

文章编号: 2095-2163(2023)06-0078-07

中图分类号: TP399

文献标志码: A

基于改进 YOLOv7 算法的葡萄叶病斑检测方法研究

李天, 唐璐

(西华师范大学 计算机学院, 四川 南充 637000)

摘要: 针对传统葡萄叶病斑检测漏检率高、小目标检测精度低等问题, 提出一种基于 YOLOv7 改进模型。该算法重新设计锚框以适应葡萄叶片病斑数据集; 将全局注意力机制 GAM 嵌入到 YOLOv7 模型的颈部中, 获取更加丰富的跨通道信息和位置信息; 针对个别病斑存在粘连情况, 引入柔性非极大值机制 (Soft-NMS) 策略提高检测框回归精度。为了验证改进后的模型效果, 选取 3 种常见葡萄叶病害进行测试及其验证。验证结果表明, 改进后的 YOLOv7 模型在此数据集上的平均精度 (mAP) 达到 87.2%, 比 YOLOv7 模型提高了 3.1%。改进算法能够适用于自然环境下的葡萄叶病害检测, 为葡萄病害的防控提供了新方法。

关键词: 目标检测; YOLOv7; 注意力机制; 机器学习

Research on the detection method of molar leaf disease spot based on the improved YOLOv7 algorithm

LI Tian, TANG Lu

(College of Computing, China West Normal University, Nanchong Sichuan 637000, China)

[Abstract] Aiming at the problems of low missed detection rate and low detection accuracy of traditional grape leaf disease spots, an improved model based on YOLOv7 is proposed. The algorithm redesigns the anchor box to fit the grape leaf disease spot dataset. The global attention mechanism GAM was embedded in the neck of the YOLOv7 model to obtain richer cross-channel information and position information. In view of the adhesion of individual disease spots, the flexible non-maximum mechanism (Soft-NMS) strategy was introduced to improve the regression accuracy of the detection frame. In order to verify the effect of the improved model, three common grape leaf diseases were selected for testing and verification. It is verified that the average accuracy (mAP) of the improved YOLOv7 model on this dataset reaches 87.2%, which is 3.1% higher than the YOLOv7 model. The experimental results show that the improved algorithm can be applied to the detection of grape leaf diseases in the natural environment, which provides a new method for the prevention and control of grape diseases.

[Key words] target detection; YOLOv7; attention mechanisms; machine learning

0 引言

中国是葡萄生产大国, 产量居世界首位。近年来, 中国的葡萄产业正处于关键时期, 葡萄种植农户不断调整和优化品种结构以及栽培模式, 种植面积和产量稳步增长, 向现代化、高质量的产业转型。然而, 葡萄生长过程中经常受到黑腐病、轮斑病、褐斑病、叶枯病等多种病害的干扰, 严重妨碍了葡萄的正常生长, 危害产量和质量, 为果农的经济带来了损失。

随着人工智能技术的发展, 计算机视觉、图像处理等技术已经广泛应用于农业领域, 如水稻病害识别、果蔬品质检测等, 这些技术的应用使得农业生产

朝着更加智能化和高效化的方向发展。近年来, 在农作物病害识别领域, CNN 展现出了优越的性能表现^[1-3]。尤其是以单阶段和双阶段目标检测算法为代表的深度学习技术, 更是在农业病虫害检测上大展身手, 获得了广泛的应用。双阶段目标检测算法首先以候选框的方式确定目标位置, 然后再对这些候选框进行分类和回归, 其代表模型有 Faster-RCnn。虽然双阶段目标检测算法具备较高准确率, 但是在推理速度和部署能力上有所欠缺^[4]。单阶段目标检测算法是直接从图像中提取物体的位置和类别信息, 不需要进行物体候选框的生成。其中比较著名的算法有 YOLO^[5-6] 和 SSD 等。YOLO 使用卷积神经网络来提取目标特征, 最终将提取到的信

