

基于改进 YOLOv7 的玉米作物害虫检测研究

宫 妍, 翟俊杰, 王 凯, 李 玉

(哈尔滨商业大学 轻工学院, 哈尔滨 150028)

摘要: 面对玉米作物害虫检测中目标体积较小、形态多变且种类分布不均的情况, 现有检测器会出现误检、漏检等问题; 针对以上问题, 提出了基于 YOLOv7 的玉米作物害虫检测算法 SPD-YOLOv7; 制作收集玉米害虫数据集, 采用数据增强方法扩充数据集; 引入 SPD-Conv 模块, 替换原先骨干和头部网络中的部分跨步卷积层, 减少随着网络加深细节信息的丢失, 提高模型获取小目标特征和位置信息的能力; 将 ELAN-W 模块与 CBAM 注意力机制结合, 使网络更好地学习害虫特征, 抑制背景信息, 关注目标本身; 改进后的 YOLOv7 网络模型准确率达到了 98.38%, 平均精度均值达到了 99.4%; 相较于原始的 YOLOv7 模型, 准确性和平均精度均值分别提高了 2.46、3.19 个百分点, 与 Faster-RCNN、YOLOv3、YOLOv4、YOLOv5 和 YOLOv6 主流算法的检测精度相比更具优势, 且满足实时性; 实验结果说明改进算法有利于快速识别玉米作物的虫害分布, 可用于实际农田间的害虫实时监测。

关键词: YOLOv7; SPD-Conv 模块; CBAM 注意力机制; 虫害检测; 小目标检测

Study on Corn Crop Pest Detection Based On Improved YOLOv7

GONG Yan, ZHAI Junjie, WANG Kai, LI Yu

(College of Light Industry, Harbin University of Commerce, Harbin 150028, China)

Abstract: Faced with the challenges of small target volumes, diverse morphologies, and uneven distributions of pests in maize crop pest detection, existing detectors suffer from issues such as false positives and false negatives. In response to above challenges, a maize crop pest detection algorithm SPD-YOLOv7 based on YOLOv7 is proposed. The maize pest dataset is created and collected, and the data augmentation technique is used to enlarge the data set. The SPD-Conv module is introduced to replace some step convolution layers in the original backbone and head networks, and mitigate the loss of detailed information as the network deepens, thereby enhancing the model's ability to capture the features and positional information of small targets. By integrating the ELAN-W module with the CBAM attention mechanism, it enables the network to better learn pest features, suppress background noise, and focus on the target itself. The improved YOLOv7 network achieves an accuracy of 98.38% and mean average precision of 99.4%. Compared to the original YOLOv7 model, the accuracy and mean average precision improve by 2.46 and 3.19%, respectively. The enhanced algorithm is superior to mainstream algorithms such as Faster-RCNN, YOLOv3, YOLOv4, YOLOv5, and YOLOv6 in the detection accuracy, while maintaining real-time performance. Experimental results indicate that the proposed algorithm has the rapid identification of maize crop pest distributions and can be applied for real-time pest monitoring in practical agricultural fields.

Keywords: YOLOv7; SPD-Conv module; CBAM attention mechanism; pest detection; small target detection

0 引言

作为农业大国, 玉米是我国重要的粮食作物之一。其产量的高低直接关系到国民经济和社会民生^[1]。玉米作物害虫, 如玉米蚜虫、草地贪夜蛾、扁盾蝽、稻蝗、双斑萤叶甲, 是威胁玉米粮食安全的主要因素^[2]。据文献^[3]研究表明: 在未采取药物防治措施的情况下, 草地贪夜蛾对玉米产量造成的损失高达 48.35%, 果穗受害率更是高达 98.91%。双斑萤叶甲的侵害会导致玉米田地大面积减产, 减产率可达 10%~30%^[4]。传统的玉米虫害监测主要依赖于专业农业人员和种植人员对农业知识的熟练运用和经验的积累^[5]。人工识别过程繁琐复杂、耗时耗力, 可以在小规模农田实现监测, 但难以满足大规模和快速检测虫害的

实际生产需求。

随着人工智能技术的不断发展, 机器学习与深度学习方法在植物病虫害检测领域已被广泛应用^[6-7]。例如, 文献^[8]采用图像分割, 特征提取和分类器设计的技术路线, 提出了一种基于图像特征提取的虫害识别方法。文献^[9]针对玉米象虫进行研究, 运用基于 K-means 硬聚类算法的图像分割技术, 成功提取了玉米象虫的周长、面积、复杂度等 3 个关键特征进行研究。深度学习方法通过网络训练提取目标的多尺度特征, 有效提升了模型精度和泛化能力。在植物害虫检测领域, 这种方法表现出卓越的效果^[10-12]。当前主流的目标检测网络包括 SSD^[13]、Faster R-CNN^[14-15] (Faster region with CNN) 系列和 YOLO^[16] (You only look

收稿日期: 2024-02-27; 修回日期: 2024-03-29。

基金项目: 黑龙江省哲学社会科学规划项目(23YSD245); 哈尔滨商业大学博士启动项目(2019DS087)。

作者简介: 宫 妍(1981-), 女, 工学博士, 讲师。

引用格式: 宫 妍, 翟俊杰, 王 凯, 等. 基于改进 YOLOv7 的玉米作物害虫检测研究[J]. 计算机测量与控制, 2024, 32(9): 58-65.

once) 系列等。文献 [17] 提出了改进的 SSD 的水稻虫害识别方法, 其平均精度均值达到 79.3%。文献 [18] 基于 Fast R-CNN 模型和草地贪夜蛾对玉米叶片的摄食痕迹, 提出了一种名为 Pest Region-CNN 的端到端目标检测模型。文献 [19] 构建了 YOLOv4-Corn 模型, 该模型有效地解决了由于玉米害虫体积小、容易重叠等因素导致的不易识别问题, 但检测精度仍有提升空间。

