

# 基于优化 CBAM 改进 YOLOv5 的农作物病虫害识别<sup>①</sup>



王 昕, 董 琴, 杨国宇

(盐城工学院 信息工程学院, 盐城 224051)

通信作者: 王 昕, E-mail: 1315639947@qq.com

**摘 要:** 针对多种农作物病虫害图像, 在自然环境下因虫害种类繁多, 小目标特征相似的技术问题, 导致检测困难难以达到令人满意的精度. 本文提出了一种自然背景下加强局部特征和全局特征信息融合的害虫检测识别模型 YOLOv5-EB, 在公开的大规模害虫数据集 IP102 上进行实验, 结果表明该研究比 YOLOv5 模型精确度提高了 5 个百分点. 引入一维卷积替换 CBAM 中通道注意力的 MLP 操作, 优化了通道注意力经过全局处理后容易忽略通道内信息交互的问题; 其次使用 6×6 卷积替换 Focus 操作, 来增强提取害虫特征的能力. 实验结果表明, 对害虫进行检测时, YOLOv5-EB 的平均精度值达到了 87%, 与 Faster R-CNN、EfficientDet、YOLOv3、YOLOv4、YOLOv5 模型相比, 不仅有效提高了作物害虫图像的识别性能, 而且有效提高了检测速度. 研究表明, YOLOv5-EB 算法满足对多种农作物病虫害目标检测的准确性和实时性要求.

**关键词:** YOLOv5 算法; CBAM; 病虫害检测; IP102; 目标识别; 深度学习

引用格式: 王昕, 董琴, 杨国宇. 基于优化 CBAM 改进 YOLOv5 的农作物病虫害识别. 计算机系统应用, 2023, 32(7): 261-268. <http://www.c-s-a.org.cn/1003-3254/9175.html>

## YOLOv5 Improved by Optimized CBAM for Crop Pest Identification

WANG Xin, DONG Qin, YANG Guo-Yu

(School of Information Engineering, Yancheng Institute of Technology, Yancheng 224051, China)

**Abstract:** For a variety of crop disease and pest images, it is difficult to achieve satisfactory accuracy due to the technical problems of various diseases and pests and similar characteristics of small targets in the natural environment. In this study, a pest detection and identification model, namely YOLOv5-EB that enhances the fusion of local feature and global feature information in the natural background is proposed, and experiments are carried out on the published large-scale pest dataset IP102. The results show that the accuracy of this study is improved by five percentage points compared with the YOLOv5 model. The MLP operation of replacing channel attention in CBAM with one-dimensional convolution is introduced, which optimizes the problem that channel attention is easy to ignore the information interaction in the channel after global processing. Secondly, the Focus operation is replaced by 6×6 convolution to enhance the ability to extract pest features. The experimental results show that the average accuracy of YOLOv5-EB reaches 87% in detecting pests, which not only effectively improves the identification performance of crop pest images but also increases the detection speed compared with Faster R-CNN, EfficientDet, YOLOv3, YOLOv4, and YOLOv5 models. The study reveals that the YOLOv5-EB algorithm meets the accuracy and real-time requirements of target detection of various crop diseases and pests.

**Key words:** YOLOv5 algorithm; convolutional block attention module (CBAM); pest detection; IP102; target identification; deep learning

<sup>①</sup> 基金项目: 江苏省产学研合作项目 (BY2022460)

