推理并生成预测结果所需的时间; PS 为模型输出之后进行的后续处理时间。

2 结果与分析

2.1 不同版本 YOLOv8 模型对数据集中炭疽病害识别效果

5 个版本 YOLOv8 模型对数据集中辣椒炭疽病害识别效果如表 2 所示。从表可以看出,YOLOv8x 模型对辣椒炭疽病害识别的精确率(Precision)、召回率(Recall)、平均精度均值 mAP_{50} 和 $mAP_{50:95}$ 比 YOLOv8n 模型提升 0.2 个百分点、0.9 个百分点、0.2 个百分点和 1.5 个百分点,但其浮点运算量、模型大小和模型参数量分别是 YOLOv8n 模型的 31.5 倍、22.0 倍、22.7 倍,每 1 s 传输帧数(FPS)亦远低于 YOLOv8n 模型。YOLOv8s 模型的精确率比 YOLOv8n 高 0.5 个百分点,但其平均精度均值 mAP_{50} 和 $mAP_{50:95}$ 均比 YOLOv8n 模型低,且其浮点运算量、模型大小和模型参数量分别是 YOLOv8n 模型的 3.5 倍、3.6 倍、3.7 倍。上述结果说明,复杂的模型结构虽然能提升识别精度,但会极大地增加计算量,降低模型效率。因此,从模型轻量化的角度,本研究选择 YOLOv8n 作为基线模型是适宜的。

2.2 消融试验结果

不同改进模块对模型识别性能影响的消融试验结果如表 3 所示。从表中可以看出,3 个改进模块均能有效提高模型的检测性能,提升模型识别效率。利用 A_{down} 下采样模块替换 YOLOv8n 模型骨干网络(Backbone)中原有的 Conv 下采样模块的方案 1,模型对辣椒病害识别的召回率下降 0.9 个百分点,但 mAP_{50} 、 $mAP_{50:95}$ 和精确率较基线模型(方案 0)提升 0.1 个百分点、0.9 个百分点和 2.2 个百分点,且浮点运算量和模型大小分别降低 8.5% 和 11.9%,说明 A_{down} 模块能够有效简化模型结构,优化计算效率,加速模型处理速度。引入 SlimNeck 模块的方案 2,模型对辣椒病害识别的召回率下降 0.7 个百分点,但 mAP_{50} 、 $mAP_{50:95}$ 和精确率较基线模型分别提升 0.5 个百分点、1.5 个百分点和 0.8 个百分点,浮点运算量和模型大小分别降低 9.8% 和 5.1%,说明 GSConv 和 VoVGSCSP 模块能够提高模型的感受野和特征提取能力,有效地学习辣椒病害的特征。引入 Aux Head 模块的方案 3,模型对辣椒病害识别的召回率、 mAP_{50} 、 $mAP_{50:95}$ 和精确率较基线模型分别提升 0.4 个百分点、1.5 个百分点、1.1 个百分点和 0.2 个百分点,而浮点运算量和模型大小分别增

加 36.6% 和 25.4%,说明 Aux Head 模块的应用能提升模型的精确率、 mAP 等,但也一定程度上增加了模型的复杂度,主要是因为它引入了额外的网络分支和计算流程,以及需要额外的资源来计算和反向传播额外的损失函数。同时引入 A_{down} 模块和 SlimNeck 模块的方案 4、引入 A_{down} 模块和 Aux Head 的方案 5 以及引入 SlimNeck 模块和 Aux Head 的方案 6,模型的复杂度虽有一定程度的提升,但模型亦能提取出更多的特征信息,避免特征丢失,从而使得模型对辣椒病害识别的召回率、 mAP_{50} 、 $mAP_{50:95}$ 和精确率均有不同程度的提升。同时引入 A_{down} 模块、SlimNeck 模块和 Aux Head 模块的方案 7(YOLOv8n-ATA 模型)对辣椒病害识别的召回率、 mAP_{50} 、 $mAP_{50:95}$ 和精确率较基线模型分别提升 3.5 个百分点、2.9 个百分点、2.9 个百分点和 2.6 个百分点,结果表明在模型浮点运算量和模型大小分别增加 19.5% 和 10.2% 条件下, mAP_{50} 、 $mAP_{50:95}$ 、精确率和召回率分别达到 85.2%、44.7%、87.3% 和 79.9%。综上所述,改进后的 YOLOv8n-ATA 模型虽然浮点运算量和模型大小有所增加,但模型提取病害特征信息的能力得到了较大程度的提升。即改进方案能有效均衡模型的精度和效率。

