

基于改进 YOLOv9 的黄瓜病害识别^①

邵佳慧, 姚百蔚, 田 宏

(大连交通大学 轨道智能工程学院, 大连 116052)

通信作者: 田 宏, E-mail: th@djtu.edu.cn



摘 要: 为解决黄瓜病害种类繁多且症状相似导致的识别困难问题, 本文提出一种改进的 YOLOv9 黄瓜病害识别模型 BiFEL-YOLOv9, 以提高自然背景下黄瓜病害的检测精度. 首先在关键网络层引入加权双向特征金字塔网络模块 (bidirectional feature pyramid network, BiFPN), 增强了模型对多尺度特征的融合能力; 其次结合特征增强模块 (feature enhancement) 和大核选择性注意力机制 (large selective kernel block, LSKBlock) 对原始的 RepNCSPeLAN4 模块进行改进得到 RNFEL 模块, 增强了模型的特征表示能力及对复杂背景的鲁棒性. 实验结果表明, BiFEL-YOLOv9 模型准确率达到 97.96%、召回率达到 95.51%、平均精度均值 $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 分别达到 98.21% 和 95.12%, 均优于原 YOLOv9 模型, 有效实现了黄瓜病害的检测与识别.

关键词: YOLOv9; 病害识别; 深度学习; 特征提取; 目标检测

引用格式: 邵佳慧,姚百蔚,田宏.基于改进 YOLOv9 的黄瓜病害识别.计算机系统应用,2025,34(7):208–214. <http://www.c-s-a.org.cn/1003-3254/9899.html>

Cucumber Disease Recognition Based on Improved YOLOv9

SHAO Jia-Hui, YAO Bai-Wei, TIAN Hong

(College of Rail Intelligent Engineering, Dalian Jiaotong University, Dalian 116052, China)

Abstract: To address the difficulty in identifying cucumber diseases caused by numerous varieties and similar symptoms, this study proposes an improved YOLOv9 model for cucumber disease recognition, named BiFEL-YOLOv9, to enhance detection accuracy in natural backgrounds. Initially, a weighted bidirectional feature pyramid network module is incorporated into critical network layers to enhance the model's multi-scale feature fusion capability. Following that, the original RepNCSPeLAN4 module is enhanced by integrating a feature enhancement module and a large selective kernel block (LSKBlock) to obtain the RNFEL module, which improves the model's feature representation capability and robustness to complex backgrounds. Experimental results indicate that the BiFEL-YOLOv9 model achieves an accuracy of 97.96%, a recall rate of 95.51%, and mean average precision scores of 98.21% for $mAP_{0.5}$ and 95.12% for $mAP_{0.5:0.95}$, all of which surpass the performance of the original YOLOv9 model. The proposed model effectively accomplishes the detection and recognition of cucumber diseases.

Key words: YOLOv9; disease recognition; deep learning; feature extraction; object detection

黄瓜在中国各地广泛种植, 包括北方寒冷地区及南方温室大棚^[1]. 然而, 黄瓜在其生长周期内可能会遭受多种病害的侵扰. 这些病害在各产区发生普遍, 且呈

现加重趋势, 严重影响了黄瓜的产量和品质^[2]. 因此, 及时有效地解决这些病害问题对于减少农业损失至关重要. 传统病害检测依赖视觉观察^[3], 效率低且准确性差.

① 基金项目: 国家自然科学基金重点项目 (32130085)

