

基于 CBP-YOLO 的玉米田间草地贪夜蛾侵染痕迹检测

齐江涛^{1,2,3}, 王 硕^{1,2,3}, 熊悦淞^{1,2,3}, 高芳芳^{1,2,3}, 王发赢^{1,2,3},
邹宗峰⁴, 刘英智⁴, 刘慧力^{1,2,3*}

(1. 吉林大学工程仿生教育部重点实验室, 长春 130022; 2. 吉林大学生物与农业工程学院, 长春 130022; 3. 吉林省智慧农业装备与技术重点实验室, 长春 130022; 4. 烟台市农业技术推广中心, 烟台 264001)

摘 要: 针对无人机检测玉米田间草地贪夜蛾侵染痕迹时, 因目标尺寸小、尺度变化大、特征表征不明显导致识别精度低的问题, 该研究提出了一种面向无人机多尺度图像的侵染痕迹检测模型 CBP-YOLO (coordinated-BiFPN-P2-YOLO, CBP-YOLO)。通过引入真实增强超分辨率生成对抗网络 (real-enhanced super-resolution generative adversarial network, Real-ESRGAN) 重建原始低分辨率图像, 恢复叶片啃食区域的纹理, 在 YOLOv8 主干网络嵌入协调注意力机制 (coordinated attention, CA), 提升小目标局部细节提取与全局位置信息感知能力, 在颈部网络引入双向特征金字塔网络 (Bi-directional feature pyramid network, BiFPN), 降低跨层级特征信息损耗, 增强多尺度目标特征表达一致性, 新增了尺寸为 160×160×64 的小目标检测头, 拓展模型对小尺度目标的特征感知能力。结果表明, 地面采样距离 (ground sampling distance, GSD) 为 0.38cm/像素时检测性能最佳, Real-ESRGAN 有效弥补了分辨率不足引起的特征信息缺失。CBP-YOLO 在无人机多尺度叶片数据集上 mAP@0.5 达到 76.5%, 较原模型提升 3.4 个百分点, 较 YOLOv9m、YOLOv10m、YOLOv11m、Faster R-CNN 和 RetinaNet 分别提升 10.1、7.2、5.1、9.3 和 17.9 个百分点, 验证了其对于玉米田间草地贪夜蛾侵染痕迹的检测能力。该研究为无人机检测大田作物病害提供了参考。

关键词: 无人机; 深度学习; YOLOv8; 草地贪夜蛾; 智慧农业; 玉米叶片

doi: 10.11975/j.issn.1002-6819.202509276

中图分类号: S24

文献标志码: A

文章编号: 1002-6819(2026)-12-0195-09

齐江涛, 王硕, 熊悦淞, 等. 基于 CBP-YOLO 的玉米田间草地贪夜蛾侵染痕迹检测[J]. 农业工程学报, 2026, 42(12): 195-203. doi: 10.11975/j.issn.1002-6819.202509276 <http://www.tcsae.org>
QI Jiangtao, WANG Shuo, XIONG Yuesong, et al. Detecting *Spodoptera frugiperda* infestation traces in maize fields using CBP-YOLO[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2026, 42(12): 195-203. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202509276 <http://www.tcsae.org>

0 引 言

草地贪夜蛾 (*Spodoptera frugiperda*) 自入侵中国以来已成为玉米生产的主要害虫, 常导致玉米大幅减产, 并对全球粮食安全构成潜在威胁^[1-3]。及时、准确地检测玉米叶片是否受草地贪夜蛾侵染, 有利于草地贪夜蛾的防治和筛选优质抗虫品种, 对于提高农业生产效率至关重要^[4]。然而, 当前大多数虫害检测仍依赖人工田间观察, 存在效率低、主观性强等问题。不同观察者因经验和环境差异可能给出截然不同的结果^[5]。相比之下, 基于图像的自动化监测手段有望克服人工巡查的局限, 为农业有害生物监测提供更稳定可靠的技术支撑。无人机由于其覆盖广泛和适应性强的优势, 被广泛用于农田病虫害监测^[6]。

ISHENGOMA 等^[7]采用基于无人机航拍与卷积神经网络的识别方案, 并结合 Shi Tomasi 角点检测进行图像增强优化, 显著改善了受草地贪夜蛾侵染玉米叶片的细节表达能力, 有效提高模型对虫害区域的聚焦能力, 并实现近乎 100% 的识别率。张国忠等^[8]针对藕田病虫害诊治问题, 采用分支结构对模型的浅层特征提取模块进行改进, 在模型中引入了挤压与激励机制与锐化余弦卷积, 并基于 Plantvillage^[9]数据集进行迁移学习实现了 91.34% 的莲藕病虫害识别准确率。目标检测模型作为计算机视觉中的重要任务之一, 能够同时实现病虫害的类别识别与空间位置定位。当前主流的目标检测模型主要可分为两类: 双阶段目标检测算法和单阶段目标检测算法。双阶段目标检测算法首先通过一个区域提议网络从输入图像中生成一系列可能包含目标的候选区域, 在这些候选区域的基础上, 进一步提取更精细的特征, 以此执行类别判别与边界框的精确定位, 主要代表为 Faster R-CNN^[10]。但由于其双阶段处理机制, 整体推理速度较慢, 限制了其在高实时性要求的场景应用。单阶段目标检测算法则无需显示生成候选区域, 一次前向推理直接输出图像中所有的目标的类别标签及其对应的边界框坐标。该类算法的代表为 YOLO 系列^[11-13]和 RetinaNet^[14]等。在保证较高检测精度的同时, 兼具良好的推理速度, 能够满足无人机等资源受限设备的部署需求, 并契合病虫

收稿日期: 2025-09-29 修订日期: 2026-03-31

基金项目: 吉林省科技发展规划项目 (吉林省智慧农业装备与技术重点实验室: YDZJ202502CXJD004)

作者简介: 齐江涛, 博士, 教授, 博士生导师, 研究方向为仿生智能农机装备。Email: qijiangtao@jlu.edu.cn

*通信作者: 刘慧力, 高级工程师, 研究方向为智能农机装备。

Email: hlliu@jlu.edu.cn

中国农业工程学会会员: 齐江涛 (E041201131S)

害防治中对时效性的严格要求。

YOLO 在面对无人机航拍病虫害图像面临着很多挑战,翁海勇等^[15]为了提升 YOLOv7 模型捕捉穗颈瘟斑细粒度特征的能力,引入了压缩注意力机制 SE 和可变形卷积,改进后的模型能够自适应调整感受野,较好地实现了不同水稻品种穗颈瘟抗性的区分,为水稻黄熟期育种中穗颈瘟抗性评估提供了参考。BAO 等^[16]利用残差通道注意力网络 RCAN 对无人机航拍图像进行超分辨率增强,并结合改进的 DDMA-YOLO 模型实现了茶叶病害的检测,平均精度提升了 7.8 个百分点。侯文慧等^[17]提出 G-YOLOv8 模型,结合 SRGAN 超分辨率增强、三重注意力机制及小目标检测层设计,并利用迁移学习缓解样本不足问题,在多类番茄病害检测中将平均精度提升至 97.8%,显著增强了小目标与相似病害的识别能力。马超伟等^[18]提出了一种基于改进 YOLOv8 的轻量化小麦病害检测方法,引入深度可分离卷积结构和全局注意力机制,提高了模型对重要特征的提取能力并减小了模型的参数量,为小麦病害检测无人机等移动端检测装备的部署和应用提供了参考。

