

李 龙,李梦霞,李志良. 基于改进 YOLO v8 的水稻害虫识别方法[J]. 江苏农业科学,2024,52(20):209-219.  
doi:10.15889/j.issn.1002-1302.2024.20.025

# 基于改进 YOLO v8 的水稻害虫识别方法

李 龙,李梦霞,李志良

(长江大学计算机科学学院,湖北荆州 434000)

**摘要:**实时监测稻田害虫泛滥情况是预防水稻产量降低的重要手段之一。针对当前的目标检测算法在实际稻田环境下检测精度较低且模型计算量较大、难以实现实时检测等问题,提出一种基于 YOLO v8 的改进的水稻害虫识别算法 YOLO v8 - SDPS。首先在主干网络中用 SD\_Conv 卷积替代标准卷积,重构特征提取模块,在降低参数量的同时尽可能保留害虫目标的边缘特征信息,提升对害虫目标的特征提取能力;其次在颈部引入基于 Slim - Neck 范式的 GSCov 模块和 VoV - GSCSP 模块,在减少模型计算量的同时提升模型的检测精度;最后在 SPPF 层前引入 PSA 注意力模块,降低背景的噪声干扰,使模型更加关注个体的空间位置信息。用本研究提出的算法在经数据增强后的自建水稻害虫数据集上进行试验,结果表明,YOLO v8 - SDPS 获得 86.6% 的平均检测精度,相较于原始 YOLO v8n 模型提升 4.1 个百分点。同时改进后的模型参数量为 2.62 M,计算量为 7.5 GFLOPs,相较于基准模型分别降低 16.8% 和 15.7%,实现了模型轻量化和较高检测精度的平衡。在害虫小且密集、背景干扰严重、光照强烈等复杂环境下,YOLO v8 - SDPS 均能较好地识别出目标个体,有效地降低漏检率和误检率,具有较好的鲁棒性,可为稻田实时巡检提供有效技术支持。

**关键词:**目标检测;水稻害虫;深度学习;YOLO v8;极化自注意力

**中图分类号:**S126;TP391.41 **文献标志码:**A **文章编号:**1002-1302(2024)20-0209-11

水稻是我国主要的农作物之一,种植面积约占粮食作物总种植面积的 1/3<sup>[1]</sup>。我国水稻病虫害种类种类繁多,农药用量居高不下,因此及时检测稻田害虫不仅有利于预防虫害爆发而造成水稻产量降低,而且能避免农药使用过量而污染环境<sup>[2]</sup>。早期的害虫检测主要依赖人工判断,但是田间害虫情况复杂,对于农业人员的专业水平要求较高,且人工检测费时费力,存在较强的主观性<sup>[3]</sup>。

随着计算机视觉技术的不断更新迭代,越来越多的学者将该技术与农业生产结合起来。Zou 等结合虫洞的颜色特征、形状特征等设计出一种基于支持向量机(SVM)的分类器,用于判断西兰花幼苗图像中的虫洞大小,进而为精准控制农药喷洒用量提供参考<sup>[4]</sup>。田冉等利用红外传感器和 SVM 图像处理方法,在果树害虫检测上取得较好的结果<sup>[5-6]</sup>。上述基于机器学习的方法虽然取得一定的成果,但

较为依赖人工手动地设计特征,且实际农田环境较为复杂,成像条件较为严苛,在实时监测方面存在一定的不足<sup>[7-10]</sup>。基于深度学习的方法近年来逐渐受到研究者的青睐,以卷积神经网络(CNN)、循环神经网络(RNN)、Transformer 等为代表的深度学习方法可对数据进行无监督学习,并从大量的数据中自动学习多种层次的特征信息,较传统机器学习方法展现出更好的性能<sup>[11-16]</sup>。例如,余颢等提出一种基于 SSD 网络模型的水稻害虫识别方法,用特征金字塔模型替换 SSD 原有的多尺度特征图,同时优化激活函数,使得模型对目标的检测精度得到提升,在自建数据集上取得 79.3% 的平均检测精度<sup>[17]</sup>。范春全等针对数据集较小、害虫种类不足导致的模型识别精度下降问题,构建涵盖 16 种水稻病虫害的近 2 万张水稻病虫害数据集,以 ResNet50 为主干网络进行试验,验证模型的性能<sup>[18]</sup>。