本文针对玉米害虫体积小、相似度高导致检测结果不佳的问题, 提出了一种基于改进的 YOLOv7 算法。本文依托北大荒哈尔滨智慧农业产业园项目构建了一个包含常见玉米作物害虫的数据集。以该数据集为基础, 使用 SPD-Conv 模块^[20] 替换传统算法中的代步卷积, 减少了检测目标的细节损失, 进一步提高了对小目标昆虫的检测精度, 增强了网络的特征提取能力。引入混合注意力机制 CBAM^[21] 与 ELAN-W 模块结合, 使神经网络分别在通道和空间维度上扩大感受野, 捕捉到更多的特征信息, 灵活有效地提高特征信息表示能力。改进后的模型能够快速而准确地复杂的自然环境中检测玉米作物害虫, 为玉米虫害检测提供了有效的技术支持。

1 YOLOv7 网络结构

YOLOv7^[22] 是 YOLO 系列中近两年推出的网络模型。该模型是目前在 PASCAL VOC 数据集^[23] 上具有最快推理速度和最佳识别效果的 YOLO 模型。它在速度和准确性方面超越了先前的版本和其他目标检测模型, 能够满足实时准确检测自然条件下玉米作物害虫的需求。然而, 尽管原始的 YOLOv7 算法已经实现了较高的检测精度, 但针对降低相似特征、目标尺度和形状变化对检测精度的影响, 仍有提升空间。因此, 本文基于 YOLOv7 作为基础检测模型, 并对其进行了改进。YOLOv7 网络架构如图 1 所示。

YOLOv7 网络架构由 3 个核心部分组成: 输入端 (Input)、主干网络 (Backbone) 和检测头网络 (Head)。输入端通常处理 $640 \times 640 \times 3$ 尺寸的图像, 在经过预处理后送入主干网络。该主干网络基于 YOLOv5 架构, 融入了 Efficient Layer Aggregation Networks (ELAN) 结构和 Max Pooling 1 (MP1) 结构, 同时结合 CBS 模块以提取输入图像的特征。检测头网络则包括 Spatial Pyramid Pooling 和 Cross Stage Partial Channel (SPPCSPC) 结构、ELAN-H

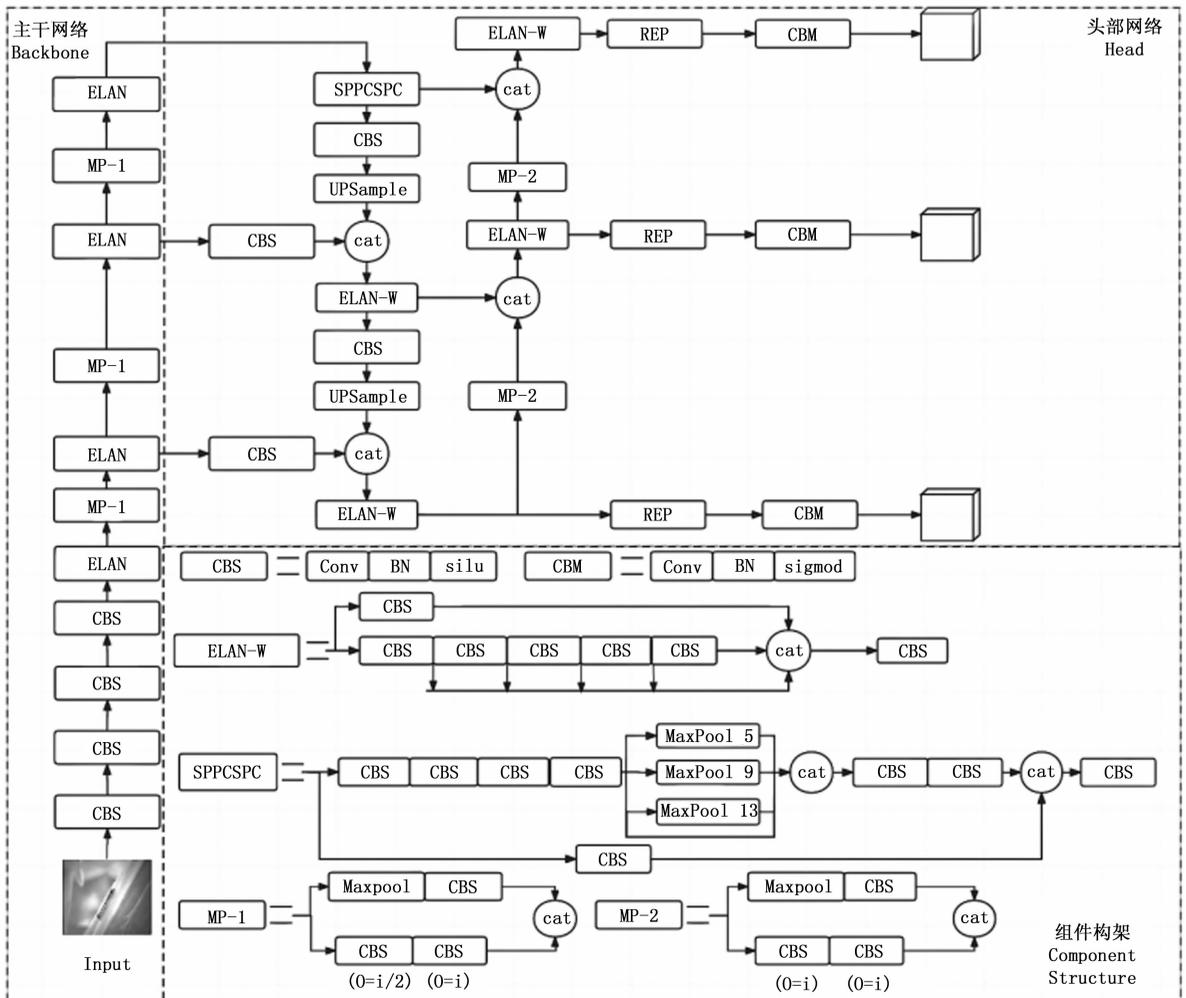


图 1 YOLOv7 网络结构

和 MP2 特征提取结构, 以及 RepConv 结构。主干网络输出的 3 个特征层在检测头网络中进一步训练, 整合这些输出以实现目标的多尺度检测^[24], 从而输出 3 个不同尺寸的预测结果。

