

承达瑜, 赵伟, 何伟德, 等. 基于改进 YOLOv5n 模型的农作物病虫害识别方法[J]. 江苏农业学报, 2024, 40(11): 2021-2031.  
doi: 10.3969/j.issn.1000-4440.2024.11.005

# 基于改进 YOLOv5n 模型的农作物病虫害识别方法

承达瑜<sup>1</sup>, 赵伟<sup>1</sup>, 何伟德<sup>1</sup>, 武择鹏<sup>1</sup>, 王建东<sup>2</sup>

(1. 河北工程大学矿业与测绘工程学院, 河北 邯郸 056038; 2. 中国农业科学院农业环境与可持续发展研究所, 北京 100081)

**摘要:** 针对模型对复杂场景下农作物病虫害的识别精度低、模型参数量大的问题, 本研究对轻量级 YOLOv5n 模型进行改进。首先, 在 YOLOv5n 模型的骨干网络中加入坐标注意力模块, 使模型关注检测目标及其位置, 减少复杂背景对模型的影响。其次, 引入加权的双向特征融合金字塔网络 (BiFPN), 减少小目标信息丢失, 提高了模型的特征学习能力。最后, 用损失函数 SIoU 代替损失函数 CIoU, 在不改变模型参数量的情况下, 提升了目标检测精度。在无人机采集到的玉米病虫害数据集上, 本研究提出的 AgriPest-YOLOv5n 模型的  $mAP@0.50$  达 81.32%, 在 Jetson Xavier 开发板上检测速度达到 77 FPS, 模型大小为 1.63 MB。改进后的 YOLOv5n 模型能够满足轻量化的要求, 能够实时、准确地识别复杂背景下的农作物病虫害, 本研究结果可为病虫害精准防治提供技术支持。

**关键词:** 农业病虫害; 目标检测; 轻量级模型; 注意力机制

**中图分类号:** S435 **文献标识码:** A **文章编号:** 1000-4440(2024)11-2021-11

## Identification method of crop diseases and insect pests based on improved YOLOv5n model

CHENG Dayu<sup>1</sup>, ZHAO Wei<sup>1</sup>, HE Weide<sup>1</sup>, WU Zepeng<sup>1</sup>, WANG Jiandong<sup>2</sup>

(1. School of Mining and Geomatics Engineering, Hebei University of Engineering, Handan 056038, China; 2. Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences, Beijing 100081, China)

**Abstract:** In order to solve the problems of low recognition accuracy for crop diseases and insect pests in complex scenes and large model parameters of the model, the lightweight YOLOv5n model was improved in this study. Firstly, a coordinate attention module was added to the backbone network of YOLOv5n model to make the model focus on the detection target and its location and reduce the influence of complex background on the model. Secondly, the weighted bi-directional feature fusion pyramid network (BiFPN) was introduced to reduce the information loss of small targets and improve the model's feature learning ability. Finally, the loss function SIoU was used to replace the loss function CIoU, which improved the target detection accuracy without changing the parameters of the model. In the dataset of corn pests and diseases collected by unmanned air vehicle, the AgriPest-YOLOv5n model  $mAP@0.50$  proposed by this study reached 81.32%, and the detection speed reached 77 FPS on the Jetson Xavier development board. The size of the model was 1.63 MB. The improved YOLOv5n model can meet the requirement of light weight, and can identify crop diseases and insect pests in real time and accurately under complex background. The results of this study provide technical support for the precision control of crop diseases and insect pests.

**Key words:** agricultural pest and disease; object detection; lightweight model; attention mechanism

收稿日期: 2024-01-29

基金项目: 河北省重大科技成果转化专项 (22287401Z); 国家自然科学基金项目 (42071246)

作者简介: 承达瑜 (1980-), 男, 江苏常州人, 博士, 副教授, 主要从事地理大数据挖掘、作物长势遥感监测及计算机视觉技术研究。(E-mail) yuyumails@126.com

通讯作者: 赵伟, (E-mail) davis@163.com

农作物病虫害防治是农业生产的重要环节, 而高效的农作物病虫害检测与识别技术是实现有效防治的前提<sup>[1]</sup>。传统的作物病虫害检测主要依赖农

业专家或农业人员的经验,不仅费时、费力,还会受检查人员主观判断的影响,容易造成不必要的经济损失<sup>[2]</sup>。近年来,随着计算机视觉技术的迅速发展,基于视觉信息处理的病虫害检测方法已被广泛应用于农业领域,计算机视觉技术结合无人机巡航成为主流的大田虫情监控方法<sup>[3]</sup>。然而,无人机视角下拍摄的大田农作物图像背景较为复杂,除了农作物外,还有杂草、田间杂物等其他干扰,而且害虫的大小、颜色和形状各异,感染病虫害的农作物症状也多种多样,同时,无人机搭载的移动设备或边缘设备算力较低。因此,需要探索在复杂大田背景和低算力设备条件下有效检测和识别病虫害的方法。

目前,基于视觉信息处理的农作物病虫害检测方法包括图像处理技术和深度学习技术<sup>[4]</sup>。图像处理技术通常涉及手动提取各种图像特征,包括颜色、纹理、形状、大小等<sup>[5-8]</sup>。然而这些图像特征分析技术存在一些局限性,如操作繁琐、鲁棒性低、适应性低、准确性差和耗时长等。随着计算机性能的不断提高和深度学习技术的突破性发展,基于深度学习的目标分类和检测方法在农业中的应用越来越多<sup>[9-11]</sup>。在病虫害检测领域,李就好等<sup>[12]</sup>在 ResNet-50 的基础上融入特征金字塔网络(FPN),提出了一种基于改进的更快速区域卷积神经网络(Faster R-CNN)的苦瓜叶部病害目标检测方法,提升了对苦瓜叶部病害的检测精度。Wang 等<sup>[13]</sup>重新构建了区域建议网络(RPN),提出了一种正负样本均衡的 RPN 结构,能够生成更多高质量的锚框,从而更容易检测到小目标。宋中山等<sup>[14]</sup>将原始 Faster R-CNN 全连接层模块转变为二进制全卷积神经网络,提升了对柑橘病虫害的识别准确率。这些病虫害检测和识别算法大多基于 R-CNN 深度学习模型,该模型网络结构复杂、参数量和计算量大,难以应用于移动端或边缘计算设备。张善文等<sup>[15]</sup>在 VGG16 模型的基础上,提出一种基于多尺度注意力的卷积网络,将模型的识别准确率提升到 92.44%。Redmon 等<sup>[16]</sup>提出的 YOLO 算法在模型参数量和检测速度方面具有明显的优势。吕石磊等<sup>[17]</sup>提出一种 YOLOv3-LITE 轻量级网络模型,使用 MobileNet-v2 作为模型的骨干网络,该模型检测速度比 Faster-RCNN 提升了近 20 倍。储鑫等<sup>[18]</sup>在 YOLOv4 网络的基础上,提出了 MobileNetv1-YOLOv4 模型,模型参数量减少 80%。周维等<sup>[19]</sup>结合 GhostNet 网络,提出了一种基于改进的 YOLOv4-GhostNet 的水稻病

虫害识别方法,模型权重大小缩减为 42.45 MB。骆润玫等<sup>[20]</sup>使用 YOLOv5s 网络模型作为基础网络,引入多尺度特征融合模块,提高了模型对广佛手病虫害识别的准确率。Xue 等<sup>[21]</sup>将自注意力机制和卷积块注意力模块集成到 YOLOv5 中,使模型更加关注茶叶病虫害。

