

基于改进 YOLO v7 的农田复杂环境下害虫识别算法研究

赵 辉^{1,2} 黄 鏢^{1,2} 王红君^{1,2} 岳有军^{1,2}

(1. 天津理工大学电气工程与自动化学院, 天津 300384; 2. 天津市复杂系统控制理论与应用重点实验室, 天津 300384)

摘要: 为使巡检机器人能够对体积小且密集、形态多变、数量多且分布不均的害虫进行高效精准识别, 提出了一种基于改进 YOLO v7 的害虫识别方法。该方法将 CSP Bottleneck 与基于移位窗口 Transformer (Swin Transformer) 自注意力机制相结合, 提高了模型获取密集害虫目标位置信息的能力; 在路径聚合部分增加第 4 检测支路, 提高模型对小目标的检测性能; 将卷积注意力模块 (CBAM) 集成到 YOLO v7 模型中, 使模型更加关注害虫区域, 抑制背景等一般特征信息, 提高被遮挡害虫的识别精确率; 使用 Focal EIou Loss 损失函数减少正负样本不平衡对检测结果的影响, 提高识别精度。采用基于实际农田环境建立的数据集的实验结果表明, 改进后算法的精确率、召回率及平均精度均值分别为 91.6%、82.9% 和 88.2%, 较原模型提升 2.5、1.2、3 个百分点。与其它主流模型的对比实验结果表明, 本文方法对害虫的实际检测效果更优, 对解决农田复杂环境下害虫的精准识别问题具有参考价值。

关键词: 农田复杂环境; 害虫识别; 目标检测; 深度学习; YOLO v7; Swin Transformer

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2023)10-0246-09

OSID:



Pest Identification Method in Complex Farmland Environment
Based on Improved YOLO v7

ZHAO Hui^{1,2} HUANG Biao^{1,2} WANG Hongjun^{1,2} YUE Youjun^{1,2}

(1. School of Electrical Engineering and Automation, Tianjin University of Technology, Tianjin 300384, China

2. Tianjin Key Laboratory of Complex System Control Theory and Application, Tianjin 300384, China)

Abstract: In order to enable the inspection robot to efficiently and accurately identify small, dense, morphologically variable, numerous and unevenly distributed pests, a pest recognition method based on the improved YOLO v7 was proposed. CSP Bottleneck was combined with a self-attentional mechanism based on shift window transformer (Swin Transformer), which improved the ability of the model to obtain the location information of dense pests. A fourth detection branch was added to the path aggregation part to improve the detection performance of the model on small targets. The convolutional attention module (CBAM) was integrated into the YOLO v7 model to make the model pay more attention to the pest area, suppress the background and other general feature information, and improve the recognition accuracy of blocked pests. Focal EIou Loss function was used to reduce the influence of positive and negative sample imbalance on detection results and improve the recognition accuracy. According to the experimental results, the accuracy rate, recall rate and mAP of the improved algorithm were 91.6%, 82.9% and 88.2%, respectively by using the data set established based on the actual farmland environment, which was 2.5, 1.2 and 3 percentage points higher than that of the original model. Compared with other mainstream models, the experimental results showed that the proposed method was more effective in the actual detection of pests, and it had practical application value in solving the problem of accurate identification of pests in complex farmland environment.

Key words: complex farmland environment; pest identification; target detection; deep learning; YOLO v7; Swin Transformer

0 引言

农作物的健康生长决定着农业生产的产量及品质,作物健康监测是精准农业的一个重要组成部分。由于害虫种类繁多且为害特点复杂,人工识别效率低下,无法及时对农田虫害进行全面有效防控^[1-2]。农田巡检机器人可实现害虫的自动监测^[3-4],在降低人工成本的情况下实现全天候的实时监测,其中巡检机器人的视觉系统是高效、准确识别害虫的关键。因此,研究一种能够在复杂农田环境下快速准确识别害虫的方法,对提高农业生产效率和质量具有重要意义。

随着人工智能技术的发展,机器学习和深度学习的方法已经被大量应用于农作物的病虫害检测。DENG 等^[5]研究了一种基于 SIFT-HMAX 和局部配置模型(Local configuration pattern)提取不变特征的方法,并利用支持向量机(Support vector machine, SVM)对害虫进行识别,缺点是检测图像的时间较长,实时性差;KASINATHAN 等^[6]利用 GrabCut 算法对图像进行前景与背景的分割,并利用直方图均衡化对图像的对比度进行增强来提取害虫特征,对害虫具有较好的检测性能,但无法对数量较多的害虫进行检测。传统检测方法需要人工提取作物害虫特征,然而害虫的颜色和纹理与田间背景的颜色和纹理相似,通过传统机器学习方法对农田复杂环境下的害虫进行检测,易受复杂环境影响造成检测精度差,难以满足巡检机器人的实际工作要求。

深度学习方法能够通过数据集训练的方式提取目标多尺度特征,提高了模型精度及泛化能力,在害虫检测方面得到广泛应用^[7-10]。主流的目标检测网络主要有 SSD^[11]、YOLO (You only look once) 系列^[12-14]和 Faster R-CNN (Faster region with CNN) 系列^[15-17]等。在害虫检测方面,DONG 等^[18]提出一种具有信道重新校准特征金字塔网络(CRFPN)和自适应锚(AA)模块的 CRA-Net,提高了小目标害虫检测的准确性;JIAO 等^[19]设计了自适应特征融合金字塔网络(AFFP-Net),可提取害虫最优多尺度特征,改善了病害检测性能;NANNI 等^[20]利用显著图方法突出图像中与目标最相关的像素,并与 CNN 相结合的方式对害虫进行识别,但是对小目标的识别效果差;姚青等^[21]基于改进的 CornerNet,利用重叠滑动窗和检测框抑制的方法,实现了对不同密度下稻飞虱的检测;郭庆文等^[22]提出一种基于显著图分析技术的深度网络害虫检测模型,有效提高了诱虫板图像害虫检