作者简介: 李天(1995-), 男, 硕士研究生, 主要研究方向: 图像处理、计算机视觉; 唐璐(1998-), 女, 硕士研究生, 主要研究方向: 图像处理。

收稿日期: 2023-04-08

哈尔滨工业大学主办 ◆ 学术研究与应用

息转化为回归问题,SSD 也采用类似的思路。

综上,本文采用 YOLOv7 目标检测算法为基础模型,在此基础上,通过重新设计锚框以适应葡萄叶病斑数据集。并且在模型颈部添加 GAM 注意力机制来获取更丰富的跨通道信息,提高模型的特征提取能力。针对密集且具有粘连性质的葡萄叶病斑,将原始 YOLOv7 中的 NMS 极大值抑制策略更换为对密集目标更敏感的柔性极大值抑制策略 Soft-NMS^[7]。在公开数据集 PlantVillage 上对 3 种常见葡萄叶病斑进行识别,以验证模型效果。

1 YOLOV7 模型简介

YOLOv7 网络模型主要包含了输入(Input)、骨干网络(Backbone)、颈部(Neck)、头部(Head)等 4 部分,其结构图如图 1 所示。

骨干网络模块由多个卷积层组成,包括 BConv 卷积层、E-ELAN 卷积层和 MPConv 卷积层。其中 BConv 卷积层由卷积层、BatchNormalization (BN)层和 LeakyReLU 激活函数构成,用于提取不同尺度的图像特征;E-ELAN 卷积层采用原始的 ELAN^[8] 设计架构,引导不同特征组的计算块学习更多样化的特征,提高网络的学习能力,而不破坏原有梯度路径;MPConv 卷积层在 BConv 层的基础上,通过添加 Maxpool 层,形成上下两个分支,上分支通过 Maxpool 将图像长宽减半,通过 BConv 层将图像通

道减半。下分支则通过第一个 BConv 层将图像通道减半。两个分支提取到的特征经过 Cat 操作融合,旨在进一步优化卷积层的效果,提高了网络的特征提取能力。

头部模块采用路径聚合特征金字塔网络^[9] (Path Aggregation Feature Pyramid Network,PAFPN) 结构,引入自底向上的路径,使得底层信息更容易传递到高层,从而实现了不同层次特征的高效融合。Prediction 模块通过 REP (RepVGG Block) 结构对 PAFPN 输出的 P3、P4 和 P5 等 3 个不同尺度的特征进行图像通道数调整,最后经过 1×1 卷积用于置信度、类别和锚框的预测。虽然 YOLOv7 算法在常见任务场景(如行人、车辆检测)中表现出色,但将其直接应用于葡萄叶病斑检测仍然存在许多问题,如:

- (1) 受生长时间影响,葡萄叶病斑存在出现位置随机、病斑尺寸不一等问题;有的病斑像素很小,容易产生漏检情况。
- (2) 不同种类病斑存在一定相似度。以黑腐病和黑麻疹为例,叶枯病病斑初期症状与黑腐病类似,系统很容易把像素相对较低的黑腐病误识为叶枯病。
- (3) 个别种类病斑还存在粘连性,容易在视觉上造成一种不属于任何一类病斑的效果。针对这 3 种问题,本文从检测锚框、注意力机制、极大值抑制 3 方面对 YOLOV7 进行改进。