上述研究虽然取得了一定的效果,但存在如下问题:(1)真实稻田场景与实验室环境有较大差异,导致训练出来的模型泛化能力不强;(2)实际稻田场景中害虫分布不均、尺度不一、背景噪声较大,容易出现误检、漏检等情况。针对上述问题,本研究基于 YOLO v8 设计一种新的水稻害虫识别模型,以期对稻田害虫实时巡检提供技术支撑。

收稿日期:2024-04-09

基金项目:国家自然科学基金(编号:62173049、62273060);湖北省教育厅科学研究计划(编号:D20211302)。

作者简介:李 龙(2000—),男,湖北武汉人,硕士研究生,研究方向为深度学习与目标检测。E-mail:2022710628@yangtzeu.edu.cn。

通信作者:李梦霞,博士,副教授,硕士生导师,研究方向为油气田软件开发、最优化理论与算法。E-mail:limengxia@yangtzeu.edu.cn。

## 1 材料和方法

### 1.1 数据来源

本试验所用数据集由 2 个部分组成,即 IP102 公开数据集中的部分水稻害虫子集和从 Roboflow 网站(<https://universe.roboflow.com>)中获取的水稻害虫图像数据<sup>[19]</sup>,其中 IP102 包含 102 种害虫的共计 75 222 张图像数据,本研究在上述 2 个数据集来源中综合选取对水稻生产危害较大的 6 种水稻害

虫,即稻褐飞虱、黑尾叶蝉、稻纵卷叶螟、稻蜡、水稻二化螟、稻水蝇作为研究对象。

### 1.2 数据预处理

将获取的图像分辨率统一调整为 640 像素 × 640 像素,并保存为 JPG 格式,剔除其中质量较差的图像后采用 Labelimg 标注工具进行标注,共计获得 1 828 张图像数据,并按照 7 : 2 : 1 的比例将其划分成训练集、验证集和测试集。部分图像数据如图 1 所示。

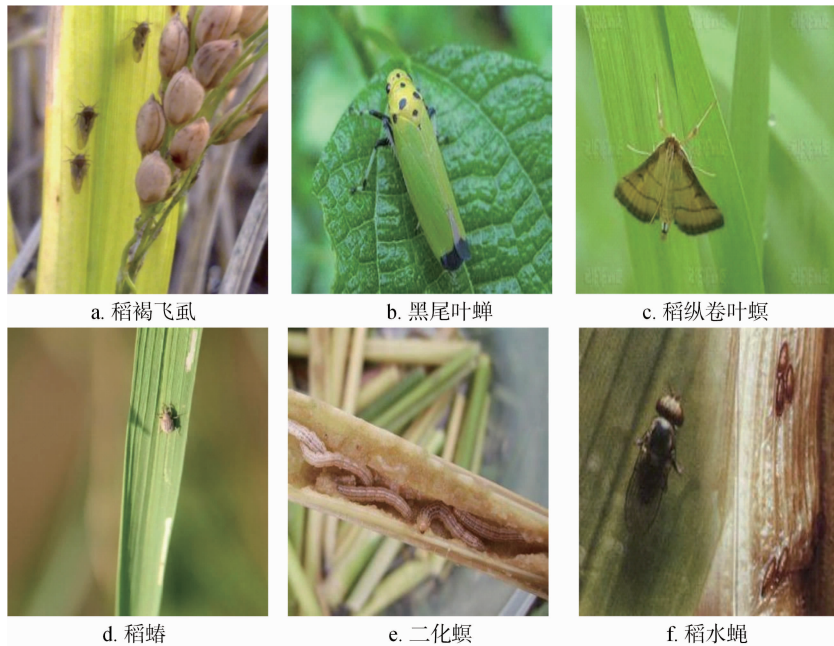


图1 原始数据集图像示例

为提高模型在不同场景下的泛化能力和鲁棒性,本研究对训练集中的数据进行增强,训练集图像数据增强前后各类害虫图像数量如表 1 所示。

表 1 训练集图像数据增强前后害虫图像分布情况

害虫类别	增强前数量 (张)	增强后数量 (张)
稻褐飞虱	152	633
黑尾叶蝉	204	642
稻纵卷叶螟	167	637
稻蜡	315	639
二化螟	103	646
稻水蝇	339	643
总计	1 280	3 840