测精度,mAP 达到 86.40%;肖德琴等^[23]提出一种改进的 Faster R-CNN 算法,对不同天气下不同密度的害虫进行检测,采用 ResNet101 作为特征提取网络时,mAP 达 87.14%;彭红星等^[24]利用改进后的 YOLOF 模型,对田间害虫进行检测,平均精度均值达到 93.7%。上述方法能够从训练数据中有效学习到目标特征,弥补了传统机器学习人工提取特征的不足,并且在准确率和检测效率上也有一定程度的提高,但还存在以下问题尚待解决:①当前害虫检测算法是在实验室环境下进行检测的,当在农田复杂环境下对害虫进行检测时,害虫图像不仅受背景因素影响,而且害虫形态、尺度多变,很容易造成害虫目标误检和漏检。②已有算法主要针对特定群体的害虫进行目标检测,农田复杂环境下害虫体积小、数量多、分布不均且会存在作物遮挡的情况,从而影响识别准确率。

针对上述问题,本文通过融合注意力模块、改进多尺度特征融合模块和优化损失函数,提出一种基于改进 YOLO v7 的农田复杂环境下害虫识别方法,以提高模型在农田复杂环境下对密集小目标害虫的检测效果。

1 实验数据集

实验样本数据集共包含 2 部分:采用公开害虫图像数据集 IP102^[25]和在佐治亚大学等机构设立的网址(<https://www.insectimages.org/index.cfm>)下载的数据集,包括蝼蛄、白星花金龟、麦长管蚜、红蜘蛛、蝗虫、玉米螟、飞蛾、蓟马和麦叶蜂 9 种常见农田害虫图像。为了符合巡检机器人在复杂场景下的实际工作要求,共筛选出包含农田复杂环境的 3 824 幅图像作为实验数据集,并按照比例 8:1:1 创建训练集、验证集和测试集。然后将标签格式保存为 PASCAL VOC 数据集格式,并将图像统一裁剪为 640 像素×640 像素,调整后的部分害虫数据集图像如图 1 所示,训练集、验证集和测试集各类害虫详细信息如表 1 所示。

2 作物害虫检测方法

2.1 YOLO v7 网络结构

YOLO v7^[26]是 YOLO 系列最新推出的结构,是目前在 MSCOCO 数据集上推理速度和识别效果最优的 YOLO 模型,该模型兼顾了检测速度与精度,满足复杂环境下巡检机器人识别快速、精准的要求,因此,本文以 YOLO v7 为基础检测模型并对其进行改进。YOLO v7 由骨干网络(Backbone network)、颈部网络(Neck network)和预测网络(Prediction

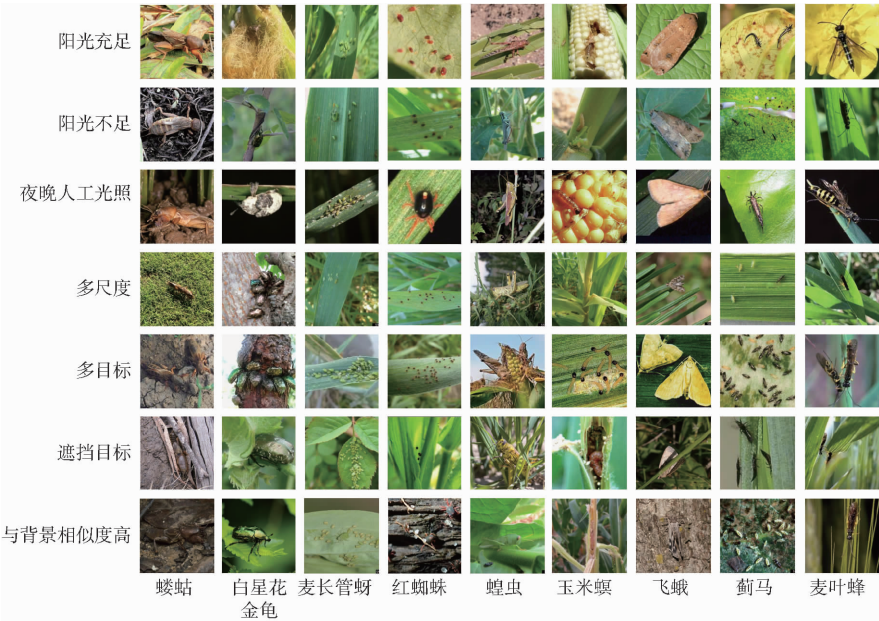


图 1 复杂场景下的害虫图像

Fig. 1 Image of pest in complex scenes

表 1 害虫数据集信息明细

Tab. 1 Pest dataset information details

待检测标	训练集 数量	验证集 数量	测试集 数量	标签
蝼蛄	410	50	50	mole cricket
白星花金龟	350	44	44	protaetia brevitarsis
麦长管蚜	350	38	38	grain aphid
红蜘蛛	290	40	40	red spider
蝗虫	300	38	38	locust
玉米螟	347	44	44	corn borer
飞蛾	420	54	54	moth
蓟马	270	38	38	thrips
麦叶蜂	313	41	41	wheat sawfly
总数	3 050	387	387	

network)3 部分组成。Backbone 部分由集成卷积单元 (CBS)、E-ELAN 架构以及 MPConv 模块组成,实现对特征的下采样,生成目标的语义信息;E-ELAN 模块通过利用 CBS 以不同方式的残差连接来增强网络对目标的特征提取能力;MPConv 模块通过利用 Maxpooling 与 CBS 并行连接的方式分别对输入特征图进行 2 倍下采样并进行信息融合,可以得到更多的特征信息,提升算法的检测精度。Neck 部分利用 CBS、Cat-ELAN 模块和 MPConv 将不同特征层之间的特征信息相互融合,输出到 Prediction 部分以生成检测不同尺度的检测框并预测出含有置信度的目标类别。