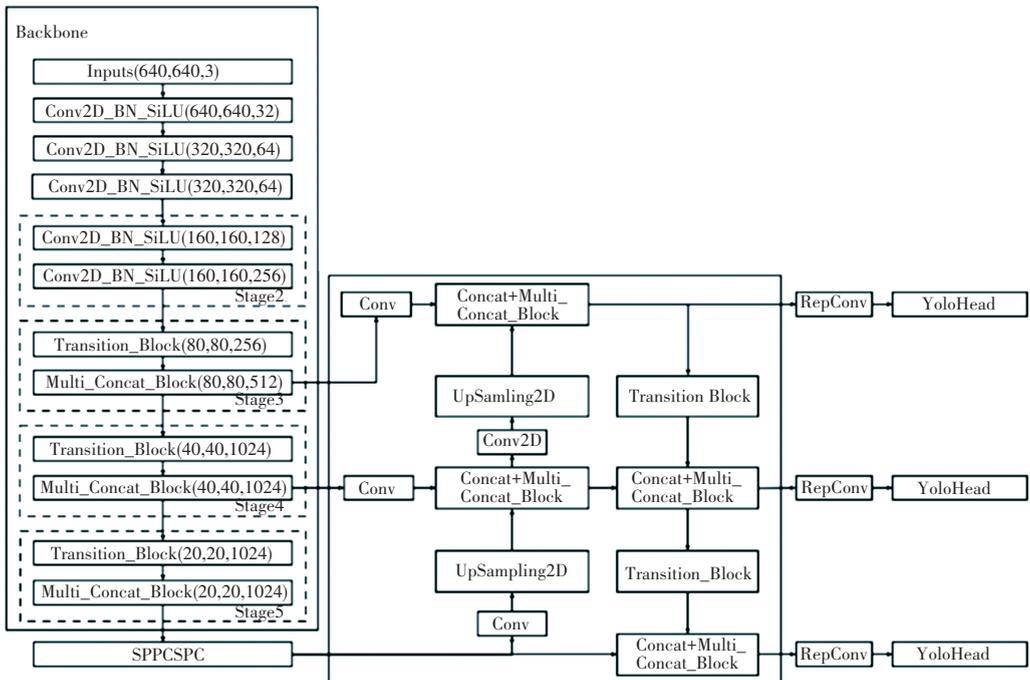


图 1 YOLOv7 网络结构
Fig. 1 YOLOv7 network structure

2 改进 YOLOV7 模型

2.1 GAM 注意力机制

全局注意力机制 (Global Attention Mechanism,

GAM)^[10],可以起到减少网络信息缩减并放大大局维度交互特征的作用。该机制在 CBAM 中的顺序通道-空间注意机制的基础上,对其子模块进行了优化设计,整体模块如图 2 所示。

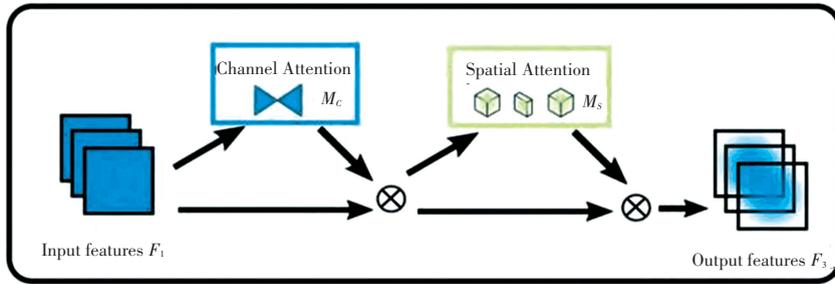


图 2 Gam 注意力机制

Fig. 2 GAM attention mechanism

其中,输入特征由 F_1 表示,将图 2 中的一系列中间操作定义为中间状态 F_2 ,输出状态定义为 F_3 ,则三者之间的关系如公式(1)所示:

$$\begin{aligned} F_2 &= M_c(F_1) \otimes F_1 \\ F_3 &= M_s(F_2) \otimes F_2 \end{aligned} \quad (1)$$

2.1.1 通道注意力子模块

如图 3 所示,原始输入特征 F_1 维度为 $C \times W \times H \times W$

H ,通道注意力子模块首先对其进行三维通道置换,将信息保存为 $W \times H \times C$ 的形式;利用一个两层的 MLP,第一层进行编码操作将通道数 C 缩减至 C/R ,第二层进行解码操作以获取与输入特征具有相同通道数的结果。最终,通过对结果进行 *Sigmoid* 激活函数处理,得到权重系数 M_c ,可以有效地扩大跨维通道空间的依赖性。