表 3 消融试验结果

Table 3 Results of ablation experiments

方案	精确率 (%)	召回率 (%)	平均精度均值 mAP_{50} (%)	平均精度均值 $mAP_{50:95}$ (%)	浮点运算量 (G)	模型大小 (M)	参数量 ($\times 10^6$)	传输帧数 (帧, 1 s)
0	84.7	76.4	82.3	41.8	8.2	5.9	3.0	31.2
1	86.9	75.5	82.4	42.7	7.5	5.2	2.6	34.0
2	85.5	75.7	82.8	43.3	7.4	5.6	2.8	34.7
3	84.9	76.8	83.8	42.9	11.2	7.4	3.8	33.4
4	86.0	77.8	83.5	43.0	6.9	5.1	2.5	35.1
5	87.2	76.3	82.8	43.6	10.5	6.6	3.4	34.1
6	85.3	78.3	84.2	43.4	10.4	7.1	3.6	35.2
7	87.3	79.9	85.2	44.7	9.8	6.5	3.3	35.9

方案 0~7 见表 1。

不同消融试验方案的模型识别结果热力图如图 7 所示。从图中可以看出,消融试验方案 1 至方案 6,模型在识别时都能成功关注到炭疽病害,但普遍存在将无病害的背景或叶片误识别为病斑的情况。相比之下,原模型错误地将部分叶片识别为细菌性叶斑病病斑,漏检了真正的炭疽病病斑。而采用 3 个改进模块的模型能够正确识别病斑,不存在误检情况。

2.3 改进模型 YOLOv8n-ATA 对辣椒病害的识别效果

2.3.1 YOLOv8n 模型改进前后的识别效果比较

YOLOv8n 模型改进前后对辣椒病害的识别效果如表 4 所示。从表中可以看出,改进后的模型(YOLOv8n-ATA)对辣椒病害的识别效果比原 YOLOv8n 模型有显著的提升。虽然 YOLOv8n-ATA 的浮点运算量和模型大小比原 YOLOv8n 模型增加

19.5%和 10.2%,但辣椒病害的识别精确率、召回率、 mAP_{50} 和 $mAP_{50:95}$ 分别提升 2.6 个百分点、3.5 个百分点、2.9 个百分点和 2.9 个百分点,同时每 1 s 传输帧数增加 15.1%,提升 4.7s。由此可以看出,本研究改进后的模型能更好地实现模型复杂度和准确率的平衡。YOLOv8n 模型和 YOLOv8n-ATA 模型

在验证集上的 mAP_{50} 和损失变化曲线如图 8 所示。从图中可以看出,随着迭代次数的增加,2 个模型的平均精度均值快速增加并达到稳定,而损失则快速下降并达到稳定。但在相同的迭代次数下,YOLOv8n-ATA 模型拥有更高的平均精度均值和更低的损失值。