2.2 改进的 YOLO v7 害虫检测模型

2.2.1 害虫检测模型构建

虽然 YOLO v7 网络模型能够很好地权衡速度与精度,但是在农田复杂环境中,由于作物枝叶遮

挡等环境因素造成许多目标特征提取困难,特征信息容易丢失,不利于对作物害虫的检测。为了提高检测精度与检测效率,提出改进的 YOLO v7 害虫检测模型,其网络结构如图 2 所示。Backbone 部分提取的特征 B1 与经过 3 次上采样的特征进行特征融合后得到 N4,最后经过 Rep 卷积模块后得到 P1,用于检测极小目标;进行特征提取后的特征图 B2 与经过 2 次上采样的特征图进行特征融合后得到 N3,再将 N3 与 N4 进行级联操作后,经过 Rep 卷积模块后得到 P2,用于检测小目标;同理可得到特征图 P3 和 P4 分别用于对中目标和大目标的检测。

2.2.2 基于 Swin Transformer 的特征提取模块

由于在农田复杂环境下的害虫呈现密集且分布不均现象,容易造成模型无法准确定位害虫目标位置,出现漏检问题。YOLO v7 网络模型为 CNN 网络,具有局部相关性和平移不变性,缺乏全局建模的能力。为进一步提高模型对密集目标的特征提取能力,在 YOLO v7 网络模型的骨干网络顶层引入 Swin Transformer 特征提取模块 (SwinCSPX),通过 Swin Transformer 自注意力机制^[27],有效地融合局部特征和全局特征,获取全局上下文信息,从而指导模型更好地完成对密集目标的检测。因此,本研究将 Swin Transformer Block 与 CSP Bottleneck 结合,构成 SwinCSPX 模块,其结构如图 3 所示。

Swin Transformer Block 利用基于窗口多头自注意力 (Windows multi-head self-attention, W-MSA) 和基于移位窗口多头自注意力 (Shifted windows multi-head self-attention, SW-MSA) 相结合的方式,指导