收稿时间: 2022-12-02; 修改时间: 2023-02-03; 采用时间: 2023-02-27; csa 在线出版时间: 2023-05-19

CNKI 网络首发时间: 2023-05-22

在农业病虫害防治领域,计算机视觉应用图像识别技术加强对害虫的及时预防措施,例如病叶片识别<sup>[1-3]</sup>,水果检测<sup>[4-6]</sup>和水稻病虫害识别<sup>[7-9]</sup>,病虫害图像识别逐渐成为时代的发展趋势.早期发现和识别农作物中的害虫主要以人工观察的方法,根据经验总结结论,而作物的症状有非常相似的视觉特征,如颜色纹理等,其在时间消耗、准确度上受到限制.要克服这些局限性,目前的工作是以计算机视觉与深度学习相结合的方式,利用优化CBAM<sup>[10]</sup>注意力机制和改进YOLOv5主干网络等技术,用于检测农作物病虫害,为后续的增产农作物、减少农药滥用等具有重要意义.随着人工智能技术的发展,国内外学者对害虫识别研究已经取得了十足的进展.Kasinathan等人<sup>[11]</sup>认为害虫的纹理、颜色、形状、HOG和GIST特征的组合可以最大化提升识别性能.Liu等人<sup>[12]</sup>利用集成学习思想,将迁移学习后的CNN网络系列害虫图像识别和分析模型有效融合,极大提高模型对农作物害虫识别和分类的准确性.曾伟辉等人<sup>[13]</sup>利用ResNet网络融合注意力机制,并通过自校准卷积模块的方式,提升了对低分辨率病虫害图像的识别精度.这些方法虽然可以增强检测识别害虫的性能,但是大多是针对单个农作物

病虫害进行检测对比实验,对多种农作物病虫害检测实验未有涉及.

为了提升网络性能,学者先后提出了SENet<sup>[14]</sup>、CBAM、ECA<sup>[15]</sup>等注意力机制,在公开数据集上,这些研究增加注意力机制可以有效提升网络性能.但对多种农作物病虫害数据集实验时,没有达到理想的效果.因此本文提出一种优化CBAM注意力机制的方法,增强通道信息之间的特征融合能力,和改进YOLOv5主干网络的害虫特征提取能力,融合两种方法,构建多种农作物病虫害识别网络.

## 1 数据来源

本研究中的农作物病虫害数据集来源于IP102数据集<sup>[16]</sup>.IP102是一个用于目标分类和检测任务的农作物病虫害数据集,由于原数据集包含了102个害虫类别的75000多幅图像,且拥有水稻、玉米、小麦、甜菜、苜蓿、葡萄、柑橘、芒果共8种作物.本文研究中的数据集是IP102数据集的缩减版,保留原数据集特性的同时,从8种作物共计选取94个害虫类别,共计2485幅图像.各类农作物的部分害虫图像示例如图1所示.



图1 部分害虫图像示例

## 2 农作物病虫害图像预处理

为了扩充数据集,提高识别的准确性,采用裁剪、

平移、改变亮度、加噪声、旋转角度、镜像、cutout<sup>[17]</sup>操作转换训练集中原有图像,对多个样本进行数据扩

充,增加数据集的多样性.同时为了模型更好地识别病虫害,采用了马赛克数据增强方法提高网络检测的泛化性能.

### 2.1 数据扩充

由于获取的原数据集数量少且类别不均衡,直接用来训练容易导致平均精确度过低.葡萄作物内部分病虫害图像数据过多,约为3种害虫(十星叶甲、大青叶蝉、盲蝽科)的图像分别为65、79、110幅图像,其数据量达到了其他害虫的3倍,因此只对除却这4种害虫外的农作物害虫进行数据扩充随机操作,以达到数据类别均衡.预计扩充样本3倍,而十星叶甲、大青叶蝉、盲蝽科的图像数量保持不变,数据集均衡过后,网络模型训练会极大提升识别精确度,有效缓解病虫害类别不均衡的问题.各类别农作物和虫害类别数,样本数量及数据扩充后得到的样本数,如表1所示.

表1 农作物害虫种类和样本数量

农作物	虫害类别数	样本数量	数据扩充
水稻	14	328	984
玉米	13	291	873
小麦	9	198	594
甜菜	8	177	531
苜蓿	13	358	1074
葡萄	12	545	1127
柑橘	15	388	1164
芒果	10	200	600
总数	94	2485	6947

### 2.2 数据增强

由于原数据集的农作物害虫样本繁多,包含94个虫害类别,但是图像却少样本,这对于获得良好的模型性能是远远不够的.因此,根据数据集其中的特性,使用马赛克在线数据增强和传统离线数据增强的方法组合,对图像进行位移、大小变换,并将其9种虫害图像随机拼接裁剪成一张图像.通过上述方式对数据集进行处理,能够间接提高模型的每次训练样本数,降低计算开销.

