

基于改进YOLOv8的水稻病虫害检测方法研究：融合SCSA注意力机制与WiOUv3损失函数

沈珈毅¹, 周泽润¹, 沈云飞¹, 朱 宣¹, 黄金城^{1*}, 宋芷欣², 杨舒雅³

¹盐城工学院信息工程学院, 江苏 盐城

²盐城工学院材料科学与工程学院, 江苏 盐城

³南通大学数学与统计学院, 江苏 南通

收稿日期: 2026年9月23日; 录用日期: 2026年10月20日; 发布日期: 2026年10月27日

摘 要

传统的水稻病虫害检测方法通常存在小目标识别精度低、背景干扰强以及特征提取能力有限的问题。为解决这些挑战, 本研究提出了一种基于YOLOv8 (You Only Look Once version 8)的改进水稻病虫害检测框架。该方法引入SCSA (Spatial-Channel-Spatial Attention)注意力机制, 用来增强特征表达能力。本研究采用WiOUv3 (Wise Intersection over Union Version 3)替代传统IoU损失。该损失函数综合考虑预测框与真实框之间的重叠区域和中心距离, 引入动态梯度权重机制。本研究在公开水稻病虫害数据集上进行评估, 性能指标包括mAP@0.5、mAP@0.5:0.95、Precision和Recall。结果表明, YOLOv8-SCSA-WiOU在mAP@0.5指标上达到0.834, mAP@0.5:0.95为0.689, 相比基线YOLOv8n的0.800和0.655有显著提升, Precision和Recall同样有所提高。

关键词

YOLOv8, SCSA注意力机制, WiOUv3损失函数, 小目标检测

Improved YOLOv8 for Rice Pest and Disease Detection with SCSA Attention and WiOUv3 Loss

Jiayi Shen¹, Zerun Zhou¹, Yunfei Shen¹, Xuan Zhu¹, Jincheng Huang^{1*}, Zhixin Song², Shuya Yang³

*通讯作者。

文章引用: 沈珈毅, 周泽润, 沈云飞, 朱宣, 黄金城, 宋芷欣, 杨舒雅. 基于改进 YOLOv8 的水稻病虫害检测方法研究: 融合 SCSA 注意力机制与 WiOUv3 损失函数[J]. 计算机科学与应用, 2026, 16(10): 171-181.

DOI: 10.12677/csa.2026.1610320

¹School of Information Engineering, Yancheng Institute of Technology, Yancheng Jiangsu²School of Materials Science and Engineering, Yancheng Institute of Technology, Yancheng Jiangsu³School of Mathematics and Statistics, Nantong University, Nantong Jiangsu

Received: September 23, 2026; accepted: October 20, 2026; published: October 27, 2026

Abstract

Traditional rice pest and disease detection methods often suffer from low accuracy in detecting small targets, strong background interference, and limited feature extraction capabilities. To address these challenges, this study proposes an improved rice pest and disease detection framework based on YOLOv8 (You Only Look Once version 8). The proposed method incorporates the SCSA (Spatial-Channel Attention) attention mechanism to enhance feature representation. In addition, WiOUv3 (Wise Intersection over Union Version 3) is employed to replace the conventional IoU loss function. The proposed loss function jointly considers the overlap region and center distance between the predicted and ground-truth bounding boxes and introduces a dynamic gradient weighting mechanism. The proposed method is evaluated on a publicly available rice pest and disease dataset using mAP@0.5, mAP@0.5:0.95, Precision, and Recall as performance metrics. Experimental results show that YOLOv8-SCSA-WiOU achieves an mAP@0.5 of 0.834 and an mAP@0.5:0.95 of 0.689, which are significantly higher than 0.800 and 0.655, respectively, for the baseline YOLOv8n. Precision and Recall are also improved.

Keywords

YOLOv8, SCSA Attention Mechanism, WiOU Version 3 Loss Function, Small Object Detection

Copyright © 2026 by author(s) and Hans Publishers Inc.

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

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

Open Access

1. 引言

水稻是全球重要的粮食作物之一，其生长过程中易受多种病虫害的影响。传统病虫害监测方法主要依赖人工巡查，但是效率低下，容易受到主观因素影响。Anirudha Ghosh 等人提出传统害虫防治方法主要依赖人工田间巡查和化学农药使用，这些方式往往成本高昂、耗时费力，且可能对环境造成负面影响[1]。Poonam Adhikari 等人认为目前农作物害虫检测主要采用人工采样、传感器和雷达等多种技术手段。但这些方法存在覆盖范围有限、成本高昂且操作复杂等局限性[2]。近年来，随着深度学习技术的发展，基于计算机视觉的自动化病虫害检测方法成为研究热点。尤其是目标检测算法，如 YOLO (You Only Look Once) 系列，因其高效的实时性和较好的检测精度，在农业病虫害监测中得到了广泛应用。

然而，现有的 YOLO 模型在处理小目标检测任务时仍面临诸多挑战。小目标在图像中的占比小，特征信息稀疏，容易被背景噪声淹没，导致检测精度下降。为了解决这些问题，研究者们提出了多种改进策略，包括引入注意力机制和优化损失函数等。Ji 等人提出了一种基于 YOLOv5s 架构的创新检测方法 YOLO-PEST。该方法通过多通道采集水稻害虫图像，并随机裁剪图像以遮挡检测框，有效模拟害虫重叠场景。在特征融合过程中，集成 ConvNeXt 模块通过多尺度特征提取提升小目标检测精度，同时引入 CoTAttention 机制增强模型在复杂环境条件下的鲁棒性[3]。