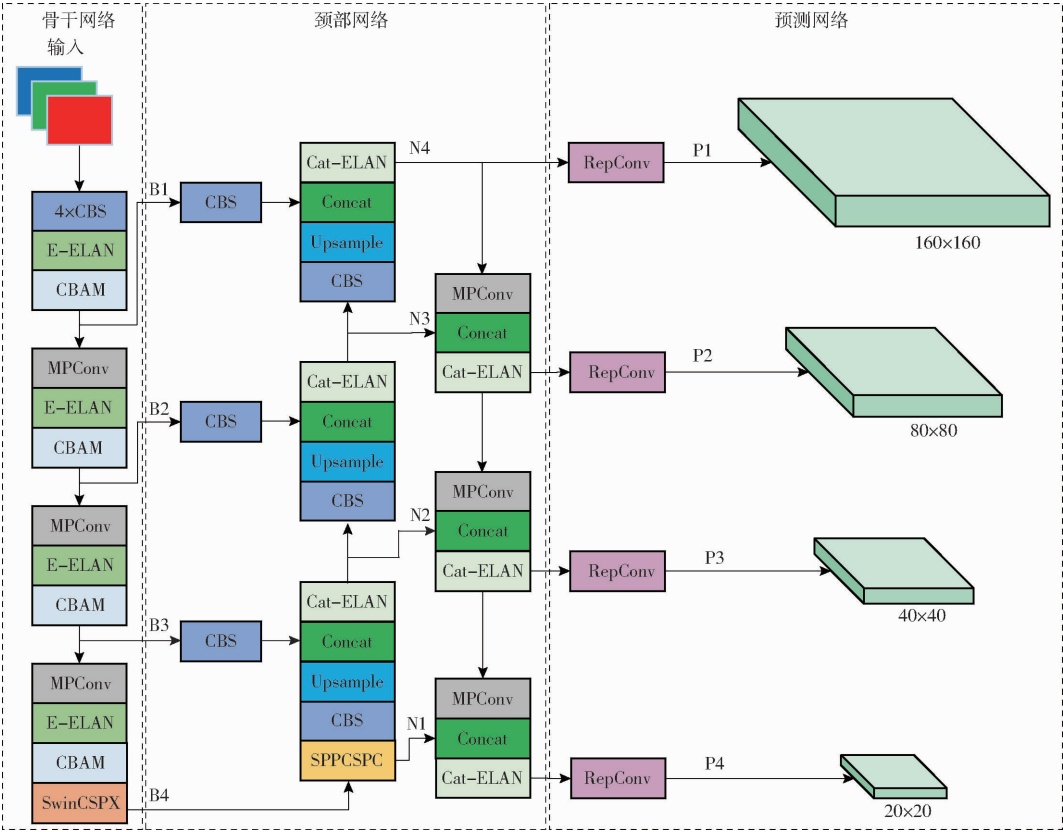


图 2 改进的 YOLO v7 网络结构图

Fig. 2 Improved YOLO v7 network structure diagram

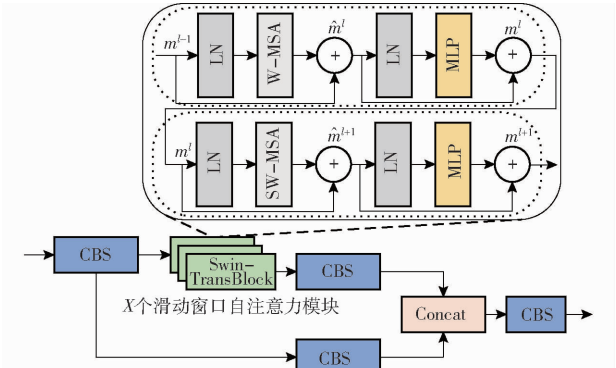


图 3 SwinCSPX 模块

Fig. 3 SwinCSPX module

模型建立害虫特征图通道与害虫目标位置信息的关系,利用全局上下文信息进行特征通道与害虫目标的信息交互,提高模型对密集型目标位置信息的识别能力,即 W-MSA 将输入特征图分割成 4 个尺寸相同的窗口,在 W-MSA 之后使用移位窗口自注意力(SW-MSA)将 4 个不同窗口间的信息融合。其具体实现方法是将 W-MSA 的 4 个窗口利用滑动窗口得到的编码为 A~I 的 9 个窗口,通过循环移位的方法对 9 个窗口进行信息重组,保证了移位后窗口的尺寸与原 W-MSA 窗口一致,最后进行窗口内 Transformer 自注意力^[28]计算,不仅减少了计算量,而且实现了全局建模。

SW-MSA 移动窗口的重组示意图如图 4 所

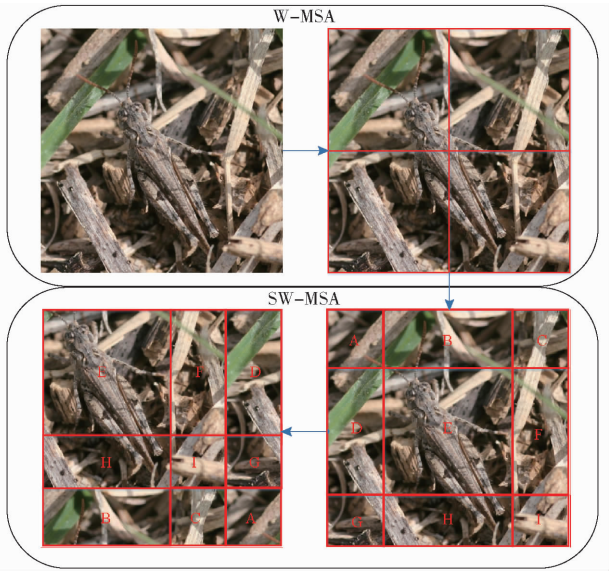


图 4 移动窗口重组示意图

Fig. 4 Diagram of relocating shifted windows

示。图 4 上方矩形框内执行的操作为 W-MSA,将图像划分为 4 个窗口;图 4 下方矩形框内执行操作为 SW-MSA,利用滑动窗口将得到的 4 个窗口划分为尺寸不一的 9 个窗口(A~I),通过循环移位操作将(F,D)、(H,B)及(I,G,C,A)构成的小窗口重新组合成 4 个同样尺寸的窗口,从而完成不同窗口间的信息交流。Swin Transformer Block 可以表示为

$$\begin{cases} \hat{m}^l = W - \text{MSA}(\text{LN}(m^{l-1})) + m^{l-1} \\ m^l = \text{MLP}(\text{LN}(\hat{m}^l)) + \hat{m}^l \\ \hat{m}^{l+1} = SW - \text{MSA}(\text{LN}(m^l)) + m^l \\ m^{l+1} = \text{MLP}(\text{LN}(\hat{m}^{l+1})) + \hat{m}^{l+1} \end{cases} \quad (1)$$

式中 $W - \text{MSA}$ ——窗口多头自注意力
 $SW - \text{MSA}$ ——移位窗口多头自注意力
 m^{l-1} ——输入
 m^l, m^{l+1} —— $W - \text{MSA}$ 和 $SW - \text{MSA}$ 的输出
 $\hat{m}^l, \hat{m}^{l+1}$ —— $W - \text{MSA}$ 和 $SW - \text{MSA}$ 与上一
层输入进行特征融合后的输出
 $\text{MLP}()$ ——多层感知机
 $\text{LN}()$ ——层标准化操作

2.2.3 多尺度路径聚合模块

为了更好地利用底层特征,从而提升模型对小目标检测性能,在原 YOLO v7 网络模型的基础上,构建 4 种不同尺度的检测头来识别害虫多尺度目标。浅层特征图中包含更丰富的颜色、纹理等信息。随着网络层数的加深,深层网络感受野变大,得到的特征图更具全局性,可以得到表达能力更强的语义信息,但是小目标的一些细粒度的特征信息容易丢失^[28]。通过在路径聚合部分增加第 4 检测支路,将浅层网络中的底层信息特征与深层网络的高语义特征进行信息融合,从而提升模型在实际农田环境下对小目标害虫的检测精度。

2.2.4 注意力机制模块

CBAM 利用通道注意力和空间注意力对输入特征图进行自适应特征优化,其模型结构如图 5 所示。在通道注意力机制中,其输入的特征图 F 在分别经过平均池化和最大池化进行映射后,将得到的特征传递到多层感知机 (MLP) 中,获得每个通道的权重,将权重进行 Add 操作后,通过 Sigmoid 激活函数得到系数矩阵 M_c ,原始特征图 F 与系数矩阵 M_c 相乘得到优化后的特征图 F' 。 M_c 计算公式为

$$M_c = \sigma(\text{MLP}(\text{Avgpool}(F))) + \sigma(\text{MLP}(\text{Maxpool}(F))) \quad (2)$$

式中 σ ——Sigmoid 激活函数
Avgpool——平均池化操作
Maxpool——最大池化操作

经过通道注意力机制优化的特征图 F' 在空间注意力机制中,经过全局平均池化和最大池化得到权重映射并对其进行堆叠形成特征图空间权重。经过一个 7×7 卷积层并利用 Sigmoid 进行激活后得到系数矩阵 M_s 。最后将特征图 F' 和系数矩阵 M_s 相乘,对特征图上的特征进行重要程度的重分配,得到优化后的特征图 F'' ,将 F'' 与原始特征图 F 进行特征融合得到自适应特征图 F^* ,从而使网络能够从更多的维