尽管国内外学者已围绕 YOLO 系列模型的无人机航拍病虫害检测开展了大量的研究,并在提升检测精度与模型鲁棒性方面取得了一定成果,但在复杂农田场景下因小目标检测、多尺度特征融合和特征表征导致检测精度低的研究仍存在不足。针对上述问题本研究基于 YOLOv8 提出一种面向多尺度航拍图像的草地贪夜蛾侵染叶片痕迹检测模型 CBP-YOLO。旨在提升无人机在玉米田间检测草地贪夜蛾侵染痕迹的精度。

1 材料与方法

1.1 材料

1.1.1 数据来源

本研究于 2024 年 7 月 20 日在山东省烟台市烟台农业科学院采集数据。使用搭载固定焦距相机的 DJI P4 多光谱无人机采集可见光图像,相机分辨率为 1 600×1 300 像素,水平视场角 62.7°,焦距 5.74 mm。为兼顾空间分辨率与地面覆盖,飞行高度设置 3、5 和 7 m 三个高度,数据采集时间为 2024 年 8 月 24 日 09:00-12:00,光照稳定。地面采样距离 (ground sampling distance, GSD) 是从无人机的元数据中推导出来的,代表每个像素对应的地面距离^[19],该参数与飞行高度、传感器分辨率及相机视场角关联,按式 (1) 计算,可精确计算各高度下采集图像的空间分辨率。

$$GSD = \frac{2H \tan\left(\frac{\theta}{2}\right)}{N} \quad (1)$$

式中 H 为飞行高度, m; θ 为水平视场角, (°); N 为图像水平像素数。

本研究采集无人机航拍图像包括 3 种 GSD 分别为 0.22、0.38 和 0.53 cm/像素,图像数量分别为 490、269 和 120 幅。玉米田间草地贪夜蛾侵染斑块在不同的 GSD 图像中呈现显著的像素尺寸差异。田间玉米正处于大喇叭口期至抽雄期之间,叶片幼嫩多汁,易受草地贪夜蛾幼虫啃食。侵染后叶片形成具有典型识别特征的侵染痕

迹,主要呈现为叶片上出现不规则孔洞、锯齿状缺刻和大面积啃食区域,严重时可导致叶片断裂,数据集中包含土壤裸露区、杂草、田间阴影以及光照不均匀等多种复杂背景干扰因素 (图 1)。

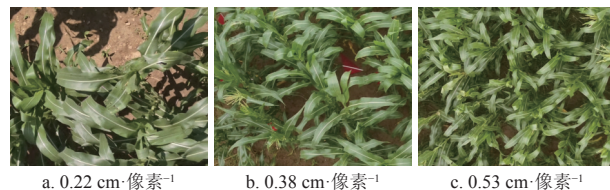


图 1 不同地面采样距离的典型图像样本

Fig.1 Typical image samples at different ground sampling distances

为扩充数据集并便于标注,将每幅图像均匀划分为 4 个不重叠的 800×650 像素子图,确保无重复区域。随后利用开源工具 LabelImg 对草地贪夜蛾侵染的叶片区域进行人工框选标注。剔除无标注目标的图像后,最终保留 879 张有效标注图像,按照 8:1:1 的比例划分为训练集、验证集以及测试集,如表 1。

表 1 无人机航拍图像数据集
Table 1 UAV aerial image dataset

数据集 Dataset	地面采样距离 Ground sampling distance/ (cm·像素 ⁻¹)	图像数量 Number of images	共计 Total	标注数量 Number of annotations
训练集 Train set	0.22	393	703	2 749
	0.38	213		
	0.53	97		
验证集 Val set	0.22	46	88	305
	0.38	31		
	0.53	11		
测试集 Test set	0.22	51	88	336
	0.38	25		
	0.53	12		

1.1.2 超分辨率图像预处理

GSD=0.53 cm/像素的图像虽然覆盖范围较大,能够在一次飞行中获取更大面积的农田信息,但其固有的低分辨率特性使得目标对象在图像中的像素数量明显减少。由此导致草地贪夜蛾侵染玉米叶片损伤位置目标微小、边缘模糊、纹理细节缺失,影响后续目标检测模型对关键特征的提取,为缓解低分辨率带来的信息损失,本研究引入了超分辨率重建技术对原始图像进行预处理,以提升输入数据的空间细节表达能力。

早期超分辨率 (super-resolution, SR) 算法多基于插值或浅层模型,其重建结果往往过于平滑,难以恢复高频细节。为提升视觉真实性,ESRGAN^[20]引入生成对抗网络 (generative adversarial network, GAN) 框架,通过感知损失与对抗训练生成更具纹理细节的图像。但其仍基于简单退化模型,且在处理复杂真实退化时表现受限,包括模糊、噪声、压缩伪影共存等。针对上述问题,本文采用真实增强超分辨率生成对抗网络 (real enhanced super-resolution generative adversarial network, Real-ESRGAN)^[21]作为超分辨率重建核心算法,Real-ESRGAN 是一种面向真实场景的单图像超分辨率方法,其结合残差密集块 (residual in residual dense blocks, RRDBs) 生成器与 U-Net 判别器,在恢复高频细节的同时有效抑制伪影,从而实现对真实低分辨率图像的高质

量重建。并引入多次模糊、下采样、噪声和 JPEG 压缩等阶退化过程合成更贴近实际的训练数据，计算式如下（式中分别以 I_b 、 I_n 、 I_s 、 I_{LR} 表示）。

$$I_b = \text{Blur}(I; k, \sigma) \quad (2)$$

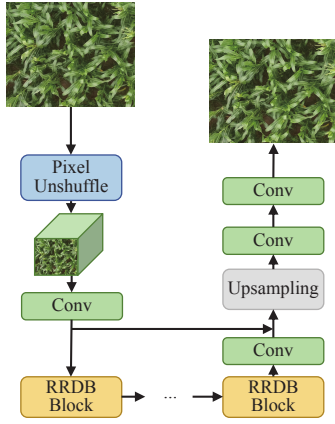
$$I_n = \text{Noise}(I_b) \quad (3)$$

$$I_s = \text{Down}(I_n; r_k) \quad (4)$$

$$I_{LR} = \text{JPEG}(I_s; Q) \quad (5)$$

式中 I 表示输入图像， $\text{Blur}(I; k, \sigma)$ 表示模糊操作，其中 k 为模糊核大小， σ 为核参数。 $\text{Noise}(I)$ 表示添加加性高斯白噪声。 $\text{Down}(I_n; r_k)$ 表示按比例 r_k 进行下采样。 $\text{JPEG}(I_s; Q)$ 表示质量因子为 Q 的 JPEG 压缩操作。

Real-ESRGAN 支持 $\times 2$ 放大倍率（图 2），在输入送入 ESRGAN 主干网络前，引入 Pixel Unshuffle 操作^[22]对输入图像进行预处理，该操作通过将空间分辨率降低为原来的 $1/2$ ，同时将通道数扩展为原来的 4 倍，使主干网络的大部分卷积运算在低维特征空间中完成，显著减少 GPU 显存占用与计算成本。生成器主干采用 RRDB 结构，RRDB 通过多层残差与密集连接的嵌套设计，强化了特征的逐层复用能力与梯度传播效率。为了缓解 GAN 训练中常见的过度锐化与振铃伪影问题，在判别器中引入了谱归一化^[23]正则化对卷积权重进行约束，在细节增强与视觉真实性之间取得良好平衡。



注：Pixel Unshuffle 为像素反排列操作模块；Conv 为卷积层；RRDB Block 为残差中的残差密集块；Upsampling 为上采样模块。
Note: Pixel Unshuffle denotes the pixel unshuffle operation module; Conv denotes the convolution layer; RRDB Block denotes the Residual-in-Residual Dense Block; Upsampling denotes the upsampling module.