2 YOLOv7 检测算法改进

2.1 SPD-Conv (Space To Depth) 模块

SPD-Conv 模块最早由 Raja 等人提出, 其主要由 space-to-depth (SPD) 和 non-strided convolution 层两部分组成的卷积操作。它被用来代替传统的跨步卷积, 以减少在小物体检测中因占据小像素比例而导致的细节信息损失。通过使用 SPD-Conv, 可以有效地提高小物体检测的准确性, 并保留更多的细节信息。

SPD-Conv 的操作流程如图 2 所示。尽管特征图经过下采样, 但仍保留了所有通道维度的信息。具体步骤包括将特征图沿着通道方向分割 N 次, 每个子特征图的分割后长度和宽度等于 S/N , 同时通道数保持不变, 并将子特征图沿通道方向合并到特征图 X_1 中, 即按照式 (1) 进行操作。这一过程有助于有效地保留特征图的信息并提高模型的性能。

$$X(S \times S \times C_1) \rightarrow X_1\left(\frac{S}{N} \times \frac{S}{N} \times N^2 C_1\right) \quad (1)$$

使用非跨步卷积层 (步幅=1) 对特征图进行卷积, 以获得特征, 尽可能保留更多的特征信息, 即式 (2):

$$X_1\left(\frac{S}{N} \times \frac{S}{N} \times N^2 C_1\right) \rightarrow X_2\left(\frac{S}{N} \times \frac{S}{N} \times C_2\right) \quad (2)$$

其中: $C_2 < N^2 C_1$ 。这种操作保留了判别特征信息的同时, 还具有向下采样特征图的优点。

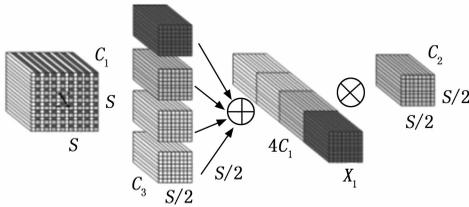


图 2 SPD-Conv 结构

在本文中, 为了应对跨步卷积导致的细节损失等问题^[25], MP-1 和 MP-2 模块在 Backbone 和 Head 中都经过了修改。具体来说, 最大池化层后的 CBS 卷积被替换为 SPD-Conv 模块, 这一改动可以在图 3 中清晰地展示出来。与此同时, 本文还对 Backbone 中输入端的第三个 CBS 卷积进行了替换, 采用了 SPD-Conv 模块。通过这些改进, 模型的性能得到了有效提高, 同时也成功避免了细节丢失的问题, 使得模型在小物体检测等任务中表现更加出色。

2.2 CBAM 模块

CBAM (Convolutional Block Attention Module) 是一种轻量且高效的注意力模块, 如图 4 (a) 所示。CBAM 注意力机制由通道注意力和空间注意力两个模块组成。在接收到中间特征图后, 该模块分别对通道和空间维度进行处理。经过处理后的特征图与原始特征图融合, 从而实现了

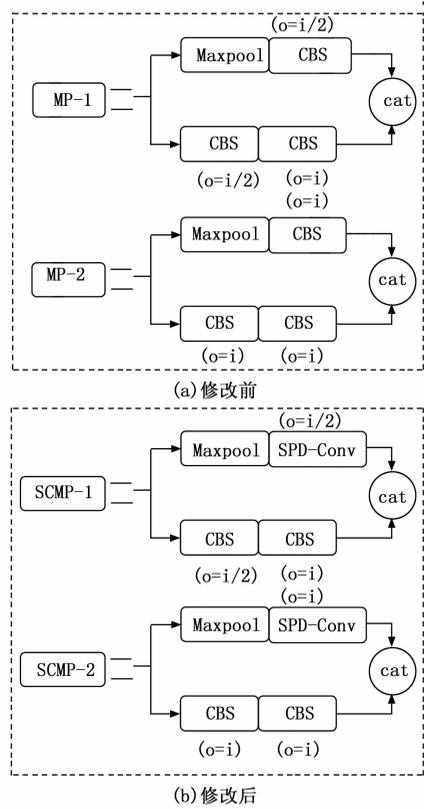


图 3 MP-1, MP-2 模块改进前后

特征的自适应优化。CBAM 模块的引入可以有效提升模型的表现, 并在各种任务中展现出良好的性能。

通道注意力机制^[26]如图 4 (b) 所示, 首先, 输入的特征图通过全局平均池化 (GAP) 和全局最大池化 (GMP) 进行处理, 以获取局部和全局信息。接着, 这些特征图经过共享感知机 (MLP) 的升维和降维处理后, 融合为一个特征图。最后, 通过激活函数得到通道注意力机制。这种处理方式可以使 CBAM 模块更好地关注重要的特征, 提高模型对目标的检测能力和鲁棒性。

空间注意力机制如图 4 (c) 所示。输入的特征图同样经过全局平均池化和全局最大池化操作, 得到两种不同的通道特征描述算子, 并将它们进行拼接。随后, 这些特征图通过一个 7×7 大小的卷积层进行处理, 最终通过 sigmoid 激活函数处理, 得到空间注意力向量。这种处理方式有助于模型更好地理解 and 利用特征图中的空间信息, 从而提升检测和分类任务的性能。

为了在复杂环境中及时检测小目标害虫, 本文对 ELAN-W 模块进行了改进。将其最后一个卷积层替换成了 CBAM 注意力机制, 更有助于提取害虫的关键局部细节特征, 从而显著增强了神经网络的学习和表达能力。通过该改进, 能够更准确地检测到小目标害虫, 提高了模型在实际应用中的鲁棒性和准确性。