增强方式:(1)以 50% 的概率在水平和竖直方向上进行翻转;(2)调整亮度,随机增强或降低 25% 的图片亮度,以贴合实际稻田场景不同时间段的光照情景;(3)随机加入 2.5 px 高斯模糊,模拟实时检

测中设备的抖动情况。将 3 种图像增强方法随机组合,得到 3 840 张图像数据。增强后的部分数据集图像如图 2 所示。

### 1.3 方法

1.3.1 网络模型 YOLO v8 是 Ultralytics 于 2023 年 1 月发布的 YOLO 系列最新版本,融合了众多的 SOTA 技术,相较于先前的版本,在扩展性方面有明显提升,可用于目标检测、图像分类、实例分割等多个领域<sup>[20-23]</sup>。相较于先前的 YOLO v5,YOLO v8 提出以新的 C2f 模块替换原本的 C3 模块,丰富了梯度流,同时使模型进一步轻量化;在 Head 部分,YOLO v8 采用目前主流的解耦头结构(decoupled-head),将分类和检测头分离开来,同时遵循 Anchor-Free 思想,舍弃了先前使用的 Anchor-Based<sup>[24]</sup>。在损失函数的计算上,YOLO v8 采用了 Task-Aligned Assigner 正样本分配策略,并引入了 Distribution Focal Loss,用于目标框的回归损失计算,其结构如图 3 所示。