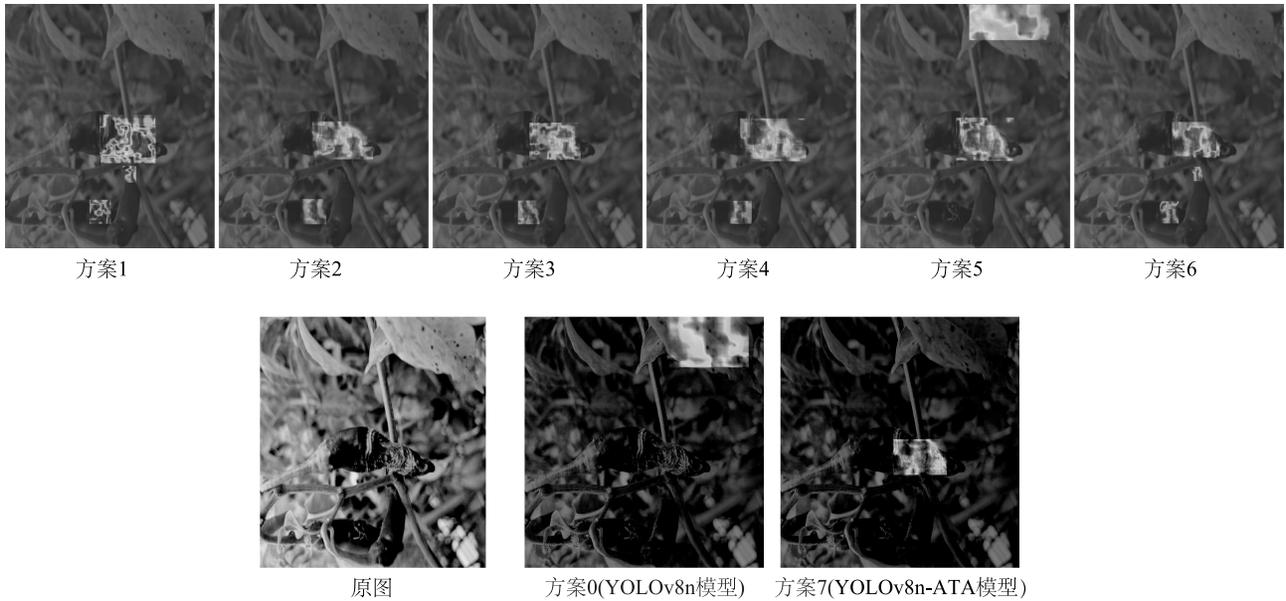


图 7 不同消融试验方案的模型识别结果热力图

Fig.7 Heat maps of model recognition results for different ablation experiment schemes

表 4 模型改进效果对比

Table 4 Comparison of model improvement effects

模型	精确率 (%)	召回率 (%)	平均精度均值 mAP_{50} (%)	平均精度均值 $mAP_{50:95}$ (%)	浮点运算量 (G)	模型大小 (M)	参数量 ($\times 10^6$)	传输帧数 (帧, 1 s)
YOLOv8n	84.7	76.4	82.3	41.8	8.2	5.9	3.0	31.2
YOLOv8n-ATA	87.3	79.9	85.2	44.7	9.8	6.5	3.3	35.9

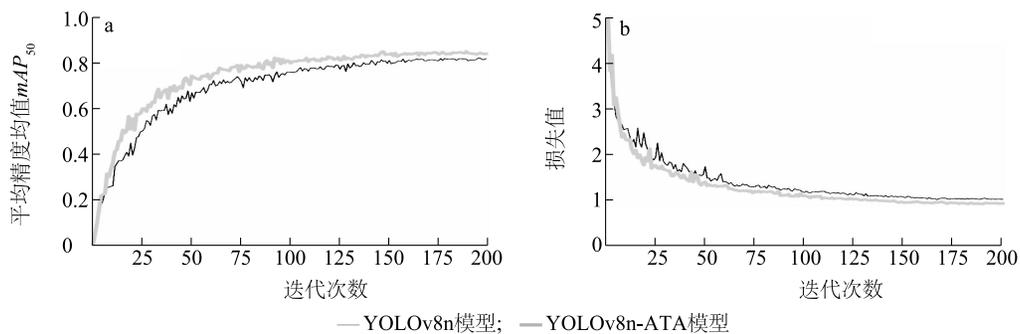


图 8 YOLOv8n 模型与 YOLOv8n-ATA 模型对辣椒病害的识别效果比较

Fig.8 Comparison of the recognition performance of YOLOv8n and YOLOv8n-ATA models for chili diseases

YOLOv8n 模型与 YOLOv8n-ATA 模型对测试集

辣椒病害识别结果的可视化效果如图 9 所示。从图

中可以看出,改进前的模型(YOLOv8n 模型)在较复杂环境下对辣椒病害存在误检,将非病害信息误判为炭疽病和脐腐病,显示 YOLOv8n 模型的学习能力和检测准确度相对较低。此外,原模型对病害的边