## 3 优化CBAM改进YOLOv5算法

针对生活在复杂场景和识别具有类似特征的农作物病虫害,如果农作物不同的异常现象表现出相似的视觉特征,YOLOv5分类器可能无法区分它们.为了提高小目标的检测效果,利用图像中的全局信息,加强主

干网络中局部特征和全局特征的交互性能,进一步增强复杂背景中正确识别小目标的能力,以适应不同农田环境下的目标检测.

本文首次通过将CBAM注意力中的通道注意力机制改进为ECAM模块,来解决通道注意力经过全局处理后容易忽略通道内信息交互的问题;同时将YOLOv5原主干网络Darknet改进,增大卷积核尺寸,且融合ECAM注意力机制.通过这些改进,来提升网络模型对复杂场景中具有类似特征的农作物病虫害的目标检测识别能力.

### 3.1 YOLOv5的结构模型

YOLOv5目标检测框架在中大型目标物体的检测上具有很大的性能优势,实时监测速度达到了140 fps,且具有较高的识别精度,应用范围广泛.YOLOv5由3个主要部分组成:骨干网络、瓶颈层和检测头,模型的结构如图2所示.

初始的YOLOv5检测框架对农作物病虫害数据集并不理想,在融合优化CBAM注意力机制的基础上在主干网络做出改进,使其适应复杂环境下的农作物病虫害.

### 3.2 基于CBAM的改进注意力模块

CBAM是一种结合通道和空间的注意力机制模块,有着轻量级的参数,在公开数据集上的分类检测性能方面持续改进.但是在本文的农作物病虫害数据集中,虫害分类多,少样本和部分病虫害症状特征相似等问题,分类检测性能没有达到理想效果.图3是CBAM的结构图.

原通道注意力机制在输入特征图后,分别经过基于宽度和高度的全局池化和最大池化操作,然后经过MLP,将MLP输出特征进行基于element-wise的加权操作,再经过Sigmoid操作后,生产最终的通道注意力特征图.该操作在将一个通道内信息直接进行全局池化和最大池化处理,容易忽略通道内的信息交互.本文对CBAM的通道注意力部分进行优化改进,弃掉最大池化操作,通过平均池化操作聚合一个特征图的空间信息,这个空间信息表示被压缩后的特征.然后,借鉴ECA的思想,用卷积核大小为 $K$ 的一维卷积代替MLP网络.经过一维卷积处理,加强不同通道之间信息交互能力,之后用Sigmoid函数简化输出.简言之,实现以 $K$ 为大小的局部跨通道交互信息,如图4是改进后的通道注意力模块ECAM.

