

基于改进C3模块的YOLOv5三种注意力机制对水稻害虫检测性能比较

沈珈毅¹, 冉伟豪¹, 黄金城^{1*}, 韩刘婧¹, 焦天慧¹, 王欣悦², 李锦怡³, 王杰文⁴

¹盐城工学院信息工程学院, 江苏 盐城

²盐城工学院经济管理学院, 江苏 盐城

³盐城工学院设计艺术学院, 江苏 盐城

⁴宾夕法尼亚州立大学计算机科学与工程系, 美国 斯泰特科利奇

收稿日期: 2025年7月2日; 录用日期: 2025年7月31日; 发布日期: 2025年8月6日

摘要

中国是全球最大的水稻生产国, 种植面积常年约占全球总量的20%。虫害是制约水稻产量和品质的主要因素之一, 传统虫害检测方法主要依赖人工观察和性诱捕法, 但两者均存在局限性, 难以满足现代农业对虫害的精准防控需求。随着农业智能化的发展与深度学习的广泛应用, 基于深度学习的农作物病虫害智能检测已成为研究热点方向。本文针对上述问题, 依据实际场景采用目标检测方法对水稻害虫进行检测, 对YOLOv5 (You Only Look Once version 5)检测算法进行改进, 在其基础上引入三种主流注意力机制——ECA (Efficient Channel Attention)、CBAM (Convolutional Block Attention Module)和SE (Squeeze-and-Excitation), 分别构建改进模块C3ECA、C3CBAM和C3SE, 分别替代原主干网络中的C3模块, 从而形成三种改进模型: YOLOv5s-C3ECA、YOLOv5s-C3CBAM与YOLOv5s-C3SE。通过在公开稻田害虫图像数据集上的训练与测试, 实验结果表明, 与原始的YOLOv5s相比, YOLOv5s-C3CBAM和YOLOv5s-C3ECA改进模型在mAP@0.5、mAP@0.5:0.95分别有不同程度的提升, 其中YOLOv5s-C3CBAM在mAP@0.5、mAP@0.5:0.95上分别提升1.2%和0.3%; YOLOv5s-C3ECA在mAP@0.5、mAP@0.5:0.95上分别提升2.5%和1.1%。结果表明, 所提出的改进策略在保持模型轻量化的同时提升了水稻害虫目标的检测精度与稳定性, 适用于资源受限场景下的农业智能终端部署, 具有良好的实际应用价值。

关键词

YOLOv5, SE, CBAM, ECA, 目标检测, 水稻害虫

Comparative Study on Rice Pest Detection Performance Using YOLOv5 Enhanced with Three C3 Attention Modules

*通讯作者。

文章引用: 沈珈毅, 冉伟豪, 黄金城, 韩刘婧, 焦天慧, 王欣悦, 李锦怡, 王杰文. 基于改进 C3 模块的 YOLOv5 三种注意力机制对水稻害虫检测性能比较[J]. 计算机科学与应用, 2025, 15(8): 41-49. DOI: 10.12677/csa.2025.158196

Jiayi Shen¹, Weihao Ran¹, Jincheng Huang^{1*}, Liujing Han¹, Tianhui Jiao¹, Xinyue Wang², Jinyi Li³, Jiewen Wang⁴

¹School of Information Engineering, Yancheng Institute of Technology, Yancheng Jiangsu

²School of Economics and Management, Yancheng Institute of Technology, Yancheng Jiangsu

³School of Design and Art, Yancheng Institute of Technology, Yancheng Jiangsu

⁴Department of Computer Science and Engineering, The Pennsylvania State University, State College, USA