2.3 SPD-YOLOV7 目标检测模型

卷积神经网络在网络层数加深时, 由于采用了卷积步

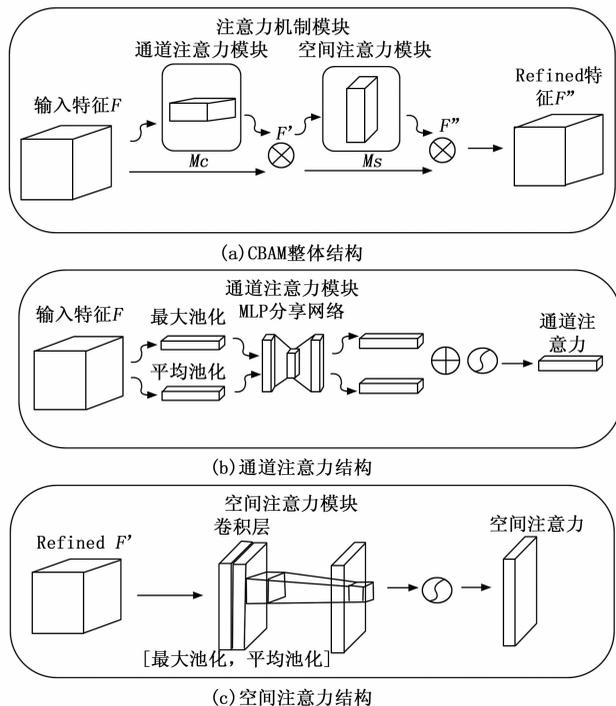


图 4 CBAM 结构

长和池化层, 容易导致小目标害虫的细节信息丢失。为了更充分地提取有效特征, 本文对 YOLOv7 的 Backbone 和 Head 中的 MP-1 和 MP-2 结构进行了改进, 将 CBS 模块替换为 SPD-Conv 模块。这一改进显著提升了小目标害虫的检测精度。此外, 为了增强玉米作物害虫的特征提取和学习能力, 本文将 ELAN-W 模块的最后一个卷积层替换为 CBAM 注意力机制。通过在通道和空间两个维度上调节注意力, 突出了玉米作物害虫的关键特征信息, 如颜色、形状和纹理, 并抑制了冗余特征, 确保准确提取害虫图像中的关键特征。改进后的 SPD-YOLOv7 模型如图 5 所示, 其中黑色虚线框表示本模型的改进部分。这些改进使得网络更适应复杂环境下的目标检测任务, 提高了模型的鲁棒性和准确性。

3 实验设置

3.1 实验数据集

本研究的实验数据来源于位于黑龙江省北大荒哈尔滨智慧农业产业园的地点。为确保数据的真实性和普遍性, 采集地点选在大约 1 亩的试验田进行数据采集。拍照设备采用了佳能 D600 型数码相机, 使用了 35 mm~135 mm 的中长焦镜头和 100 mm 的微距镜头对玉米作物害虫进行近距离拍摄。数据采集时间为 2023 年 6 月中旬至 8 月初, 主要涵盖了蝗虫、草地贪夜蛾和双斑萤叶甲等 3 种主要的玉米害虫。在数据筛选过程中, 剔除了因抖动、过度曝光等问题而导致成像质量差的照片, 最终形成了包含 1 340 张照片的数据集, 其中每种昆虫约有 400 多张样本。图 6 展示了样本实例, 第 1 至第 3 行分别展示了蝗虫、草地贪夜蛾和双斑

萤叶甲的样本。通过这些精心采集和筛选的数据, 可以对玉米害虫进行准确的识别和分析。

3.2 数据集图像处理与数据增强

为了缩短训练时长、提升效率, 本文采用了压缩和归一化处理对图像进行处理, 以适配 YOLOv7 网络模型所需的显存和计算资源。原始玉米害虫数据集中的图像分辨率为 5 472 像素×3 648 像素, 经过统一压缩处理后变为 1 080 像素×720 像素。在数据标注方面, 本文使用了 Labelimg 数据标注工具对这 3 种害虫进行了标注, 标签格式采用了 VOC, 并以 xml 文件格式保存。为了与 YOLO 模型兼容, 使用 python 转换脚本将标签格式转化为 YOLO 标签格式, 并以 txt 文件格式保存。

在数据采集过程中, 尽管本文尽量保证玉米 3 种作物害虫样本数量的平衡, 但不同生长阶段的害虫样本数量仍存在较大差距。例如, 草地贪夜蛾成虫样本仅约 100 张, 而幼虫样本却多达 380 张。为了模拟真实的玉米害虫生长环境, 并解决样本数据不均衡导致模型性能下降的问题, 本文进行了进一步的处理。本研究采用了多种数据增强方法来扩充样本数据集^[27]。首先通过加入高斯噪声的方法对图片进行处理, 其次分别使用暗处理和亮处理的方法来模拟不同光照条件下的玉米害虫图像。这些增强方法使得害虫图像总数达到 3 000 张。最后按照 8: 1: 1 的比例划分训练集, 验证集和测试集。

3.3 实验环境及超参数设置

实验环境: 实验平台采用 Windows11 操作系统, CPU 处理器为 AMD Ryzen 7 7735H with Radeon Graphics, GPU 为 NVIDIA GeForce RTX 4060 (8G), 深度学习框架为 Pytorch-1.11.0, 编程语言为 Python-3.8.18。集成开发环境为 PyCharm2022.3.3。实验训练中超参数设置如表 1 所示。

表 1 实验超参数设置

参数	参数值
学习率	0.01
动量	0.937
权重	0.000 5
迭代周期	300
批量尺寸	16
图像尺寸	640×640
学习率衰减方法	余弦退火算法

3.4 模型训练评价指标

本研究使用了目标检测中常用的评价指标, 包括准确率 (precision, P)、召回率 (recall, R)、三类玉米作物害虫平均准确率 (mAP, mean average precision) 以及检测速度 (FPS, frames per second) 和损失函数 (Loss Function), 来全面评估模型在性能上的表现。这些指标能够客观地度量模型在目标检测任务中的精确度、召回率、多类别平均准确率以及实时检测速度。