综上所述,当前深度学习算法已经被广泛应用于农作物病虫害检测,特别是以 YOLO 为代表的目标检测算法<sup>[22-24]</sup>。但是,上述方法中所使用的数据集场景简单,而真实农田场景较为复杂。并且上述方法中的模型不适用于算力较低的移动端或边缘设备。为此,本研究拟对 YOLOv5n 模型进行改进,提升病虫害检测模型的性能,实现对农业病虫害的高效检测和精准防治。

## 1 材料与方法

### 1.1 农业病虫害数据集

本研究以农业领域的农作物病虫害作为研究对象,旨在利用目标检测算法对常见农作物病虫害进行高效而准确的识别。为了评估不同农业场景中的模型性能,采用 IP102 数据集<sup>[25]</sup>和自制玉米病虫害数据集。IP102 是一个大型的用于害虫识别的公开数据集,包含了 102 个害虫类别和 75 222 张图像。本研究从中选取 30 种有关小麦、玉米和水稻农作物的病虫害,共 18 976 张图像,并将原始数据集的 XML 格式标签转换为 YOLO 数据集所需的 TXT 格式,部分样本如图 1a 所示。同时,利用高空无人机搭载变焦相机在河北省邯郸市的一处中型农场中航拍实际生产中的玉米病虫害,结合手机摄像头和 Kaggle 公开数据集共收集到 2 232 张图像,并使用随机上下左右翻转、高斯模糊、图像缩放和旋转等数据增强的方法,将数据集扩充至 3 356 张图像,病虫害类型包括早期玉米的黄叶病和虫害导致的烂叶病,自制玉米病虫害数据集部分图像如图 1b 所示。

### 1.2 试验方法

1.2.1 YOLOv5n 模型结构及其改进 YOLOv5 网络模型是一种基于深度学习的单阶段目标检测算法,网络结构主要分为骨干网络(Backbone)、颈部(Neck)、头部(Head)3 个模块。骨干网络部分使用跨阶段局部网络(CSPNet)和快速空间金字塔池化(SPPF)进行特征提取,颈部部分使用路径聚合网络(PANet)进行特征融合,头部网络部分有 3 个不同比例的特征图,

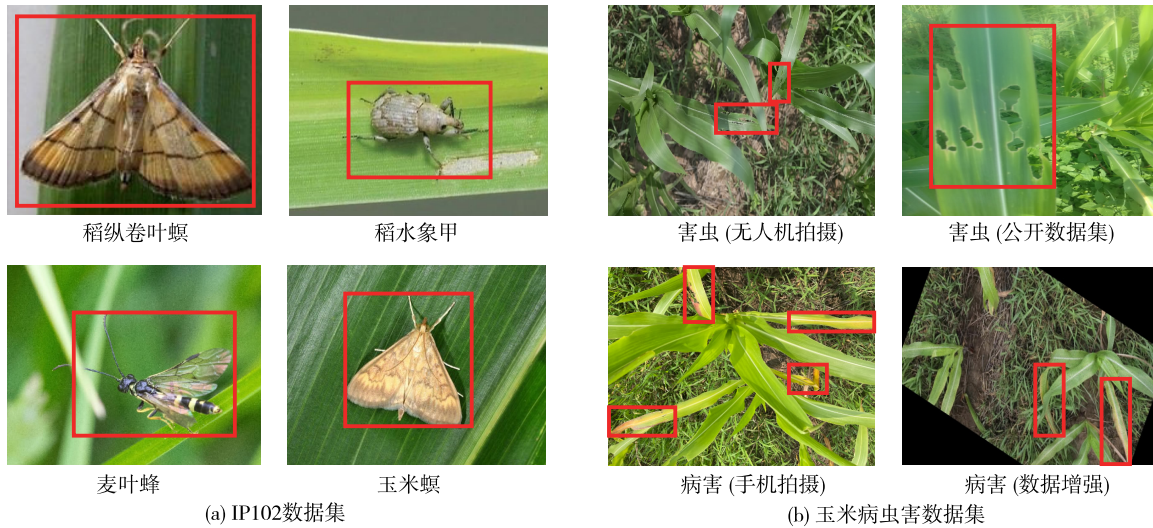
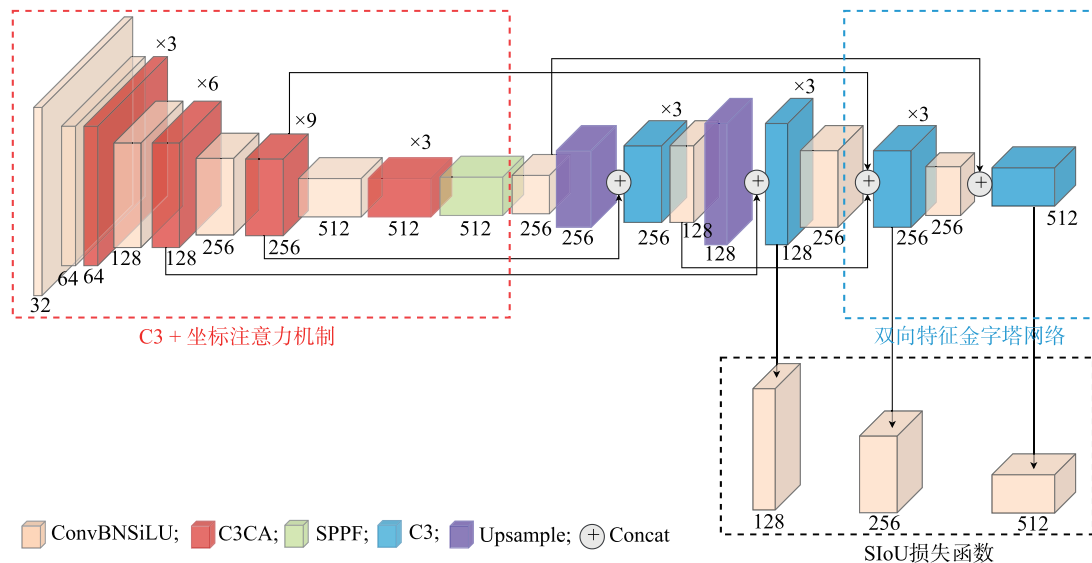


图 1 农业病虫害数据集

Fig.1 Dataset for agricultural pests and diseases

用于目标检测。YOLOv5 网络模型的特点是速度快、通用性强,近年来在农业病虫害检测领域被广泛应用。根据网络的深度和宽度,YOLOv5 模型分为 5 个版本,YOLOv5x、YOLOv5l、YOLOv5m、YOLOv5s、YOLOv5n,其中 YOLOv5n 的模型参数量最小,适用于内存小和算力低的边缘计算设备。目前 YOLO 系列模型的最新版本是 YOLOv8,但是在英伟达 Jetson Xavier 开发板的图形处理器(GPU)环境中,TensorRT 推理框架对 YOLOv5 模型的支持相对更加稳定。同

时,YOLOv5 已经被广泛应用于各项研究中,这使得本研究的结果更易于与现有结果进行比较和验证,确保试验的可重复性。针对农业病虫害图像中目标密集且小目标居多的特点,本研究提出了 AgriPest-YOLOv5n 网络模型,该模型通过对 YOLOv5n 模型中骨干网络部分的 C3 模块以及颈部的路径聚合网络模块进行改进,增强空间和语义信息,并用损失函数 SIoU 替换原始损失函数 CIoU,进而提升检测精度,模型结构如图 2 所示。