Received: Jul. 2nd, 2025; accepted: Jul. 31st, 2025; published: Aug. 6th, 2025

Abstract

China is the largest rice producer worldwide, with a planting area accounting for approximately 20% of the global total. Pest infestation is one of the primary factors limiting rice yield and quality. Traditional pest detection methods mainly rely on manual inspection and sex pheromone trapping, both of which have limitations and fail to meet the precise pest control requirements of modern agriculture. With the development of agricultural intelligence and the widespread application of deep learning, deep learning-based intelligent detection of crop pests and diseases has become a research hotspot. To address these issues, this study employs an object detection approach for rice pest identification and improves the YOLOv5 (You Only Look Once version 5) detection algorithm by introducing three mainstream attention mechanisms—ECA (Efficient Channel Attention), CBAM (Convolutional Block Attention Module), and SE (Squeeze-and-Excitation). Corresponding improved modules C3ECA, C3CBAM, and C3SE are constructed to replace the original C3 modules in the backbone network, forming three improved models: YOLOv5s-C3ECA, YOLOv5s-C3CBAM, and YOLOv5s-C3SE. Experiments conducted on a publicly available rice pest image dataset demonstrate that compared with the original YOLOv5s, the YOLOv5s-C3CBAM and YOLOv5s-C3ECA models achieve varying degrees of improvement in mAP@0.5 and mAP@0.5:0.95, with YOLOv5s-C3CBAM improving by 1.2% and 0.3%, and YOLOv5s-C3ECA improving by 2.5% and 1.1%, respectively. The results indicate that the proposed improvement strategy enhances detection accuracy and stability of rice pest targets while maintaining model lightweight characteristics, making it suitable for deployment on resource-constrained agricultural intelligent terminals and offering significant practical value.

Keywords

YOLOv5, SE, CBAM, ECA, Object Detection, Rice Pest

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

水稻是世界最重要的粮食作物之一,2020年中国稻米产量达到21,186 kt,产量在全球占比为37.31%。但近年来水稻各类病害的发生直接影响了水稻的产量和质量,给农业带来了严重的经济损失,制约着中国水稻产业的健康稳定发展[1]。然而在水稻生长过程中,稻飞虱、螟虫、叶蝉等害虫的频繁爆发严重影响产量与品质,给农业生产带来损失。长期以来,大多数传统农作物的虫害识别均以人工检测为主,不仅效率低,成本高,还容易受到主观因素影响。近年来,研究人员开始将图像处理和机器学习技术应用于农作物病虫害的检测[2]。随着计算机视觉技术的迅速发展,图像处理与机器学习方法被广泛应用于农

作物病虫害识别任务中。已有研究表明, 基于支持向量机(SVM)与 K 均值聚类等方法可以实现一定程度的病害分类识别。例如, Zhang 等[3]提出了一种基于叶片图像的黄瓜病害识别方法, 结合 K-means 聚类与稀疏表示分类器; Sun 等[4]将超像素分割方法与 SVM 相结合用于病害图像识别; Sethy 等[5]则基于深度特征与 SVM 完成了水稻病害分类。深度学习技术在农业领域中逐渐展现出强大的特征提取与端到端识别能力, 尤其是目标检测算法在农作物病虫害自动识别中的应用取得了显著进展。尽管上述方法取得一定成果, 但过于依赖手工特征提取, 难以满足稻田场景下害虫检测对实时性和精度的双重要求。YOLO (You Only Look Once)系列作为当前目标检测领域的主流算法兼具检测精度与实时性。其中, YOLOv5 作为该系列的代表, 在农业图像识别中具有广泛应用前景。然而, YOLOv5s 在应对稻田场景中密集、小目标害虫检测任务时, 存在检测精度下降的问题。对于水稻害虫这一特定场景, 专门优化的小目标检测方法相对缺乏, 注意力机制在该场景中缺乏系统性的对比分析, 模型轻量化与部署性考虑不足, 难以兼顾性能和实用性。针对上述问题, 本文在 YOLOv5s 的基础上, 引入三种主流轻量注意力机制——SE (Squeeze-and-Excitation)、CBAM (Convolutional Block Attention Module)和 ECA (Efficient Channel Attention), 分别构建 C3SE、C3CBAM 与 C3ECA 模块, 替代 YOLOv5s 主干网络中的 C3 结构, 构建三种改进模型。通过在公开稻田害虫图像数据集上的对比实验, 本文验证了所提出模型在小目标检测精度与模型轻量化方面的有效性与可行性, 为水稻害虫智能监测提供了一种高效、可部署的解决方案。

2. YOLOv5 神经网络

YOLOv5 是由 Ultralytics 团队在 2020 年发布的一种单阶段(One Stage)目标检测算法。YOLOv5 采用了一种新的模型结构, 将原来的模型简化了, 并加强了模型的鲁棒性、速度和精度。相对于其他目标检测算法, YOLOv5 具有响应速度快、精确度高等优势, 并且可以在较低的硬件条件下进行物体检测任务。YOLOv5 广泛应用于工业检测、农业识别和安防监控等领域。其整体网络结构由 Backbone、Neck 和 Head 三个部分组成, 如图 1 所示。