本研究在 YOLOv8n 的 SPPF 后、Neck 前引入 SCSA 注意力模块,采用 WIoUv3 替换 CIoU 边界框回归损失,同时在公开水稻病虫害数据集上验证改进模型的有效性。在农业病虫害检测中,注意力机制的引入有助于抑制背景干扰,突出病虫害区域,提高小目标的检测精度。WiOUv3 (Wise Intersection over Union Version 3)损失函数通过动态调整权重,增强了模型对小目标的关注,提高了定位精度和收敛稳定性。

2. 相关研究工作

早期研究多依赖卷积神经网络(CNN)实现图像分类或语义分割,对病虫害进行整体识别,但在小目标、多尺度及复杂背景条件下仍表现出精度不足的问题。为提高检测性能,基于目标检测的算法逐渐成为研究重点,其中 YOLO 系列因其端到端训练、实时推理能力以及较高的精度,在农业图像分析中得到广泛应用。针对小目标特征信息稀疏的问题,多尺度特征融合方法被提出,例如 FPN (Feature Pyramid Network)与 BiFPN (Bidirectional Feature Pyramid Network),通过跨层特征交互与加权融合,有效缓解了小目标信息在高层特征中的丢失现象。与此同时,注意力机制的发展为目标检测提供了新的优化手段。通过模拟人类视觉聚焦的特性,注意力模块能够显著增强模型对关键区域的响应,同时抑制背景噪声。典型方法包括通道注意力(SE)、空间-通道联合注意力(CBAM)以及双重注意力网络(DANet),这些方法在复杂农业图像场景下对小目标均表现出良好的捕获能力。诸如 CBAM (Convolutional Block Attention Module)和 BiFPN (Bidirectional Feature Pyramid Network)等注意力机制已被成功应用于 YOLOv8 模型中,显著提升了检测性能。Yi 等人基于改进版 YOLOv8 提出一种轻量级且精准的水稻害虫检测算法。

在 YOLOv8 网络中构建多分支卷积块注意力模块(M-CBAM),增强害虫目标的特征提取能力,从而获得更优的检测效果。引入最小点距离交并比(MPDIoU)作为边界框损失度量,加快模型收敛速度并提升检测精度。采用轻量级 Ghost 卷积模块显著减少模型参数,同时保持最佳检测性能[4]。SCSA (Spatial-Channel-Spatial Attention)模块通过分层空间-通道-空间建模,实现局部与全局特征的高效交互,为小目标检测提供了新的特征增强策略。针对多尺度特征捕获的需求,研究者们提出了利用不同卷积核或特征金字塔结构进行特征融合的方法。

传统 IoU (Intersection over Union)损失虽然能够度量预测框与真实框之间的重叠区域,但当两者无交集时梯度为零,限制了模型在训练初期的收敛效率。为解决该问题,研究者提出了 GIoU、DIoU 和 CIoU 等改进方法,通过引入外接矩形距离、中心点距离及宽高比惩罚等约束增强定位能力。然而,这类基于几何关系的优化方式仍存在两个不足:其一,损失权重不具备动态性,难以根据学习难度自适应调整;其二,定位精度对复杂目标(如形状变化剧烈或小目标)仍然不足。为此,Tong 等人设计了具有动态非单调调频(non-monotonic FM)的 WIoUv3。得益于动态非单调调频的智能梯度增益分配策略[5]。

3. 改进的 YOLOv8 Network

YOLOv8 是 Ultralytics 于 2023 年推出的一种目标检测模型。YOLOv8 对 YOLOv5 提出的理念进行了优化与扩展,重点提升了模型在实时目标检测任务中的准确性和可用性[6]。YOLOv8 采用端到端的深度学习框架,其整体网络结构可划分为三个核心模块:骨干网络、颈部网络和检测头。骨干网络基于改进的 C2f 结构,负责从输入图像中提取多尺度特征,既能保留丰富的空间结构信息,还能获取高层语义特征,有效降低模型参数量与计算复杂度。颈部网络采用 PAN-FPN 特征融合结构,将不同尺度的特征图进行自顶向下与自底向上的双向融合。检测头采用解耦式结构设计,将分类任务与回归任务分离处理,分别预测目标类别概率和边界框参数。在训练阶段,YOLOv8 对输入图像进行 Mosaic 数据增强、尺度变换等预处理操作,并统一图像转换为 RGB 格式作为网络输入,随后依次经过骨干网络特征提取、颈部网络特征融合以及检测头预测,在三个不同尺度的特征层上分别完成对大、中、小目标的回归与分类。