图2 增强后数据集图像示例

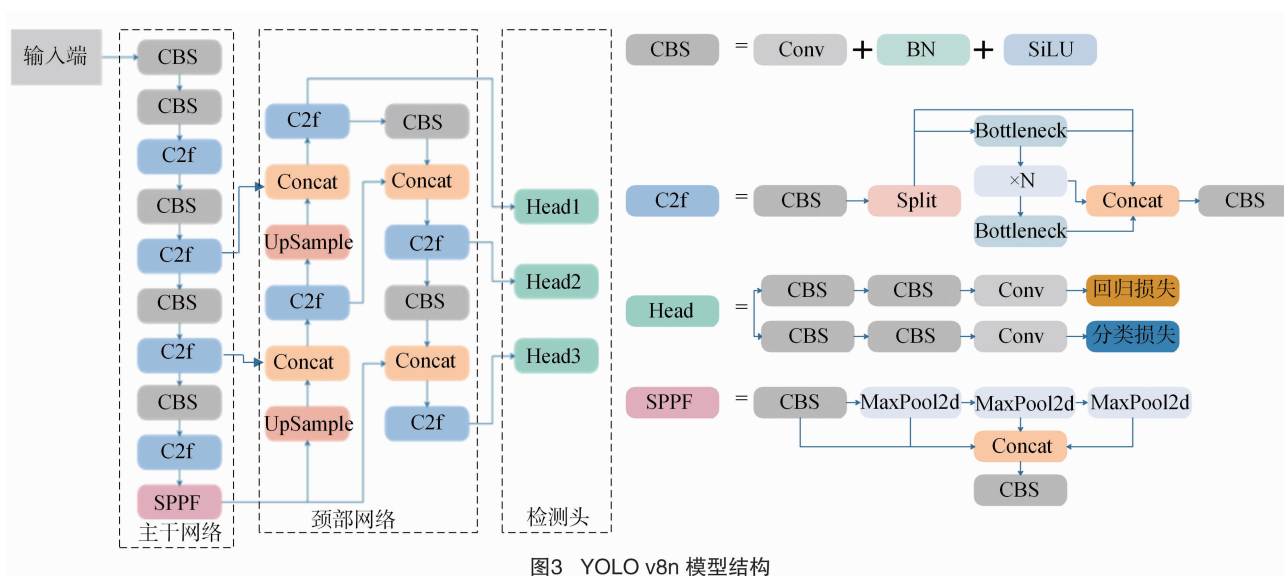


图3 YOLO v8n 模型结构

**1.3.2 改进 YOLO v8 模型** 本研究以 YOLO v8n 为基准模型,提出改进的 YOLO v8 - SDPS,主要改进内容如下:(1)保留主干网络第 1 层的  $3 \times 3$  卷积,当图像分辨率较大时,利用大卷积核能更好地获取全局特征,并同时步长缩短为 1,以尽可能地保留特征图的细粒度信息,另外将 SPD - Conv 模块和深度可分离卷积(DSConv)相结合,重构特征提取模块,称为 SD\_Conv,在降低参数数量的同时,提升模型对低分辨率下小目标的提取能力;(2)遵循 Slim - Neck 范式的设计思想,引入 GSConv 和 VoV - GSCSP

模块,分别替换颈部的卷积操作和 CSP 模块,降低模型的复杂度,综合考虑全局和局部特征信息,更好地捕捉节点的上下文信息;(3)在主干网络的 SPPF 层前添加极化自注意力模块 PSA (polarized self - attention),抑制背景噪声,使得网络模型更加关注个体的空间坐标信息。

改进后的网络结构如图 4 所示。

**1.3.3 SD\_Conv 模块** 在实际害虫检测任务中,害虫个体往往表现出小目标的特点,只占整张图像的较小部分,在卷积操作中特征容易丢失,是造成模

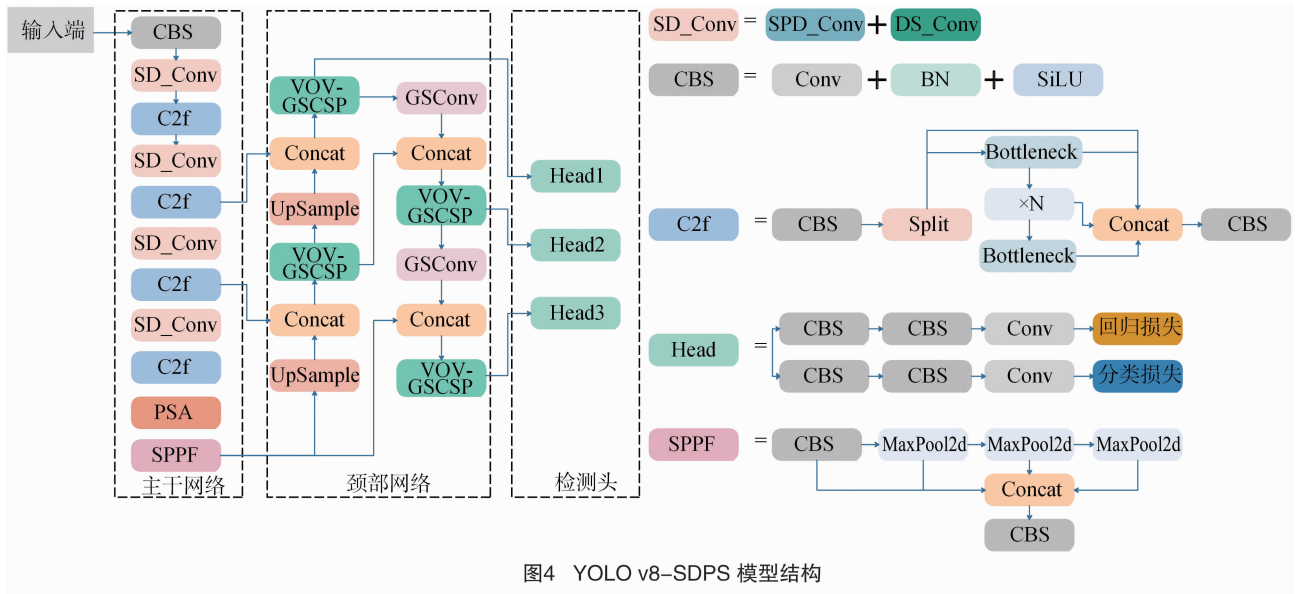


图4 YOLO v8-SDPS 模型结构

型检测性能降低的原因之一,而原始的 YOLO 系列对于此类小目标的检测并不友好,因此本研究在原始 YOLO v8n 的基础上,在主干网络中首先引入 SPD\_Conv 模块,以提升模型的特征提取能力<sup>[25]</sup>。