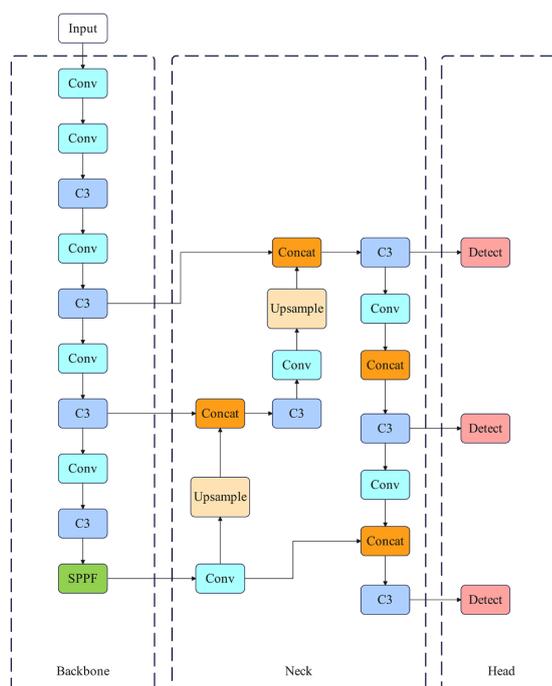


Figure 1. Overall structure of the YOLOv5 network

图 1. YOLOv5 网络整体结构

Backbone 是 YOLOv5 的特征提取模块,负责从输入图像中提取多层次的视觉特征。YOLOv5 采用改进的 CSPDarknet 作为主干网络结构,其中 Focus 等模块用于初步处理图像,Conv 模块用于提取基本特征,C3 模块通过引入跨阶段部分连接结构提高了特征表达能力并降低计算开销。此外,SPP (Spatial Pyramid Pooling)模块通过多尺度池化增强了感受野,使模型更好地捕捉不同尺度的目标信息,为后续检测提供丰富的特征基础。

Neck 部分用于对来自 Backbone 的不同层次的特征图进行融合,以增强模型对不同尺寸目标的检测能力。YOLOv5 的 Neck 结合了 FPN (Feature Pyramid Network)与 PAN (Path Aggregation Network)结构,FPN 自上而下地整合高层语义特征与低层空间特征,而 PAN 则自下而上补充底层细节信息。这样的双向融合结构使得模型在保持语义表达能力的同时,提升了对小目标和边缘目标的检测效果。

Head 是 YOLOv5 用于最终预测的位置,其功能是输出每个候选框的位置、置信度以及类别信息。YOLOv5 的检测头在三个尺度(通常为 P3/8、P4/16、P5/32)上进行输出,分别对应小、中、大目标的检测需求。在每个尺度的特征图上,模型在每个网格位置回归出多个锚框的边界框坐标、目标置信度以及所属分类概率。最终通过非极大值抑制(NMS)筛除冗余框,得到高质量的检测结果,实现对多目标的精确识别。其整体网络结构由 Backbone、Neck 和 Head 三个部分组成,如图 1 所示。

3. YOLOv5 网络引入注意力机制

3.1. SE 注意力机制

SE 注意力机制(Squeeze-and-Excitation Attention Mechanism)是一种有效的通道注意力机制。其核心思想是通过显式建模通道间的依赖关系,自适应地为不同通道分配权重,从而增强关键信息、抑制冗余特征,提高神经网络的表达能力与性能。该机制主要包括两个阶段:压缩(Squeeze)和激励(Excitation)。在压缩阶段,利用全局平均池化将每个通道的空间特征压缩为一个单一数值,形成通道级的全局描述向量;在激励阶段,借助一个包含两个全连接层的非线性结构学习各通道的重要性,并通过 Sigmoid 函数生成归一化的权重系数,进而对原始特征图进行加权重标定。其结构如图 2 所示,具体操作所用运用的公式如下:

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_{2\delta}(W_1 z)) \quad (2)$$