缘和整体信息学习不足,导致脐腐病出现多个检测框,同时对褐斑病和细菌性叶斑病存在漏检等现象。相比之下,改进后的 YOLOv8n-ATA 模型能显著减少误检和漏检。

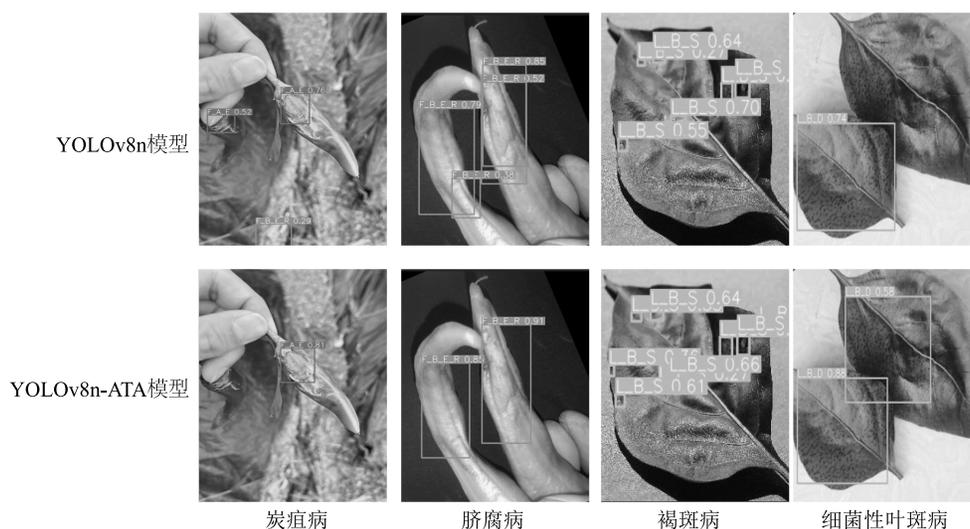


图9 模型改进效果可视化对比

Fig.9 Visualization comparison of model improvement effects

2.3.2 混淆矩阵 利用 YOLOv8n 模型与 YOLOv8n-ATA 模型对验证集中辣椒病害识别的混淆矩阵如图 10 所示。由图中可以看出, YOLOv8n-ATA 模型在辣椒病害识别方面表现出色,与基线模型(YOLOv8n 模型)相比,改进后的模型对各类别病害的检测准确率均有所提升,正确分类病害样本的数量分别从原来的 217 个、189 个、161 个和 96 个增加

到 218 个、194 个、161 个和 96 个。另外, YOLOv8n-ATA 模型还消除了基线模型对于脐腐病的 3 个错检点,且误判为背景的次数减少。总的来说,改进后的模型(YOLOv8n-ATA 模型)在所有类别病害识别上均展现了更好的性能,在细粒度的类别区分上有了更高的识别能力。

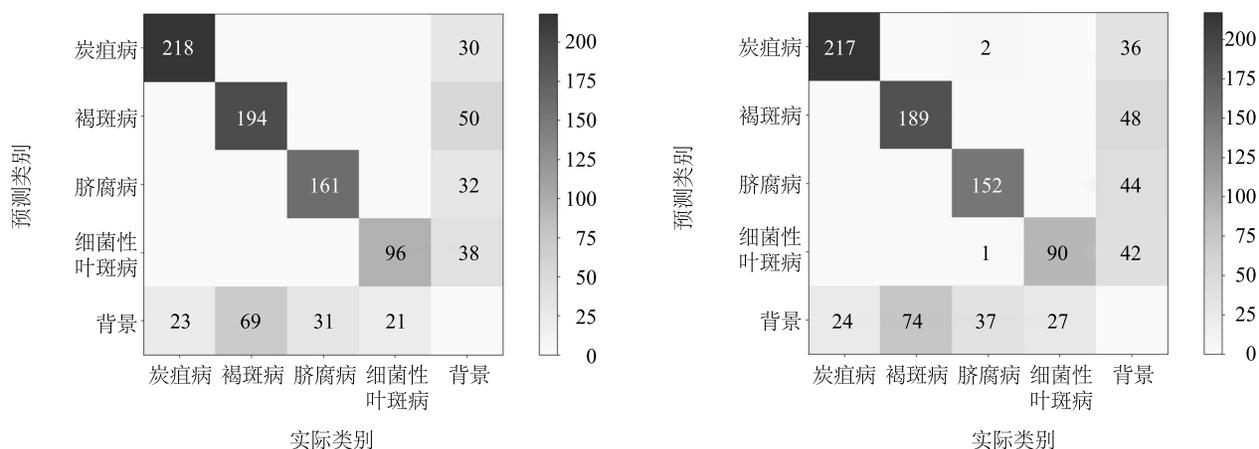


图10 YOLOv8n 模型与 YOLOv8n-ATA 模型对验证集中辣椒病害数据识别的混淆矩阵

Fig.10 Confusion matrix of YOLOv8n and YOLOv8n-ATA models for the identification of chili disease data in the validation set