ConvBNReLU:激活函数为 ReLU 的卷积模块;C3CA:加入坐标注意力机制的卷积模块;SPPF:快速空间金字塔池化模块;C3:卷积模块;Upsample:上采样模块;Concat:特征融合;SIoU:可扩展轻量级自适应神经网络的交并比算法。

图 2 AgriPest-YOLOv5n 模型结构

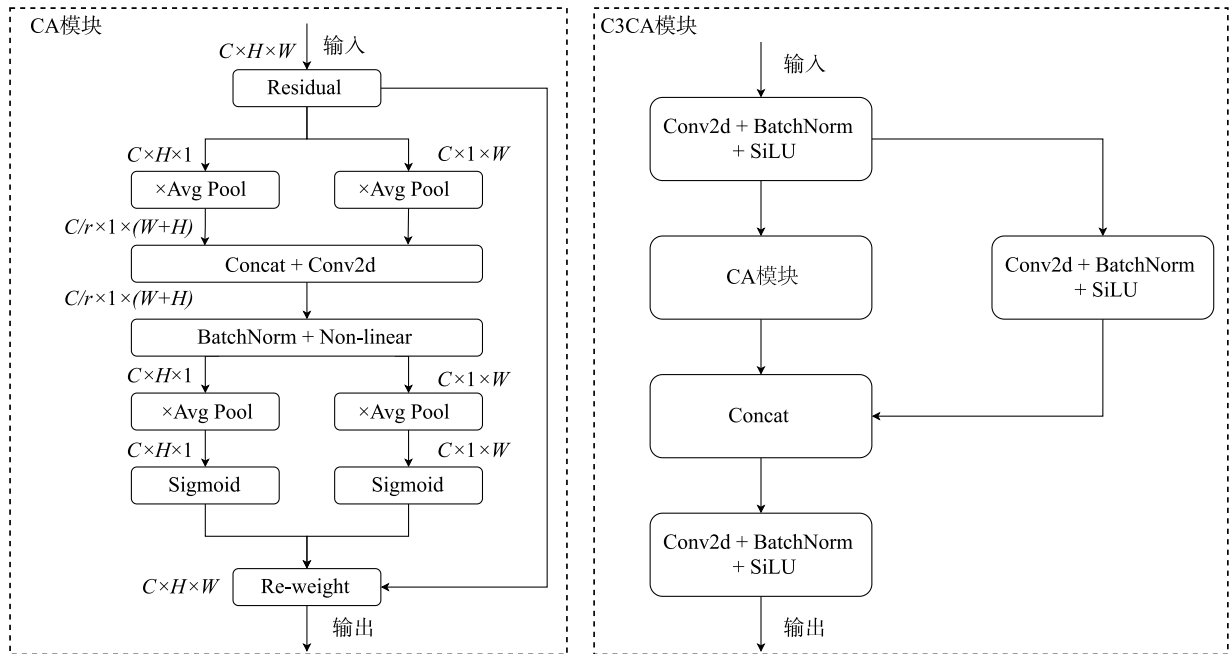
Fig.2 AgriPest-YOLOv5n model architecture



**1.2.2 坐标注意力机制** 在传统的轻量级神经网络模型中,通常采用 SE 模块<sup>[26]</sup>作为注意力机制,然而该模块仅关注通道之间的关系,却忽略了位置信息在目标检测任务中的作用,从而使得模型在目标检测任务中的精度降低。为解决这一问题,Hou 等<sup>[27]</sup>提出了一种新颖有效的注意力机制——坐标注意力机制(CA)。CA 将图像中水平和垂直方向上的坐标信息融合到通道信息中,能够捕捉更广泛的空间特征,避免了大量的计算,使得轻量级模型能够更有针对性地关注不同位置的特征,从而提高目标检测的精确性。坐标注意力机制不仅克服了传统 SE 模块在位置信息利用方面的不足,还提升了轻量级模型在目标检测任务中的性能。

C3CA 模块与 CA 的相似之处在于池化操作过程中对图像高度( $H$ )和宽度( $W$ )进行加权计算。但

不同的是,C3CA 模块输入的特征图来自 C3 模块,特征通过多层卷积操作提取。在使用 CA 计算通道注意力之前,C3CA 模块使用 $1\times 1$ 和 $3\times 3$ 的卷积核提取特征,并通过残差连接有效整合图像的通道和空间信息。接下来,分别在高度和宽度维度上执行自适应平均池化。得到每个子张量的平均值后,用 2 个全连接层对这些平均值进行编码,以获得每个子张量的权重。最后,将这些权重组合并对特征图进行加权。通过重新组合特征,将加权后的特征图与原始输入特征图进行逐元素相加,最终得到新的特征图。C3CA 注意力模块通过强化模型对目标特征的关注,降低了背景噪声对目标的干扰,有效提高了模型对目标检测的准确性。图 3 为 C3CA 注意力模块的结构,这种设计使得轻量级模型在目标检测任务中能够有效利用位置信息,提高模型性能。



$C$ 、 $H$ 、 $W$ :输入图像的通道数、高度和宽度;CA 模块:坐标注意力模块;C3CA:加入坐标注意力机制的卷积模块;Residual:残差模块;Avg Pool:平均池化;Concat:特征融合;Conv2d:卷积模块;BatchNorm:批量正则化层;Non-linear:非线性层;Sigmoid:激活函数;SiLU:卷积函数;Re-weight:特征缩放。

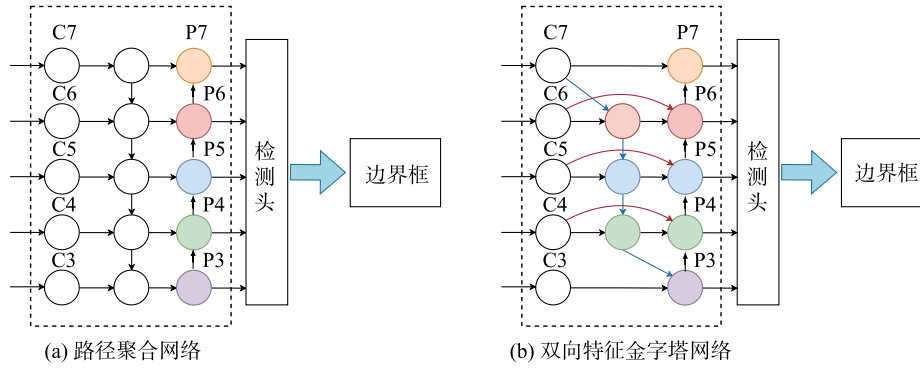
图 3 C3CA 模块结构

Fig.3 C3CA module structure

**1.2.3 双向特征金字塔网络** 在 AgriPest-YOLOv5n 模型中,YOLOv5s 中的路径聚合网络(PAN)被替换为 Tan 等<sup>[28]</sup>提出的加权双向特征金字塔网络(BiFPN),以提高信息挖掘深度和特征提取能力。BiFPN 的引入实现了深浅层特征的多尺度融合,加强了特征信息

在不同网络层之间的传递,提升了模型的检测能力。如图 4 所示,BiFPN 相对于 PAN 的改进主要体现在以下几个方面:第一,通过删除只有 1 条输入边的网络节点,BiFPN 简化了双向网络结构,减少了对融合不同特征贡献较小的节点,提高了计算效率。第二,