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (3)$$

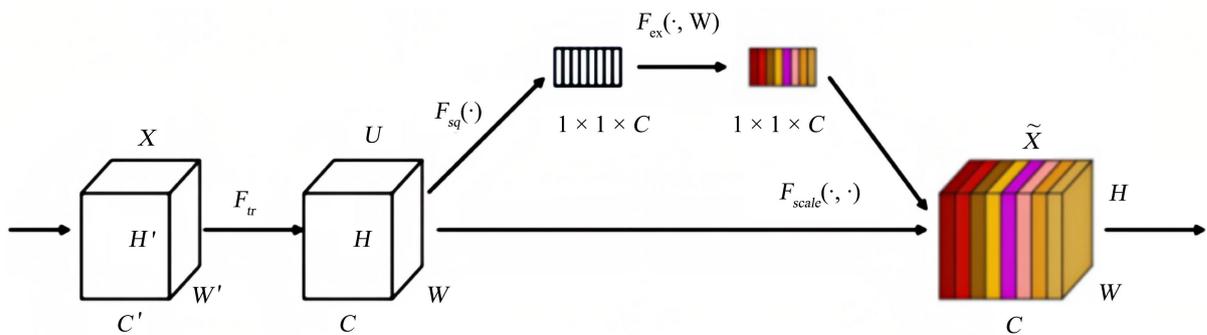


Figure 2. Structure of the SE attention mechanism

图 2. SE 注意力机制结构

3.2. CBAM 注意力机制

CBAM 注意力机制(Convolutional Block Attention Module)是一种结合了通道注意力与空间注意力的轻量级注意力机制。其核心思想是通过顺序地建模通道和空间两个维度上的注意力信息,使网络能够自适应地关注“重要的特征通道”和“关键的空间位置”,从而增强有用信息、抑制冗余干扰,进一步提升模型的特征表示能力与检测性能。CBAM 模块主要包含两个子模块:通道注意力模块(Channel Attention Module)和空间注意力模块(Spatial Attention Module),按顺序串联而成。在通道注意力模块中,分别使用全局最大池化和全局平均池化对输入特征图进行压缩,获得两个通道描述向量。随后,这两个向量通过一个共享的多层感知机(MLP)进行处理并相加,再经 Sigmoid 函数生成通道注意力权重,对输入特征图在通道维度上进行加权增强。在空间注意力模块中,首先在通道维度上对特征图进行最大池化和平均池化操作,将其压缩为两个空间注意力图。然后将二者拼接后送入一个 7×7 卷积层,并通过 Sigmoid 激活函数生成空间注意力图,用于对通道增强后的特征图在空间位置上进一步加权调整。该结构如图 3 所示。

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (4)$$

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \quad (5)$$

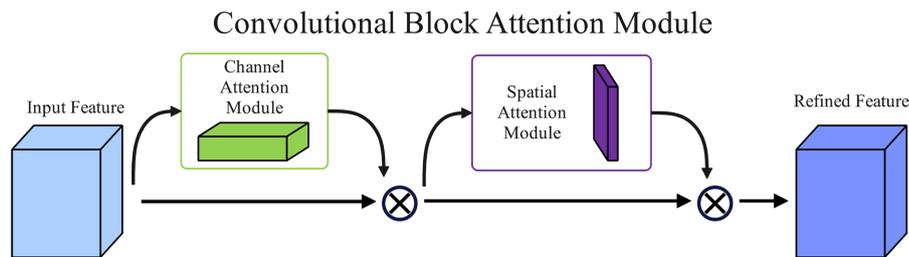


Figure 3. Structure of the CBAM attention mechanism
图 3. CBAM 注意力机制结构

3.3. ECA 注意力机制

ECA 注意力机制(Efficient Channel Attention Mechanism)是一种高效轻量的通道注意力机制。该方法在不引入维度压缩和复杂结构的前提下,通过局部一维卷积对通道间的依赖关系进行建模,从而实现更简洁高效的特征增强。ECA 的核心思想是避免传统 SE 模块中由于维度压缩可能带来的信息损失,采用无压缩策略直接在通道维度上建模依赖,提升了注意力学习的效率与表达能力,其结构如图 4 所示。