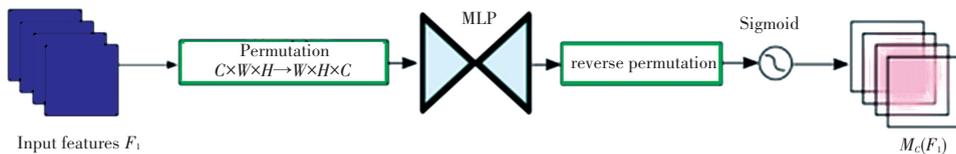


图 3 通道注意力

Fig. 3 Channel attention

2.1.2 空间注意力子模块:

如图 4 所示,输入特征 F_2 采用了双重卷积,每个卷积层都使用了 7×7 的卷积核,以达到空间信息融合的效果。通道注意力子模块根据特征的重要性

进行缩减并得到缩放后的新特征。最后,对于这个特征的权重系数 M_s ,采用 *Sigmoid* 激活函数进行处理,以获得更加准确的权重值。

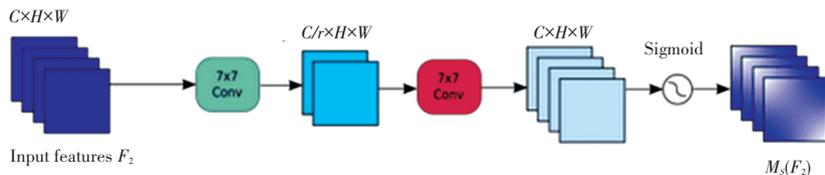


图 4 空间注意力

Fig. 4 Spatial attention

2.2 柔性非极大值抑制算法

在使用 YOLO 模型进行检测时,一张图片可能会产生多个候选框,每个候选框代表着可能存在的目标区域。然而,由于算法本身的不确定性,会导致同一个目标在不同的候选框中多次出现。非极大值抑制 (Non-Maximum Suppression, NMS) 是目标检测领域用于消除重叠的边界框或者其他类型区域的一

种策略。其基于一定的规则,通过筛选出最优的检测结果来消除冗余的候选框。

NMS 的主要思想,是对于所有候选框计算其得分(例如置信度分数),并选择得分最高的框作为输出。然后,将与该框高度重叠的所有候选框从集合中删除。通过不断重复这个过程,直到集合中不存在重叠的候选框。具体见公式(2)。

$$S_i = \begin{cases} S_i, IoU(M, b_i) < N_i \\ 0, IoU(M, b_i) \geq N_i \end{cases} \quad (2)$$

式中: M 代表当前置信度最高的边界框, b_i 代表邻域内的相邻边界框, S_i 是当前边界框得分值, IoU 代表两个边界框的阈值, N_i 代表该边界框的阈值。

NMS 在对相邻边界框进行处理时, 会判断 M 和 b_i 的 IoU 阈值是否小于 N_i , 如果小于, 则保持 S_i 不变, 否则就设置为 0。这种处理方式对于稀疏目标的识别效果很好。然而, 个别种类病斑(如黑麻疹)存在密集粘连情况, 此时 NMS 方法会让两个检测框之间产生抑制效果, 从而导致置信度较低的目标漏检。

为解决上述问题, 本文使用一种将得分和重合度全部纳入考虑的非极大值抑制方法 Soft-NMS^[11]。该方法不会直接剔除置信度较低检测框, 而是以权重衰减的方式决定置信度较低检测框的去留, 特别适用于密集目标检测。具体操作如式(3)所示。当重叠度小于给定阈值时, 分类置信度分数保持不变; 而当重叠度大于等于给定阈值时, 分类置信度分数按线性规则进行衰减。

$$S_i = \begin{cases} S_i & IoU(M, b_i) < N_i \\ S_i(1 - IoU(M, b_i)) & IoU(M, b_i) \geq N_i \end{cases} \quad (3)$$

2.3 锚框重设

YOLOv7 模型的每个检测头都可以产生 3 个初

始锚框, 这些锚框是通过对 COCO 数据集进行 k-means 聚类生成的, 适用于该数据集中常见类别的查找。然而, 在葡萄叶病斑数据集小目标以及密集病斑较多时, 原始锚框并不适用于该数据集。因此本文采用 K-means 聚类算法, 对葡萄叶病斑数据集重新进行锚框设计。