SPD\_Conv 模块由 2 个部分构成,分别为 space - to - depth 部分 (SPD 层) 和 non - strided convolution 部分 (Conv 层)。SPD 模块会对输入的特征图先进行下采样,并保留通道维度中的信息。对于任何尺度为  $S \times S \times C_1$  的中间特征图  $X$ , SPD 可将其切分为公式(1)至公式(3)的子特征序列,子图  $f_{x,y}$  由  $i + x$  和  $i + y$  按比例整除所有条目  $X(i + y)$  得到。因此,每个子图按比例因子对  $X$  进行下采样。

$$f_{0,0} = X[0:S:scale, 0:S:scale],$$

$$f_{1,0} = X[1:S:scale, 0:S:scale], \dots,$$

$$f_{scale-1,0} = X[scale-1:S:scale, 0:S:scale]; \quad (1)$$

$$f_{0,1} = X[0:S:scale, 1:S:scale], f_{1,1}, \dots,$$

$$f_{scale-1,1} = X[scale-1:S:scale, 1:S:scale]; \quad (2)$$

$$f_{0,scale-1} = X[0:S:scale, scale-1:S:scale], f_{1,scale-1}, \dots,$$

$$f_{scale-1,scale-1} = X[scale-1:S:scale, scale-1:S:scale]。 \quad (3)$$

以  $scale$  为 2 时为例,模块首先对大小为  $S \times S \times C_1$  的特征图  $X$  进行下采样操作,在特征图横向和纵向上每隔 1 个像素进行跳跃采样,得到 4 个宽高减半、通道数保持不变,即大小为  $S/2 \times S/2 \times C_1$  的子特征图  $f_{0,0}, f_{1,0}, f_{0,1}, f_{1,1}$ ,接着将 4 个子特征图在通道维度上依次进行拼接,得到 2 倍下采样中间特征图  $X_2$ 。为尽可能保留所有的判别性特征信息,在 SPD 特征转换层后对中间特征图连接一个包含有

$C_2$  个大小为  $1 \times 1$  的卷积核的非跨步卷积层 ( $stride = 1, C_2 = scale^2 C_1$ ),有效减少细粒度信息的损失,此外使用  $1 \times 1$  的卷积可以减少参数量、扩大模型的感受野,SPD\_Conv 模块示意如图 5 所示。

考虑到 SPD\_Conv 模块在卷积过程中会造成通道维度成倍增加,进而引起参数量的翻倍增长,不利于模型的轻量化实现,本研究在 SPD\_Conv 模块后添加一层深度可分离卷积 (DSConv) 进行跨通道信息融合,在保证卷积核学习能力的同时减少参数量,其结构示意图如图 6 所示。

DSConv 包含逐通道卷积和逐点卷积两部分,在逐通道卷积中,首先对输入特征图的每个通道进行分组卷积,得到中间特征图,然后对中间特征图用  $1 \times 1$  卷积核进行卷积操作,对同一空间坐标上不同层的信息进行有效利用,以进一步提取特征<sup>[26]</sup>。设输入特征图大小为  $W \times H \times C$ ,卷积核大小为  $K \times K$ ,则标准卷积参数量 ( $Conv_{params}$ )、DSConv 参数量 ( $DSConv_{params}$ ) 及两者参数量对比 (Comparison) 的计算公式分别为公式(4)、(5)、(6)。