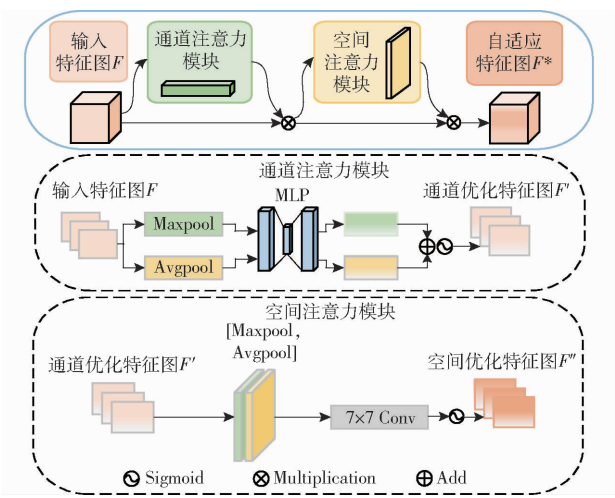


图 5 卷积注意力模块

Fig. 5 Convolution block attention module

度关注到感兴趣的目标区域。

2.2.5 损失函数

由于农田复杂环境下拍摄的害虫图像背景复杂,前景与背景容易出现类失衡,训练过程中高质量的锚框远少于误差大的锚框,质量差的样本会产生较大梯度从而影响模型优化方向。

为了解决以上问题,本文引入 Focal EIou Loss (Focal and efficient IoU loss) 替换原损失函数中的 CIou Loss,其计算公式为

$$L_{\text{Focal-EIoU}} = \text{IoU}^\gamma L_{\text{EIou}} \quad (3)$$

式中 γ ——控制异常值抑制程度的参数
 L_{EIou} ——有效交并比损失
相当于对高质量的回归目标进行加权,提高锚框的回归精度。

EIoU 计算公式为

$$L_{\text{EIou}} = L_{\text{IoU}} + L_{\text{dis}} + L_{\text{asp}} = 1 - \text{IOU} + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(\omega, \omega^{gt})}{c_\omega^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (4)$$

式中 b, b^{gt} ——预测框和真实框的中心点
 ω, ω^{gt} ——预测框和真实框的宽
 h, h^{gt} ——预测框和真实框的高
 c_ω, c_h ——包含预测框和真实框最小外接框的宽度和高度
 c ——预测框与真实框最小外接矩形的对角线距离
 ρ ——欧几里得距离

EIoU Loss 将损失函数分为预测框和真实框的重叠损失、位置损失以及宽高损失。

Focal EIou Loss 在 CIou Loss 的基础上,不仅优化了边界框回归,挖掘了高质量样本,还对宽高损失进行改进,直接计算预测框与真实框宽高的差值作为边长损失函数,提升了边界框收敛速度,优化了模

型检测性能。

3 结果与分析

3.1 实验运行平台及模型测试指标

本实验在基于 Ubuntu 18.04、Python 3.8 与 PyTorch 1.7 搭建的深度学习环境中完成,在 GTX1080Ti × 2 的服务器上进行实验,同时使用 CUDA v11.0、cuDNN v7.6.5 对运算进行加速。

模型对蝼蛄、白星花金龟、麦长管蚜、红蜘蛛、蝗虫、玉米螟、飞蛾、蓟马和麦叶蜂 9 种常见农田害虫识别结果进行分析,用精确率 (Precision, P)、召回率 (Recall, R)、平均精度均值 (mAP) 及检测速度评价模型性能。

3.2 结果评估

3.2.1 不同识别模型结果对比

基于制作的害虫数据集,将改进算法与其他 6 种目标检测网络算法在同一计算平台进行对比实验,检测结果如表 2 所示。

表 2 不同模型实验结果

Tab.2 Experiment results of different models				
模型	精确率/	召回率/	平均精度	检测速度/
	%	%	均值/%	(f·s ⁻¹)
SSD	88.5	69.8	75.8	25.6
Faster R-CNN	59.0	77.1	69.2	17.9
YOLO v3	82.2	72.6	77.4	21.7
YOLO v4	87.3	72.1	78.3	24.7
YOLO v5	91.7	78.5	83.6	26.1
YOLO v7	89.1	81.7	85.2	28.0
改进 YOLO v7	91.6	82.9	88.2	27.3

改进算法的精确率为 91.6%,召回率为 82.9%,平均精度均值为 88.2%,在保证精确率与其它模型基本一致的情况下,召回率和平均精度均值均高于其他 6 种模型,比 YOLO v5 模型的召回率和平均精度均值分别高了 4.4、4.6 个百分点,比 YOLO v7 模型的召回率和平均精度均值分别高了 1.2、3 个百分点。改进算法的平均检测速度为 27.3 f/s,虽然略低于检测速度最快的 YOLO v7 模型,但是检测算法的其他性能指标均优于该模型。因此,综合模型整体的检测性能指标,本文算法在识别精度与速度方面均具有较大优势。

图 6 为改进算法的检测效果图,改进算法能够在不同光照、多尺度以及不同程度遮挡的复杂环境下很好地检测出不同种类害虫,且漏检和误检情况较少,模型的鲁棒性强。

模型能够精准检测出害虫的种类及位置,表明模型具有较强的特征提取及推理能力。本文将 7 种算法对 9 种害虫的识别结果 (平均精确率和召回率) 进行对比,如表 3 所示。改进的 YOLO v7 模型