图 2 Real-ESRGAN 超分辨率重建算法架构

Fig.2 Architecture of the Real-ESRGAN super-resolution reconstruction algorithm

本研究对 GSD 为 0.53 cm/像素的无人机图像进行 Real-ESRGAN 处理后，图像尺寸由 800×650 像素提升至 $1\,600 \times 1\,300$ 像素，实现了空间分辨率的增强。在无人机飞行高度 H 、水平视场角 θ 均不变的条件下，成像几何关系未发生改变，仅图像水平像素数 N 提升为原来的 2 倍，根据 GSD 的定义，对应的 GSD 减小为原来的一半，即 $\text{GSD}=0.26 \text{ cm/像素}$ （图 3）。

随后，将重建后的图像均匀划分为分辨率为 800×650 像素的 4 个子图像。在划分过程中，部分标注框因切割线被分割而生成新的标注，经人工核查并剔除不合

理的标注，同时删除了无标注框的图像，最终生成了一个 GSD 为 0.26 cm/像素的新数据子集，并将其定义为 $\text{GSD}=0.26 \text{ cm/像素}$ 的超分辨率重建数据集，该数据集与原始采集的 $\text{GSD}=0.22$ 、0.38 和 0.53 cm/像素图像一同进行后续处理与分析，具体数据划分如表 2 所示。

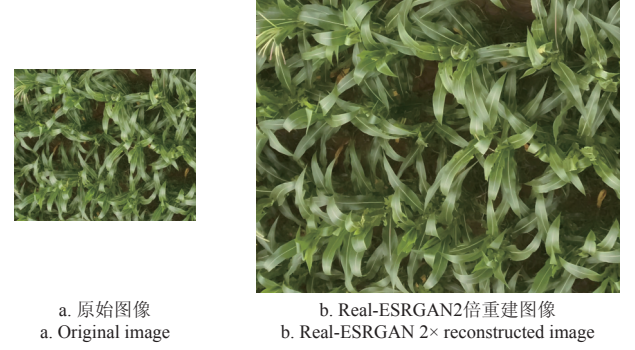


图 3 Real-ESRGAN 重建高分辨率图像

Fig.3 Real-ESRGAN reconstructed high-resolution image

表 2 超分辨率重建数据集统计

数据集 Dataset	地面采样距离 Ground sampling distance/ (cm·像素 ⁻¹)	图像数量 Number of images	标注数量 Number of annotations
训练集 Train	0.26	363	682
验证集 Val	0.26	39	64
测试集 Test	0.26	41	87

1.1.3 数据增强

样本的数量与质量对模型效果具有直接影响，充足且多样化的训练数据是提升模型泛化能力与检测性能的重要基础。草地贪夜蛾属于迁飞性重大害虫，具有暴发性强、扩散速度快的特点，通常会提前采取化学防治以控制危害程度，因此田间受自然侵染的典型侵染痕迹样本留存时间短、数量有限。为缓解样本规模不足对模型训练的不利影响，本研究采用数据增强策略以真实地模拟自然环境下无人机成像条件的变化。本研究对数据集的 1 066 张训练集图像（703 张原始图像和 363 张超分辨率重建图像）进行了旋转、翻转、亮度与对比度调整、模糊及锐化等操作，将数据集扩展至 12 792 张图像。

1.2 方法

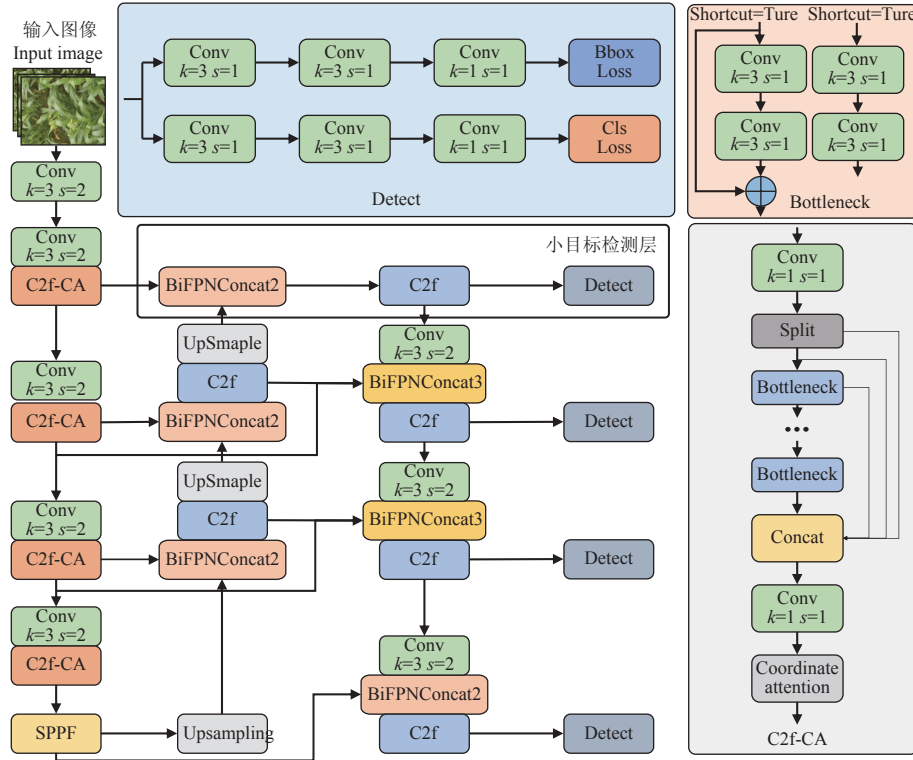
1.2.1 模型结构

YOLOv8 是 Ultralytics 公司于 2023 年发布的深度学习目标检测算法，其汲取了 YOLOv5、YOLOv7 等先前版本的优点，具有优异的性能^[24]。该版本提供了 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 和 YOLOv8x 等不同规模的模型，随着网络深度与宽度的增加，在检测精度和推理速度之间实现灵活平衡。综合考虑检测精度与推理速度后选取 YOLOv8m 作为基础模型，应用于无人机多高度航拍图像中的玉米田间草地贪夜蛾侵染痕迹检测。

YOLOv8 的整体结构由 Backbone、Neck 和 Head 共 3 个部分构成，分别承担特征提取、特征融合与目标检测的功能。针对草地贪夜蛾侵染玉米叶片图像中目标尺寸

小、尺度变化大、特征对比度低等问题,在 YOLOv8 m 模型上进行了 3 方面改进,在 Backbone 的 C2f 模块中引入协调注意力机制 (coordinated attention, CA) [25],形成 C2f-CA 结构,通过嵌入方向感知和通道加权策略,有效削弱背景噪声对模型的干扰,突出叶片损伤痕迹等关键区域特征,进而提升模型对细粒度特征的表达能力。

在 Neck 部分引入 BiFPN,通过加权双向融合机制强化深层次特征的信息互补,有效提升了模型在多尺度目标上的检测性能。最后在原有 3 个检测头的基础上新增一个浅层检测头,提高对小尺度目标的感知能力,以此减少漏检和误检的发生。基于上述结构改进,本文提出 CBP-YOLO 模型 (图 4)。



注: k 为卷积核的尺寸、 s 为步长; C2f-CA 为带坐标注意力的跨阶段部分连接模块; SPPF 为快速空间金字塔池化模块; BiFPNConcat2 为双向特征金字塔融合模块第二级; C2f 为跨阶段部分连接模块; BiFPNConcat3 为双向特征金字塔融合模块第三级; Detect 为检测头; Bbox Loss 为边界框损失; Cls Loss 为分类损失; Shortcut 为跳跃连接; Split 为特征分割操作; Concat 为特征拼接操作。

Note: k and s are the kernel size and stride, respectively; C2f-CA is the cross stage partial module with coordinate attention; SPPF is the spatial pyramid pooling fast module; BiFPNConcat2 is the second-level bidirectional feature pyramid fusion module; C2f is the cross stage partial module; BiFPNConcat3 is the third-level bidirectional feature pyramid fusion module; Detect is the detection head; Bbox Loss is the bounding box regression loss; Cls Loss is the classification loss; Shortcut is the skip connection; Split is the feature splitting operation; and Concat is the feature concatenation operation.