如图 5 所示, K 均值(K-means)算法是一种无监督学习算法, 用于将数据集划分为 K 个不同的组或簇, 其基本思想是通过最小化数据点与其所属簇中心点之间的距离, 来确定数据点所属的簇。K 均值算法的基本实现步骤如下:

- (1) 初始化簇中心点: 随机选择 K 个数据点作为初始簇中心点。
- (2) 分配数据点到最近的簇: 对于每个数据点, 计算其与每个簇中心点之间的距离, 并将其分配给距离最近的簇。
- (3) 重新计算簇中心点: 对于每个簇, 计算其所有数据点的平均值, 并将平均值作为新的簇中心点。
- (4) 重复步骤(2)、步骤(3), 直到簇中心点不再改变或达到最大迭代次数。
- (5) 输出结果: 输出聚类结果, 即每个数据点所属的簇。

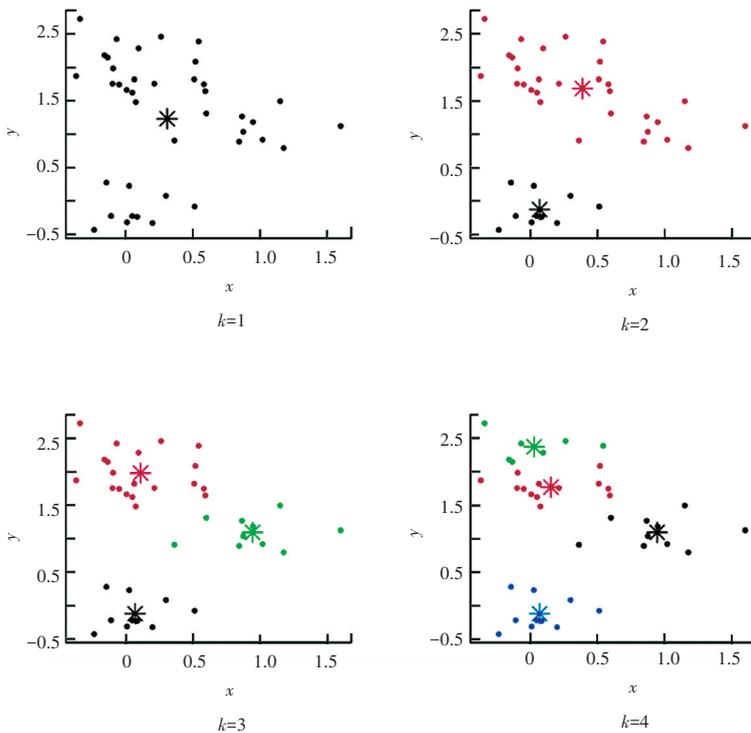


图 5 K-means 聚类

Fig. 5 K-means clustering

3 实验

3.1 数据集

本节所用数据集分为两部分,一部分来源于网络爬取,另一部分来自于大型公开植物病害数据集 PlantVillage, PlantVillage 数据集由植物病理学家对叶片病害进行确诊后得到的,具有权威性。实验选取的3种常见葡萄叶片病害分别为黑腐病、叶枯病和黑麻疹,具体病斑特征如图6所示。



(a) 黑麻疹 (b) 黑腐病 (c) 叶枯病

图6 不同病斑图像

Fig. 6 Images of different diseases

上述数据集共包含葡萄叶病斑图像2400张,其中黑麻疹、黑腐病、叶枯病各为800张。本文使用 Labellmg 标注软件进行葡萄叶病斑的标注工作,将图像中不同类型病斑目标的位置进行标注后,将数据集按照训练集:验证集=8:2的比例随机划分,并构建 VOC 格式的数据集。