P 表示模型正确识别的目标总数与总识别目标 (包括

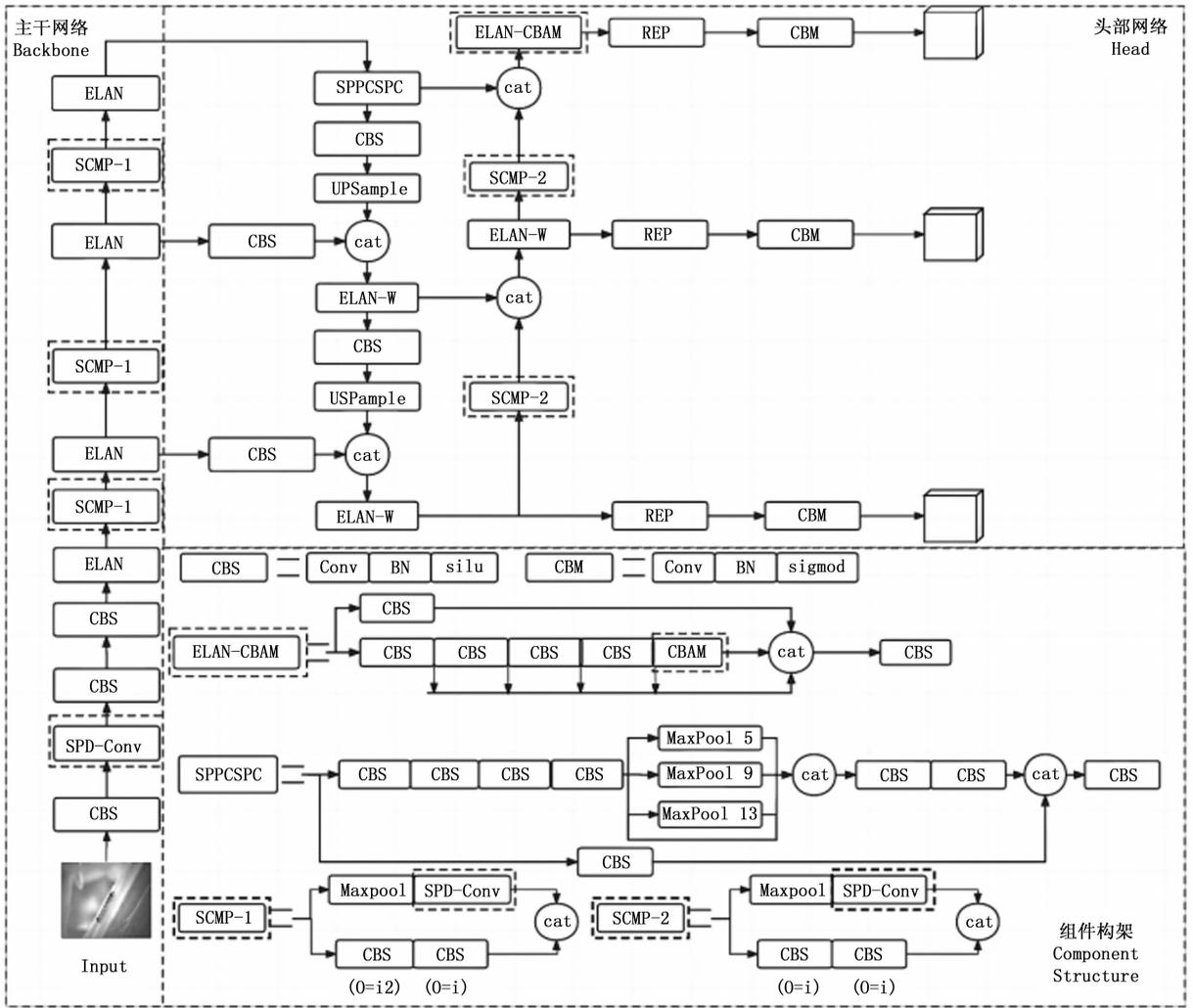


图 5 SPD-YOLOv7 网络模型

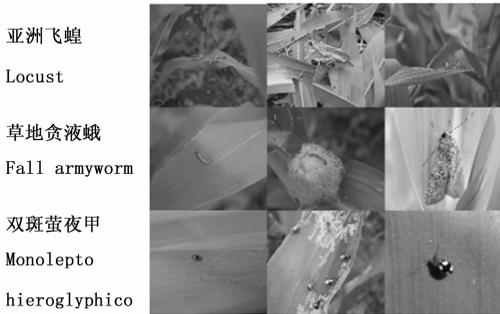


图 6 玉米害虫数据集

正确和错误识别的) 之间比例。其数学表达式表示为:

$$P = \frac{TP}{TP + FP} \quad (3)$$

式中, TP 表示成功检测到的玉米害虫图片数量, FP 表示检测错误的害虫的图片数量。

R 表示模型检测到的正确目标与实际存在的所有目标 (包括未被模型检测到的) 之间的比例。其数学表达式表示为:

$$R = \frac{TP}{TP + FN} \quad (4)$$

式中, FN 表示没有检测到的玉米作物害虫图片数量。

mAP 是指对不同召回率阈值下的精度进行计算并对这些精度取平均来评估模型在不同召回率水平上的表现。它是一种常用的评价指标, 用于客观地评估目标检测模型的性能。 mAP 的数学表达式表示为:

$$mAP = \frac{1}{N} \sum_{k=i}^N AP(k) \quad (5)$$

式中, N 为玉米作物害虫类别的数量, k 为阈值, $AP(k)$ 为检测到的第 k 类害虫的 AP 值。

FPS 为模型每秒可以处理或者检测害虫图像的帧数, 用来衡量模型的检测速度。

本文使用分类损失来衡量模型的稳定性能, L_{CE} 表示二分类损失, 一般可通过 sigmoid 或者

softmax 得到, 本文通过 sigmoid 激活函数得到二分类损失, L_{CE} 数学表达式表示为:

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C g_i^c \log s_i^c \quad (6)$$

式中, N 是采集的图片样本数量, g 是检测到的所含害虫的

样本个体, c 是检测错误的玉米作物害虫图片数量。

4 实验结果与分析

4.1 改进前后实验对比分析

改进前后的训练结果如图 7 所示。从图 7 (a) 可以看出, 在基本模型训练参数相同的情况下, 随着训练轮数的不断增加, YOLOv7 和 SPD-YOLOv7 模型的损失值均逐渐减小, 表明模型在训练过程中没有出现过拟合现象。这两个模型在经过 300 轮训练后达到了最佳状态。此时, SPD-YOLOv7 的损失值比 YOLOv7 模型低 0.008。而且, SPD-YOLOv7 网络模型的置信度损失收敛^[28]速度更快, 并且始终低于 YOLOv7 的置信度损失。从图 7 (b) ~ 7 (d) 可以看出改进后的 SPD-YOLOv7 模型的准确率、召回率和 mAP 分别提高了 2.9 个百分点、6.8 个百分点和 3.19 个百分点, 训练过程曲线更加平稳。损失值是衡量模型训练是否优异的一个重要指标, 损失值越低说明预测框与真实框之间差别越小, 检测效果就越好。这证明使用 SPD-YOLOv7 模型可以更好地提高检测性能。

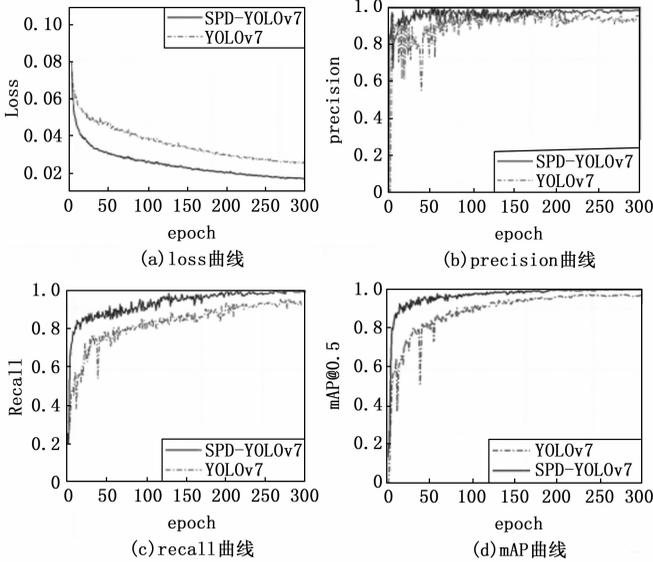


图 7 实验结果曲线图

为了直观地描述和分析每个类别的分类情况, 绘制了混淆矩阵, 如图 8 所示。在混淆矩阵中, 每行代表预测的类别, 每列代表实际的类别, 而对角线上的数据表示正确分类的比例。通过观察图中的结果, 可以看出大部分目标都被正确预测, 这表明该模型具有良好的性能。

4.2 检测结果对比

为直观展示模型检测性能, 分别用 YOLOv7 与 SPD-YOLOv7 对相同的害虫图像进行检测, 由图 9 可以看出改进后 SPD-YOLOv7 模型减少了原模型误检, 漏检的问题, 准确率方面 SPD-YOLOv7 模型要比原始模型的检测的精度高。实验证明改进的 SPD-YOLOv7 模型检测效果更好, 检测能力更强。

4.3 消融实验结果分析

为验证改进部分对玉米害虫检测模型的影响, 采用消

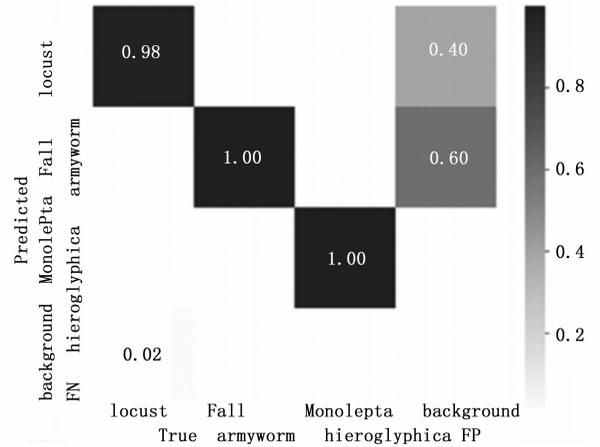


图 8 混淆矩阵图

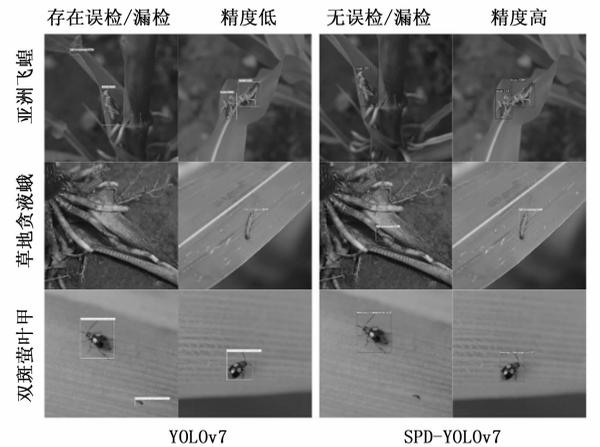


图 9 检测效果图

融实验方法验证各改进策略的有效性。本研究共设计 4 组不同模型的消融实验, 实验结果如表 2 所示。

表 2 消融实验性能对比

Test No.	model	P	R	mAP	FPS
1	YOLOv7	95.92	92.27	96.21	71.7
2	YOLOv7-SPD	98.50	94.32	97.89	68.3
3	YOLOv7-CBAM	96.41	98.23	99.20	72.0
4	YOLOv7-SPD-CBAM	98.38	99.51	99.40	69.0