原始 YOLOv8n 模型在 Backbone 末端通过 SPPF 模块输出 P5 层级的高层语义特征后,直接将其输入

至 Neck (PAN-FPN)进行多尺度特征融合与检测,而本研究改进的模型在 SPPF 模块之后、Neck 之前,专门插入了一层 SCSA 注意力模块(- [-1, 1, SCSA, [1024]]),特征图在进入多尺度融合前,会先经过 SCSA 模块的处理。该 SCSA 模块由可共享多语义空间注意力(SMSA)与渐进式通道自注意力(PCSA)协同构成,其核心作用在于对 SPPF 输出的 1024 通道高层语义特征先进行空间与通道维度的联合重校准,通过协同融合空间与通道特征,抑制复杂背景(如稻叶纹理、光影)的干扰,并增强微小害虫目标的特征表征能力,为后续的 Neck 和检测头提供信噪比更高的特征图。

原始 YOLOv8n 在 BboxLoss 中采用 CIoU 作为边界框回归损失,本研究把损失函数替换为 WIoUv3 (Wise-IoU v3),在 loss.py 的 BboxLoss 类中将参数改为 WIoU = True, scale = True,并在 metrics.py 中新增了 WIoU_Scale 类以支持动态非单调聚焦机制。WIoU v3 不再单纯依赖 IoU 来评估锚框质量,而是引入了“离群度”(outlier degree)这一度量标准,通过动态非单调聚焦机制。(即_scaled_loss 函数中的 beta 与 alpha 调节),对高质量锚框降低梯度增益,同时大幅削弱低质量标注样本产生的有害梯度,使模型在训练过程中能够将注意力聚焦于普通质量的锚框上,实现更精准的边界框回归。

4. SCSA 注意力机制

本研究引入了 SCSA (Spatial-Channel-Spatial Attention)注意力模块,该注意力机制通过空间 - 通道 - 空间的分层注意力建模,同时结合多尺度深度可分卷积和多头通道注意力机制。SCSA 模块在特征提取阶段采用分层“空间 - 通道 - 空间”建模策略,通过双向空间映射,采用卷积核大小为 3, 5, 7, 9 的多尺度深度可分卷积。其通道自注意力机制通过 Q 、 K 、 V 矩阵建模通道间的全局特征依赖,实现语义层面的信息重构。SCSA 由 Si 等人于 2025 年提出,其核心贡献在于揭示了空间注意力与通道注意力之间的协同效应[7]。现有混合注意力方法(如 CBAM、DANet)通常将空间注意力和通道注意力简单串联或并联,但忽略了二者之间的语义不平等问题:通道注意力在重加权通道时,对不同通道施加同等程度的抑制或增强,却不区分这些通道内部的空间语义差异;空间注意力则专注于定位,却不区分被定位区域属于哪个语义类别。当两者简单组合时,这种语义不平等会导致跨通道的语义冲突。该机制的设计目标是通过“空间先验引导通道重校准”来消除这种冲突。SMSA (可共享多语义空间注意力)的作用不是简单的空间注意力,而是向 PCSA 注入可共享的多语义空间先验,使通道自注意力在建模通道间依赖关系时能够“知道”每个通道中的空间信息对应什么语义,从而做出更合理的通道重校准决策。论文应明确表述:SCSA 的本质是一种以多语义空间先验为引导的通道语义重构机制,而非“空间注意力 + 通道注意力”的简单叠加。SCSA 注意力机制的具体结构如图 1 所示。

SCSA 虽以 SMSA→PCSA 串联,但并非“先空间、后通道”的机械顺序,而是协同设计:SMSA 在空间维度以多尺度一维卷积生成多语义空间先验,PCSA 在通道维度做自注意力,二者维度解耦、避免资源竞争;SMSA 的先验注入使 PCSA 在建模通道依赖时“知道”各通道的空间语义,实现轻量引导;PCSA 除通道重校准外,还通过通道交互弥合 SMSA 多尺度分组带来的子特征语义分歧。因此,SCSA 的“协同”是 SMSA→PCSA 的引导与 PCSA 内部弥合的双向作用,而非单向空间注意力作用于通道注意力。具体而言,给定输入特征图: $X \in R^{B \times C \times H \times W}$, 其中 B 、 C 、 H 、 W 分别表示批次大小、通道数、高度和宽度。

SMSA 首先沿空间维度进行平均池化得到水平方向和垂直方向的投影特征: $X \in R^{B \times C \times H \times W}$, SCSA 首先沿空间维度进行平均池化得到水平方向和垂直方向的投影特征,将二维空间注意力分解为两个一维方向上的序列建模,避免了直接在 $H \times W$ 空间上做卷积或池化带来的高计算量,保留了空间结构信息:

$$X_H = \text{GAP}_W(X) \in R^{B \times C \times H}$$

$$X_W = \text{GAP}_H(X) \in R^{B \times C \times W}$$

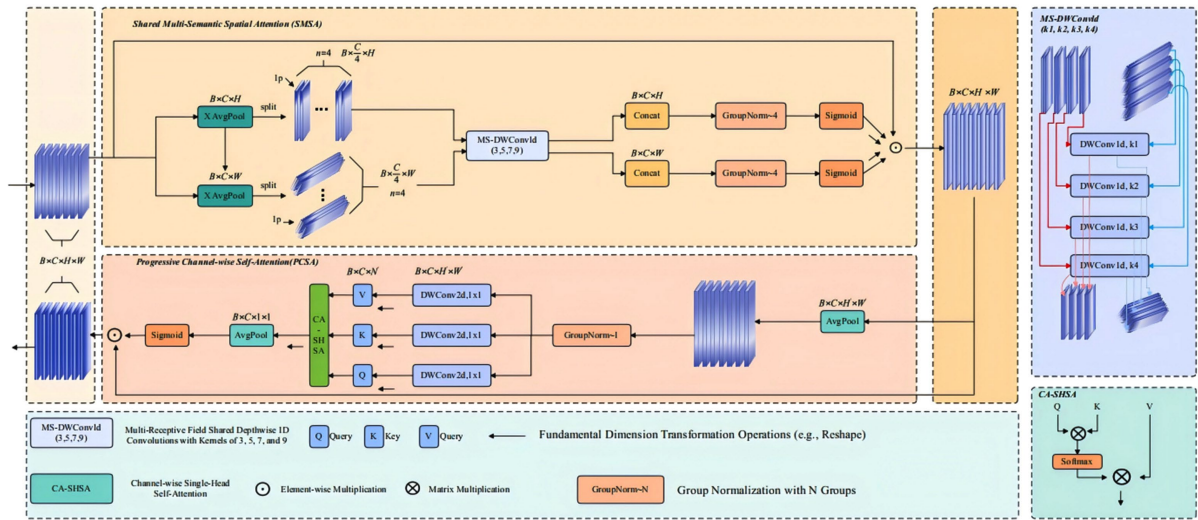


Figure 1. Structural diagram of the spatial-channel-spatial attention

图 1. SCSA 注意力机制结构图

随后将特征沿通道维度拆分为 K 个独立子特征，每个子特征通道数为 C/K ：

$$X_H^i = X_H \left[:, (i-1) \times \frac{C}{K} : i \times \frac{C}{K} :, i=1, \dots, K \right]$$

$$X_W^i = X_W \left[:, (i-1) \times \frac{C}{K} : i \times \frac{C}{K} :, i=1, \dots, K \right]$$

对每个子特征应用不同卷积核大小的深度可分一维卷积：

$$(k_i \in \{3, 5, 7, 9\}) : \tilde{X}_H^i = \text{DWConvId}_{k_i}(X_H^i), \tilde{X}_W^i = \text{DWConvId}_{k_i}(X_W^i)$$

多核并行使 SMSA 在同一个注意力层内同时建模从局部到全局的多语义空间信息——小核(3, 5)捕获局部纹理(如稻叶边缘、虫体轮廓)，大核(7, 9)捕获全局结构(如害虫在稻田中的分布模式)，而非让网络在后续层中逐级学习这些信息。

将各子特征拼接后经组归一化(GN)和 Sigmoid 激活，分别得到高度方向和宽度方向的注意力图：

$$\text{Attn}_H = \sigma \left(\text{GN} \left(\text{Concat} \left(\tilde{X}_H^1, \dots, \tilde{X}_H^4 \right) \right) \right)$$

$$\text{Attn}_W = \sigma \left(\text{GN} \left(\text{Concat} \left(\tilde{X}_W^1, \dots, \tilde{X}_W^4 \right) \right) \right)$$

空间增强特征为两个方向注意力图的逐元素乘积与原始输入的乘积：

$$X_S = \text{Attn}_H \times \text{Attn}_W \times X$$

组归一化的选择保证了不同子特征在拼接后的尺度一致性，同时避免了批归一化在小批次下引入的统计噪声，保留了各子特征之间的语义差异(而非将其强行归一化为相同分布)。

PCSA 的输入是 SMSA 输出的空间增强特征 X_S ，为降低通道自注意力的计算复杂度，首先通过平均池化将空间分辨率从 (H, W) 压缩至 (H', W') ：

$$X_p = \text{AvgPool}_{(H', W')} (X_S) \in R^{B \times C \times H' \times W'}$$

“渐进式”的含义体现在两个层面：其一，压缩是渐进进行的，先经过 SMSA 的空间增强，再进入

PCSA 的压缩,而非在原始特征上直接压缩;其二,压缩保留了 SMSA 注入的空间先验信息—— X_s 中的空间语义结构在池化后以全局上下文的形式被保留,而非被丢弃。通过 1×1 深度卷积从 X_p 生成查询、键、值矩阵:

$$Q = F_{proj}^Q(X_p), K = F_{proj}^K(X_p), V = F_{proj}^V(X_p)$$

注意力计算沿通道维度进行,注意力矩阵 $A \in R^{C \times C}$:

$$X_{\text{attn}} = \text{Softmax}\left(\frac{QK^T}{\sqrt{C}}\right)V$$

原论文明确指出这一选择的理论依据:单头注意力在通道维度上产生更强的通道间交互,能够更有效地弥合 SMSA 多尺度处理引入的语义分歧;而多头注意力将通道维度拆分为多个子空间,反而削弱了跨通道语义对齐的能力。对 X_{attn} 进行全局平均池化并经过 Sigmoid 激活,生成通道注意力权重:

$$\text{Attn}_C = \sigma(\text{GAP}(X_{\text{attn}}))$$

最终 PCSA 输出为通道权重与空间增强特征的乘积:

$$\text{PCSA}(X_s) = \text{Attn}_C \times X_s$$

最终,得到完整的 SCSA 公式:

$$\text{SCSA}(X) = \text{PCSA}(\text{SMSA}(X))$$

5. WiOUv3 损失函数

传统的 IoU (Intersection over Union) 损失定义为:

$$L_{\text{IoU}} = 1 - \text{IoU} = 1 - \frac{|B_p \cap B_g|}{|B_p \cup B_g|}$$

本研究引入了基于权重重分配策略的 WiOU (Weighted IoU) 损失函数。其关键思想是通过动态权重机制使模型将更多的学习关注分配给难以回归的样本。WiOU 利用预测框与真实框中心点距离和框体尺寸差异构建难度系数,并通过指数函数映射生成惩罚项,从而为 IoU 损失赋予自适应调节能力:难回归样本获得更大的梯度权重,而易回归样本权重则被抑制,提升了模型对误差敏感区域的拟合能力。