收稿时间: 2024-11-08; 修改时间: 2025-01-15; 采用时间: 2025-02-11; csa 在线出版时间: 2025-05-29

CNKI 网络首发时间: 2025-05-29

近年来,深度学习技术在植物病害检测领域发展迅速,农作物病虫害识别准确率大幅提升^[4],展现出强大的泛化能力,为农业自动化开辟了广阔的应用前景。

基于深度学习的植物病害检测技术涵盖了 CNN^[5]、RNN^[6]、Faster R-CNN^[7]以及 YOLO^[8]系列等多种方法。其中, YOLO 系列算法以其快速和准确的检测能力被广泛应用于包括农作物病害检测在内的各种场景。董雁凯等人^[9]通过改进 YOLOv5 的激活函数和引入注意力机制,优化了黄瓜叶片面积的估算,使得模型的准确率达到 81.71%。李淑菲等人^[10]提出了一种结合可见光谱的 YOLOv5 目标检测模型,在黄瓜病害识别上的平均精度均值达到了 80.9%。乔琛等人^[11]通过引入特征提取模块和注意力机制提出了一种基于 YOLOv5 的黄瓜霜霉病菌孢子检测模型,其准确率达到 88.3%。这些研究证明 YOLO 系列算法在黄瓜病害检测中有效,但是并没有考虑黄瓜实际种植场景中叶片图像的多样性及复杂背景问题。因此,在复杂背景下实现黄瓜病害快速、准确的检测仍是一项具有挑战性的任务。

针对上述问题,本文提出一种基于 YOLOv9^[12]的黄瓜病害识别模型 BiFEL-YOLOv9。该模型通过引入 BiFPN^[13]加权双向特征金字塔网络模块,增强了多尺度特征融合,提升了局部细节的捕捉能力。结合特征增强模块和 LSKBlock^[14]大核选择性注意力机制对原始的 RepNCSPeLAN4 模块进行改进,得到了 RNFEL 模块,进一步增强了模型的特征表示能力,提高了模型对复杂背景的鲁棒性。实验结果表明, BiFEL-YOLOv9 模型有效地提高了自然背景下黄瓜病害的识别准确率。

1 BiFEL-YOLOv9 黄瓜病害识别模型

1.1 YOLOv9 模型

YOLOv9 以高精度和快速检测的能力,被广泛应用于视频监控、自动驾驶和农业监测等各种场景。该模型通过引入可编程梯度信息 (programmable gradient information, PGI) 和泛化高效层聚合网络 (generalized efficient layer aggregation network, GELAN) 等先进技术,显著提升了模型的准确性和效率。PGI 技术解决了神经网络在训练过程中容易出现的信息丢失问题。GELAN 则使得网络结构更加灵活和高效,能够在保持轻量级的同时达到更高的准确度和速度。

尽管 YOLOv9 在物体检测的精度和速度方面展示了较高的性能,但在处理复杂场景和高分辨率图像时

其检测精度仍有提升空间。

1.2 BiFPN 加权双向特征金字塔网络

为了提升目标检测精度和效率, YOLOv9 集成了特征金字塔网络 (FPN)^[15]和路径聚合网络 (PANet)^[16]。FPN 专注于从不同层次提取多尺度特征以识别各种尺寸的目标,而 PANet 通过增强特征层间的信息交换来改善特征传递。尽管这种结合在特定数据集上表现良好,但其泛化能力到其他数据集或复杂场景下有所下降。因此本文引入了 BiFPN 加权双向特征金字塔网络替换原有的 FPN 和 PANet 结构,旨在通过增强模型在复杂场景下的目标定位和识别能力,提升模型的整体性能和泛化能力。

图 1 展示了 FPN、PANet 和 BiFPN 的网络结构对比。FPN 仅支持自上而下的信息流, PANet 在 FPN 的基础上增加了自下而上的路径。与 PANet 相比, BiFPN 将双向路径作为可重复使用的特征网络层,以便融合更多特征。如图 1(c) 所示, BiFPN 包括自顶向下的语义信息传递路径;自底向上的位置信息传递路径;以及通过跳跃连接和双向路径实现的加权融合和双向跨尺度连接。

BiFPN 使用加权特征融合方法如式 (1) 所示,其中 w_i 通过应用 *ReLU* 函数确保非负, ϵ 是避免数值不稳定的小数值。以第 6 层为例,式 (2)、(3) 描述了 BiFPN 的两个融合特征。其中, P_6^{td} 表示从上到下的中间特征, P_6^{out} 是自下而上的输出特征。Resize 表示上采样或下采样操作, w 是通过训练学习得到的参数, ϵ 是偏差项。

$$O = \sum_i \frac{w_i}{\epsilon + \sum_j w_j} \cdot I_i \quad (1)$$

$$P_6^{\text{td}} = \text{Conv} \left(\frac{w_1 \cdot P_6^{\text{in}} + w_2 \cdot \text{Resize}(P_7^{\text{in}})}{w_1 + w_2 + \epsilon} \right) \quad (2)$$

$$P_6^{\text{out}} = \text{Conv} \left(\frac{w'_1 \cdot P_6^{\text{in}} + w'_2 \cdot P_6^{\text{td}} + w'_3 \cdot \text{Resize}(P_5^{\text{out}})}{w'_1 + w'_2 + w'_3 + \epsilon} \right) \quad (3)$$