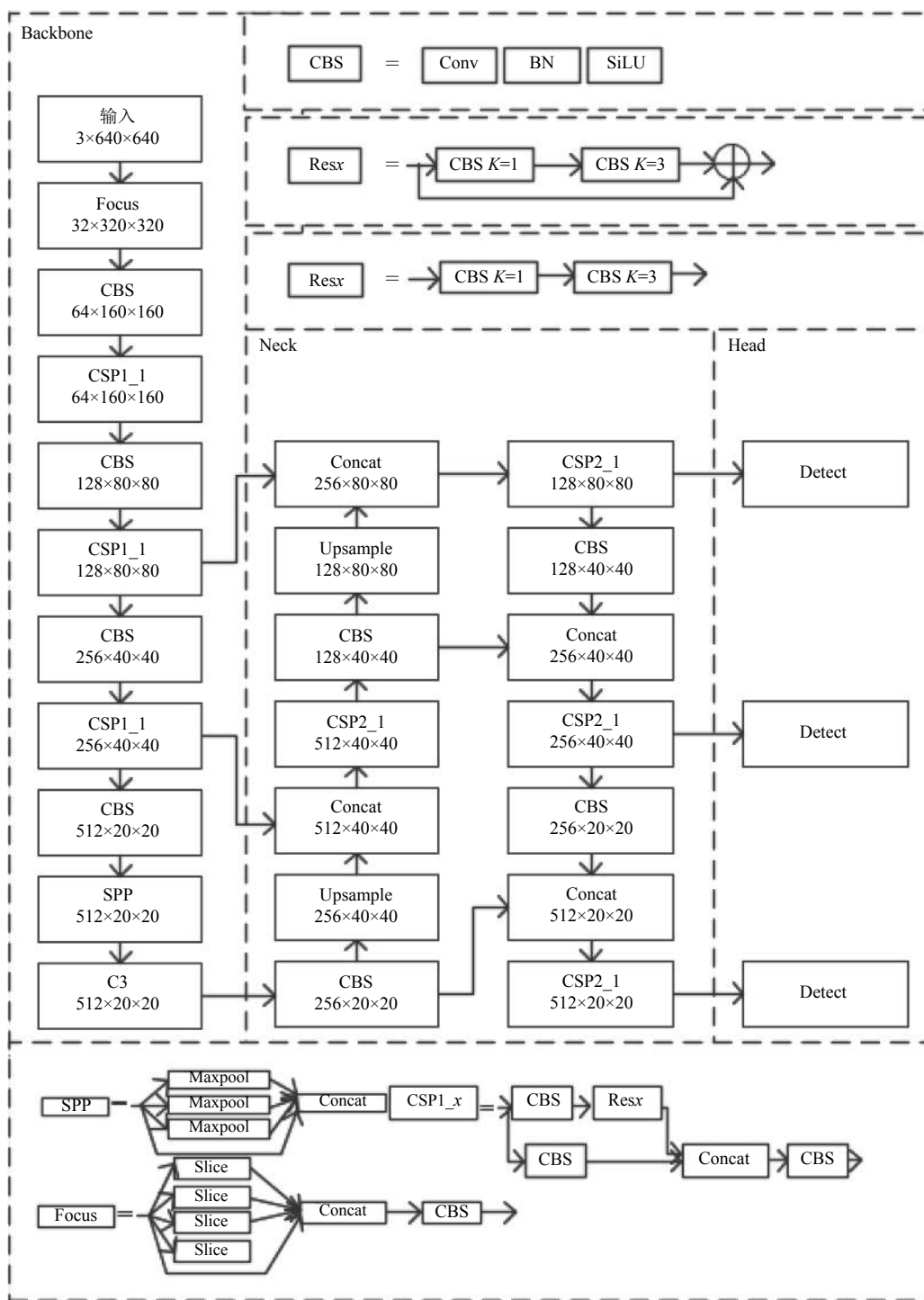


图2 YOLOv5 结构模型

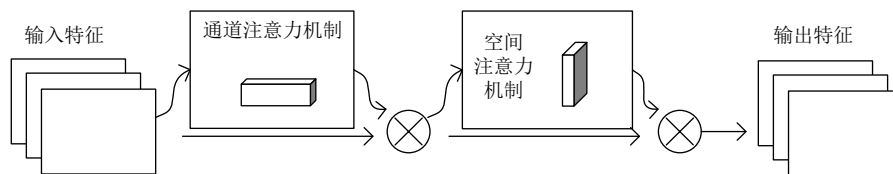


图3 CBAM 结构图

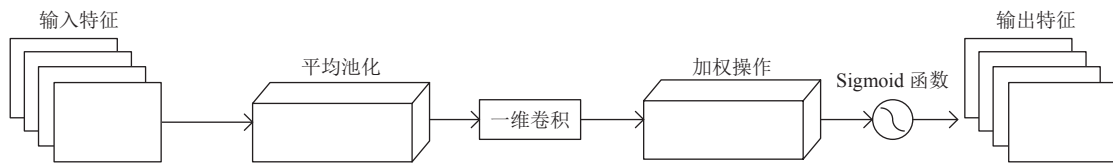


图4 ECAM 结构图

通道注意力计算方式如式 (1). 其中,  $\sigma$  为 Sigmoid 函数,  $\text{conv1d}$  为一维卷积,  $k$  表示卷积核大小,  $\text{AvgPool}$  表示平均池化操作. 通过交叉验证确定卷积核  $k$  的最优大小为 3.

$$M_{EC}(F) = \sigma(\text{conv1d}_{k=3}(\text{AvgPool}(F))) \quad (1)$$

### 3.3 骨干网络改进

在 YOLOv5 骨干网络引入优化的 CBAM 注意力机制, 优化的 CBAM 的通道注意力机制加强了通道之间信息的融合, 空间注意力机制重点检测目标的位置, 通过两者结合可以使输出信息更加聚焦于重点信息, 从而达到模型准确率的效果. 本文将原始的 YOLOv5 中 stem 部分的 Focus 切片操作换成  $6 \times 6$  的卷积, 使用 SPPF 模块替换了 SPP 和后面的 C3 结构, 因为 SPP 的并行操作会影响速度. 如图 5 所示.