在同一层的输入和输出节点之间增加了额外的边,这样的设计可以有效集成更多功能,同时不增加计算量,进一步丰富了特征表示。第三,PAN 仅有 1 个自顶向下和自底向上的路径,而 BiFPN 将每个双向路径看作 1 个特征网络层,并且多次重复同一层,以实现更高层次的特征融合,从而提高模型的表达能力。BiFPN 通过下采样、上采样和双向路径,强化了不同尺度之间的特征融合效果。这种设计有助于模型利用检测目标在不同尺度下的特征,提升识别精度。



C3、C4、C5、C6、C7 表示对应卷积层输出的特征图;P3、P4、P5、P6、P7 表示对应特征金字塔网络融合后的特征图。

图 4 路径聚合网络 (PAN) 与加权双向特征金字塔网络 (BiFPN) 对比

Fig.4 Comparison between path aggregation network (PAN) and bi-directional feature pyramid network (BiFPN)

1.2.4 SIoU 损失函数 交并比 (Intersection over union,  $IoU$ ) 用于计算预测框面积和真实框面积的交集与并集之比,根据  $IoU$  的定义,2 个矩形的重叠面积即交集越大, $IoU$  值越接近 1,反之  $IoU$  值接近 0,一般认为当  $IoU$  值大于某个阈值时预测正确。损失函数通过判别模型的预测值与真实值的差值,可以为训练提供正确的方向。 $IoU$  计算方法如公式(1)所示:

$$IoU = \frac{R_1 \cap R_2}{R_1 \cup R_2} \quad (1)$$

式中, $R_1$  表示预测框的面积, $R_2$  表示真实框的面积。

在基础的 YOLOv5n 模型中,使用  $CloU$  损失函数作为边界损失函数, $CloU$  计算方法如公式(2)所示:

$$CloU = IoU - \left[ \frac{p^2(b, b^{gt})}{c^2} + \alpha v \right] \quad (2)$$

式中, $b$  和  $b^{gt}$  分别表示预测框和真实框的中心坐标, $c$  为最小封闭区域的对角线长度; $\alpha$  为权重参数,计算方法如公式(3)所示; $v$  表示纵横比的一致程度,计算方法如公式(4)所示。

BiFPN 结合特征融合 (Concat) 能够有效结合不同层级的特征图,确保重要信息不会被忽略或丢失,同时不受特征图中通道数量的限制。采用 Concat 方法可以更好地保留各层的信息,有利于提高模型的特征表达能力,尤其是在处理复杂场景和多尺度目标时表现更为出色。BiFPN 的另一优势在于其能更好地利用不同尺度深度图之间的相关性。通过引入更复杂的路径结构,模型能够获取并感知不同尺度的信息,并通过融合这些信息来提高模型检测性能。

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (3)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (4)$$

式中, $w$ 、 $h$  分别表示预测框的高度、宽度; $w^{gt}$ 、 $h^{gt}$  分别表示真实框的高度、宽度。

$CloU$  损失函数考虑了形状对结果的影响,特别是由于输入图像的长和宽而造成的检测损失,但忽略了角度对结果的影响。因此,本研究引入  $SIoU$ <sup>[29]</sup> 来加速 2 个预测框之间距离的收敛,从而减少推理时间和计算损失。 $SIoU$  定义了角度损失函数  $\Lambda$ ,计算方法如公式(5)所示; $\varphi$  是 2 个矩形中心之间的锐角,计算方法如公式(6)所示。

$$\Lambda = 1 - 2\sin^2\left(\varphi - \frac{\pi}{4}\right) \quad (5)$$

$$\varphi = \arcsin \left[ \frac{\max(b_{cy}^{gt} - b_{cy}) - \min(b_{cy}^{gt} - b_{cy})}{\sqrt{(b_{cx}^{gt} - b_{cx})^2 + (b_{cy}^{gt} - b_{cy})^2}} \right] \quad (6)$$

式中, $(b_{cx}, b_{cy})$  为预测框的中心坐标, $(b_{cx}^{gt}, b_{cy}^{gt})$  为真实框的中心坐标。

$SIoU$  还定义了距离损失函数  $\Delta$ , 计算方法如公式 (7) 所示。 $\rho_x$  表示归一化后的预测框中心  $x$  坐标与真实框中心  $x$  坐标之间的相对差,  $\rho_y$  表示归一化后的预测框中心  $y$  坐标与真实框中心  $y$  坐标之间的相对差, 计算方法如公式 (8) 所示。

$$\Delta = [1 - e^{-(A-2)\rho_x}]^\theta + [1 - e^{-(A-2)\rho_y}]^\theta \quad (7)$$

$$\rho_x = \left( \frac{b_{cx}^{gt} - b_{cx}}{c_w} \right)^2, \rho_y = \left( \frac{b_{cy}^{gt} - b_{cy}}{c_h} \right)^2 \quad (8)$$

式中,  $A$  为角度损失函数,  $\theta$  为超参数,  $c_w$ 、 $c_h$  分别为真实框和预测框最小外接矩形的宽、高。

$SIoU$  还使用 2 个矩形形状相似度的差异作为计算损失函数的标准, 定义形状损失为  $\Omega$ , 计算方法如公式 (9) 所示:

$$\Omega = (1 - e^{-\omega_w})^\theta + (1 - e^{-\omega_h})^\theta \quad (9)$$

式中,  $(w, h)$ 、 $(w^{gt}, h^{gt})$  分别为预测框、真实框的宽和高,  $\omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})}$ 、 $\omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})}$  和  $\theta$  是超参数。

最后得到  $SIoU$  损失函数, 计算方法如公式 (10) 所示:

$$SIoU = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (10)$$

表 1 试验环境及参数

Table 1 Experimental environment and parameters

软硬件	训练环境参数	推理环境参数
操作系统	Ubuntu 20.04 LTS	Ubuntu 20.04 LTS
CPU	Intel (R) Xeon (R) CPU E5-2680	NVIDIA Carmel ARM® v8.2
GPU	NVIDIA Tesla P100 16 GB	NVIDIA Jetson Xavier NX 16 GB
CUDA	12.2	12.3
cuDNN	8.2.0	8.6.0
编程语言	Python	Python
深度学习框架	PyTorch	PyTorch

CPU: 中央处理器; GPU: 图形处理器; CUDA: 统一计算设备架构; cuDNN: 深度神经网络加速库; Ubuntu 20.04 LTS: Linux 发行版操作系统; Intel (R) Xeon (R) CPU E5-2680: 英特尔 CPU 型号; NVIDIA Carmel ARM® v8.2: 英伟达 CPU 型号; NVIDIA Tesla P100: 英伟达 GPU 型号; NVIDIA Jetson Xavier NX: 英伟达开发板 GPU 型号; Python: 编程语言; PyTorch: 深度学习框架。

为了评估模型检测的性能, 使用精度 ( $Precision$ )、均值平均精度 ( $mAP$ ) 和每秒检测帧数 (Frame per second, FPS) 作为评价指标。如表 2 所示, 数据集中每个被检测到的边界框可以分为 4 种情况: 真阳性 (TP) 表示阳性样本中正确预测为阳性的样本数量; 真阴性 (TN) 表示阴性样本中正确预测为阴性的样本数量; 假阳性 (FP) 表示阴性样本中错误预测为阳性的样本数量; 假阴性 (FN) 表示阳性样本中错误预测为阴性的样本数量。