通过引入 BiFPN 模块, YOLOv9 的结构得到了优化,能够更有效地融合不同层次的特征,从而提升了模型的检测精度和鲁棒性。

1.3 RNFEL 模块

针对原模型在复杂背景下特征提取能力不足的问题,本文通过融合特征增强模块和 LSKBlock 注意力机制,对原有的 RepNCSPeLAN4 模块进行了改进,从而得到了 RNFEL 模块。

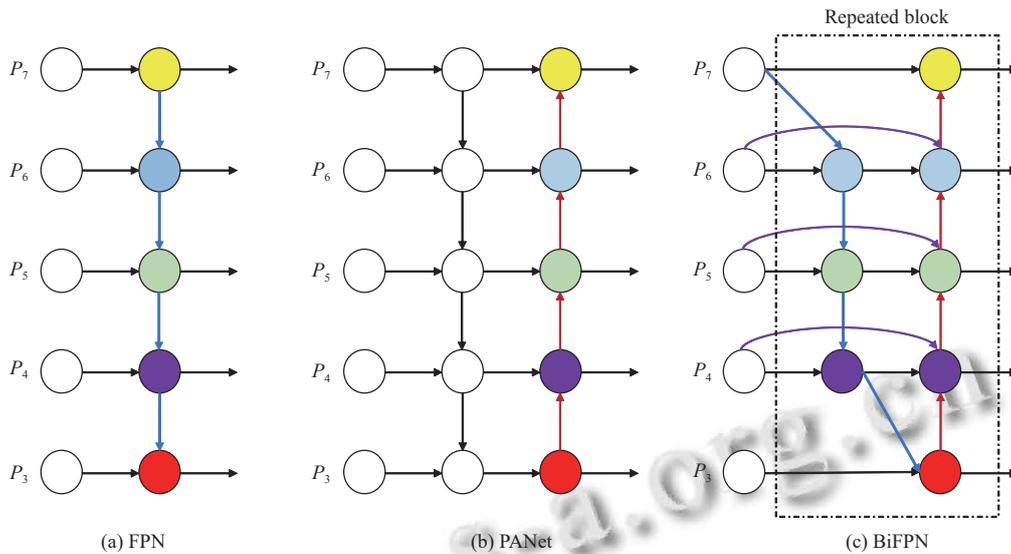


图1 FPN、PANet 和 BiFPN 的网络结构对比图

RepNCSPELAN4 是 YOLOv9 中的特征提取-融合模块, 通过结合 CSPNet^[17]和 ELAN^[18]形成泛化高效层聚合网络 GELAN, 从而提高目标检测的准确性和效率. 图 2 展示了 RepNCSPELAN4 模块的结构细节. 其中, Conv 层负责特征图的初步处理以及通道数的调整, RepNCSP 模块实现特征的深度提取和融合.

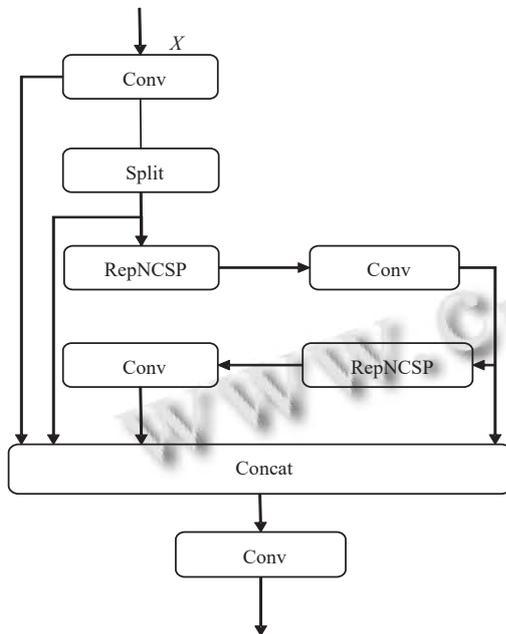


图2 RepNCSPELAN4 结构图

图 3 所示为 LSKBlock 结构, 包括大核选择子块 (LK selection) 和前馈网络子块 (FFN). LK selection 允许网络动态调整其感受野, 以适应不同大小和形状的目标, FFN 可以增加网络的非线性表达能力, 帮助学习

更复杂的特征表示.