图 4 CBP-YOLO 模型结构
Fig.4 CBP-YOLO model structure

1.2.2 C2f-CA 特征提取模块

在光照变化、阴影及遮挡等复杂环境下,引入注意力机制能够提升模型对输入图像中关键特征的感知能力,减弱背景噪声和无关信息的干扰。挤压与激励模块 (squeeze-and-excitation, SE) 是一类常用的注意力机制,通过全局平均池化 (global average pooling, GAP) 计算各通道的重要性权重,并利用自适应方式调整通道间的相对权重,以突出关键特征。GAP 操作将空间维度压缩为单一标量,细粒度空间信息丢失,不同空间位置之间的关联也被忽略。

卷积块注意力模块 (convolutional block attention module, CBAM) 在通道注意力的基础上引入大尺度卷积核提取空间特征,实现通道注意力与空间注意力的结合。CBAM 主要关注局部特征关联,对中远距离之间的相关性的建模能力有限,并且依赖全连接层计算注意力,参数规模和计算开销显著增加,实时检测效率受限。

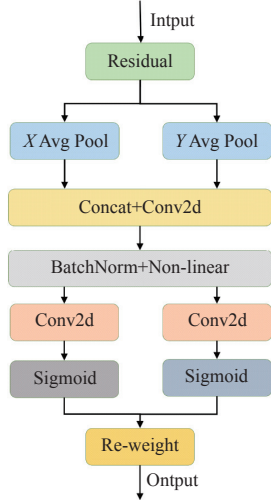
CA 将通道注意力分解为两个独立的一维特征编码过程,一条路径沿空间维度捕捉长程依赖,另一条路径

保留精确位置信息 (图 5)。CA 注意力机制通过两步过程同时建模通道信息与位置信息。第一步为坐标信息嵌入与坐标注意力生成。输入特征图尺寸为 $C \times H \times W$, 在水平方向和垂直方向进行特征聚合,在保留位置信息的同时捕捉长程依赖。该过程分别沿高度和宽度维度对特征进行编码,生成两个独立的特征图 z_c^h 和 z_c^w 如式 (6) 和式 (7) 所示。该分解方式使注意力模块在保持空间定位精度的同时获取全局上下文信息,有效弥补传统通道注意力在压缩过程中导致的空间细节缺失。

$$z_c^h = \frac{1}{W} \sum_{i=0}^{W-1} X_c(h, i) \quad (6)$$

$$z_c^w = \frac{1}{H} \sum_{j=0}^{H-1} X_c(j, w) \quad (7)$$

式中 $X_c(h, i)$ 和 $X_c(j, w)$ 分别表示通道 c 在空间坐标 (h, i) 和 (j, w) 处的取值, W 和 H 分别为特征图的宽度和高度, z_c^h 与 z_c^w 为沿高度和宽度维度聚合得到的上下文描述子。



注：Residual 为残差连接模块；X Avg Pool 为 X 方向全局平均池化；Y Avg Pool 为 Y 方向全局平均池化；Concat+Conv2d 为拼接与二维卷积；BatchNorm+Non-linear 为批归一化与非线性激活；Conv2d 为二维卷积层；Sigmoid 为激活函数；Re-weight 为通道重标定。
Note: Residual is the residual connection module; X Avg Pool is the global average pooling along the X-axis; Y Avg Pool is the global average pooling along the Y-axis; Concat+Conv2d is the concatenation and 2D convolution operation; BatchNorm+Non-linear is the batch normalization and non-linear activation operation; Conv2d is the 2D convolution layer; Sigmoid is the activation function; and Re-weight is the channel re-weighting operation.

图 5 CA 注意力机制模块结构

Fig.5 CA Attention mechanism module structure

将聚合后的特征图拼接后输入共享的 1×1 卷积层 F_1 进行变换，并经非线性激活函数处理，得到中间特征矩阵 $f \in R^{(C/r \times H \times W)}$ ，其中 r 为压缩比例系数， z_w 和 z_h 分别为宽度和高度方向的特征张量，如式 (8) 所示。

$$f = \delta(F_1([z_h, z_w])) \quad (8)$$

将中间特征矩阵 f 沿空间维度分解为两个张量 f_h 和 f_w ，采用不同的 1×1 卷积变换 F_h 和 F_w ，并经 Sigmoid 激活函数处理，最后得到水平与垂直方向的注意力权重 g_h 和 g_w ，如式 (9) 和 (10) 所示。

$$g_h = \sigma(F_h(f_h)) \quad (9)$$

$$g_w = \sigma(F_w(f_w)) \quad (10)$$

最后，该注意力模块的输出可表示为

$$y_c(i, j) = x_c(i, j) \cdot g_c^h(i) \cdot g_c^w(j) \quad (11)$$

式中 $y_c(i, j)$ 表示经过 CA 注意力加权后输出特征图在 (i, j) 的特征值； $x_c(i, j)$ 表示输入特征图在 (i, j) 处的特征值； $g_c^h(i)$ 和 $g_c^w(j)$ 分别表示在高度和宽度方向上的第 i 和第 j 个位置的注意力权重。

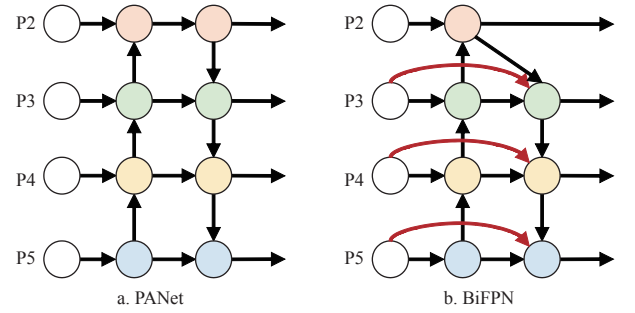
1.2.3 BiFPN 架构

本研究构建的模型在不同尺度上设置了 4 个检测头，分别对应 $160 \times 160 \times 64$ 、 $80 \times 80 \times 128$ 、 $40 \times 40 \times 256$ 和 $20 \times 20 \times 512$ 的特征图，以适应无人机图像中目标尺度差异显著的检测需求。特征金字塔网络 (feature pyramid network, FPN) [26] 最初用于缓解单尺度特征提取器在处理不同尺寸目标时的不足，通过融合低层细节与高层语义特征，提高模型对多尺度信息的表征能力。YOLOv8 模型采用路径聚合网络 (path aggregation network, PANet) [27]，能够捕获不同层级的特征并促进跨尺度融合。PANet 实现了跨尺度信息连接，提升了模型