图 6 改进算法检测效果

Fig.6 Detection effects of improved algorithm

对麦长管蚜和红蜘蛛这两种类别的小目标样本的识别精确率分别为 88.9% 和 76.2%,比 YOLO v7 模型分别提升 0.4、9.4 个百分点,表明改进后的模型能够提取到更底层的特征,使模型学习到了更多的细节特征,从而提高了模型对小目标的识别能力。改进的 YOLO v7 模型对全类别害虫的召回率较原模型提升 1.2 个百分点,对体型小且分布较为密集的麦长管蚜、红蜘蛛这两类害虫的召回率比原 YOLO v7 模型分别提高 1.0、5.7 个百分点,说明改进的模型对密集情况下的害虫也有很好的检测效果。

由图 7 可知,改进的 YOLO v7 模型不仅漏检和误检极低,而且在白天和夜间环境下对密集小目标的精确率和召回率更高。在图像 1 中,Faster R-CNN 模型对遮挡目标的检测效果较差,且存在误检情况,而 YOLO v3 模型存在更多的漏检情况,YOLO v5 模型检测效果优于以上 2 种模型,但是左上方被遮挡的白星花金龟未检测到;图像 2 中,YOLO v4 模型检测与背景相似的多目标漏检程度较高,Faster R-CNN 模型作为二阶段目标检测算法,容易将背景识别成目标,检测精度较差;在对密集小目标的检测结果中,SSD 模型在图像 3 和图像 4 中仅检测出一小部分的麦长管蚜和红蜘蛛,YOLO v3 和 YOLO v4 模型虽然检测效果优于 SSD,但仍有小部分目标未检测到。改进的 YOLO v7 模型不仅能够检测出被枝叶遮挡的害虫,同样,该算法也能对密集且分布不均的小目标有较高的检测精度。

3.2.2 消融实验

本文共设计 5 组不同模型的消融实验,实验结果如表 4 所示。

表 3 不同种类害虫识别的平均精确率与召回率对比

Tab.3 Comparison of average precision rate and recall rate for different kinds of pests identification																		%	
模型	蝼蛄		白星花金龟		麦长管蚜		红蜘蛛		蝗虫		玉米螟		飞蛾		蓟马		麦叶蜂		
	平均	召回	平均	召回	平均	召回	平均	召回	平均	召回	平均	召回	平均	召回	平均	召回	平均	召回	
	精确率	率	精确率	率	精确率	率	精确率	率	精确率	率	精确率	率	精确率	率	精确率	率	精确率	率	
SSD	90.5	91.3	80.6	79.2	52.6	37.8	54.5	46.8	90.1	87.8	82.3	76.1	91.4	80.1	50.2	43.7	89.7	79.6	
Faster R-CNN	77.9	92.6	71.4	90.7	53.6	62.7	50.8	56.1	82.7	88.0	75.1	80.0	80.3	90.9	52.6	51.4	78.4	81.7	
YOLO v3	92.4	83.8	86.0	77.2	74.3	70.8	55.0	50.9	72.9	74.4	77.5	76.4	92.3	85.7	53.2	49.3	92.7	85.0	
YOLO v4	96.3	90.1	90.4	81.3	75.2	68.8	54.2	52.2	78.9	75.6	78.5	70.7	94.8	86.5	54.8	46.5	81.2	77.5	
YOLO v5	96.9	91.2	88.2	84.6	79.9	72.3	64.7	50.9	93.7	89.9	85.2	84.1	95.7	92.1	54.1	50.7	93.2	90.9	
YOLO v7	96.2	95.1	94.4	88.5	88.5	86.2	66.8	58.2	91.9	95.3	86.3	80.0	94.9	90.5	59.8	56.7	88.3	84.8	
改进 YOLO v7	99.1	94.3	94.6	86.5	88.9	87.2	76.2	63.9	94.3	93.0	86.5	81.8	99.4	91.6	63.2	56.7	91.6	90.9	