FE 特征增强模块是深度学习网络中用于提升特征表示能力的组件. 首先使用一个卷积层提取关键特征; 其次通过 LSKBlock 模块来突出重要特征并抑制噪声; 最后通过另一个卷积层将经过增强处理的特征映射回原始的通道数, 以便这些增强的特征能够无缝集成到网络的后续部分.

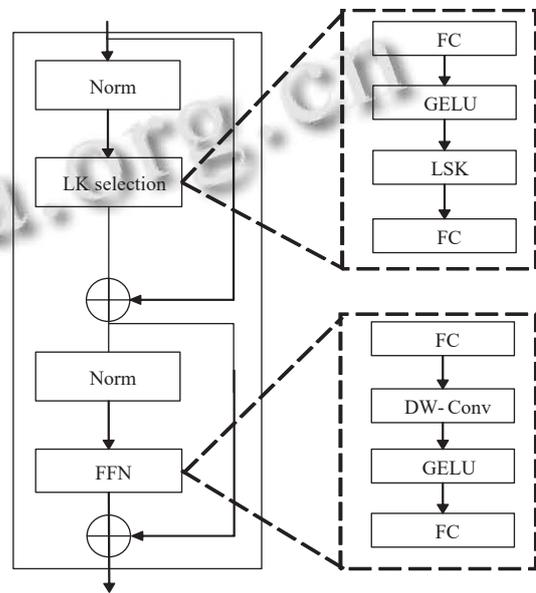


图3 LSKBlock 结构图

通过动态调整感受野以适应不同目标所需的背景范围. 在这个过程中, 式 (4) 定义了深度可分离卷积核大小和扩张率的增长条件, 确保感受野快速增长. 感受野的计算方式为式 (5), 其中 k 为第 i 个卷积核的大小,

d 为扩张率, RF 为感受野. 这种设计能够让模型根据需要调整感受野, 提高目标检测的准确性.

$$RF_1 = k_1, RF_i = d_i(k_i - 1) + RF_{i-1} \quad (4)$$

$$k_{i-1} \leq k_i; d_1 = 1, d_{i-1} < d_i \leq RF_{i-1} \quad (5)$$

图 4 为 RNFEL 模块结构. 特征增强模块提取关键特征, LSKBlock 通过动态调整感受野, 突出重要特征并抑制噪声. RNFEL 通过融合特征增强模块和 LSKBlock 注意力机制, 进一步优化了 BiFPN 融合后的特征.

1.4 BiFEL-YOLOv9 模型结构

由 RNFEL 模块和 BiFPN 模块融合到 YOLOv9 模型得到的 BiFEL-YOLOv9 模型, 其结构如图 5 所示.

首先 BiFPN 模块通过双向信息流和加权融合机制, 将不同层次的特征进行有效整合, 为 RNFEL 模块提供了更丰富的多尺度特征. RNFEL 模块则通过特征增强和 LSKBlock 注意力机制, 进一步优化这些特征, 增强模型对复杂背景的鲁棒性和对小目标的识别能力.