从表 2 的数据可以看出, 实验 1 采用了原始的 YOLOv7 模型, 在玉米害虫识别方面取得了较好的性能。准确率为 95.92%, 召回率为 92.27%, mAP 为 96.21%, FPS 为 71.7。在实验 2 中, 本文在初始模型的基础上, 分别在 Backbone 和 Head 中添加了 SPD_Conv 模块。结果显示, 相较于实验 1, mAP 提升了 1.68 个百分点, 达到了 98.50%, 准确率提升了 2.58 个百分点。这说明 SPD_Conv 模块能够减少细节信息的损失, 丰富小目标物体的特征信息。在实验 3 中, 本文在初始模型中引入了 CBAM 注意力机制。结果显示, 相较于实验 1, mAP 提升了 2.99 个

百分点，达到了 99.20%，召回率提升了 3.91 个百分点。这表明 CBAM 注意力机制能够在通道和空间两个维度上分配注意力，提升了对小目标害虫的检测的召回率和准确率。实验 4 综合了实验 2 和实验 3 的改进方法，虽然检测速度略有降低，但准确率、召回率和 *mAP* 分别提升了 2.46、7.24 和 3.19 个百分点。因此，本文提出的 SPD-YOLOv7 模型在综合性能上表现最优，准确率达到 98.38%，召回率达到了 99.51%，*mAP* 达到了 99.40%，FPS 为 69.5。通过消融实验证明，本研究对 YOLOv7 目标检测模型的改进具有积极的意义。

4.4 主流目标检测算法对比实验

为了进一步验证所改进的 YOLOv7 算法性能的优越性。在完全相同的数据集和配置环境下，将改进后的 YOLOv7 算法与几种主流目标检测算法（包括 YOLOv5、YOLOv4、YOLOv3 和 Faster R-CNN）进行了比较。通过对这些算法在各种评价指标上的表现，证明了改进的 YOLOv7 算法在目标检测方面相较于其他主流算法具有明显的性能优势。实验结果如表 3 所示。

表 3 主流目标检测算法的结果对比 %

model	<i>P</i>	<i>R</i>	<i>mAP</i>	FPS
Faster-RCNN	93.48	83.89	91.07	46
YOLOv3	88.97	81.84	92.33	59
YOLOv4	92.06	84.14	90.84	65
YOLOv5	94.60	90.19	95.25	74
YOLOv6	94.26	87.73	94.87	98
YOLOv7	95.92	92.27	96.21	71.7
SPD-YOLOv7	98.38	99.51	99.40	69.0

从表 3 的数据可以看出，相比于 Faster-RCNN、YOLOv3、YOLOv4 和 YOLOv5，改进后的 YOLOv7 算法在准确率、召回率和平均精度均值方面表现出更为突出的检测性能。虽然改进后的 YOLOv7 在检测速度上略低于 YOLOv5，但仍然具有较高的性能。与其他主流模型相比，改进后的 YOLOv7 在准确度上提升了 3.78~9.41 个百分点，在召回率上提升了 9.32~17.67 个百分点，在平均精度均值上提升了 4.15~8.56 个百分点。这些结果进一步证明了改进后的 YOLOv7 在目标检测方面的优越性。

4.5 模型实用性分析

为了验证所改进的 YOLOv7 算法的实用性，本文首先从参数量角度将改进的 YOLOv7 与其他算法参数量进行对比。

表 4 主流目标检测算法的参数量对比

model	Parameter quantity/M	Model_Size /Mb
Faster-RCNN	3.8	54.3
YOLOv3	6.0	25.6
YOLOv4	5.6	28.3
YOLOv5	5.8	14.4
SPD-YOLOv7	3.6	12.8

从表 4 的参数量对比分析可以看出，本研究使用的改进的 YOLOv7 模型的参数量较小，相比其他几个模型来说更为紧凑。同时，该模型的模型大小也较低，这使得它在存储和传输方面更加高效。尽管改进的 YOLOv7 模型参数较少，但仍然对硬件要求较高。为了实现模型的可移植性，可以先在使用的数据集上进行预训练，然后将预训练的模型应用到 app 中进行运行。这种方法可以有效地降低硬件要求，并保证模型在不同设备上的适用性。如图 10 所示，将模型移植到手机设备上，并使用该嵌入式设备实现了对玉米作物害虫的精准检测。

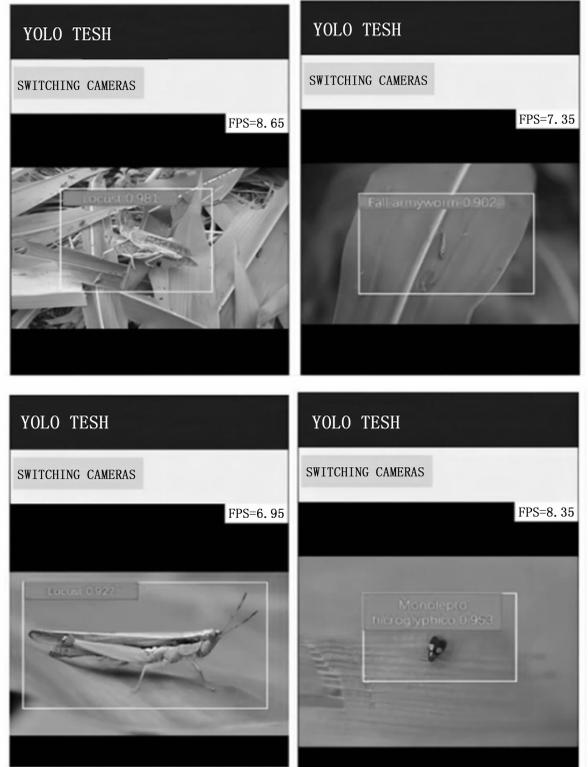


图 10 手机 app 检测效果图

图 10 显示的图片为高配置手机客户端的显示结果，可以观察到系统在手机 app 中使用预训练模型进行目标检测的结果，与在硬件条件充足的情况下相比，实际的检测准确度稍低。然而，手机 app 上运行模型的轻量化实现了模型的高效性。在今后的研究中，可以进一步对模型进行轻量化处理，包括剪枝、量化和使用 ROI 池化层来加速计算速度。同时，为了提高准确率，可以将模型部署在具有高配置服务器的云端，将采集到的图像数据上传到云端进行目标检测，并将最终分类结果返回到手机端，以适应不同配置的嵌入式手机设备。这种方式可以更方便地将本研究应用于实际农业生产中。