农作物病虫害形状和大小存在多样性, 为了提高模型性能, 减少输出时预测框与真实框的偏差, 本研究在目标检测模块中用更精确的  $SIoU$  损失函数替换原来的  $CIoU$  损失函数, 增强整体框架的检测能力。

## 2 结果与分析

### 2.1 试验环境及评价指标

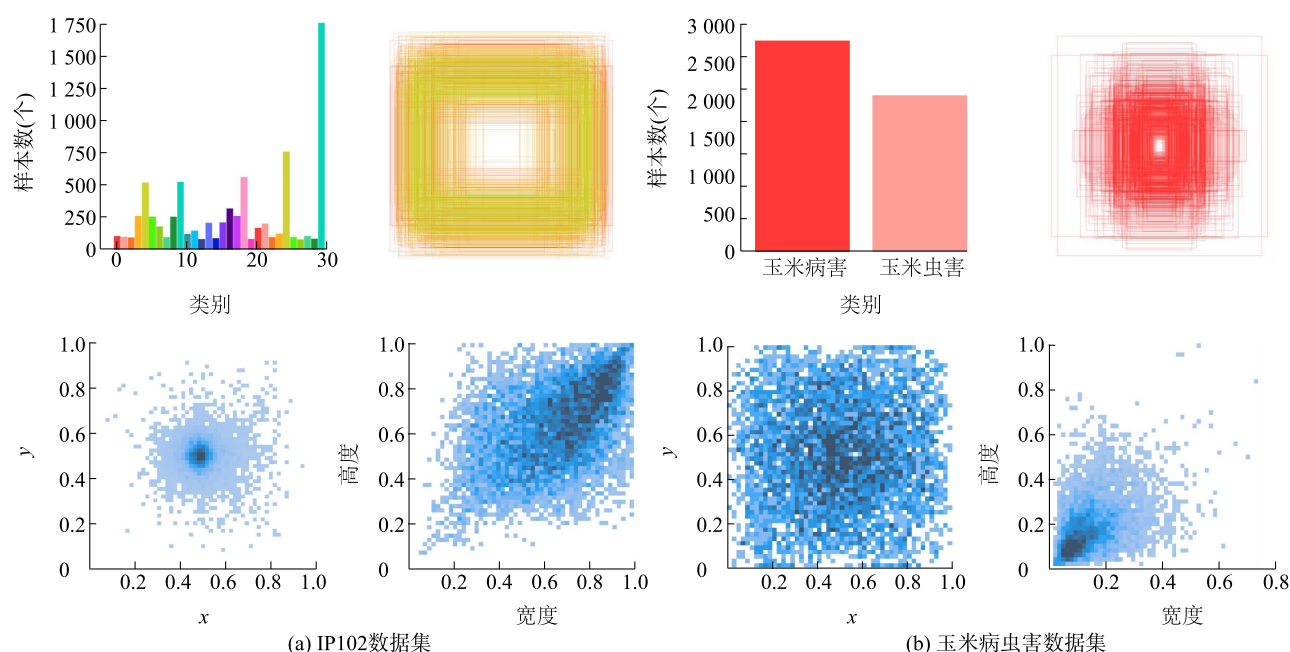
本研究使用包含深度学习框架 PyTorch 和 GPU 的服务器对农作物病虫害数据集进行训练, 试验的软硬件环境及详细参数情况如表 1 所示。试验过程中利用 YOLOv5n 预训练权重来更好地初始化神经网络参数, 设置训练次数为 600 epoch (训练轮次), 初始学习率为 0.01, 批处理大小设置为 16, 图像尺寸为 640×640 像素, 进行训练的 IP102 公开数据集和自制玉米病虫害数据集图片数量分别为 18 976 张和 3 356 张。每类数据集数量按照训练集: 验证集: 测试集 = 6 : 2 : 2 的比例进行划分。训练集中的样本类别分布、边界框的尺寸分布、中心点的位置分布和宽高分布信息如图 5 所示。在推理阶段将训练好的模型部署在 Jetson Xavier 开发板上, 结合 TensorRT 推理框架测试检测速度, 训练时置信度阈值设置为 0.25, 推理时阈值设置为 0.50。

表 2 验证数据集上的混淆矩阵

Table 2 Confusion matrix on the validation dataset

混淆矩阵	预测值	
	阳性样本	阴性样本
阳性样本 (真实值)	TP	FN
阴性样本 (真实值)	FP	TN

TP: 阳性样本中正确预测为阳性的样本数量; TN: 阴性样本中正确预测为阴性的样本数量; FP: 阴性样本中错误预测为阳性的样本数量; FN: 阳性样本中错误预测为阴性的样本数量。



左上图中,0: 稻纵卷叶螟;1: 二化螟;2: 水稻蝇蛆;3: 稻水象甲;4: 蛴螬;5: 蝼蛄;6: 金针虫;7: 小地老虎;8: 红蜘蛛;9: 玉米螟;10: 麦叶蜂;11: 白星花金龟;12: 桃蛀螟;13: 麦红吸浆虫;14: 跳甲;15: 甜菜象虫;16: 牧草盲蝽;17: 蝗总科;18: 豆芫菁;19: 芫菁;20: 十星叶甲;21: 葡萄天蛾;22: 斑衣蜡蝉;23: 四带虎天牛;24: 大青叶蝉;25: 盲蝽科;26: 吹绵蚧;27: 枯叶夜蛾;28: 褐缘蛾蜡蝉;29: 脊胸天牛。

图5 训练集中的样本信息

Fig.5 Sample information in the training set

精度 (*Precision*) 反映了模型在所有检测到的边框中的正确性,计算方法如公式 (11) 所示:

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

均值平均精度 (*mAP*) 计算方法如公式 (12) 所示:

$$mAP = \frac{1}{C} \sum_{i=1}^C Precision(i) \quad (12)$$

式中,  $C$  表示病虫害总类别数。

$mAP@0.50$  的定义为在交并比 (*IoU*) 阈值为 0.50 的情况下,对每一类分别计算平均精度 (*AP*),然后取平均值,计算方法如公式 (13) 所示。 $mAP@0.50 : 0.95$  的定义为在交并比阈值从 0.50 到 0.95 的范围内,以 0.05 的步长递增,计算每个阈值下的 *mAP*,然后取平均值,计算公式如 (14) 所示。

$$mAP@0.50 = \frac{\sum_{i=1}^K AP_i}{K} \times 100\% \quad (13)$$

$$mAP@0.50 : 0.95 = \frac{mAP@0.50 + mAP@0.55 + \dots + mAP@0.95}{10} \times 100\% \quad (14)$$

## 2.2 消融试验

在 IP102 公开数据集上,利用消融试验验证改进的 YOLOv5n 模型对  $mAP@0.50$  的影响。结果如表 3 所示,相较于 YOLOv5n 模型,引入了 BiFPN 注意力机制的 YOLOv5n+BiFPN 模型  $mAP@0.50$  提高了 0.26 个百分点,引入了 C3CA 注意力机制的 YOLOv5n+C3CA 模型  $mAP@0.50$  提高了 1.05 个百分点,使用损失函数 SIoU 的 YOLOv5n+C3CA+BiFPN+SIoU 模型的  $mAP@0.50$  提升了 1.44 个百分点。试验结果证明了该方法的鲁棒性。