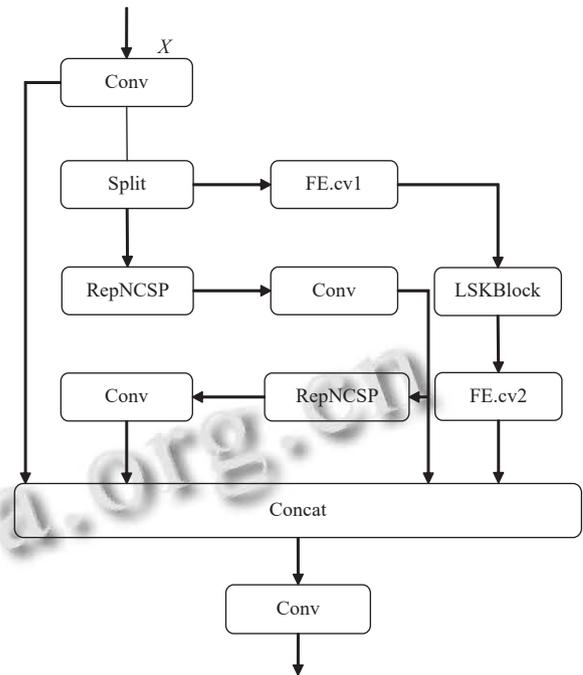


图 4 RNFEL 模块结构图

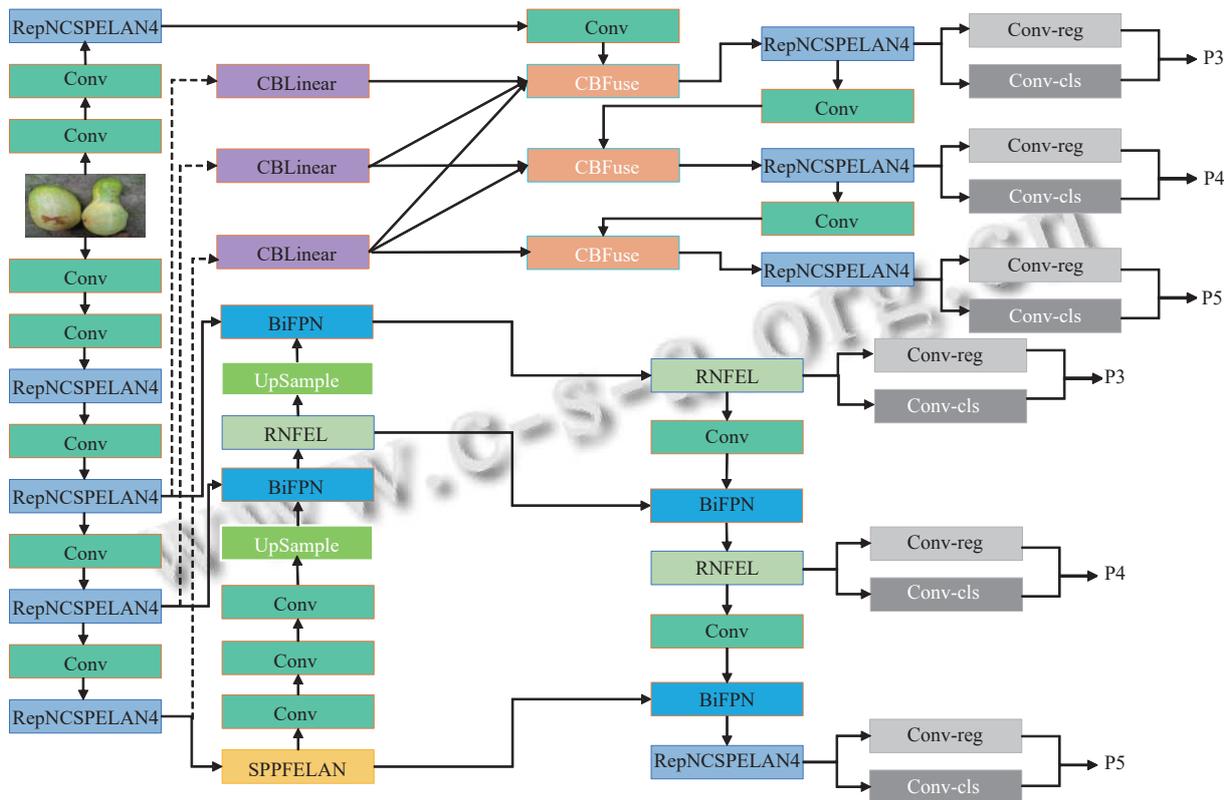


图 5 BiFEL-YOLOv9 模型结构图

通过这种协同工作方式, BiFEL-YOLOv9 模型在自然背景下的黄瓜病害图像处理中, 尤其在复杂背景

和多尺度目标情况下, 展现出更强的特征提取能力和更高的识别精度.