本研究采用的是 WiOUv3 版本,即引入非单调映射的权重调整策略,突破了以往权重随 IoU 单调递减的限制。当预测框位置出现偏差但仍具备较高回归潜力时, WiOUv3 能适当提升其梯度贡献,避免模型在训练中丢失潜在可学习样本;而对于高 IoU 但边界误差较大或不稳定的框, WiOUv3 则降低其权重,防止过拟合。相比于 v1 和 v2, WiOUv3 更加平衡了训练过程中的稳定性与敏感性,使优化方向更加符合实际目标定位需求。

WiOUv1 通过预测框与真实框的中心点距离以及外接矩形对角线长度构造动态权重项。设预测框中心为 $c_p = (x_p, y_p)$, 真实框中心为 $c_g = (x_g, y_g)$, 则中心点距离为

$$\rho^2 = \|c_p - c_g\|^2 = (x_p - x_g)^2 + (y_p - y_g)^2$$

同时使用外接矩形对角线长度作为尺度归一化因子, c 为真实框外接矩形对角线长度:

$$c^2 = W_g^2 + H_g^2$$

两者结合构成回归难度度量:

$$D = \frac{\rho^2}{c^2}$$

于是 WIoUv1 的动态权重项为

$$R_{\text{WIoU}} = \exp(D) = \exp\left(\frac{\rho^2}{c^2}\right)$$

对应的 WIoUv1 损失为:

$$L_{\text{WiOUv1}} = R_{\text{WIoU}} L_{\text{IoU}} = \exp\left(\frac{\rho^2}{c^2}\right) \cdot (1 - \text{IoU})$$

WIoUv3 在 WIoUv1 的基础上引入非单调聚焦系数 r , 使模型能够根据样本的离群程度动态分配梯度增益:

$$\beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}}$$

其中 L_{IoU}^* 为当前 IoU 损失, $\overline{L_{\text{IoU}}}$ 为 IoU 损失的滑动平均值。进一步定义:

$$r = \frac{\beta}{\delta \alpha^{\beta-\delta}}$$

其中 α 和 δ 为超参数, 常见取值为 $\alpha=1.9, \delta=3$ 。由此得到 WIoUv3 的损失函数为:

$$L_{\text{WiOUv3}} = r L_{\text{WiOUv1}} = r \exp\left(\frac{\rho^2}{c^2}\right) (1 - \text{IoU})$$

因此, WIoUv3 通过动态非单调聚焦机制, 在保留 WIoUv1 中心距离加权思想的同时, 进一步调节不同质量锚框的梯度贡献, 使模型在训练过程中更加关注普通质量的锚框, 从而提升边界框回归的定位精度与收敛稳定性。

6. 实验结果与分析

6.1. 硬件及其参数设置

本次实验利用 PyTorch 深度学习框架以及 vscode 开源集成开发环境进行模型构建与训练。硬件方面, 采用 Intel Core i5-12400F 处理器、24 GB RAM 内存, 并搭载 NVIDIA GeForce RTX 3090 GPU, 支持 CUDA 11.8 加速。软件方面采用 PyTorch 2.0.0 与 Python 3.8.10。完成了 YOLOv8n 原始模型与改进模型(YOLOv8n-SCSA-WIoU)的训练与测试工作, 确保了实验的高效性与模型在高性能计算设备下的训练稳定性。如下表 1 所示:

Table 1. Configuration of experimental software and hardware equipment

表 1. 实验软硬件设备配置

CPU	RAM	GPU	Cuda	Pytorch	python
I5-12400f	24.0 G	NVIDIAGeForceRTX3090	11.8	2.0.0	3.8.10

6.2. 评价指标

采用 $\text{mAP}@0.5$ 、准确率(P)、召回率(R)、 $\text{mAP}@0.5:0.95$ 来评价训练结果。本次实验采用 $\text{mAP}@0.5$ 、 $\text{mAP}@0.5:0.95$ 、准确率(Precision, P)和召回率(Recall, R)四项指标对模型的检测性能进行综合评估。准确

率(P)和召回率(R)的公式如下:

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

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

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

6.3. 对比实验结果

本次实验收集了 11 种常见的水稻害虫,从百度图片、Google Images、Bing Images、农业病虫害专题网站等互联网公开渠道检索收集图像,经人工筛选、去重、类别统一与复核后,最终保留 4565 张图像,由于图像来源多样,原始拍摄时间、地点、设备及环境条件无法统一追溯,数据集具有非受控采集和多源异构特征。通过 labelImg 软件对图像中小目标进行标注,并按 6:2:2 的比例将标注数据集随机划分为训练集、测试集和验证集。本实验对比分析了原始 YOLOv5s 与 YOLOv5n 模型及其引入不同注意力机制(SE, CBAM, ECA)后的改进模型在稻田害虫检测任务中的性能表现,评价指标包括:准确率(Precision)、召回率(Recall)、mAP@0.5 以及 mAP@0.5:0.95。实验结果如下表 2 所示:

Table 2. Comparison of detection performance of different models

表 2. 不同模型的检测性能对比

Model	Precision	Recall	map@0.5	map@0.5:0.95
YOLOv8n-SCSA-WIoUv3	0.879	0.754	0.834	0.689
YOLOv8n	0.818	0.728	0.800	0.655
YOLOv8s	0.812	0.719	0.792	0.646
YOLOv5n	0.798	0.706	0.781	0.631
YOLOv5s	0.785	0.694	0.766	0.612
YOLOv5n-C3CBAM	0.588	0.566	0.577	0.370
YOLOv5n-C3ECA	0.633	0.588	0.609	0.394

本研究提出的 YOLOv8n-SCSA-WIoUv3 模型在所有评价指标上均取得了最优结果,其 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 分别达到 0.879、0.754、0.834 和 0.689,全面超越了原始 YOLOv8n 基线(0.818, 0.728, 0.800, 0.655),四项指标分别提升了 6.1%、2.6%、3.4%和 3.4%,充分验证了 SCSA 注意力机制与 WIoUv3 损失函数协同改进的有效性。在与其他模型的横向对比中,改进后的模型不仅优于同系列的 YOLOv8s(mAP@0.5 为 0.792),也显著高于 YOLOv5 系列模型(如 YOLOv5n 的 0.781 和 YOLOv5s 的 0.766)。YOLOv5 系列模型所采取的数据集与本研究是同一数据集,是 Shen 等人在一年前所做的研究[8]。值得注意的是,在 YOLOv5n 上直接嵌入 CBAM 或 ECA 注意力机制的变体表现极差,其 mAP@0.5 分别骤降至 0.577 和 0.609, mAP@0.5:0.95 更是跌至 0.370 和 0.394,这说明并非所有注意力机制的引入都能带来正向收益,盲目叠加反而可能导致模型容量过载或特征退化,而本研究通过合理的位置选择与机制设计,成功规避了这一问题。

为了进一步验证改进模型的实际检测效果,本研究分别调用了 YOLOv8n-SCSA-WIoUv3 (图 2)和原始 YOLOv8n (图 3)进行可视化推理对比。从直观检测结果来看,改进模型在所有测试样本上均表现出优异的性能,各类害虫的置信度均保持在 0.88 以上(最高达 0.95),检测框精准锁定目标主体。反观原始 YOLOv8n 模型,其检测结果暴露了基线模型在小目标检测和复杂背景下的局限性:其一,特征提取能力不足导致分类错误,例如将 Thripidae 误判为 Philothripidae (置信度仅 0.46),将 Ephydriidae 误判为 Cicadellidae;其二,背景干扰抑制能力弱,导致部分目标(如 Hesperidae、Delphacidae)的置信度大幅下降至 0.85 左右;其三,边界框回归不够稳定,未能很好地适应害虫形态的变化。