在自制玉米病虫害数据集上,利用消融试验验证改进的 YOLOv5n 模型对  $mAP@0.50$  的影响。结果如表 4 所示,相较于 YOLOv5n 模型,引入了 BiFPN 注意力机制的 YOLOv5n+BiFPN 模型  $mAP@0.50$  提高了 0.31 个百分点,引入了 C3CA 注意力机制的 YOLOv5n+C3CA 模型  $mAP@0.50$  提高了 3.85 个百分点。试验结果表明,加权双向特征融合金字塔网络 (BiFPN) 可以减少小目标信息丢失,提高了模型的特征学习能力。在 YOLOv5n 原始 C3 模块中引入 CA 注意力机制能够准确定位对象的位置,减少分类和对象损失,从而提高模型的检测性能。相比于 *CIoU*,*SIoU* 损失函数考虑到了



预测帧和真实帧之间的矢量角度,在模型参数几乎不变化的情况下,加速了模型的收敛,提升了对病虫害检测的精度。试验结果表明,相较于 YOLOv5n 模型, YOLOv5n+C3CA+BiFPN 模型  $mAP@0.50$  提升了 4.53 个百分点,使用损失函数  $Siou$  的 YOLOv5n+C3CA+BiFPN+ $Siou$  模型的  $mAP@0.50$  提升了 5.80 个百分点,证明该组合模型效果较好。

表 3 在 IP102 公开数据集上的模型验证(验证集)  
Table 3 Validation of the model on the IP102 public dataset (validation set)

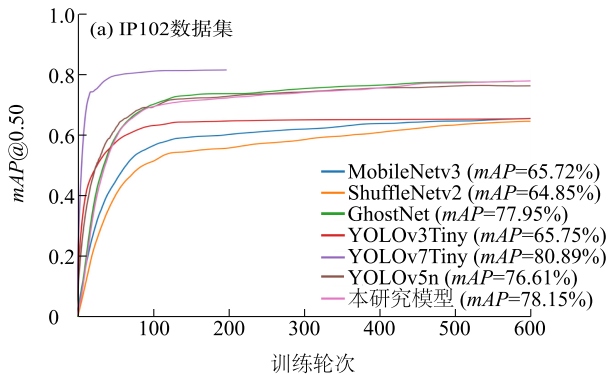
模型	参数量 (MB)	浮点量	$mAP@0.50$ (%)
YOLOv5n	1.68	4.23	76.71
YOLOv5n+BiFPN	1.70	4.28	76.97
YOLOv5n+C3CA	1.62	4.10	77.76
YOLOv5n+C3CA+BiFPN	1.63	4.15	77.87
YOLOv5n+C3CA+BiFPN+ $Siou$	1.63	4.15	78.15

YOLOv5n; YOLOv5 模型的最小参数版本; BiFPN; 加权双向特征融合金字塔网络; C3CA; 加入坐标注意力机制的卷积模块;  $Siou$ ; 损失函数。  $mAP@0.50$ ; 在交并比阈值为 0.50 的情况下,对每一类分别计算平均精度(AP),然后取的平均值。

真实场景下收集到的玉米病虫害数据集场景相对于 IP102 数据集更加复杂,对比表 3 和表 4 可以看出,在复杂场景下的农作物病虫害数据集上,本研究提出的模型 YOLOv5n+C3CA-BiFPN+ $Siou$  效果更佳。

2.3 与先进的轻量级模型对比

对比本研究提出的模型 YOLOv5n+C3CA+BiFPN+ $Siou$  与轻量级模型 MobileNetv3、ShuffleNetv2、GhostNet、YOLOv3Tiny、YOLOv5n 和 YOLOv7Tiny 在 2 类农作物病虫害数据集上的检测性能,所有模型的数据均在 Ultralytics 版 YOLOv5 框架下训练得出。

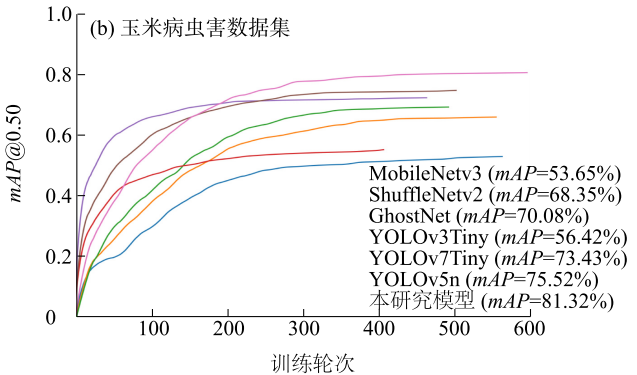


训练结果如图 6 所示,图 6a 为各轻量级模型在 IP102 公开数据集上的  $mAP@0.50$  训练曲线。如表 5 所示,在简单场景下, YOLOv5n+C3CA+BiFPN+ $Siou$  模型的  $mAP@0.50$  仅低于 YOLOv7Tiny 模型,但相较于 YOLOv7Tiny 模型, YOLOv5n+C3CA+BiFPN+ $Siou$  模型参数量更少,帧率更高。图 6b 为各轻量级模型在自制玉米病虫害数据集上的  $mAP@0.50$  训练曲线,在轻量级模型 MobileNetv3、ShuffleNetv2、GhostNet、YOLOv3Tiny、YOLOv5n 和 YOLOv7Tiny 中, YOLOv5n 模型的  $mAP@0.50$  最高, YOLOv5n+C3CA+BiFPN+ $Siou$  模型在 YOLOv5n 模型基础上进行了改进, YOLOv5n+C3CA+BiFPN+ $Siou$  模型的  $mAP@0.50$  高于 YOLOv5n 模型。如表 6 所示,在复杂场景下, YOLOv5n+C3CA+BiFPN+ $Siou$  模型的  $mAP@0.50$  为 81.32%,在 Jetson Xavier 开发板上的帧率为 76.92 FPS。综上可知,本研究提出的 YOLOv5n+C3CA+BiFPN+ $Siou$  模型更适用于实际农业生产中的农作物病虫害实时检测。

表 4 在自制玉米病虫害数据集上的模型验证(验证集)  
Table 4 Validation of the model on the self-made corn pest dataset (validation set)

模型	参数量 (MB)	浮点量	$mAP@0.50$ (%)
YOLOv5n	1.68	4.23	75.52
YOLOv5n+BiFPN	1.70	4.28	75.83
YOLOv5n+C3CA	1.62	4.10	79.37
YOLOv5n+C3CA+BiFPN	1.63	4.15	80.05
YOLOv5n+C3CA+BiFPN+ $Siou$	1.63	4.15	81.32

YOLOv5n; YOLOv5 模型的最小参数版本; BiFPN; 加权双向特征融合金字塔网络; C3CA; 加入坐标注意力机制的卷积模块;  $Siou$ ; 损失函数。  $mAP@0.50$ ; 在交并比阈值为 0.50 的情况下,对每一类分别计算平均精度(AP),然后取的平均值。



$mAP@0.50$ ; 在交并比阈值为 0.50 的情况下,对每一类分别计算平均精度(AP),然后取的平均值。

图 6 2 类数据集下的轻量级模型  $mAP@0.50$  曲线(验证集)

Fig.6  $mAP@0.50$  curves for lightweight models under two different datasets (validation set)