2 黄瓜病害数据集及模型训练

2.1 数据集及预处理

实验中的数据采用 cucumber disease recognition dataset^[19]数据集中的 1280 张黄瓜病害图像, 包含 8 种类别, 包括炭疽病、细菌性枯萎病、腹腐病、霜霉病、腐霉病、软茎枯萎病、新鲜叶子和新鲜黄瓜。数据集中的图像均来自真实田地, 反映了黄瓜在自然条件下可能遭受的各种病害情况。部分数据集样本如图 6。

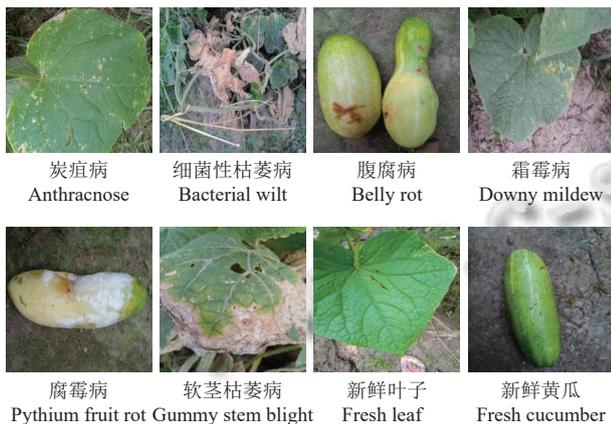


图 6 部分数据集样本

为了提高模型准确性和泛化能力, 对原始图像进行旋转、亮度调整、加噪及镜像等数据增强操作, 并调整图像尺寸为 640×640 像素。扩充后的数据集共有 8402 张图片, 按照 7:2:1 的比例划分为训练集、验证集和测试集。

2.2 模型训练

本实验平台基于 Ubuntu 22.04 操作系统, 软件环境为 Python 3.8, 深度学习框架为 PyTorch 2.3.0。GPU 为 NVIDIA RTX 4080, CUDA 版本为 12.3。

实验参数设置: 图像输入尺寸为 640×640, 批量 Batch size 为 4, 初始学习率为 0.01, 动量参数为 0.937, 权重衰减为 0.0005, 模型迭代次数 Epoch 为 400。

3 实验结果与分析

3.1 评价指标

为全面评估模型性能, 本文采用精确率 (precision, P)、召回率 (recall, R)、平均精度 (average precision, AP)、平均精度均值 (mean average precision, mAP) 作为模型的评价指标。其中, P 用于衡量模型正确识别病害样本的比例; R 用于评估模型能够识别出的实际病害样本占有所有实际病害样本的比例; AP 是在每个类别

上计算精确率和召回率曲线下面积的平均值, 反映了模型对不同病害类别的识别能力; mAP 用于评估模型在黄瓜多类别病害检测任务中的整体性能, 具体计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$R = \frac{TP}{TP + FN} \quad (7)$$

$$AP = \int_0^1 PdR \quad (8)$$

$$mAP = \frac{1}{c} \sum_{i=1}^c AP \quad (9)$$

其中, TP 表示真正类, 即模型正确预测为病害的样本数量; FP 表示假正类, 即模型错误预测为病害的样本数量; FN 表示假负类, 即模型未能检测到的实际病害样本数量。

3.2 消融实验

为了验证 BiFPN 模块和 RNFEL 模块对黄瓜病害检测模型性能的贡献, 本研究以 YOLOv9 模型为基础网络模型开展消融实验, 主要以测试集的平均精度均值作为评判标准, 表 1 为消融实验结果。

表 1 消融实验结果 (%)

实验	+BiFPN	+RNFEL	$mAP_{0.5}$	$mAP_{0.5:0.95}$
1	—	—	96.28	91.14
2	√	—	96.52	91.82
3	—	√	97.02	92.41
4	√	√	98.21	95.12