## 4 农作物病虫害识别结果与分析

### 4.1 实验环境

模型在服务器上搭建和训练, 该服务器配备 Intel(R) Xeon(R) Gold 6258R CPU 和 GeForce RTX 3060 32 GB GPU. 软件环境为 Python 3.9, 深度学习框架为 PyTorch 1.12.1.

### 4.2 识别结果

采用了本文改进的 YOLOv5-EB 模型对 IP102 病虫害进行识别, 结果如图 6, 图 7, 可以看出, YOLOv5-EB 能够对害虫进行准确的定位和识别.

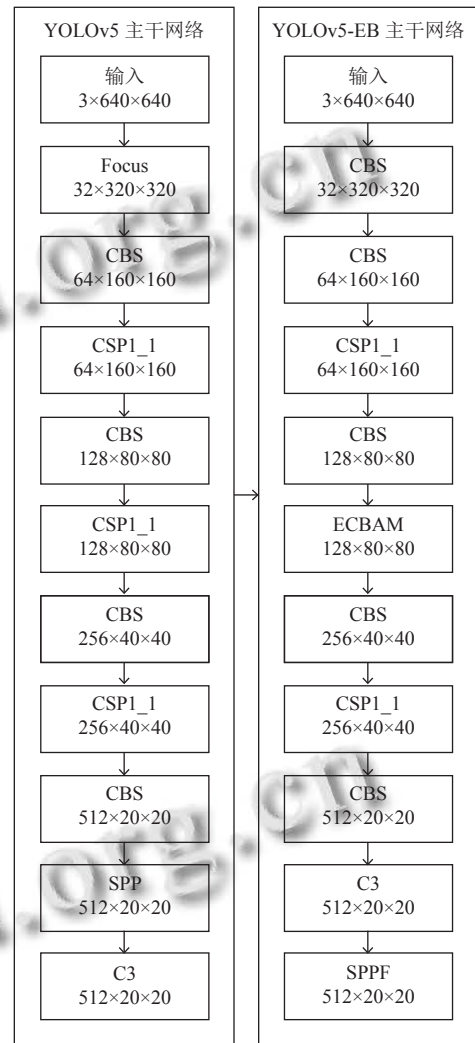


图5 融合优化 CBAM 的改进 YOLOv5 主干网络

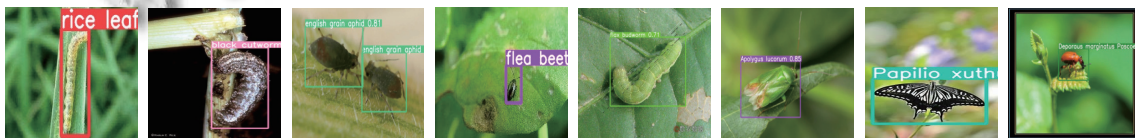


图6 YOLOv5 识别结果

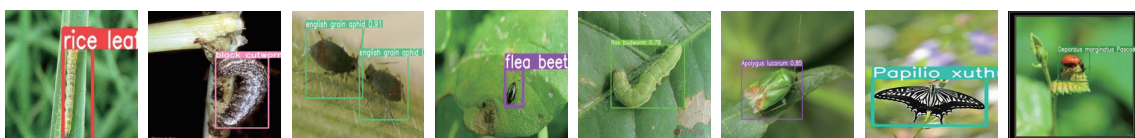


图7 YOLOv5-EB 识别结果

### 4.3 网络模型实验对比

为了客观评价网络模型性能,对 Faster R-CNN、EfficientDet、YOLOv3、YOLOv4、YOLOv5 以及本文改进的 YOLOv5-EB 模型进行了比较,评估指标包括精确率、召回率、F1 值和平均精度 (MAP)。由表 2 可见,在对 IP102 的多农作物病虫害进行识别时,YOLOv5-EB 在网络性能上优于 EfficientDet、YOLOv3、YOLOv4 和 Faster R-CNN,而 YOLOv5 的初始性能则略优于它们。对于农作物虫害的识别,YOLOv5-EB 的各项指标都是最高的,包括精确率、召回率、F1 值和平均精度均值。其中,如图 8 所示,YOLOv5-EB 的平均精度值分别比 Faster R-CNN 高出 20 个百分点,比 EfficientDet 高出近 7 个百分点,比 YOLOv3 高出 8 个百分点,比 YOLOv4 高出 13 个百分点,比 YOLOv5 高出 5 个百分点,对比其他的目标检测模型加快了模型训练时的收敛速度。