对不同尺寸目标的适应性 (图 6a)，但其主要依赖低层细节与高层语义的简单拼接，融合效果有限。为进一步增强不同尺度目标之间的特征交互能力，本研究在 Neck 部分引入加权双向特征金字塔网络 (Bi-directional feature pyramid network, BiFPN) [28] 进行特征融合。BiFPN 采用双向跨尺度连接，并引入可学习权重对不同尺度特征进行动态调节，可实现高层语义信息与低层细节信息的高效融合 (图 6b)，其计算过程如式 (12) 所示。

$$\hat{P}_i = \frac{\sum_j w_j \cdot P_j}{\sum_j w_j + \epsilon} \quad (12)$$

式中 $\hat{P}_i$ 为第 i 层融合后的特征图； P_j 为第 j 层的输入特征图 w_j 为对应 P_j 的非负可学习权重； ϵ 是为避免除零而加入的常数。



注：P2 是尺寸为 $160 \times 160 \times 16$ 的特征图；P3 是尺寸为 $80 \times 80 \times 128$ 的特征图；P4 是尺寸为 $40 \times 40 \times 256$ 的特征图；P5 是尺寸为 $20 \times 20 \times 512$ 的特征图。
Note: P2 is a feature map with a size of $160 \times 160 \times 16$; P3 is a feature map with a size of $80 \times 80 \times 128$; P4 is a feature map with a size of $40 \times 40 \times 256$; P5 is a feature map with a size of $20 \times 20 \times 512$.

图 6 PANet 和 BiFPN 的特征网络结构

Fig.6 Feature network architecture of FPN and BiFPN

1.2.4 面向小目标的检测头扩展

无人机航拍图像在不同 GSD 下的目标尺寸差异较大，原始模型有 3 个检测头：P3 检测头提取尺寸 $80 \times 80 \times 128$ 的浅层特征图，P4 检测头提取尺寸 $40 \times 40 \times 256$ 的中层特征图，P5 检测头提取尺寸为 $20 \times 20 \times 512$ 的深层特征图。小目标通常缺乏纹理和边界特征，在 Backbone 网络的多次下采样过程中，其细节逐渐丢失，检测性能受到影响。浅层特征图具有较高分辨率和较小感受野，更适合小目标检测。为提高检测能力，在 Neck 部分增加了尺寸为 $160 \times 160 \times 64$ 的 P2 小目标检测头，该检测头可检测对应于 640×640 图像中 4×4 像素感受野小目标 (图 7)。

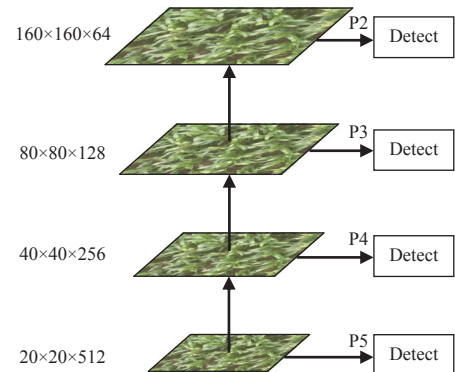


图 7 增加小目标检测头示意图

Fig.7 Schematic diagram of adding a small target detection head

1.3 试验环境及评价指标

本研究的试验硬件平台为一台运行 Windows 10 操作系统的工作站, 配置 Intel(R) Core(TM) i9-13900K CPU、128 GB 内存和 NVIDIA GeForce RTX 4090 显卡。深度学习框架采用 PyTorch 2.2.1, 并结合 cuDNN 12.1 进行加速以提升训练效率。模型输入图像大小设置为 640×640 像素, 训练轮数设置为 100, Batch size 为 8, 训练优化器为 AdamW, 初始学习率为 0.01, 动量为 0.937, 权重衰减系数为 0.0005。

本研究采用精确率 (precision)、召回率 (recall)、平均精度 (average precision, AP) 以及 F1-score 对所提模型的检测性能进行评价。其中, AP 是评估深度学习中目标检测模型的重要评估指标。当检测边界框与真实标注框的交并比 (intersection over union, IoU) 超过指定阈值时, 检测结果记为正确, AP@0.5 表示在 IoU 阈值为 0.5 时计算的平均精度, AP@0.5:0.95 表示在 IoU 阈值从 0.5 到 0.95、步长为 0.05 时的平均精度。

2 结果与分析

2.1 不同地面采样距离对比试验分析

为评估 GSD 对模型检测性能的影响, 本研究采用多尺度数据集进行统一训练。改进的 CBP-YOLO 模型分别

在 GSD 为 0.22、0.38 和 0.53 cm/像素的图像上进行了训练, 各 GSD 水平下的总体检测结果见表 3。

表 3 CBP-YOLO 模型在不同 GSD 图像上的检测结果

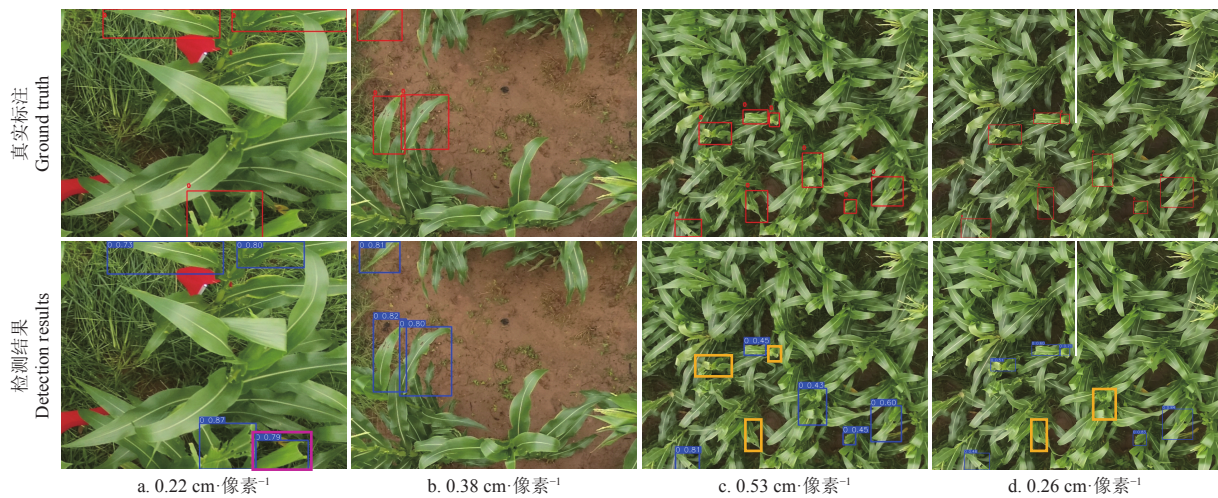
Table 3 Detection results of CBP-YOLO model for different GSD images

GSD/ (cm·像素 ⁻¹)	精确率 Precision/%	召回率 Recall/%	AP@0.5/ %	AP@0.5:0.95/%	F1 分数 F1-score/%
0.22	69.2	72.8	74.5	43.1	70.9
0.38	78.5	70.5	79.8	44.4	74.3
0.53	57.0	54.8	56.8	31.5	55.9
0.26	68.2	61.2	64.1	35.8	64.5

注: AP@0.5 为交并比阈值为 0.5 时的平均精度; AP@0.5:0.95 为交并比阈值从 0.5 到 0.95 (步长为 0.05) 时的平均精度。

Note: AP@0.5 is the average precision when the IoU threshold is 0.5; AP@0.5:0.95 is the average precision when the IoU thresholds ranges from 0.5 to 0.95 with a step size of 0.05.