2.4 模型对比试验结果

采用 YOLOv8n-ATA 模型和 YOLOv5s 模型、YOLOv7-Tiny 模型、YOLOv7-W6 模型、YOLOv9-Tiny 模型、SSD 模型、RetinaNet 模型对辣椒病害的识别效果如表 5 所示。从表中可以看出, YOLOv5s 模型的浮点计算量和模型大小分别是 YOLOv8n-ATA 模型的 1.6 倍和 2.2 倍, 且模型对病害的识别精确率、召回率、平均精度均值均低于 YOLOv8n-ATA 模型。YOLOv7-Tiny 模型对病害识别的精确率、召回率和 mAP_{50} 、 $mAP_{50:95}$ 分别比 YOLOv8n-ATA 低 4.0 个百分点、4.1 个百分点、6.9 个百分点和 7.0 个百分点, 且识别效率也低。使用 Aux 检测头的 YOLOv7-W6

模型识别精度低于 YOLOv8n-ATA 模型, 且浮点运算量和参数量大幅度上升, 进而导致识别效率大幅度下降。YOLOv9-Tiny 模型虽然有较低的模型大小和参数量, 但其对辣椒病害识别的 mAP_{50} 、 $mAP_{50:95}$ 、精确率和召回率均低于 YOLOv8n-ATA 模型。同样, SSD 模型和 RetinaNet 模型的浮点运算量、参数量和模型大小均远高于 YOLOv8n-ATA 模型, 且模型对病害识别的精确率、召回率和 mAP_{50} 、 $mAP_{50:95}$ 亦低于 YOLOv8n-ATA 模型。从上述分析可知, 与其他模型相比, YOLOv8n-ATA 模型对辣椒病害不但具有较高的识别准确率, 同时也具有较高的识别效率。

表 5 不同模型对辣椒病害的识别性能比较

Table 5 Comparison of the recognition performance of different models for chili diseases

序号	模型	精确率 (%)	召回率 (%)	平均精度均值 mAP_{50} (%)	平均精度均值 $mAP_{50:95}$ (%)	浮点运算量 (G)	模型大小 (M)	参数量 ($\times 10^6$)	传输帧数 (帧, 1 s)
1	YOLOv5s	86.4	78.5	83.0	41.1	15.8	14.3	7.0	23.4
2	YOLOv7-Tiny	83.3	75.8	78.3	37.7	13.2	12.3	6.0	21.7
3	YOLOv7-W6	83.8	72.8	77.3	38.7	102.5	162.5	80.9	5.0
4	YOLOv9-Tiny	83.6	74.2	80.5	41.9	11.0	5.8	2.7	18.4
5	SSD	84.2	64.2	79.9	37.9	274.5	92.1	24.0	2.6
6	RetinaNet	85.4	58.5	71.8	30.7	164.6	139.0	36.4	3.6
7	YOLOv8n-ATA	87.3	79.9	85.2	44.7	9.8	6.5	3.3	35.9

不同模型对辣椒病害识别结果的热力图如图 11 所示。从图中可以看出, 对于辣椒炭疽病、脐腐病和细菌性叶斑病, YOLOv8n-ATA 模型具有最高的置信度, 且不存在多余的检测框及将大块病斑分为小块病斑进行识别的现象; 对于病斑较小的褐斑病, YOLOv8n-ATA 模型也能更准确、精确地定位到病斑上, 其他模型不同程度存在漏检、多余的检测框及将大块病斑分为小块病斑识别的现象。即 YOLOv8n-ATA 模型在辣椒炭疽病、脐腐病、褐斑病和细菌性叶斑病的识别中均表现出良好的性能, 具有较高的识别准确率。

3 结论

本研究基于 YOLOv8n 模型, 利用平均池化、卷积和最大池化等操作构成 Adown 下采样模块替代原模型 Backbone 中的 3 个下采样 Conv 层, 引入 GSConv