由表 1 可知, 加入 BiFPN 加权双向特征金字塔网络后, $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 分别提高了 0.24 和 0.68 个百分点, 表明模型在对目标的识别精度有所增强。引入 RNFEL 模块后, 模型的 $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 均有显著提升, 分别提高了 0.74 和 1.27 个百分点, 显示出模型对不同尺度目标的识别能力得到了提高。而 BiFEL-YOLOv9 算法在所有指标上都达到了最佳性能, $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 指标提升了 1.93 和 4.0 个百分点, 这表明 BiFPN 和 RNFEL 模块的结合不仅增强了模型的特征提取能力, 还提高了模型在复杂背景下的识别精度和鲁棒性, 对于提高黄瓜病害检测的准确性和可靠性具有重要作用。

3.3 对比实验

为了进一步验证本模型效果的优越性, 在相同的

实验环境与同一数据集下,使用YOLO主流检测模型YOLOv5、YOLOv7、YOLOv9与本文所提模型BiFEL-YOLOv9进行模型训练与测试,实验结果如表2所示。

表2 对比实验结果(%)

模型	<i>P</i>	<i>R</i>	<i>mAP</i> _0.5	<i>mAP</i> _0.5:0.95
YOLOv5	96.87	94.95	98.09	85.42
YOLOv7	94.67	95.30	97.43	88.84
YOLOv9	96.11	93.38	96.28	91.14
BiFEL-YOLOv9	97.96	95.51	98.21	95.12

根据表2可知,在对黄瓜病害进行识别时,BiFEL-YOLOv9模型的各项指标都是最高的,包括精确率、召回率和平均精度均值。其中相比于YOLOv9,BiFEL-YOLOv9在*mAP*_0.5和*mAP*_0.5:0.95上分别提高2.03和4个百分点。与YOLOv7和YOLOv5相比,BiFEL-YOLOv9在*mAP*_0.5和*mAP*_0.5:0.95值上也有不同程度的提高。

综上所述,BiFEL-YOLOv9模型通过结合BiFPN和RNFEF模块,有效地提升了模型在黄瓜病害检测任务上的性能。

3.4 泛化能力对比实验

为了评估模型在不同环境条件下的泛化能力,选择cucumber和cucumber leaf disease两个公开数据集上进行对比实验。其中cucumber拍摄于室内,包含3种类别3754张黄瓜病害图像。cucumber leaf disease拍摄于田间环境,包含5种类别4000张黄瓜病害图像。

本次实验重点对比了BiFEL-YOLOv9与YOLOv9的性能表现。实验结果如表3和表4所示。

表3 Cucumber数据集对比(%)

模型	<i>P</i>	<i>R</i>	<i>mAP</i> _0.5	<i>mAP</i> _0.5:0.95
YOLOv9	98.83	98.37	99.12	89.81
BiFEL-YOLOv9	99.86	99.79	99.50	91.36

表4 Cucumber leaf disease数据集对比(%)

模型	<i>P</i>	<i>R</i>	<i>mAP</i> _0.5	<i>mAP</i> _0.5:0.95
YOLOv9	85.74	85.28	89.16	80.07
BiFEL-YOLOv9	88.37	87.61	91.86	82.46

实验结果表明,在cucumber数据集上,BiFEL-YOLOv9的*mAP*_0.5比YOLOv9提高0.38个百分点,*mAP*_0.5:0.95提高1.55个百分点。在cucumber leaf数据集上BiFEL-YOLOv9在*mAP*_0.5和*mAP*_0.5:0.95上分别提高了2.7和2.39个百分点。

这表明BiFEL-YOLOv9模型不仅在特定数据集上表现优异,而且在不同环境条件下也展现了出色的

泛化能力。

4 结论与展望

针对黄瓜实际种植场景中叶片图像的多样性、遮挡问题以及复杂背景问题,本文提出了一种基于改进YOLOv9的黄瓜病害识别模型BiFEL-YOLOv9。该模型通过引入BiFPN双向特征金字塔网络模块增强了模型对多尺度特征的融合能力,同时结合特征增强模块和LSKBlock注意力机制对原始的RepNCSPeLan4模块进行改进得到RNFEF模块,提升了模型的特征表示能力对复杂背景的鲁棒性。实验结果表明,BiFEL-YOLOv9模型在精确率、召回率以及*mAP*指标上均优于现有的YOLOv5、YOLOv7和YOLOv9模型。BiFEL-YOLOv9模型为黄瓜病害快速检测和准确识别提供有效解决方案,对提升农业生产效率和减少损失具有重要应用价值。