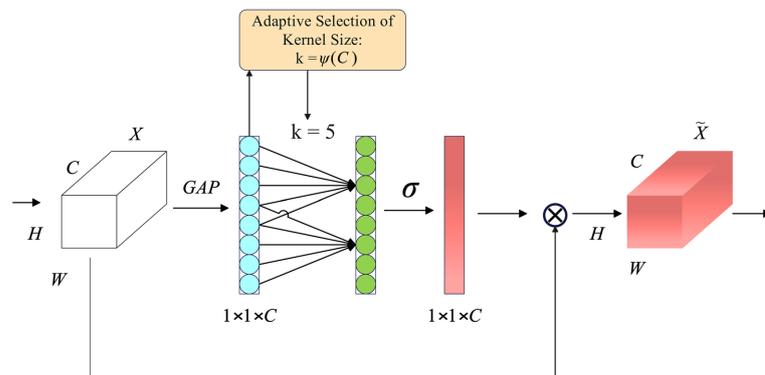


Figure 4. Structure of the ECA attention mechanism
图 4. ECA 注意力机制结构

3.4. 引入注意力机制

为进一步提升 YOLOv5 在小目标水稻害虫检测任务中的特征提取能力与检测精度, 本文将三种主流注意力机制——SE (Squeeze-and-Excitation)、CBAM (Convolutional Block Attention Module)与 ECA (Efficient Channel Attention)集成至 YOLOv5 的主干网络中, 构建改进型目标检测模型。在具体实现过程中, 本文分别定义了六个注意力模块函数: SE、C3SE、CBAM、C3CBAM、ECA 和 C3ECA。其中, SE、CBAM 和 ECA 为经典的通道/空间注意力机制模块, 直接作用于输入特征图; 而 C3SE、C3CBAM 与 C3ECA 模块则是在 YOLOv5 原有 C3 模块基础上嵌入对应注意力机制构建的新结构模块, 旨在增强深层特征表示能力。值得注意的是, 本文并非仅在局部结构中嵌入注意力机制, 而是系统性地将 YOLOv5 主干网络 (Backbone)和检测头(Head)中的所有 C3 模块, 全部替换为对应的 C3SE、C3CBAM 或 C3ECA 模块, 确保注意力机制贯穿整个特征提取与预测过程。这种全量替换式的模块替换策略, 相较于仅在主干中插入单一注意力模块, 更具一致性与结构完整性, 有助于模型在不同层级的特征图中均衡增强显著通道或关键空间区域的信息表达能力。同时, 模块替换操作保持了原有模型的整体拓扑结构与轻量级特性, 便于在资源受限设备上部署和加速推理。最终, 本文在 YOLOv5s 架构上分别构建三种改进模型: YOLOv5s-C3SE、YOLOv5s-C3CBAM 与 YOLOv5s-C3ECA, 以此为基础展开对比实验, 系统分析不同注意力机制对检测精度、模型规模与运算效率的影响。其改进的相关结构如图 5 所示。