3.2 评价指标

精确率 (Precision, P) 是指在所有检测出的目标中检测正确的概率,故其又称查准率;召回率 (Recall, R) 是度量分类模型性能的一种指标,用来衡量模型在所有正类样本中,正确识别出的样本数量与所有正类样本数量的比例;平均精度 (Average Precision, AP) 是衡量模型在检测目标时准确性的一种指标,其是在不同置信度阈值下计算的精度得分的平均值; mAP 是从类别的维度对 AP 进行平均,因此可以评价多分类器的性能。各评价指标的计算公式如下:

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (4)$$

$$R = \frac{T_p}{T_p + F_N} \times 100\% \quad (5)$$

$$AP = \int_0^1 P(r) dr \quad (6)$$

$$mAP = \frac{AP}{Num_{classes}} \quad (7)$$

3.3 实验环境及结果

本文采用 PyTorch 作为深度学习框架,操作系统选用 Ubuntu18.04,显卡为 NVIDIA GeForce RTX 2060 6 G,CUDA 版本 11.1.0。批次大小设置为 16,采用 SGD 作为优化器,初始学习率为 0.01,动量参数为 0.9,经过 200 轮次迭代后模型趋于收敛。

YOLOv7 算法改进前后的 mAP 对比曲线如图7所示。由此可见,原始 YOLOv7 算法在 3 种葡萄叶病斑上的检测 mAP 为 0.841,改进后达到了 0.872, mAP 提高 3.1 个百分点。

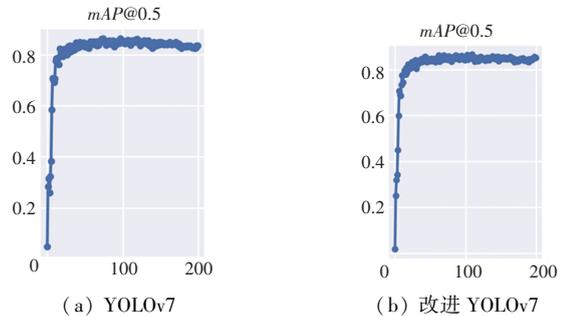


图7 改进前后 mAP 曲线

Fig. 7 Improved front and rear mAP curves of YOLOv7

3.4 对比实验

为验证改进模型在葡萄叶片病斑检测上的优势,在同一数据集上分别选用了主流目标检测模型与改进 YOLOv7 模型进行对比。图8展示了目标检测模型 Faster-rcnn、RetinaNet、SSD 与改进 YOLOv7 在同一数据集上的检测效果,不同模型检测数据详见表1。

表1 不同检测模型对比

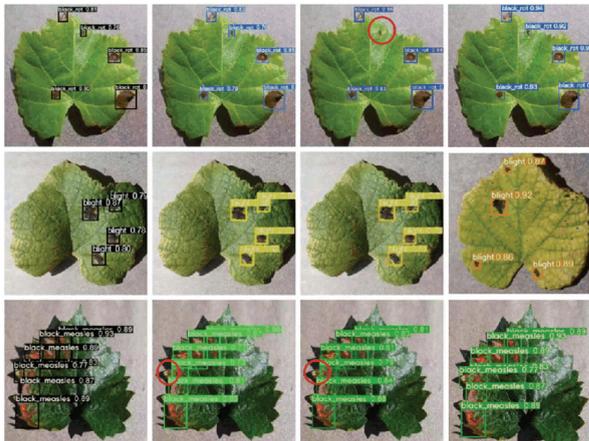
Tab. 1 Different Model comparison

模型	类别	AP	$mAP@0.5$	$mAP@0.5:0.95$
SSD	黑腐病	0.851		
	叶枯病	0.844	0.819	0.542
	黑麻疹	0.763		
Faster-Rcnn	黑腐病	0.914		
	叶枯病	0.881	0.853	0.571
	黑麻疹	0.792		
RetinaNet	黑腐病	0.871		
	叶枯病	0.846	0.832	0.55
	黑麻疹	0.778		
YOLOv7	黑腐病	0.893		
	叶枯病	0.849	0.841	0.568
	黑麻疹	0.781		
改进 YOLOv7	黑腐病	0.896		
	叶枯病	0.873	0.872	0.601
	黑麻疹	0.841		