和 VovGSCSP 模块替代原模型颈部网络中的卷积层和 C2f 模块, 并将辅助训练头 Aux 融入到原检测头中, 建立了一个辣椒炭疽病、脐腐病、褐斑病和细菌性叶斑病的检测模型——YOLOv8n-ATA 模型。测试结果表明, 改进后的 YOLOv8n-ATA 模型在对辣椒病害进行识别时, 虽然浮点计算量和模型大小比原 YOLOv8n 模型增加 19.5% 和 10.2%, 但对辣椒病害的识别精确率、召回率、 mAP_{50} 和 $mAP_{50:95}$ 分别提升 2.6 个百分点、3.5 个百分点、2.9 个百分点和 2.9 个百分点, 每 1 s 传输帧数增加 15.1%。同时, 与其他模型相比, YOLOv8n-ATA 模型能够在不过分增加模型复杂度的情况下提高检测准确率, 既保持了单阶段检测模型在计算和内存消耗方面的优势, 又融合了双阶段检测模型在检测精度方面的优势, 使得模型性能更加均衡, 满足了移动端设备的应用需求, 更适用于实际场景下的辣椒病害检测。

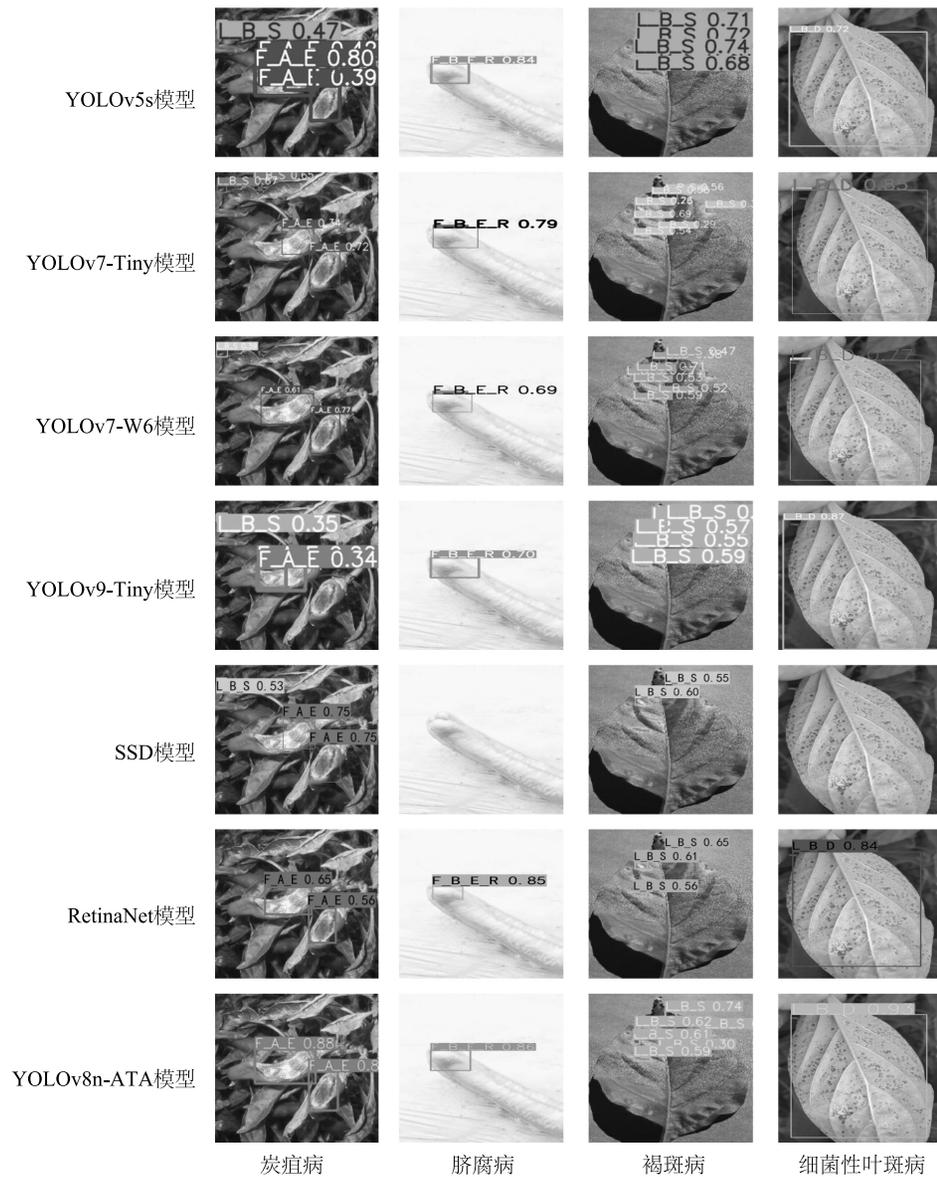


图 11 不同模型对辣椒病害的识别效果图

Fig.11 The recognition effect of different models on chili diseases

参考文献:

- [1] 张子峰. 我国辣椒产业发展现状、主要挑战与应对之策[J]. 北方园艺, 2023(14): 153-158.
- [2] 乔立娟, 赵帮宏, 宗义湘, 等. 我国辣椒产业发展现状、趋势及对策[J]. 中国蔬菜, 2023(11): 9-15.
- [3] 祝显萍. 辣椒种植技术及病害防治[J]. 种子科技, 2023, 41(21): 93-95.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once; unified, real-time object detection [C]//THE COMPUTER VISION FOUNDATION. 2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 779-788.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot Multi-Box detector [M]//LEIBE B, MATAS J, SEBE N, et al. Computer vision-ECCV 2016 Part I. Cham: Springer International Publishing, 2016: 21-37.
- [6] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(2): 318-327.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//IEEE. 2014 IEEE conference on computer vision and pattern recognition. Columbus, OH, USA: IEEE, 2014: 580-587.

- [8] 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, 2017, 39(6):1137-1149.
- [9] HE K M, ZHANG X Y, REN S Q, 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.
- [10] 储鑫,李祥,罗斌,等. 基于改进 YOLOv4 算法的番茄叶部病害识别方法[J]. 江苏农业学报, 2023, 39(5):1199-1208.
- [11] 施杰,林双双,张威,等. 基于轻量化改进型 YOLOv5s 的玉米病虫害检测方法[J]. 江苏农业学报, 2024, 40(3):427-437.
- [12] 刘诗怡,胡滨,赵春. 基于改进 YOLOv7 的黄瓜叶片病虫害检测与识别[J]. 农业工程学报, 2023, 39(15):163-171.
- [13] 周维,牛永真,王亚炜,等. 基于改进的 YOLOv4-GhostNet 水稻病虫害识别方法[J]. 江苏农业学报, 2022, 38(3):685-695.
- [14] 邹玮,岳延滨,冯恩英,等. 基于 YOLOv2 的辣椒叶部蚜虫图像识别[J]. 山东农业大学学报(自然科学版), 2023, 54(5):700-709.
- [15] 尚俊平,张冬阳,席磊,等. 基于多尺度特征融合和注意力机制的辣椒病害识别模型[J]. 河南农业大学学报, 2024, 58(6):1021-1033.
- [16] 李西兴,陈佳豪,吴锐,等. 基于改进 MaxViT 的辣椒病害识别分类方法[J]. 华中农业大学学报, 2024, 43(2):123-133.
- [17] 姚飞. 基于卷积神经网络的辣椒病害检测方法研究与应用[D]. 重庆:重庆大学, 2022.
- [18] 丁文宽. 基于卷积神经网络和机器视觉的辣椒检测与识别[D]. 天津:天津理工大学, 2017.
- [19] 王启镔,王旭. 基于改进 YOLOv5 的辣椒目标检测算法研究[J]. 建模与仿真, 2023, 12(5):4654-4662.
- [20] PRATAP V K, KUMAR N S. High-precision multiclass classification of chili leaf disease through customized EfficientNetB4 from chili leaf images[J]. Smart Agricultural Technology, 2023, 5:100295.
- [21] KANAPARTHI K R, SUDHAKAR ILANGO S. A survey on training issues in chili leaf diseases identification using deep learning techniques[J]. Procedia Computer Science, 2023, 218:2123-2132.
- [22] 刘子洋,徐慧英,朱信忠,等. Bi-YOLO:一种基于 YOLOv8n 改进的轻量化目标检测算法[J]. 计算机工程与科学, 2024, 46(8):1444-1454.
- [23] 付辰伏,任力生,王芳. 基于改进 YOLOv8 的牛只行为识别与跟踪方法[J]. 农业机械学报, 2024, 55(5):290-301.
- [24] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9: learning what you want to learn using programmable gradient information[Z/OL]. (2024-02-29). <http://arxiv.org/pdf/2402.13616>.
- [25] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures[J]. Journal of Real-Time Image Processing, 2024, 21(3):62.
- [26] WANG C Y, BOCHKOVSKIY A, MARK LIAO H Y. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//IEEE. 2023 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023:7464-7475.

(责任编辑:石春林)