5 结束语

针对玉米害虫检测中存在体积小，分辨率较低极易出现漏检，误检，以及同一物种不同生长时期的差异大等诸

多问题, 本研究提出了基于 YOLOv7 改进的玉米害虫检测算法。该算法首先采用了轻量级的混合注意力机制 CBAM, 以提升模型性能, 弥补了过去玉米害虫识别精度低、实时性差等缺陷, 增强了模型在复杂环境下提取小目标特征的能力。其次为了模拟玉米害虫实际生活中的场景, 对实地拍摄的数据集进行亮度调整处理, 增加了训练难度以获得更好的训练效果。最后引入 SPD-Conv 模块, 以替换代步卷积, 从而减少细节信息的损失, 并且丰富了小目标物体的特征信息, 提高了小目标害虫的检测精度。

经过改进的 YOLOv7 模型在玉米害虫检测中的准确率达到了 98.38%, 召回率达到了 99.51%, $mAP@0.5$ 值达到了 99.4%。与原始的 YOLOv7 模型相比, 准确性、召回率和 $mAP@0.5$ 分别提高了 2.46%, 7.24% 和 3.19%。这意味着在保持检测速度的同时, 改进后的模型明显提高了检测精度。在接下来的研究中, 将进一步进行网络模型的轻量化研究, 来提升模型的检测速度。

参考文献:

- [1] 邵 飞. 中国玉米经济: 供给与需求分析 [D]. 咸阳: 西北农林科技大学, 2011.
- [2] 查绒淇. 玉米常见虫害及综合防治要点 [J]. 新农业, 2020, 50 (3): 34-35.
- [3] 马 娇, 赵雪晴, 谭 友, 等. 草地贪夜蛾对开远市玉米产量的危害损失测定 [C] // 云南: 云南农业科技编辑, 2019: 40-42.
- [4] 伏松平. 双斑萤叶甲在天水市玉米田为害情况与防治初报 [J]. 中国植保导刊, 2022, 42 (7): 69-75.
- [5] 周素花. 商丘市玉米田草地贪夜蛾的识别与防治技术 [J]. 河南农业, 2019, 30 (35): 28-29.
- [6] 潘志国. 机器视觉技术在农作物病虫害的研究与应用 [J]. 电子测试, 2014, 21 (5): 57-58.
- [7] 韩弘炜, 张漪垠, 齐立萍, 等. 基于卷积神经网络的农作物病虫害检测综述 [J]. 智慧农业导刊, 2023, 3 (6): 6-9.
- [8] 吴 翔. 基于机器视觉的害虫识别方法研究 [D]. 杭州: 浙江大学, 2016.
- [9] 戴 宁. 图像处理在玉米虫害识别系统中的应用研究 [J]. 牡丹江大学学报, 2019, 28 (2): 104-107.
- [10] 翟肇裕, 曹益飞, 徐焕良, 等. 农作物病虫害识别关键技术研究综述 [J]. 农业机械学报, 2021, 52 (7): 1-18.
- [11] WANG F Y, WANG R J, XIE C J. Fusing multi-scale context-aware information representation for automatic in-field pest detection and recognition [J]. Computers and Electronics in Agriculture, 2020, 169: 105222.
- [12] SETIAWAN A, YUDISTIRA N, WIHANDIKA C R. Large scale pest classification using efficient Convolutional Neural Network with augmentation and regularizers [J]. Computers and Electronics in Agriculture, 2022, 200: 107204.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multi Box Detector [C] // Proceedings of the European Conference on Computer Vision (ECCV), 2016: 21-37.
- [14] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013: 580-587.
- [15] REN S, HE K, 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.
- [16] REDMON J, DIVVALA S, GIRSHICK R, et al. Youonly look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [17] 余 颢, 吴 伶, 单鲁泉. 基于 SSD 网络模型改进的水稻害虫识别方法 [J]. 郑州大学学报 (理学版), 2020, 52 (3): 49-54.
- [18] DU L, SUN Y, CHEN S, et al. A Novel Object Detection Model Based on Faster R-CNN for Spodoptera frugiperda According to Feeding Trace of Corn Leaves [J]. Agriculture, 2022, 12 (2): 248.
- [19] 段新涛, 王 伸, 赵 晴, 等. 基于改进 YOLOv4 的夏玉米主要害虫检测方法研究 [J]. 山东农业科学, 2023, 55 (10): 167-173.
- [20] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects [C] // Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD), 2022: 443-459.
- [21] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module [C] // Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [22] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [23] TONG K, WU Y. Rethinking PASCAL-VOC and MS-COCO dataset for small object detection [J]. Journal of Visual Communication and Image Representation, 2023, 93: 103830.
- [24] CUI H R, WEI Z H. Multi-Scale Receptive Field Detection Network [J]. IEEE Access, 2019, 7: 138825-138832.
- [25] WANG M, ZHAO P, LU X, et al. Fine-Grained Visual Categorization: A Spatial-Frequency Feature Fusion Perspective [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33 (6): 2798-2812.
- [26] NIU Z Y, ZHONG G Q, YU H. A review on the attention mechanism of deep learning [J]. Neurocomputing, 2021, 452 (10): 48-62.
- [27] 孙书魁, 范 菁, 孙中强, 等. 基于深度学习的图像数据增强研究综述 [J]. 计算机科学, 2023, 50 (9): 1-23.
- [28] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506 (28): 146-157.