在 GSD=0.38 cm/像素的数据集上, 模型取得最佳检测效果, 各项指标均处于较高水平, 其中 AP@0.5 和 F1-score 分别达到了 79.8% 和 74.3%, 该分辨率在空间覆盖范围与局部细节表达能力之间达到了有效平衡。图 8 展示了模型在不同 GSD 中的检测效果图, 清晰呈现叶片损伤区域的结构特征与纹理特征, 又保留了足够的局部上下文信息, 使模型能够同时利用精细局部特征与合理的空间结构信息进行准确判别。



注: 红色框表示真实标注框; 蓝色框表示模型检测框; 黄色框表示漏检; 粉色框表示误检。

Note: The red boxes are the ground truth bounding boxes; the blue boxes are the model predicted bounding boxes; the yellow boxes are the missed detections; the pink boxes are the false detections.

图 8 不同 GSD 检测结果可视化对比

Fig.8 Visualization comparison of detection results under different GSD

相比之下, GSD=0.22 cm/像素的数据集虽然包含更细致的纹理, 但高分辨率同时引入了更多噪声和无关信息, 尤其在叶片边缘和背景区域, 对检测精度产生负面影响。其精确率和召回率略有提高, 但整体 AP@0.5 和 F1-score 均低于 0.38 cm/像素水平, 说明过高的分辨率在捕捉细节的同时可能削弱模型稳定性。

而在 GSD=0.53 cm/像素的数据集上, 由于采样距离较大、分辨率较低, 图像细节丢失并伴随背景噪声增加, 导致整体检测性能明显下降, AP@0.5 和 F1-score 分别只有 56.8% 和 55.9%, 尽管较大的 GSD 提高了图像采集效率, 但对检测精度产生不利影响。

为了消除这种影响, 本研究对该 GSD 数据集应用

了 Real-ESRGAN 算法提升图像分辨率。Real-ESRGAN 超分辨率算法能够有效提升无人机航拍的低分辨率图像的细节表达能力, 进而提高草地贪夜蛾侵染叶片的检测准确性。该策略有助于缓解因高飞行高度导致的图像退化问题, 强化虫害损伤区域的结构特征和纹理细节表达, 从而在一定程度上增强模型的整体性能。由表 3 可知, 与原始 GSD=0.53 cm/像素数据集相比, GSD 为 0.26 cm/像素数据集的 AP@0.5 提高 7.3 个百分点, F1-score 提高 8.6 个百分点。上述结果表明, 超分辨率算法能够有效改善图像质量, 特别在纹理与边缘特征方面的增强, 使模型对草地贪夜蛾啃咬叶片痕迹的检测能力得到提升。

综合来看, GSD=0.38 cm/像素的数据集表现最佳,

表明在玉米田间草地贪夜蛾侵染痕迹检测任务中，适当的图像分辨率能够在减少噪声的同时保留有效细节，从而提升模型性能。分辨率过高或过低均可能对检测结果产生负面影响，因此选择合理的 GSD 对于提高图像质量、模型鲁棒性和检测精度具有重要意义。

2.2 消融试验

为系统评估各模块对 YOLOv8 检测性能的贡献，本研究开展了消融试验，分别分析 CA、BiFPN 和 P2 模块的单独引入及其组合使用对模型性能的影响。试验结果如表 4 所示。

表 4 不同模块的消融试验
Table 4 Ablation experiments of different modules %

CA	P2	BiFPN	精确率 Precision	召回率 Recall	AP@0.5	AP@0.5:0.95	F1 分数 F1-score
×	×	×	71.4	69.2	73.1	41.5	69.9
√	×	×	68.5	70.38	74.1	41.7	69.4
×	√	×	71.0	71.8	75.4	42.6	71.4
×	×	√	75.5	70.4	73.8	41.9	72.9
√	√	×	71.3	72.9	75.7	42.9	72.1
×	√	√	69.8	72.0	75.5	42.4	70.9
√	×	√	69.2	70.0	73.9	42.0	69.6
√	√	√	74.7	72.4	76.5	43.2	73.5

注：“√”表示添加该模块，“×”表示未进行改动。
Note: “√” is the added module, and “×” is no modification.

表 4 展示了不同模块组合对模型检测性能的影响。在未引入任何模块时，模型表现为基准性能。仅使用 CA 模块时，AP@0.5 相比原模型提升了 1 个百分点，表明 CA 能够引导模型关注与虫害损伤相关的关键信息通道，抑制冗余特征，对特征提取过程起到了优化作用。当仅使用 BiFPN 时，模型达到了 73.8% 的 AP@0.5，表明 BiFPN 实现了不同尺度特征的高效融合，兼顾细节信息与高层语义信息，在多尺度目标检测中展现出明显优势。

当同时引入 CA 与 P2 时，精确率略降至 71.3%，但 AP@0.5 和 F1 分数有所提高，表明 CA 和 P2 两者协同增强了模型对细节信息的感知能力，但同时也增加了候选目标数量，导致误检略有增加，从而使精确率出现小幅下降。类似地，P2 提供的细粒度特征与 BiFPN 的多尺度特征融合机制相结合，使模型对不同尺度目标的检测更加均衡，精确率为 69.8%，AP@0.5 为 75.5%，F1 达到了 70.9%。最终，当启用所有三个模块时，模型取得了最佳综合性能，精确率为 74.7%，AP@0.5 提升至 76.5%，AP@0.5:0.95 为 43.2%，F1 为 73.5%，比原模型分别提升了 3.4、1.7、3.6 个百分点，检测效果对比如图 9 所示。多模块协同优化能够在增强特征表达能力的同时，实现精确率与召回率之间的有效平衡，整体性能优于其他模块组合。

2.3 不同模型对比试验分析

为验证所提出的 CBP-YOLO 模型在草地贪夜蛾侵染玉米叶片检测中的有效性，选取主流目标检测模型进行对比评估，结果见表 5。CBP-YOLO 模型在各项指标上均优于其他模型，精确率达到 74.7%、召回率为 72.4%、AP@0.5 为 76.5%、AP@0.5:0.95 为 43.2%，F1-score 为 73.5%，AP@0.5 相较于 YOLOv9 m、YOLOv10 m、YOLOv11 m、Faster R-CNN 和 RetinaNet 分别提升了 10.1、7.2、5.1、9.3 和 17.9 个百分点。结果表明，CBP-

YOLO 的精确率、AP@0.5 和 F1-score 高于 YOLO 系列及其他版本，体现了其在细节特征建模和精确定位方面的优势。

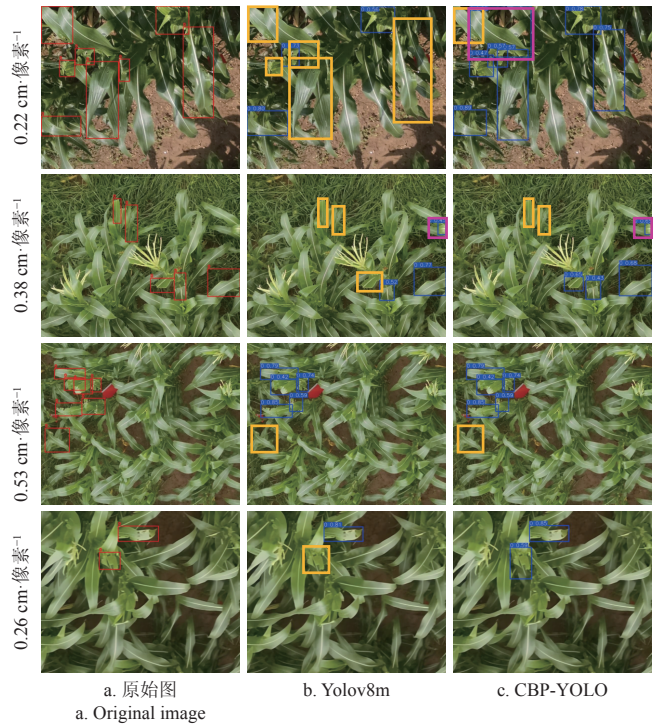


图 9 YOLOv8 m 与 CBP-YOLOv8 检测效果
Fig.9 Detection results of YOLOv8 m and CBP-YOLOv8