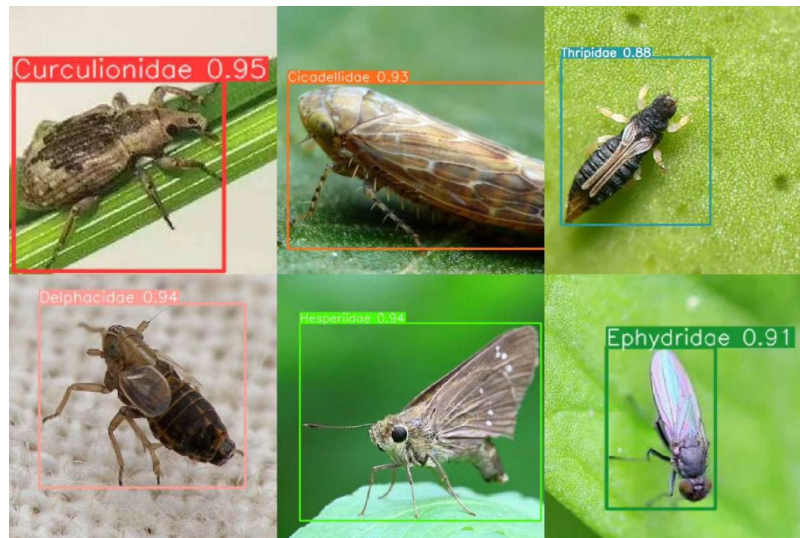


Figure 2. Detection results of the YOLOv8n-SCSA-WIoUv3 model

图 2. YOLOv8n-SCSA-WIoUv3 检测结果图

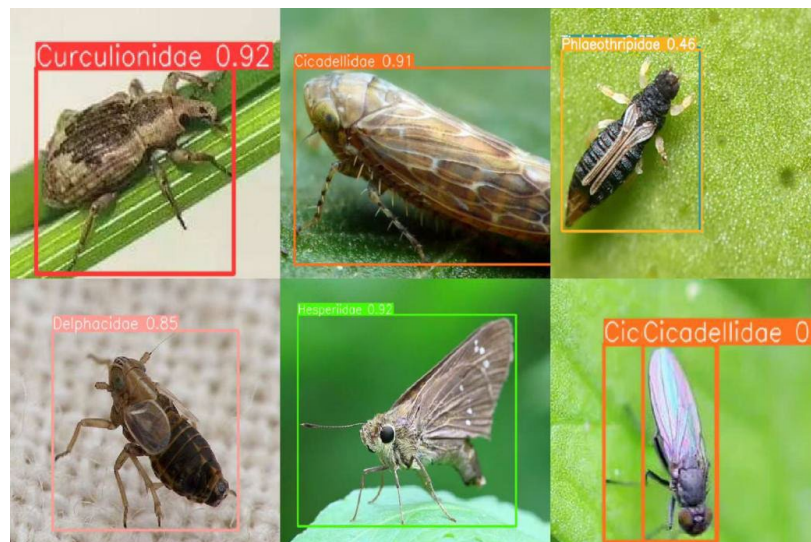


Figure 3. Detection results of the YOLOv8n model

图 3. YOLOv8n 检测结果图

本研究对比了 YOLOv8n-SCSA-WIoUv3 (图 4)与 YOLOv8n (图 5)的训练与验证曲线。从损失下降趋势观察,改进模型的三项训练损失(box_loss, cls_loss, dfl_loss)在前期展现出更陡峭的下降斜率,且曲线更

平滑,表明 SCSA 模块提供的多尺度空间与通道特征重校准机制,显著提升了模型对关键特征的捕获效率,从而加快了拟合速度。在验证阶段,改进模型的验证集损失曲线保持稳步下降且无明显过拟合迹象,这得益于 WIoUv3 损失函数引入的动态梯度权重机制。该机制通过降低高质量锚框的梯度增益并削弱低质量标注样本的有害梯度,使模型在训练后期依然能够保持稳定的优化方向。反观原始模型,其验证损失曲线存在一定震荡,说明其在面对复杂背景干扰时容易陷入局部最优或出现泛化性能下降。评价指标曲线进一步印证了上述分析:改进模型的 mAP@0.5 与 mAP@0.5:0.95 分别在约 100 和 150 个 Epoch 时达到稳定,且最终值显著高于基线模型。

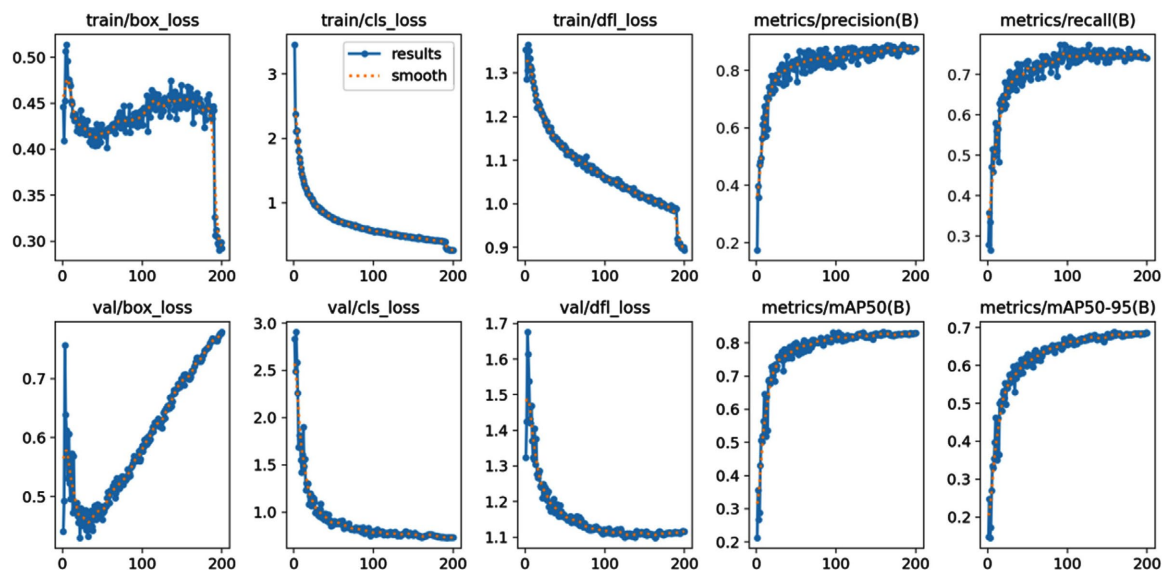


Figure 4. Training and validation metric curves of the improved YOLOv8n-SCSA-WIoUv3 model