$$Conv_{params} = K \times K \times C \times C; \quad (4)$$

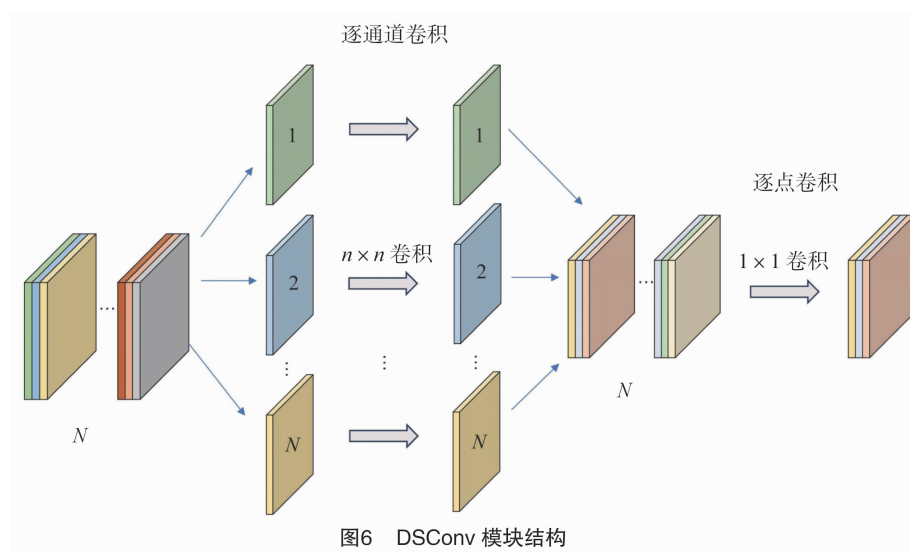
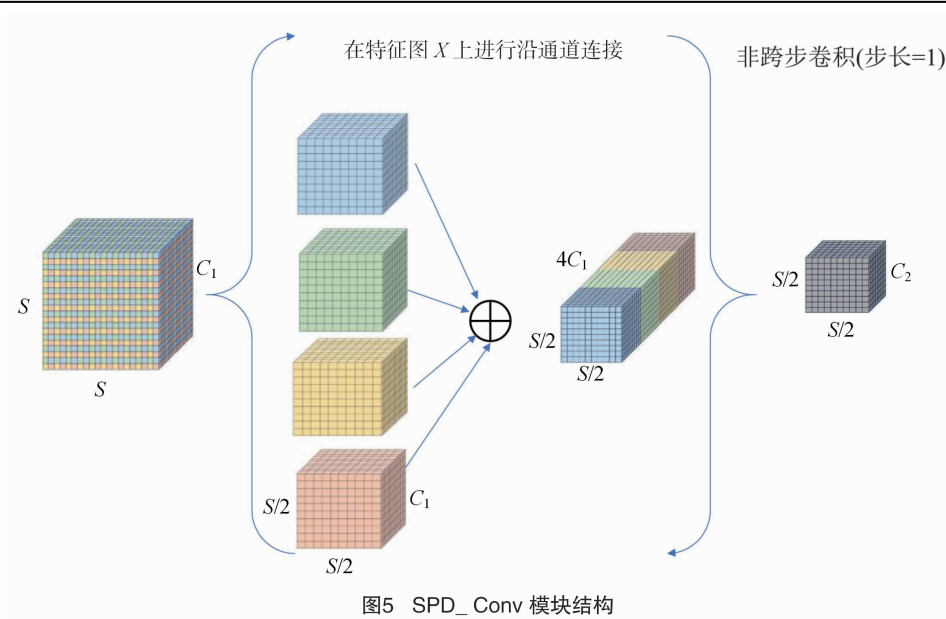
$$DSConv_{params} = K \times K \times 1 \times C + 1 \times 1 \times C \times C; \quad (5)$$

$$Comparison = \frac{DSConv_{params}}{Conv_{params}} = \frac{1}{C} + \frac{1}{K^2}。 \quad (6)$$

由公式(6)可知,标准卷积的参数量是 DSConv 的  $C$  倍,在通道数较多的情况下,DSConv 可以大幅减少模型的计算量,较好地解决了 SPD 层造成的通道数翻倍后参数量增加的问题。