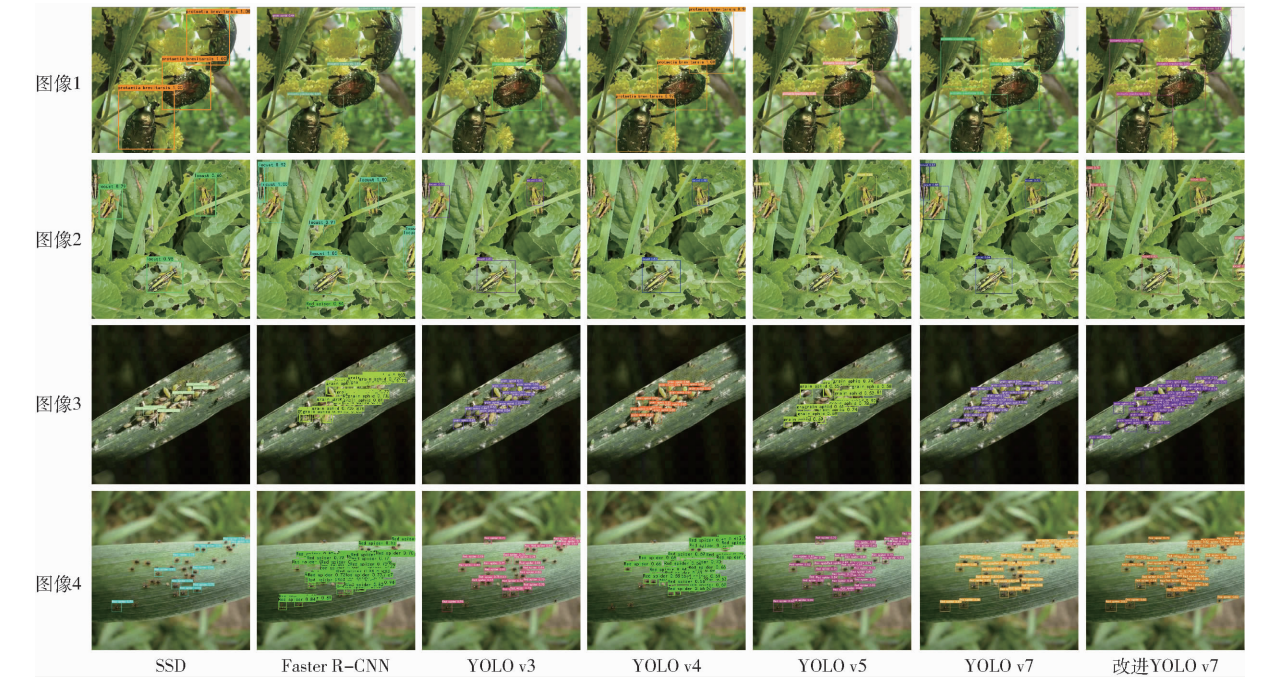


图 7 不同算法检测结果对比

Fig.7 Comparison of different algorithms results

表 4 消融实验结果

Tab.4 Ablation experiment results												
模型	精确率/%	平均精度均值/%	检测速度/($f \cdot s^{-1}$)	平均精确率/%								
				蝼蛄	白星花金龟	麦长管蚜	红蜘蛛	蝗虫	玉米螟	飞蛾	蓟马	麦叶蜂
YOLO v7	89.1	85.2	28.0	96.2	94.4	88.5	66.8	91.9	86.3	94.9	59.8	88.3
YOLO v7 + SwinCSPX	89.5	87.4	27.7	97.7	95.2	90.1	75.5	95.1	86.2	95.1	62.0	89.7
YOLO v7 + SwinCSPX + CBAM	92.2	87.6	27.4	98.1	96.2	88.2	70.6	93.7	90.8	97.3	63.0	90.5
YOLO v7 + SwinCSPX + CBAM + 多尺度	91.4	87.5	27.2	96.9	94.8	90.3	71.7	95.2	87.3	98.2	62.8	90.5
改进 YOLO v7	91.6	88.2	27.3	99.1	94.6	88.9	76.2	94.3	86.5	99.4	63.2	91.6

由表 4 可知,在骨干网络顶层添加 SwinCSPX 模块的网络模型比 YOLO v7 网络平均精度均值提升 2.2 个百分点,对麦长管蚜的检测结果提升 1.6 个百分点,说明 Swin Transformer 模块使模型能够更加关注密集的小目标,提取到了更多轮廓纹理等细节特征,使模型在层数加深的同时,依旧能够具有提

取全局上下文特征信息的能力,提升了对小目标的检测精度。YOLO v7 + SwinCSPX + CBAM 模型较 YOLO v7 + SwinCSPX 网络平均精度均值提高 0.2 个百分点,对白星花金龟和玉米螟的检测结果提升 1.0、4.6 个百分点,说明了注意力机制模块能够有效抑制背景等一般特征表达,能明显提升模型的检

测性能。在此基础上将 YOLO v7 网络路径聚合部分拓展为 4 种不同尺度的检测支路,对麦长管蚜和蝗虫的检测结果较 YOLO v7 + SwinCSPX + CBAM 模型提升 2.1、1.5 个百分点。由实验可得出,改进的 YOLO v7 算法比基准网络平均精度均值提升 3 个百分点,对蝼蛄、红蜘蛛、飞蛾、麦叶蜂的识别结果比 YOLO v7 模型提升 2.9、9.4、4.5、3.3 个百分点,检测精度显著提升。

4 结论

(1) 针对实际农田环境下存在的遮挡、形态多变和数量多且分布不均等问题,提出了一种基于改进 YOLO v7 的目标检测模型,能够在农田复杂环境下的害虫图像中准确检测出害虫种类、位置

及分布密度。首先,通过在 YOLO v7 骨干网络中加入 CBAM 模块,使模型更加关注害虫目标,抑制背景等一般特征的表达;同时将 YOLO v7 的 3 个不同尺度的检测头增加为 4 个,提高了模型对害虫目标的检测精度;其次,在骨干网络顶层添加 SwinCSPX 模块,使模型对密集且分布不均的害虫定位更加准确;最后,将原 YOLO v7 模型的损失函数替换为 Focal EIoU,提高了模型的检测性能。

(2) 改进的 YOLO v7 模型的平均精度均值为 88.2%,较原 YOLO v7 模型提升 3 个百分点,在保证检测速度的同时,提高了检测精度。面对复杂的农田环境,改进算法具有较高的稳定性和鲁棒性,为提升害虫视觉检测系统性能提供了有效手段。