为了评价本研究提出的模型 YOLOv5n+C3CA+BiFPN+SiLU 对农作物病虫害的检测性能,利用玉米病虫害数据集和 IP102 公开数据集中划分的测试集对 YOLOv5n 模型和改进后的模型检测结果进行对比。如图 7 所示,YOLOv5n 模型会漏检自制玉米病虫害数据

集中的小目标,对 IP102 公开数据集中的目标的检测精度较低。加入坐标注意力机制 CA 和加权双向特征融合 BiFPN 后,YOLOv5n+C3CA+BiFPN+SiLU 模型具有更加优越的目标检测能力和更高的准确率,并且漏检的情况明显减少,将该模型命名为 AgriPest-YOLOv5n。

表 5 不同轻量级模型在 IP102 公开数据集上的性能(验证集)

Table 5 Performance of different lightweight models on the IP102 public dataset ( validation set )

模型	参数量 ( MB )	浮点量	精度 ( % )	$mAP@0.50$ ( % )	$mAP@0.50 : 0.95$ ( % )	帧率 ( FPS )
MobileNetv3	3.38	5.97	71.87	65.72	37.23	66.67
YOLOv3Tiny	8.27	12.98	72.31	65.75	31.49	52.63
ShuffleNetv2	1.00	2.04	66.19	64.85	36.35	125.00
GhostNet	3.52	8.21	80.10	77.95	48.41	47.62
YOLOv7Tiny	5.74	13.19	81.98	80.89	52.33	45.45
YOLOv5n	1.68	4.23	77.89	76.61	47.18	90.91
本研模型究	1.63	4.10	79.28	78.15	47.69	76.92

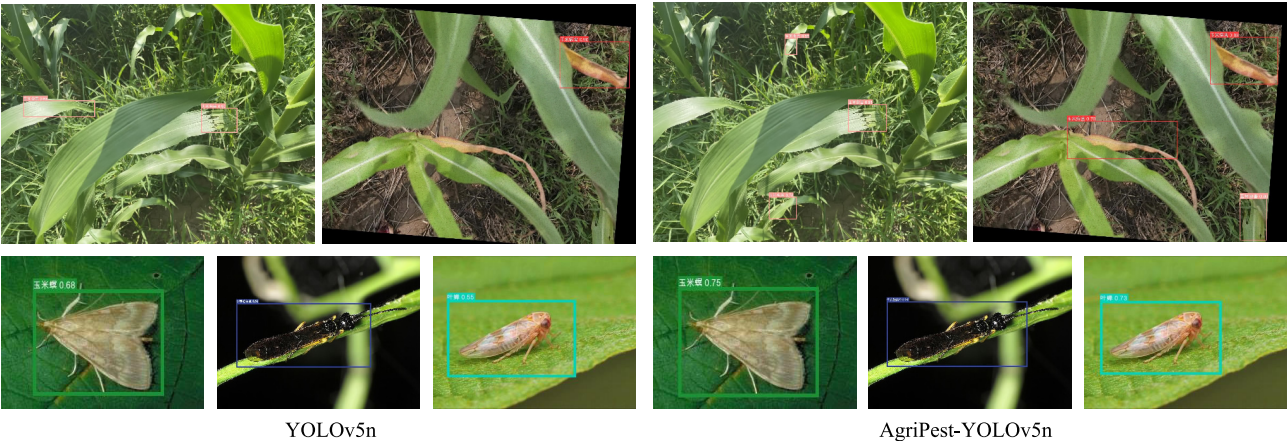
$mAP@0.50$ :在交并比阈值为 0.50 的情况下,对每一类分别计算平均精度 (AP),然后取的平均值。 $mAP@0.50 : 0.95$ :在交并比阈值从 0.50 到 0.95 范围内,以 0.05 的步长递增,计算每个阈值下的  $mAP$ ,然后得到的平均值。

表 6 不同轻量级模型在自制玉米病虫害数据集上的性能(验证集)

Table 6 Performance of different lightweight models on the self-made corn pest dataset ( validation set )

模型	参数量 ( MB )	浮点量	精度 ( % )	$mAP@0.50$ ( % )	$mAP@0.50 : 0.95$ ( % )	帧率 ( FPS )
MobileNetv3	3.38	5.97	75.26	53.65	20.51	66.67
YOLOv3Tiny	8.27	12.98	81.19	56.42	21.15	52.63
ShuffleNetv2	1.00	2.04	81.89	68.35	28.73	125.00
GhostNet	3.52	8.21	86.39	70.08	37.74	47.62
YOLOv7Tiny	5.74	13.19	88.91	73.43	38.49	45.45
YOLOv5n	1.68	4.23	88.23	75.52	38.73	90.91
本研模型究	1.63	4.10	89.94	81.32	41.75	76.92

$mAP@0.50$ :在交并比阈值为 0.50 的情况下,对每一类分别计算平均精度 (AP),然后取的平均值。 $mAP@0.50 : 0.95$ :在交并比阈值从 0.50 到 0.95 范围内,以 0.05 的步长递增,计算每个阈值下的  $mAP$ ,然后得到的平均值。



YOLOv5n:YOLOv5 模型的最小参数版本;AgriPest-YOLOv5n:本研究改进的 YOLOv5n 模型。

图 7 YOLOv5n 与改进后模型 AgriPest-YOLOv5n 识别效果对比(测试集)

Fig.7 Comparison of detection performance between YOLOv5n and the improved model AgriPest-YOLOv5n ( test set )

### 3 讨论

神经网络模型的层数和通道数直接影响计算量,参数量较大的模型精度较高,但会牺牲一定的计算速度。当图像中存在多个检测目标时,模型需要对每个目标进行定位和识别,这会增加模型的计算量。在英伟达 Jetson Xavier 开发板上利用 TensorRT 框架对验证集中的每张图像进行推理可知, YOLOv7Tiny 模型的检测速度约为 45.45 FPS。相比之下,改进后的 YOLOv5n 模型检测速度能够达到 76.92 FPS,即每 1 s 大约处理 77 张图像,远超人眼分辨图像的速度 30 FPS,说明该模型能够充分满足农业应用场景的实时性要求。

相较于 YOLOv7Tiny 算法, YOLOv5n 算法的检测性能较差,这是因为 YOLOv7Tiny 模型的尺寸和参数量相较于 YOLOv5n 模型更大,相对复杂的架构能够更有效地提取病虫害检测的关键特征,所以 YOLOv7Tiny 模型可以更有效地检测背景相对简单、目标特征较为明显的 IP102 数据集。然而在无人机采集的病虫害数据集上, YOLOv5n 比 YOLOv7Tiny 表现更好。在实际应用中, YOLOv5n 模型对复杂背景的图像中病虫害的检测能力更强,并且在资源受限的边缘计算设备和无人机实时监测场景下,检测速度也是一个关键指标。

张善文等<sup>[15]</sup>在 VGG16 模型的基础上提出了一种基于多尺度注意力卷积网络的作物害虫检测方法,然而 VGG16 模型的计算复杂,检测效率低,无法应用在计算资源有限的移动设备上实时检测。储鑫等<sup>[18]</sup>使用 Plant Village 公开数据集中的番茄叶片集作为训练数据对 YOLOv4 模型性能进行评估,并使用 MobileNet 网络对 YOLOv4 模型进行改进。然而在实验室条件下拍摄的公开数据集,不同图像之间的背景相似,检测结果易受背景噪音的干扰。在复杂场景下,本研究提出的轻量级算法 AgriPest-YOLOv5n 对玉米病虫害的识别精度较高、识别速度较快,并且识别精度和识别速度均高于 MobileNet。本研究的算法可应用于无人机系统、监控摄像机、移动设备和边缘计算盒子设备端,为实现精准农业病虫害监测提供了有效方法。