表 5 不同 YOLO 模型检测结果的比较
Table 5 Comparison of detection results of different YOLO models %

模型 Model	精确率 Precision	召回率 Recall	AP@0.5	AP@0.5:0.95	F1 分数 F1-score
YOLOv8 m	71.4	69.2	73.1	41.5	69.9
YOLOv9 m	66.3	67.9	66.4	34.8	67.1
YOLOv10 m	67.7	65.6	69.3	37.7	66.6
YOLOv11 m	73.1	60.4	71.4	38.8	66.1
Faster R-CNN	65.1	62.3	67.2	35.5	63.7
RetinaNet	58.8	66.1	58.6	32.7	62.2
CBP-YOLO	74.7	72.4	76.5	43.2	73.5

相比之下 RetinaNet、Faster R-CNN、YOLOv9 m 和 YOLOv10 m 的表现相对较差，尤其在精确率和 AP@0.5:0.95 上均明显低于 CBP-YOLO。YOLOv11 m 的精确率接近 YOLOv8 m，但其召回率较低，且 AP@0.5 也低于 CBP-YOLO，说明该模型在精确性与召回率之间尚未实现良好平衡。

总体来看，CBP-YOLO 在所有关键检测指标上均取得最优表现，较原始模型在 AP@0.5 和 F1-score 上有明显提升。该模型在叶片遮挡、光照变化及背景纹理干扰等田间场景条件下检测草地贪夜蛾侵染玉米叶片具有较高精度，可为玉米田间草地贪夜蛾侵染痕迹检测提供技术支持。

3 结 论

本研究针对无人机航拍视角下草地贪夜蛾侵染玉米叶片痕迹精准检测的需求，提出了一种面向玉米田间草地贪夜蛾的侵染痕迹检测模型 CBP-YOLO。主要结论如下：

1) 根据模型对多尺度目标识别鲁棒性的要求, 设计了无人机多高度飞行采样策略, 构建了多尺度的草地贪夜蛾侵染叶片图像数据集, 并针对模糊与细节丢失问题, 引入 Real-ESRGAN 进行超分辨率重建, 恢复了虫害侵染区域的纹理结构, 使模型对此类图像的 AP@0.5 提升了 7.3 个百分点。

2) 基于融合协调注意力 CA 机制, 对原始 C2f 模块进行重构, 以增强模型特征表征能力。引入了 BiFPN 替代传统 PANet 结构, 显著减少了多尺度特征传递中的信息损失, 同时增设 160×160、通道数为 64 的小目标检测头以提升小目标检测性能。试验结果表明, 在自建的数据集中 CBP-YOLO 检测模型 AP@0.5 达到了 76.5%, 相比于原模型, 提升了 3.4 个百分点。

[参 考 文 献]

- [1] 姜玉英, 刘杰, 朱晓明. 草地贪夜蛾侵入我国的发生动态和未来趋势分析[J]. 中国植保导刊, 2019, 39(2): 33-35.
- [2] 栗圣博, 周上朝, 董子舒, 等. 甜玉米生长后期草地贪夜蛾的为害情况调查[J]. 植物保护, 2022, 48(2): 227-231. LI Shengbo, ZHOU Shangchao, DONG Zishu, et al. Investigation on the damage characteristics of *Spodoptera frugiperda* at late developmental stage of sweet corn[J]. Plant Protection, 2022, 48(2): 227-231. (in Chinese with English abstract)
- [3] HUANG Y, LV H, DONG Y, et al. Mapping the spatio-temporal distribution of fall armyworm in China by coupling multi-factors[J]. Remote Sensing, 2022, 14(17): 4415.
- [4] OVERTON K, MAINO J L, DAY R, et al. Global crop impacts, yield losses and action thresholds for fall armyworm (*Spodoptera frugiperda*): A review[J]. Crop Protection, 2021, 145: 105641.
- [5] 张帅, 王波涛. 基于改进 YOLO11-seg 的轻量化病虫害分割模型[J]. 农业工程学报, 2025, 41(24): 173-181. ZHANG Shuai, WANG Botao. Lightweight pest segmentation model based on improved YOLO11-seg[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(24): 173-181. (in Chinese with English abstract)
- [6] 江永成, 吕蕴娴, 胡根生. 基于无人机遥感图像的茶叶枯病病斑半监督检测方法[J]. 农业机械学报, 2025, 56(11): 441-452. JIANG Yongcheng, LYU Yunxian, HU Gensheng. Semi-supervised object detection method for spots of tea leaf blight based on UAV remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(11): 441-452. (in Chinese with English abstract)
- [7] ISHENGOMA F S, RAI I A, SAID R N. Identification of maize leaves infected by fall armyworms using UAV-based imagery and convolutional neural networks[J]. Computers and Electronics in Agriculture, 2021, 184: 106124.
- [8] 张国忠, 吕紫薇, 刘浩蓬, 等. 基于改进 DenseNet 和迁移学习的荷叶病虫害识别模型[J]. 农业工程学报, 2023, 39(8): 188-196. ZHANG Guozhong, LYU Ziwei, LIU Haopeng, et al. Model for identifying lotus leaf pests and diseases using improved DenseNet and transfer learning[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(8): 188-196. (in Chinese with English abstract)
- [9] HUGHES D P, SALATHE M. An open access repository of images on plant health to enable the development of mobile disease diagnostics[EB/OL]. (2015-11-25)[2025-07-15]. <https://doi.org/10.48550/arXiv.1511.08060>.
- [10] 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.
- [11] WANG C Y, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information[EB/OL]. (2024-02-21)[2025-07-20] <https://doi.org/10.48550/arXiv.2402.13616>.
- [12] WANG A, CHEN H, LIU L, et al. YOLOv10: real time end to end object detection [EB/OL]. (2024-05-23)[2025-07-20]. <https://doi.org/10.48550/arXiv.2405.14458>.
- [13] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [EB/OL]. (2024-10-23)[2025-07-20]. <https://doi.org/10.48550/arXiv.2410.17725>.
- [14] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]// 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy: IEEE, 2017: 2999-3007.
- [15] 翁海勇, 姚越, 黄德耀, 等. 无人机低空遥感结合 YOLOv7 快速评估水稻穗颈瘟抗性[J]. 农业工程学报, 2024, 40(21): 110-118. WENG Haiyong, YAO Yue, HUANG Deyao, et al. Rapid evaluation of rice neck blast resistance using low altitude remote sensing of UAV combined with YOLOv7[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(21): 110-118. (in Chinese with English abstract)
- [16] BAO W, ZHU Z, HU G, et al. UAV remote sensing detection of tea leaf blight based on DDMA-YOLO[J]. Computers and Electronics in Agriculture, 2023, 205: 107637.
- [17] 侯文慧, 龚昌智, 曹文昊, 等. 基于超分辨率增强与改进 YOLOv8 的番茄叶片病害检测[J]. 农业工程学报, 2025, 41(16): 211-220. HOU Wenhui, GONG Changzhi, CAO Wenhao, et al. Tomato leaf disease detection method based on super-resolution enhancement and improved YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(16): 211-220. (in Chinese with English abstract)
- [18] 马超伟, 张浩, 马新明, 等. 基于改进 YOLOv8 的轻量化小麦病害检测方法[J]. 农业工程学报, 2024, 40(5): 187-195. MA Chaowei, ZHANG Hao, MA Xinming, et al. Method for the lightweight detection of wheat disease using improved YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(5): 187-195. (in Chinese with English abstract)
- [19] ZHOU H, LI Q, QIN B, et al. High-throughput wheat seedling phenotyping via UAV-based semantic segmentation and ground sample distance driven pixel-to-area mapping[J]. Computers and Electronics in Agriculture, 2025, 238: 110819.
- [20] WANG X, YU K, WU S, et al. ESRGAN: Enhanced super-resolution generative adversarial networks[C]// Computer Vision – ECCV 2018 Workshops, Cham, Switzerland: Springer, 2019: 63-79.
- [21] WANG X, XIE L, DONG C, et al. Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data[C]// 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, Canada: IEEE, 2021: 1905-1914.
- [22] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]// IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA: IEEE, 2016: 1874-1883.