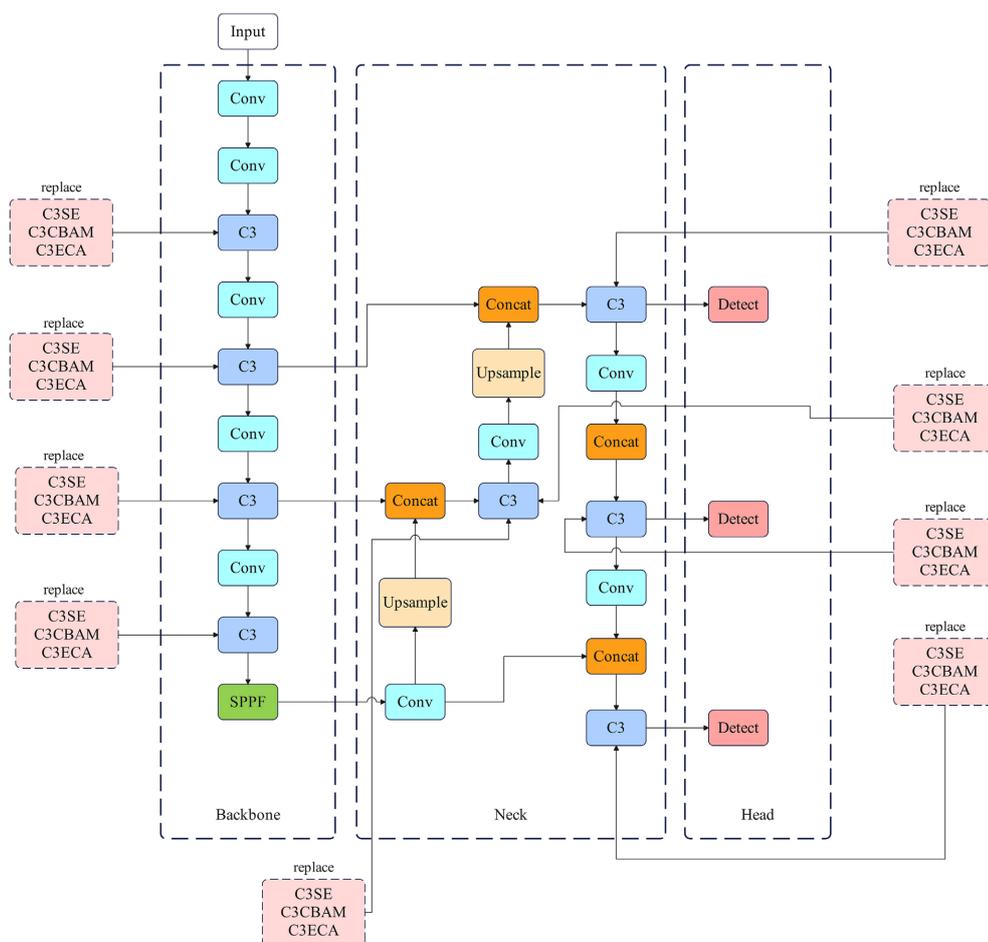


Figure 5. Schematic diagram of module replacement in the improved model

图 5. 改进模型的模块替换示意图

4. 实验结果与分析

4.1. 硬件及其参数设置

本次实验利用 PyTorch 深度学习框架以及 vscode 开源集成开发环境进行模型构建与训练。硬件方面, 采用 Intel Core i5-12400F 处理器、32GB RAM 内存, 并搭载 NVIDIA GeForce RTX 3090 GPU, 支持 CUDA 11.8 加速。软件方面采用 PyTorch 2.1.0 与 Python 3.10。完成了 YOLOv5s 原始模型与三种引入注意力机制的改进模型(YOLOv5s-C3ECA, YOLOv5s-C3CBAM, YOLOv5s-C3SE)的训练与测试工作, 确保了实验的高效性与模型在高性能计算设备下的训练稳定性。如表 1 所示:

Table 1. Configuration of experimental software and hardware equipment

表 1. 实验软硬件设备配置

CPU	RAM	GPU	Cuda	Pytorch	python
i5-12400f	32.0G	NVIDIAGeForceRTX3090	11.8	2.1.0	3.10

4.2. 评价指标

采用 $mAP@0.5$ 、准确率(P)、召回率(R)、 $mAP@0.5:0.95$ 来评价训练结果。本次实验采用 $mAP@0.5$ 、 $mAP@0.5:0.95$ 、准确率(Precision, P)和召回率(Recall, R)四项指标对模型的检测性能进行综合评估。其中, $mAP@0.5$ 反映了在 IoU 阈值为 0.5 时模型整体的平均精度, $mAP@0.5:0.95$ 更全面地考察了不同 IoU 下的检测能力, 能够体现模型对目标定位精度的要求。准确率用于衡量模型检测出的正样本中有多少是实际为正的, 召回率则表示模型成功检测出的正样本占有所有实际正样本的比例。通过这些指标的综合分析, 能够有效评估原始 YOLOv5s 模型及其改进版本(YOLOv5s-C3ECA, YOLOv5s-C3CBAM, YOLOv5s-C3SE)在水稻害虫目标检测任务中的检测精度、鲁棒性与稳定性。

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

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

$$mAP@0.5 = \frac{1}{N} \sum_{i=1}^N AP_i^{IoU=0.5} \quad (8)$$

$$mAP@0.5:0.95 = \frac{1}{10N} \sum_{i=1}^N \sum_{j=0}^9 AP_i^{IoU=0.5+0.05j} \quad (9)$$