为了进一步验证改进模型 YOLOv5-EB 的有效性,在模型收敛边框回归损失、分类概率损失、置信度损失与 Faster R-CNN、EfficientDet、YOLOv3、YOLOv4、YOLOv5 进行对比实验。

由图 9 可以得出,本文改进模型 YOLOv5-EB 在训练 100 个迭代周期后,边框回归损失前期收敛的速度略慢于 EfficientDet,在近 60 次迭代次数后收敛速度增快,边框回归损失最优;由于改进 YOLOv5-EB

引入了 ECAM 机制,导致分类概率损失表现前期收敛平缓,略慢于 YOLOv3,但充分训练后,收敛结果略优于 YOLOv3;置信度损失上对比明显,YOLOv5-EB 一直表现优于其他模型,综上所述,本文改进模型 YOLOv5-EB 对作物害虫检测的损失值对比表现最优。

表 2 网络模型评价指标 (%)

网络模型	精确率	召回率	F1 值	平均精度均值
Faster R-CNN	71.9	57.5	63.9	66.8
EfficientDet	70.6	76.0	73.2	80.3
YOLOv3	82.2	72.3	76.9	79.1
YOLOv4	74.6	69.0	71.7	74.2
YOLOv5	74.9	77.3	76.1	82.0
YOLOv5-EB	83.5	80.1	81.8	87.0

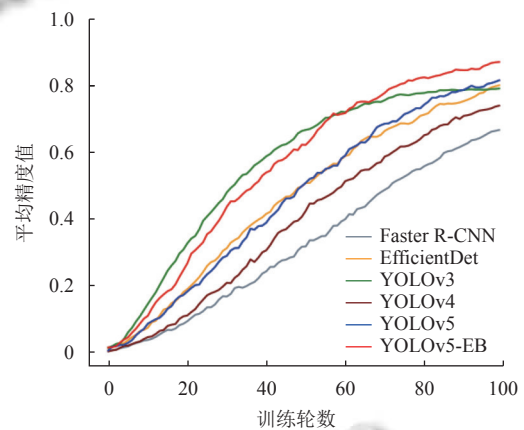


图 8 网络模型平均精确度对比

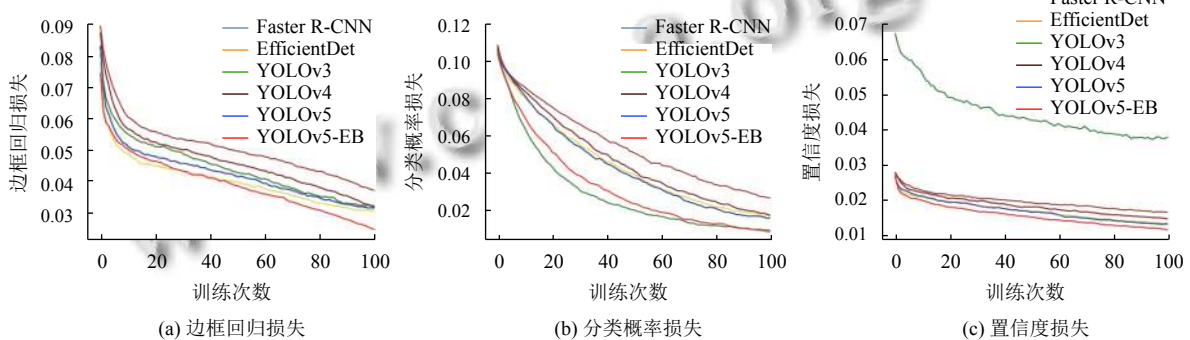


图 9 改进后模型收敛情况

为检测 YOLOv5-EB 模型的实用性,在检测时间上和 YOLOv5 进行对比实验,从数据集中随机选取 25 张图像,并将它们分成 5 组,对训练好的 YOLOv5 和 YOLOv5-EB 模型进行对比,为了消除偶然因素的影响,将 25 幅图像分成 5 组,标记为 T1-T5,求得每组