- [23] MIYATO T, KATAOKA T, KOYAMA M, et al. Spectral normalization for generative adversarial networks[EB/OL]. (2018-02-16)[2025-08-09]<https://doi.org/10.48550/arXiv.1802.05957>.
- [24] ZHENG Y, JING Y, ZHAO J, et al. LAM-YOLO: Drones-based small object detection on lighting-occlusion attention mechanism YOLO[EB/OL](2024-11-01)[2025-08-12]<https://doi.org/10.48550/arXiv.2411.00485>.
- [25] 范翔宇, 李玥, 魏霖静, 等. 面向收割机影像的轻量化 DeepLabV3+ 胡麻倒伏识别方法[J]. 农业工程学报, 2025, 41(18): 212-219.
- FAN Xiangyu, LI Yue, WEI Linjing, et al. Detecting flax lodging from harvester images using lightweight DeepLabV3+[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(18): 212-219. (in Chinese with English abstract)
- [26] 张欣冉, 王新忠, 周奎州, 等. 基于改进 YOLO11n 的复杂环境下多目标奶牛日常行为检测[J]. 农业工程学报, 2025, 41(14): 155-164.
- ZHANG Xinran, WANG Xinzhong, ZHOU Kuizhou, et al. Daily behaviour detection of multi-target dairy cows based on improved YOLO11n in complex environment[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(14): 155-164. (in Chinese with English abstract)
- [27] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA: IEEE, 2018: 8759-8768.
- [28] 任锐, 孙海霞, 张淑娟, 等. 基于改进 YOLOv8n 的不同栽培模式下玉露香梨轻量化检测[J]. 农业工程学报, 2025, 41(5): 145-155.
- REN Rui, SUN Haixia, ZHANG Shujuan, et al. Lightweight detection method for 'Yuluxiang' pear under different cultivation modes based on improved YOLOv8n[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(5): 145-155. (in Chinese with English abstract)

Detecting *Spodoptera frugiperda* infestation traces in maize fields using CBP-YOLO

QI Jiangtao^{1,2,3}, WANG Shuo^{1,2,3}, XIONG Yuesong^{1,2,3}, GAO Fangfang^{1,2,3}, WANG Faying^{1,2,3},
ZOU Zongfeng⁴, LIU Yingzhi⁴, LIU Huili^{1,2,3*}

(1. Key Laboratory of Bionic Engineering, Ministry of Education, Jilin University, Changchun 130022, China; 2. College of Biological and Agricultural Engineering, Jilin University, Changchun 130022, China; 3. Key Laboratory of Smart Agricultural Equipment and Technology of Jilin Provincial, Changchun 130022, China; 4. Yantai Agricultural Technology Extension Center, Yantai 264001, China)

Abstract: Fall armyworm (*Spodoptera frugiperda*) is one of the most serious pests in maize fields. It is often required to early and accurately detect its infestation for timely and effective pest prevention using unmanned aerial vehicle (UAV) imagery. However, reliable detection has been confined to the challenges: 1) The small and subtle feeding marks caused by the larvae, leading difficult to identify at high altitudes. 2) Consistent recognition has been limited to significant variations in object scale at different flight heights. 3) The accurate detection has also been limited to the low contrast between damaged leaf tissue and surrounding healthy foliage, especially under the different lighting and environmental conditions in fields. Collectively, advanced computer vision is necessary to robustly identify early signs of infestation at diverse scales under complex backgrounds. In this study, a robust deep learning model was developed to reliably identify the subtle infestation traces in multi-scale UAV images. A detection architecture, termed coordinated-BiFPN-P2-YOLO (CBP-YOLO), was also proposed using YOLOv8. Real-enhanced super-resolution generative adversarial network (Real-ESRGAN) was applied as a preprocessing step to reduce image degradation from low ground sampling distance. High-fidelity textures of leaf damage were reconstructed from original low-resolution inputs. The backbone of YOLOv8 was enhanced with the Coordinated attention (CA) mechanism. Spatial and channel-wise features were captured to improve the localization and discrimination of minute lesions. Furthermore, the neck component was upgraded with a Bi-directional feature pyramid network (BiFPN) for the highly efficient top-down and bottom-up cross-scale feature fusion. Information loss was minimized for consistent representation during hierarchical propagation at different scales. In addition, a detection head was added to specifically strengthen sensitivity to small targets, particularly at a 160×160 spatial resolution with 64-channel output. The improved model was trained and then evaluated on the custom UAV dataset, which was collected from maize fields naturally infested by fall armyworm under diverse lighting conditions and flight heights. Extensive experiments demonstrated that the CBP-YOLO achieved a peak performance on the imagery with a ground sampling distance (GSD) of 0.38 cm per pixel. Real-ESRGAN significantly alleviated texture blurring and edge ambiguity in low-resolution images, leading to better delineation of feeding scars. Ablation studies were conducted to evaluate the effectiveness of the improved model. There was an outstanding performance on the UAV multi-scale blade dataset. Specifically, there was an average precision (AP@0.5) of 76.5%, which increased by 3.4 percentage points, compared with the baseline model. The robustness and practical applicability of the improved model were obtained in the blades of varying scales during aerial inspection. A comparison showed that the CBP-YOLO outperformed state-of-the-art detectors—including YOLOv9 medium, YOLOv10 medium, YOLOv11 medium, Faster region-based convolutional neural network, and RetinaNet—by margins of 10.1, 7.2, 5.1, 9.3, and 17.9 percentage points in AP@0.5, respectively. Notably, the high precision was also maintained under varying illumination and partial occlusion, indicating strong generalization in agricultural environments. The improved CBP-YOLO framework effectively detected subtle, multi-scale fall armyworm infestation signals during UAV monitoring. Superior accuracy and robustness of the improved model were achieved to synergistically combine super-resolution enhancement, attention-aware feature extraction, fine-grained detection heads, and bidirectional multi-scale fusion. These findings can also provide a practical and scalable solution for early pest outbreak detection, thereby enabling timely intervention to reduce the crop losses in large-scale maize production.

Keywords: unmanned aerial vehicle(UAV); deep learning; YOLOv8; *Spodoptera frugiperda*; smart agriculture; maize leaves