4.3. 对比实验结果

本实验对比分析了原始 YOLOv5s 与 YOLOv5n 模型及其引入不同注意力机制(SE, CBAM, ECA)后的改进模型在稻田害虫检测任务中的性能表现, 评价指标包括: 准确率(Precision)、召回率(Recall)、 $mAP@0.5$ 以及 $mAP@0.5:0.95$ 。实验结果如下表 2 所示:

Table 2. Comparison of detection performance of different models

表 2. 不同模型的检测性能对比

Model	Precision	Recall	map@0.5	map@0.5:0.95
YOLOv5s	0.708	0.640	0.681	0.473
YOLOv5s-C3SE	0.716	0.625	0.671	0.467

续表

YOLOv5s-C3CBAM	0.732	0.636	0.693	0.476
YOLOv5s-C3ECA	0.769	0.653	0.706	0.484
YOLOv5n	0.654	0.589	0.606	0.392
YOLOv5n-C3SE	0.563	0.588	0.585	0.384
YOLOv5n-C3CBAM	0.588	0.566	0.577	0.370
YOLOv5n-C3ECA	0.633	0.588	0.609	0.394

图 6 中展示了原始 YOLOv5s 模型、三种引入注意力机制的改进模型(YOLOv5s-C3SE, YOLOv5s-C3CBAM, YOLOv5s-C3ECA)以及轻量化模型 YOLOv5n 及其相应变体在相同图像上的检测结果。根据表 2 可以看出,各模型在水稻害虫检测任务中的表现存在差异。原始 YOLOv5s 模型能够较为准确地检测出大部分害虫目标,但在部分区域仍存在漏检或置信度较低的情况。改进模型中, YOLOv5s-C3ECA 表现最为突出,检测框更为密集,置信度更高,能够识别出更多微小目标,表明 ECA 注意力机制在增强特征通道表达方面具有明显优势。YOLOv5s-C3CBAM 在目标定位与置信度方面也有一定提升,整体表现优于原始模型。相比之下, YOLOv5n 系列虽然具有更快的推理速度和较低的计算开销,但检测出的目标数量较少,漏检现象更为明显,尤其是在复杂背景下对小目标的检测精度较低,难以满足对精度要求较高的应用场景。