实验的平均值进行检测时间的对比,如表 3 所示。

### 4.4 消融实验

针对农作物病虫害图像识别应用需求,进一步提升图像中获取害虫信息的准确性,以 YOLOv5 模型为基础网络模型开展消融实验,由表 4 可知,引入 CBAM

模块对该研究中的数据集没有明显提升,加入 ECAM 注意力机制后准确率提高了 0.8 个百分点,采用 6×6 卷积替换原网络的 Focus 操作,大大增强了特征提取的能力,准确率提升了近 4.2 个百分点,这是因为害虫属于小目标且多尺度变换,不适合 Focus 操作将图片划分四等分,破坏了提取害虫特征的能力.选择 SPPF 是为了在不损害模型性能的同时,提升检测害虫的速度.

表 3 检测时间对比 (s)

网络模型	T1	T2	T3	T4	T5
YOLOv5	1.27	1.237	1.184	1.125	1.295
YOLOv5-EB	1.186	1.126	1.122	1.031	1.17
时间差	0.084	0.111	0.062	0.094	0.125

表 4 消融实验结果 (%)

+CBAM	+ECAM	+6×6卷积	+SPPF	准确率
—	—	—	—	82.05
√	—	—	—	81.94
—	√	—	—	82.86
—	—	√	—	86.23
—	—	—	√	82.10
—	√	√	√	87.03

#### 4.5 Grad-cam 热力图

将 YOLOv5-EB 模型产生的特征激活热力图和 YOLOv5 网络做对比,从图 10 中可以看出,原网络 YOLOv5 对目标本身检测识别更为看重,因为它更看重局部卷积,相比之下,优化后的 CBAM 模块更好地将局部卷积和全局卷积结合,有助于模型更容易判别是哪一种害虫.因为害虫本身种类较多,视觉上症状非常相似,所以更要参考自然环境下的背景信息,可以更准确和完整的定位目标对象,在识别中不仅利用害虫本身的特征,也与背景信息之间的进行有效融合.其热力图变化如图 10 所示.

#### 5 结论

改进后的 YOLOv5 算法针对害虫数据集在复杂背景下种类过多,难以检测识别的问题,优化了 CBAM 注意力机制,强化了通道注意力机制在经过全局池化后的通道信息之间交互性能,由于害虫的小目标多尺度的特性,直接采用卷积替换 Focus 操作可以有效提升检测害虫特征的能力,为增强检测害虫实时性能,引入 SPPF 模块,在保持检测识别性能的同时,大大提升检测速度.未来将继续探索如何进一步在自然背景下

提升检测农作物害虫的识别性能,提升模型的小样本识别的泛化能力.