参 考 文 献

- [1] TIAN H K, WANG T H, LIU Y D, et al. Computer vision technology in agricultural automation—a review[J]. Information Processing in Agriculture, 2020, 7(1): 1–19.
- [2] LIM S, KIM S, PARK S, et al. Development of application for forest insect classification using CNN[C] // 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV). Singapore: IEEE, 2018.
- [3] WAKCHAURE M, PATLE B K, MAHINDRAKAR A K. Application of AI techniques and robotics in agriculture: a review[J]. Artificial Intelligence in the Life Sciences, 2023, 3: 100057.
- [4] 李忠玉, 孙睿, 卢洪友. 基于模糊控制的小型农场巡检机器人系统开发与设计[J]. 现代电子技术, 2022, 45(8): 174–180.
LI Zhongyu, SUN Rui, LU Hongyou. Development and design of small-scale farm inspection robot system based on fuzzy control[J]. Modern Electronics Technique, 2022, 45(8): 174–180. (in Chinese)
- [5] DENG L M, WANG Y J, HAN Z Z, et al. Research on insect pest image detection and recognition based on bio-inspired methods[J]. Biosystems Engineering, 2018, 169: 139–148.
- [6] KASINATHAN T, SINGARAJU D, UYYALA S R. Insect classification and detection in field crops using modern machine learning techniques[J]. Information Processing in Agriculture, 2021, 8(3): 446–457.
- [7] 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.
ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1–18. (in Chinese)
- [8] WANG F W, WANG R J, XIE C J, et al. Fusing multi-scale context-aware information representation for automatic in-field pest detection and recognition[J]. Computers and Electronics in Agriculture, 2022, 169: 105222.
- [9] ZHANG L, DU J M, DONG S F, et al. AM-ResNet: low-energy-consumption addition-multiplication hybrid ResNet for pest recognition[J]. Computers and Electronics in Agriculture, 2022, 202: 107357.
- [10] ADHI S, NOVANTO Y, RANDY C W. Large scale pest classification using efficient convolutional neural network with augmentation and regularizers[J]. Computers and Electronics in Agriculture, 2022, 200: 107204.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C] // European Conference on Computer Vision. Cham: Springer, 2016: 21–37.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779–788.
- [13] REDMON J, FARHADI A. YOLO v3: an incremental improvement[Z/OL]. [2021-05-21]. <https://arxiv.org/abs/1804.02767>.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLO v4: optimal speed and accuracy of object detection[J]. arXiv Preprint, arXiv:1704.04861, 2020.
- [15] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580–587.
- [16] GIRSHICK R. Fast R-CNN[C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440–1448.
- [17] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C] // Advances in Neural Information Processing Systems, 2015: 91–99.
- [18] DONG S F, WANG R J, LIU K, et al. CRA-Net: a channel recalibration feature pyramid network for detecting small pests[J]. Computers and Electronics in Agriculture, 2021, 191: 106518.
- [19] JIAO L, XIE C J, CHEN P, et al. Adaptive feature fusion pyramid network for multi-classes agricultural pest detection[J].

Computers and Electronics in Agriculture,2022,195:106827.

[20] NANNI L, MAGUOLO G, PANCINO F. Insect pest image detection and recognition based on bio-inspired methods[J]. Ecological Informatics,2022,57:101089.

[21] 姚青,吴叔珍,蒯乃阳,等. 基于改进 CornerNet 的水稻灯诱飞虱自动检测方法构建与验证[J]. 农业工程学报,2021,37(7):183-189.

YAO Qing, WU Shuzhen, KUAI Naiyang, et al. Automatic detection of rice planthoppers through light-trap insect images using improved CornerNet[J]. Transactions of the CSAE,2021,37(7):183-189. (in Chinese)

[22] 郭庆文,王春桃,肖德琴,等. 利用显著图构建注意力深度网络检测诱虫板蔬菜害虫[J]. 农业工程学报,2021,37(19):211-219.

GUO Qingwen, WANG Chuntao, XIAO Deqin, et al. Building saliency-map-basedattention-driven deep network to detect vegetable pests of sticky trap images[J]. Transactions of the CSAE,2021,37(19):211-219. (in Chinese)

[23] 肖德琴,黄一桂,张远琴,等. 基于改进 Faster R-CNN 的田间黄板害虫检测算法[J]. 农业机械学报,2021,52(6):242-251.

XIAO Deqin, HUANG Yigui, ZHANG Yuanqin, et al. Pest detection algorithm of yellow plate in field based on improved Faster R-CNN[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(6):242-251. (in Chinese)

[24] 彭红星,徐慧明,高宗梅,等. 基于改进 YOLOF 模型的田间农作物害虫检测方法[J]. 农业机械学报,2023,54(4):285-294,303.

PENG Hongxing, XU Huiming, GAO Zongmei, et al. Insect pest detection of field crops based on improved YOLOF model[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(4):285-294,303. (in Chinese)

[25] WU X, ZHAN C, LAI Y K, et al. IP102: a large-scale benchmark dataset for insect pest recognition[C]// Long Beach: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2019: 8787-8796.

[26] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLO v7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[Z/OL]. <https://arxiv.org/abs/2207.02696>.

[27] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: hierarchical vision transformer using shifted windows[EB/OL]. (2021-03-25)[2022-02-11]. <https://arxiv.org/abs/2103.14030>.

[28] 刘树东,刘亚辉,孙叶美,等. 基于倒置残差注意力的无人机航拍图像小目标检测[J]. 北京航空航天大学学报,2023,49(3):514-524.

LIU Shudong, LIU Yahui, SUN Yemei, et al. Small object detection in UAV aerial images based on inverted residual attention [J]. Journal of Beijing University of Aeronautics and Astronautics,2023,49(3):514-524. (in Chinese)

[29] ZHU X, LYU S, WANG X, et al. TPH-YOLO v5: improved YOLO v5 based on transformer prediction head for object detection on drone-captured scenarios[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW),2021: 2778-2788.

[30] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). San Jose: Springer,2018:3-19.

(上接第 224 页)

[19] ZHAO H, QI X, SHEN X, et al. ICNet for real-time semantic segmentation on high-resolution images[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 405-420.

[20] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO Series in 2021[J]. arXiv Preprint, arXiv:2107.08430v2, 2021.

[21] AUBRY M, RUSSELL B C. Understanding deep features with computer-generated imagery[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 2875-2883.

[22] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points[J]. arXiv Preprint, arXiv:1904.07850,2019.

[23] 李科岑,王晓强,林浩,等. 深度学习中的单阶段小目标检测方法综述[J]. 计算机科学与探索,2022,16(1):41-58.

LI Kecen, WANG Xiaoqiang, LIN Hao, et al. Survey of one-stage small object detection methods in deep learning[J]. Journal of Frontiers of Computer Science and Technology,2022,16(1):41-58. (in Chinese)

[24] DING X, ZHANG X, MA N, et al. RepVGG: making VGG-style ConvNets great again[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 13728-13737.

[25] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// European Conference on Computer Vision. Berlin: Springer, 2018: 801-818.

[26] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[J]. Medical Image Computing and Computer-Assisted Intervention, 2015, 9351: 234-241.

[27] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881-2890.