### 4 结论

为了提升无人机视角下背景复杂的大田农作物

病虫害识别的检测速度和准确率,本研究提出一种基于 YOLOv5n 的改进模型 AgriPest-YOLOv5n。本研究提出的 AgriPest-YOLOv5n 主要包括以下 3 个方面的改进:第一,在原始模型骨干网络的 C3 模块中引入坐标注意力机制,使模型能够准确定位目标位置;第二,将 YOLOv5s 中的 PAN 结构替换为 BiF-PN 结构,提升模型对小目标的识别能力;第三,使用损失函数 SIoU 替代损失函数 CIoU,在不改变参数量的情况下加速模型的收敛,提升轻量级模型的检测性能。本研究提出的方法能够有效提取复杂背景下农业病虫害特征,具有较高检测精度和速度。在自制玉米病虫害数据集上, AgriPest-YOLOv5n 模型的  $mAP@0.50$  为 81.32%,在 IP102 公开数据集上, AgriPest-YOLOv5n 模型的  $mAP@0.50$  为 78.15%。 AgriPest-YOLOv5n 模型在 Jetson Xavier 开发板上的检测速度为 76.92 FPS。与其他轻量级模型相比, AgriPest-YOLOv5n 模型具有较大的优势。在今后的研究工作中,还可以进一步扩展样本库,使得模型能检测出病虫害的类型和名称,同时可以使模型进一步轻量化,并提升模型对目标的检测精度。

### 参考文献:

- [1] TIAN Y N, WANG S H, LI E, et al. MD-YOLO: multi-scale dense YOLO for small target pest detection[J]. Computers and Electronics in Agriculture, 2023, 213: 108233.
- [2] 王吉芸, 胡彦, 杨春震, 等. 云南省芒果主要病虫害发生与绿色防控措施概述[J]. 中国植保导刊, 2023, 43(11): 60-64.
- [3] 廖娟, 陶婉琰, 臧英, 等. 农作物病虫害遥感监测关键技术研究进展与展望[J]. 农业机械学报, 2023, 54(11): 1-19.
- [4] NGUGI L C, ABELWAHAB M, ABO-ZAHHAD M. Recent advances in image processing techniques for automated leaf pest and disease recognition-a review[J]. Information Processing in Agriculture, 2021, 8(1): 27-51.
- [5] PADOL P B, SAWANT S D. Fusion classification technique used to detect downy and Powdery Mildew grape leaf diseases[C]// IEEE. 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication. Jalgaon: IEEE, 2016.
- [6] 高雄, 汤岩, 陈铁英, 等. 基于图像处理的甘蓝虫害识别研究[J]. 江苏农业科学, 2017, 45(23): 235-238.
- [7] 李海, 李宜骏, 陈诗果, 等. 苹果树病虫害智能识别系统设计与实现[J]. 科学技术与工程, 2021, 21(25): 10639-10645.
- [8] DEY A K, SHARMA M, MESHRAM M R. Image processing based leaf rot disease, detection of betel vine (*Piper betle* L.)[J]. Procedia Computer Science, 2016, 85: 748-754.
- [9] TIAN Y N, YANG G D, WANG Z, et al. Instance segmentation of

- apple flowers using the improved mask R - CNN model[J]. Biosystems Engineering,2020,193:264-278.
- [10] WANG X W, ZHAO Q Z, JIANG P, et al. LDS-YOLO: a light-weight small object detection method for dead trees from shelter forest[J]. Computers and Electronics in Agriculture,2022,198: 107035.
- [11] 王铭慧,张怀清,樊江川,等. 基于深度学习网络实现番茄病虫害检测与识别[J]. 中国农业大学学报,2023,28(11):165-181.
- [12] 李就好,林乐坚,田凯,等. 改进 Faster R-CNN 的田间苦瓜叶部病害检测[J]. 农业工程学报,2020,36(12):179-185.
- [13] WANG R J, JIAO L, XIE C J, et al. S-RPN: sampling-balanced region proposal network for small crop pest detection[J]. Computers and Electronics in Agriculture,2021,187:106290.
- [14] 宋中山,汪进,郑禄,等. 基于二值化的 Faster R-CNN 柑橘病虫害识别研究[J]. 中国农机化学报,2022,43(6):150-158.
- [15] 张善文,邵彧,齐国红,等. 基于多尺度注意力卷积网络的作物害虫检测[J]. 江苏农业学报,2021,37(3):579-588.
- [16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//IEEE/CVF. 2016 IEEE conference on computer vision and pattern recognition. Las Vegas, NV: IEEE,2016.
- [17] 吕石磊,卢思华,李震,等. 基于改进 YOLOv3-LITE 轻量级神经网络的柑橘识别方法[J]. 农业工程学报,2019,35(17): 205-214.
- [18] 储鑫,李祥,罗斌,等. 基于改进 YOLOv4 算法的番茄叶部病害识别方法[J]. 江苏农业学报,2023,39(5):1199-1208.
- [19] 周维,牛永真,王亚伟,等. 基于改进的 YOLOv4-GhostNet 水稻病虫害识别方法[J]. 江苏农业学报,2022,38(3):685-695.
- [20] 骆润玫,殷惠莉,刘伟康,等. 基于 YOLOv5-C 的广佛手病虫害识别[J]. 华南农业大学学报,2023,44(1):151-160.
- [21] XUE Z Y, XU R J, BAI D, et al. YOLO-tea: a tea disease detection model improved by YOLOv5[J]. Forests,2023,14(2):415.
- [22] 张友为,王鑫鑫,范晓飞. 基于深度学习的玉米和番茄病虫害检测技术研究进展[J]. 江苏农业科学,2024,52(10):10-20.
- [23] 吴子炜,夏芳,陆林峰,等. 基于改进 YOLOv5 的水稻主要害虫识别方法[J]. 江苏农业科学,2023,51(21):218-224.
- [24] 路阳,刘婉婷,林立媛,等. CNN 与 BiLSTM 相结合的水稻病害识别新方法[J]. 江苏农业科学,2023,51(20):211-217.
- [25] WU X P, ZHAN C, LAI Y K, et al. Ip102: a large-scale benchmark dataset for insect pest recognition[C]//IEEE/CVF. 2019 IEEE/CVF conference on computer vision and pattern recognition. Long Beach, CA: IEEE,2019.
- [26] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//IEEE/CVF. 2018 IEEE conference on computer vision and pattern recognition. Salt Lake City, UT: IEEE,2018.
- [27] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//IEEE/CVF. 2021 IEEE/CVF conference on computer vision and pattern recognition. Nashville, TN: IEEE,2021.
- [28] TAN M X, PANG R M, LE Q V. Efficientdet: Scalable and efficient object detection[C]//IEEE/CVF. 2020 IEEE/CVF conference on computer vision and pattern recognition. Seattle, WA: IEEE,2020.
- [29] CHEN F X, ZHANG L X, KANG S Y, et al. Soft-NMS-enabled YOLOv5 with SIOU for small water surface floater detection in UAV-captured images[J]. Sustainability,2023,15(14):10751.

(责任编辑:成纾寒)