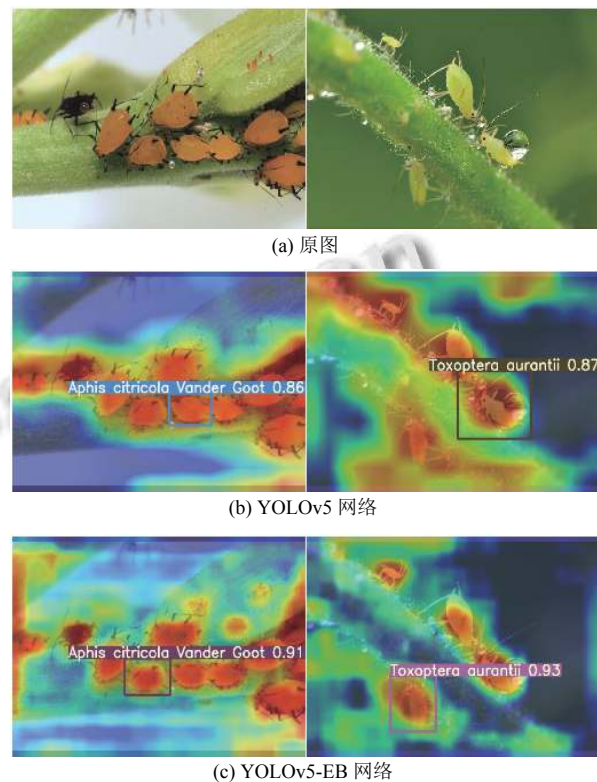


图 10 网络特征激活热力图

#### 参考文献

- 谢圣桥, 宋健, 汤修映, 等. 基于迁移学习和残差网络的葡萄叶部病害识别. 农机化研究, 2023, 45(8): 18–23, 28. [doi: 10.13427/j.cnki.njyi.2023.08.048]
- 曹跃腾, 朱学岩, 赵燕东, 等. 基于改进 ResNet 的植物叶片病虫害识别. 中国农机化学报, 2021, 42(12): 175–181. [doi: 10.13733/j.jcam.issn.2095-5553.2021.12.26]
- Koppal LB, Rajesh TM. Review on the techniques used for identification of diseased leaves. I-Manager's Journal on Pattern Recognition, 2020, 7(1): 32. [doi: 10.26634/jpr.7.1.17375]
- 陈贝文, 陈淦. 水果分类识别与成熟度检测技术综述. 计算机时代, 2022, (7): 62–65. [doi: 10.16644/j.cnki.cn33-1094/tp.2022.07.016]
- 徐印赞, 江明, 李云飞, 等. 基于改进 YOLO 及 NMS 的水果目标检测. 电子测量与仪器学报, 2022, 36(4): 114–123. [doi: 10.13382/j.jemi.B2104724]
- 朱明秀. 采摘机器人水果检测及定位研究——基于图像处理 and 卷积神经网络. 农机化研究, 2022, 44(4): 49–53. [doi: 10.13382/j.jemi.B2104724]

- 10.13427/j.cnki.njyi.2022.04.009]
- 7 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. 江苏农业学报, 2022, 38(3): 685–695. [doi: 10.3969/j.issn.1000-4440.2022.03.014]
  - 8 Rahman CR, Arko PS, Ali ME, *et al.* Identification and recognition of rice diseases and pests using convolutional neural networks. *Biosystems Engineering*, 2020, 194: 112–120. [doi: 10.1016/j.biosystemseng.2020.03.020]
  - 9 高雨亮, 徐向英, 章永龙, 等. 融合分组注意力机制的水稻病虫害图像识别算法. 扬州大学学报(自然科学版), 2021, 24(6): 53–57. [doi: 10.19411/j.1007-824x.2021.06.010]
  - 10 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018. 3–19.
  - 11 Kasinathan T, Uyyala SR. Machine learning ensemble with image processing for pest identification and classification in field crops. *Neural Computing and Applications*, 2021, 33(13): 7491–7504. [doi: 10.1007/s00521-020-05497-z]
  - 12 Liu YW, Zhang X, Gao YX, *et al.* Improved CNN method for crop pest identification based on transfer learning. *Computational Intelligence and Neuroscience*, 2022, 2022: 9709648.
  - 13 曾伟辉, 张文凤, 陈鹏, 等. 基于SCResNeSt的低分辨率水稻害虫图像识别方法. 农业机械学报, 2022, 53(9): 277–285. [doi: 10.6041/j.issn.1000-1298.2022.09.028]
  - 14 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 7132–7141.
  - 15 Wang QL, Wu BG, Zhu PF, *et al.* ECA-Net: Efficient channel attention for deep convolutional neural networks. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2022. 11531–11539.
  - 16 Wu XP, Zhan C, Lai YK, *et al.* IP102: A large-scale benchmark dataset for insect pest recognition. *Proceedings of the 32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach: IEEE, 2019. 8779–8788.
  - 17 DeVries T, Taylor GW. Improved regularization of convolutional neural networks with cutout. arXiv:1708.04552, 2017.
- (校对责编: 孙君艳)