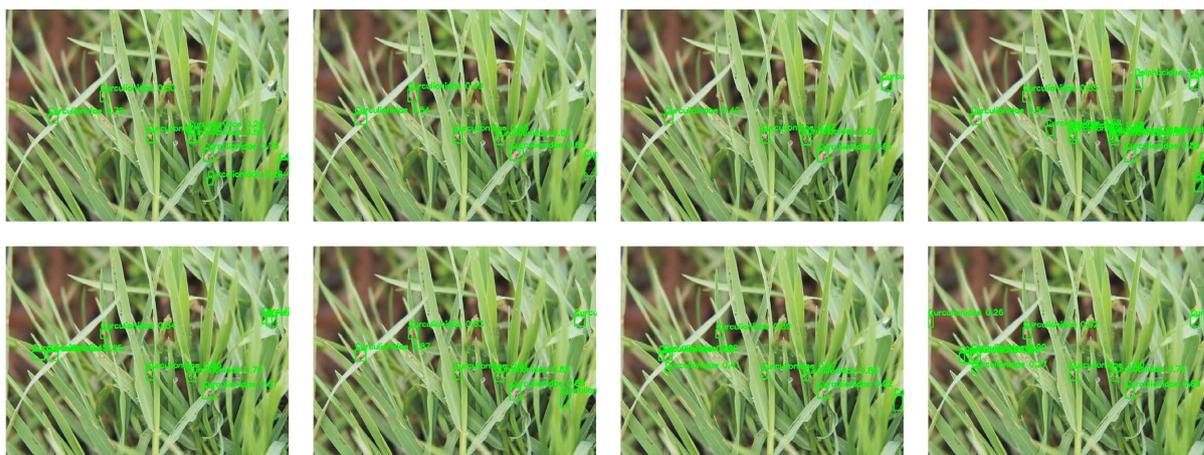


Figure 6. Visual comparison of detection results of different models

图 6. 不同模型的检测结果可视化对比

5. 讨论

本文围绕水稻害虫小目标检测任务,针对 YOLOv5 模型在移动端部署中的各种问题,提出了一种基于注意力机制的改进方法。通过将三种主流注意力机制——SE、CBAM 和 ECA,分别嵌入 YOLOv5 中的所有 C3 模块,构建了三种改进型检测模型:YOLOv5s-C3SE、YOLOv5s-C3CBAM 与 YOLOv5s-C3ECA,并与原始模型 YOLOv5s 以及更小的 YOLOv5n 系列模型进行了系统对比实验。其中, YOLOv5s-C3ECA 在在 mAP@0.5 和 mAP@0.5:0.95 指标上取得最优结果。与 SE 和 CBAM 两种注意力机制相比, ECA 注意力机制采取无全连接层的局部跨通道交互策略,参数量和计算量更小,运算速度更快,通过一维卷积动态分配通道权重,有效降低了参数量与计算开销,无需手动调参。同时避免了类如 SE 机制中全连接层

可能导致的信息损失。对于水稻害虫这类尺寸较小、形态多样的目标，ECA 能够更精准地保留特征图中细粒度的局部信息，从而提升检测性能。此外，稻田场景中存在大量背景干扰与密集小目标，ECA 在保留关键特征的同时减少了模型对无效信息的关注，有助于模型在实际应用中的泛化能力与稳定性。本研究主要采取了当前主流的三种注意力机制，并未涵盖如 Transformer 等更复杂的特征增强技术，且数据集主要为常见的十几种水稻害虫，未根据季节，地区等因素进行细致划分。未来将构建更大规模，更多样性的水稻害虫数据集，并对该模型融合特征增强技术，进一步提升模型的性能，对设备进行部署与实时检测优化，展现模型的实际运行性能。

6. 结论

本实验结果表明，大部分引入注意力机制的改进模型在精度与性能方面均有所提升，尤其为 YOLOv5s-C3ECA 模型，在 $mAP@0.5$ 和 $mAP@0.5:0.95$ 上分别达到 0.706 与 0.484，较原始 YOLOv5s 分别提升 2.5% 和 1.1%，在检测精度与稳定性方面均表现最优。同时，YOLOv5s-C3CBAM 在准确率和整体检测效果方面也具有较好表现，验证了 CBAM 模块在通道与空间特征联合建模上的有效性。此外，在轻量化版本 YOLOv5n 上的实验也显示，注意力机制在小模型中同样具备一定的性能提升作用，C3ECA 模块在 YOLOv5n 中表现相对优越，说明该机制在低计算资源条件下仍具备较强的特征建模能力。综上所述，本文所提出的 C3 注意力模块替换策略不仅提升了 YOLOv5 模型对水稻害虫小目标的检测精度和鲁棒性，也具备良好的轻量化部署适应性，为农业智能化场景中病虫害快速检测提供了技术支持与工程参考。

参考文献

- [1] Qi, L., Zhang, T., Zeng, J., *et al.* (2021) Analysis of the Occurrence of Main Diseases in Five Major Rice-Producing Areas in China in Recent Years. *China Plant Protection Guide*, **41**, 37-42, 65.
- [2] Ju, Z.Y., Yi, C., Zhou, Z.C., *et al.* (2024) YOLO-Rice: A Rice Pest Detection Method Based on YOLOv5. *Control Engineering in China*, **156**, 124-136.
- [3] Zhang, S., Wu, X., You, Z. and Zhang, L. (2017) Leaf Image Based Cucumber Disease Recognition Using Sparse Representation Classification. *Computers and Electronics in Agriculture*, **134**, 135-141. <https://doi.org/10.1016/j.compag.2017.01.014>
- [4] Sun, Y., Jiang, Z., Zhang, L., Dong, W. and Rao, Y. (2019) SLIC_SVM Based Leaf Diseases Saliency Map Extraction of Tea Plant. *Computers and Electronics in Agriculture*, **157**, 102-109. <https://doi.org/10.1016/j.compag.2018.12.042>
- [5] Sethy, P.K., Barpanda, N.K., Rath, A.K. and Behera, S.K. (2020) Deep Feature Based Rice Leaf Disease Identification Using Support Vector Machine. *Computers and Electronics in Agriculture*, **175**, Article ID: 105527. <https://doi.org/10.1016/j.compag.2020.105527>