图 4. YOLOv8n-SCSA-WIoUv3 改进模型训练与验证指标曲线图

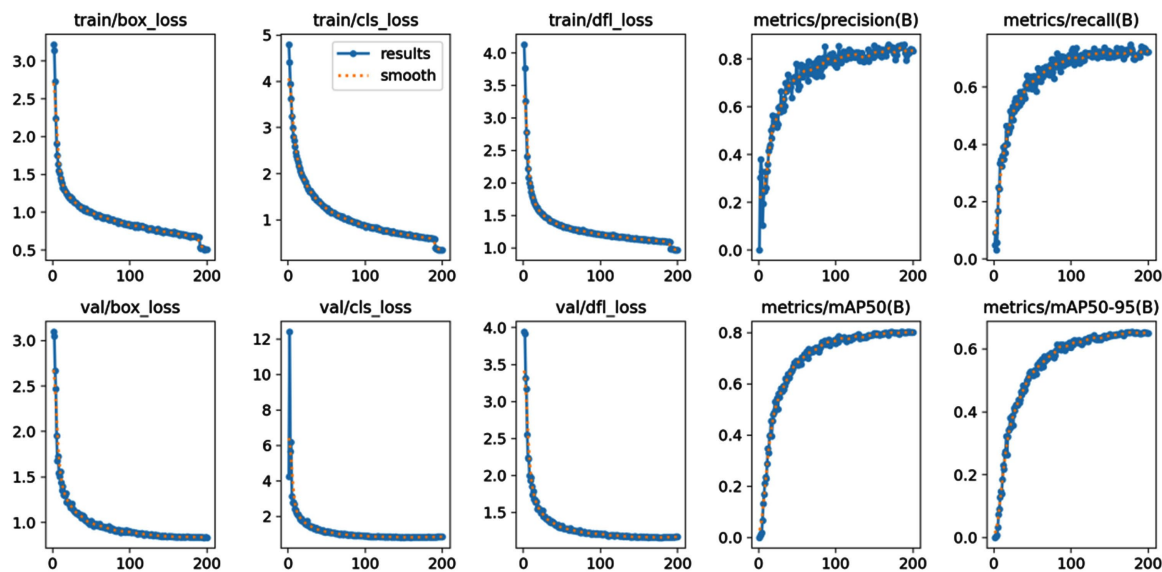


Figure 5. Training and validation metric curves of the YOLOv8n baseline model

图 5. YOLOv8n 基线模型训练与验证指标曲线图

7. 结论

本研究针对传统水稻病虫害检测中小目标识别精度低、背景干扰强以及特征提取能力有限等问题，提出了一种基于改进 YOLOv8n 的水稻病虫害检测方法。该方法在 SPPF 模块之后、Neck 之前引入 SCSA 注意力模块，对高层语义特征进行空间 - 通道 - 空间联合重校准，以增强微小害虫目标的特征表达并抑制复杂背景干扰；同时采用 WiOuv3 替换传统 CIoU 损失，通过动态非单调聚焦机制提升边界框回归精度和训练稳定性。在公开水稻病虫害数据集上的实验结果表明，改进模型 YOLOv8n-SCSA-WiOuv3 的 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 分别达到 0.879、0.754、0.834 和 0.689，相比基线 YOLOv8n 的 0.818、0.728、0.800 和 0.655 均有明显提升。与其他 YOLO 系列模型及注意力改进模型相比，本文方法在复杂稻田背景下具有更高的检测精度和更强的鲁棒性。研究表明，融合 SCSA 注意力机制与 WiOuv3 损失函数能够有效改善水稻病虫害小目标检测性能，可为农业智能化病虫害检测提供技术支持与参考。

参考文献

- [1] Ghosh, A., Sarkar, R. and Barman, D. (2026) Context-Aware Pesticide Recommendation via Few-Shot Pest Recognition for Precision Agriculture. arXiv: 2601.00243.
- [2] Adhikari, P., Kumar, R., Iyengar, S.R.S. and Kaur, R. (2021) What a Million Indian Farmers Say? A Crowdsourcing-Based Method for Pest Surveillance. arXiv: 2108.03374.
- [3] Qiang, J., Zhao, L., Wang, H., Xu, T., Jia, Q. and Sun, L. (2025) YOLO-PEST: A Novel Rice Pest Detection Approach Based on YOLOv5s. *Plant Methods*, **21**, Article No. 121. <https://doi.org/10.1186/s13007-025-01438-w>
- [4] Yin, J., Huang, P., Xiao, D. and Zhang, B. (2024) A Lightweight Rice Pest Detection Algorithm Using Improved Attention Mechanism and YOLOv8. *Agriculture*, **14**, Article 1052. <https://doi.org/10.3390/agriculture14071052>
- [5] Tong, Z., Chen, Y., Xu, Z. and Yu, R. (2023) Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. arXiv: 2301.10051.
- [6] Yaseen, M. (2024) What Is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector. arXiv: 2408.15857.
- [7] Si, Y., Xu, H., Zhu, X., Zhang, W., Dong, Y., Chen, Y., *et al.* (2025) SCSA: Exploring the Synergistic Effects between Spatial and Channel Attention. *Neurocomputing*, **634**, Article ID: 129866. <https://doi.org/10.1016/j.neucom.2025.129866>
- [8] Shen, J.Y., Ran, W.H., Huang, J.C., Han, L.J., Jiao, T.H., Wang, X.Y., Li, J.Y. and Wang, J.W. (2025) Comparative Study on Rice Pest Detection Performance Using YOLOv5 Enhanced with Three C3 Attention Modules. *Computer Science and Application*, **15**, 41-49. <https://doi.org/10.12677/csa.2025.158196>