可以看出,对于叶枯病这一种类病斑 4 种模型都表现出了较高的定位准确度和置信度水平。该类病斑虽然尺寸各异,但是特征差异化小,模型拟合度高。

SSD 在 3 个类别上的 mAP 为 0.819,表现出了较低的置信度水平。不仅如此,对于红圈圈出的小目标黑腐病病斑,以及具有粘连性质的黑麻疹病斑,SSD 均出现了漏检情况。原因在于 SSD 在特征图上的默认 anchor 较大,对于小目标病斑,其定位精度较低。对于密集病斑的漏检,SSD 使用的候选框抑制策略与 YOLO 一样,均为 NMS 策略,导致密集且相邻候选框出现了抑制效果,其中一个被剔除。

双阶段目标检测算法 Faster-Rcnn 在定位精度和置信度上均表现出了良好的效果,其在 3 个类别上的 mAP 表现为 0.853,比原始 YOLOv7 模型高出 1.2 个百分点。RetinaNet 在小目标病斑上的检测效果良好,但是对于密集病斑依然存在漏检情况。



(a) Faster-rcnn (b) RetinaNet (c) SSD (d) 改进 YOLOv7
图 8 不同模型检测效果对比

Fig. 8 Comparison of detection effects of different models

3.5 消融实验

为了详细验证改进后的 YOLOv7 性能,将实验分成 5 组,在同一数据集上通过对原始的 YOLOv7 模型、逐步加入各个模块的 YOLOv7 模型,以及合并所有改进后的 YOLOv7 模型进行对比。令原始模型为 A,B 在 A 基础上重新设计检测锚框,C 在 A 基础上加入 Soft-NMS,D 在 A 基础上加入注意力机制,E 将 3 处改进全部作用于 A,实验结果见表 2。

表 2 消融实验

Tab. 2 Ablation experiments

模型	类别	AP	$mAP@0.5$	$mAP@0.5;0.95$
A	黑腐病	0.893		
	叶枯病	0.849	0.819	0.542
	黑麻疹	0.841		
B	黑腐病	0.901		
	叶枯病	0.862	0.855	0.577
	黑麻疹	0.801		
C	黑腐病	0.880		
	叶枯病	0.859	0.858	0.584
	黑麻疹	0.836		
D	黑腐病	0.893		
	叶枯病	0.874	0.853	0.578
	黑麻疹	0.791		
E	黑腐病	0.896		
	叶枯病	0.873	0.872	0.601
	黑麻疹	0.841		

4 结束语

为了及时、准确解决常见葡萄叶片病害检测方面的问题,本文在原始 YOLOv7 模型之上进行了改进。通过重新设计检测锚框引入了全局注意力机制 GAM 并将原始 NMS 变更为 Soft-NMS 3 种策略,有效解决了小目标病斑识别精度不高以及密集病斑漏检问题。

通过大量实验验证,本文提出的改进模型其平均精度 ($mAP@0.5$) 达到了 0.872,相比原始 YOLOv7 模型提高 3.1 个百分点,能够有效识别出 3 种葡萄叶病斑。

经检测,该模型依然存在不足之处,虽然提高了检测精度,但也增加了参数量,推理速度也相对变慢。未来将从模型轻量化角度入手,探索出一种在不降低检测精度的情况下复杂度更低的模型,从而使其更易部署到移动端。

参考文献

[1] SYED-AB-RAHMAN S F, HESAMIAN M H, PRASAD M. Citrus disease detection and classification using end-to-end anchor-based deep learning model[J]. Applied Intelligence, 2022, 52(1): 927-938.

[2] PRABU M, CHELLIAH B J. Mango leaf disease identification and classification using a CNN architecture optimized by crossover-based levy flight distribution algorithm[J]. Neural Computing and Applications, 2022, 34(9): 7311-7324.