参考文献

- 徐心诚. 弱光对温室黄瓜叶片和茎中可溶性糖含量的影响. 江苏农业学报, 2015, 31(6): 1448-1450. [doi: 10.3969/j.issn.1000-4440.2015.06.040]
- 徐千惠. 黄瓜主要病害的发生规律及绿色防控技术. 上海蔬菜, 2022(2): 51-52. [doi: 10.3969/j.issn.1002-1469.2022.02.020]
- 吴泽翔, 赵杰, 胡晓. 基于特征融合和SVM的植物病害检测. 电脑编程技巧与维护, 2022(2): 39-41. [doi: 10.3969/j.issn.1006-4052.2022.02.013]
- 王弘. 基于深度学习的田间农作物病虫害识别方法研究 [硕士学位论文]. 合肥: 安徽大学, 2023.
- Pandian JA, Kumar VD, Geman O, *et al.* Plant disease detection using deep convolutional neural network. Applied Sciences, 2022, 12(14): 6982. [doi: 10.3390/app12146982]
- Seetharaman K, Mahendran T. Leaf disease detection in banana plant using Gabor extraction and region-based convolution neural network (RCNN). Journal of The Institution of Engineers (India): Series A, 2022, 103(2): 501-507. [doi: 10.1007/s40030-022-00628-2]
- Zhang Y, Song CL, Zhang DW. Deep learning-based object detection improvement for tomato disease. IEEE Access, 2020, 8: 56607-56614. [doi: 10.1109/ACCESS.2020.2982456]
- Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern

- Recognition (CVPR). Las Vegas: IEEE, 2016. 779–788.
- 9 董雁凯, 王玉超, 李博株, 等. 基于改进 YOLOv5 的黄瓜霜霉病分级方法. 江苏农业科学, 2023, 51(22): 213–220. [doi: [10.15889/j.issn.1002-1302.2023.22.029](https://doi.org/10.15889/j.issn.1002-1302.2023.22.029)]
 - 10 李淑菲, 李凯雨, 乔岩, 等. 基于可见光光谱和改进 YOLOv5 的自然场景下黄瓜病害检测方法. 光谱学与光谱分析, 2023, 43(8): 2596–2600. [doi: [10.3964/j.issn.1000-0593\(2023\)08-2596-05](https://doi.org/10.3964/j.issn.1000-0593(2023)08-2596-05)]
 - 11 乔琛, 韩梦瑶, 高苇, 等. 基于 Faster-NAM-YOLO 的黄瓜霜霉菌孢子检测. 农业机械学报, 2023, 54(12): 288–299. [doi: [10.6041/j.issn.1000-1298.2023.12.028](https://doi.org/10.6041/j.issn.1000-1298.2023.12.028)]
 - 12 Wang CY, Yeh IH, Liao HYM. YOLOv9: Learning what you want to learn using programmable gradient information. Proceedings of the 18th European Conference on Computer Vision. Milan: Springer, 2024. 1–21.
 - 13 Tan MX, Pang RM, Le QV. EfficientDet: Scalable and efficient object detection. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 10778–10787.
 - 14 Li YX, Hou QB, Zheng ZH, *et al.* Large selective kernel network for remote sensing object detection. Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris: IEEE, 2023. 16748–16759.
 - 15 Lin TY, Dollár P, Girshick R, *et al.* Feature pyramid networks for object detection. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017. 936–944.
 - 16 Liu S, Qi L, Qin HF, *et al.* Path aggregation network for instance segmentation. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 8759–8768.
 - 17 Wang CY, Liao HYM, Wu YH, *et al.* CSPNet: A new backbone that can enhance learning capability of CNN. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle: IEEE, 2020. 1571–1580.
 - 18 Wang CY, Liao HYM, Yeh IH. Designing network design strategies through gradient path analysis. arXiv:2211.04800, 2022.
 - 19 Sultana N, Shorif SB, Akter M, *et al.* A dataset for successful recognition of cucumber diseases. Data in Brief, 2023, 49: 109320. [doi: [10.1016/j.dib.2023.109320](https://doi.org/10.1016/j.dib.2023.109320)]

(校对责编: 王欣欣)