1.3.4 GSCConv 和 VoV - GSCSP 模块 Li 等在自动





驾驶领域提出 GSConv 和 Slim-Neck 设计范式,在减轻模型负担的同时实现更高的检测器计算成本效益,保证了较好的模型大小与检测精度的平衡<sup>[27]</sup>。考虑到本研究对害虫的检测,后续会部署到智能巡检设备中实现实时检测,本研究在上述改进的基础上,在 Neck 部分引入 GSConv,替换原始的标准卷积操作,引入基于 Slim-Neck 思想设计的跨级部分网络模块 VoV-GSCSP,替换原始的 CSP 模块,这在降低一定参数数量的同时,也能保证较好的精度。

GSConv 模块由标准卷积、深度可分离卷积和 Shuffle 混洗 3 个部分组成。该模块将标准卷积得到的特征信息通过 Shuffle 混洗操作渗透到深度可分离卷积生成的特征信息中,有效减少多通道信息损失,同时降低计算成本。GSConv 示意如图 7 所示。

假定输入图像通道数为  $C_1$ ,经过 1 次标准卷积得到的通道数为  $C_2/2$ ,再经过 1 次深度可分离卷积操作,得到一个通道数仍为  $C_2/2$  的输出,将 2 次输出进行 Concat 拼接和 Shuffle 混洗,得到通道数为  $C_2$  的输出结果。在 Backbone 中,特征图维度变化时会不可避免地损失部分特征信息,而 GSConv 尝试保留这些信息,当特征图处理达到 Neck 部分时,在宽高维度上已然最小,在通道维度上最大,冗余信息少,不用再进行变换,因此本研究仅在 Neck 部分用 GSConv 替换标准卷积,在减少参数数量的同时,实现更好的多尺度融合效果。GSConv 的时间复杂度可以表示为式(7)。

$$Time_{GSConv} = O\left[W \times H \times X_1 \times X_2 \times \frac{C_2}{2}(C_1 + 1)\right]。(7)$$

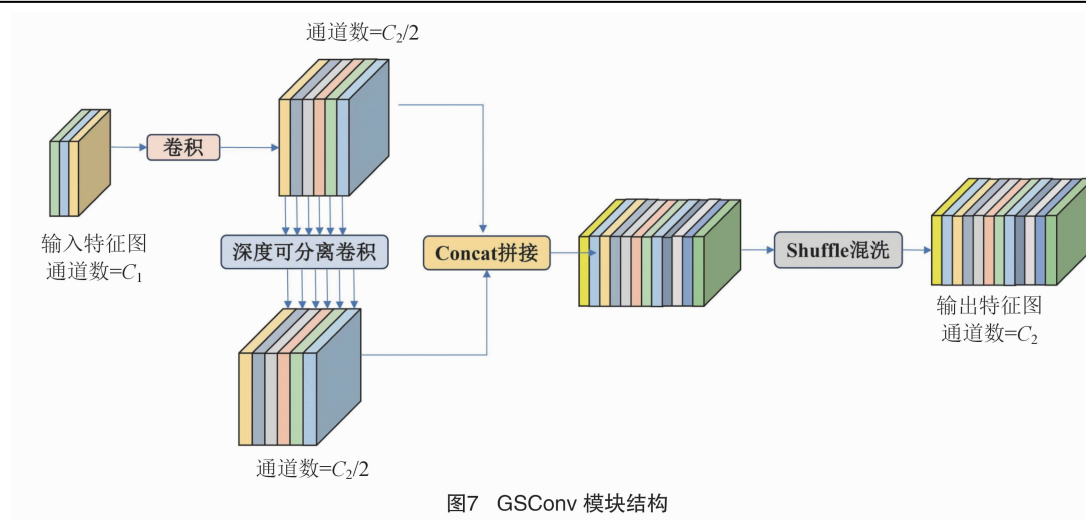


图7 GSConv 模块结构

其中,  $W$ 、 $H$  表示输入特征图的宽、高;  $X_1$  和  $X_2$  表示卷积核的大小;  $C_1$  和  $C_2$  分别表示输入特征图的通道数和输出特征图的通道数。

遵循 Slim-Neck 的设计思想, 本研究引入基于

GSConv 设计的瓶颈层 GS bottleneck 以及跨级部分网络模块 VoV-GSCSP, 并用 VoV-GSCSP 替换原有的 CSP 模块, 其示意如图 8 所示。

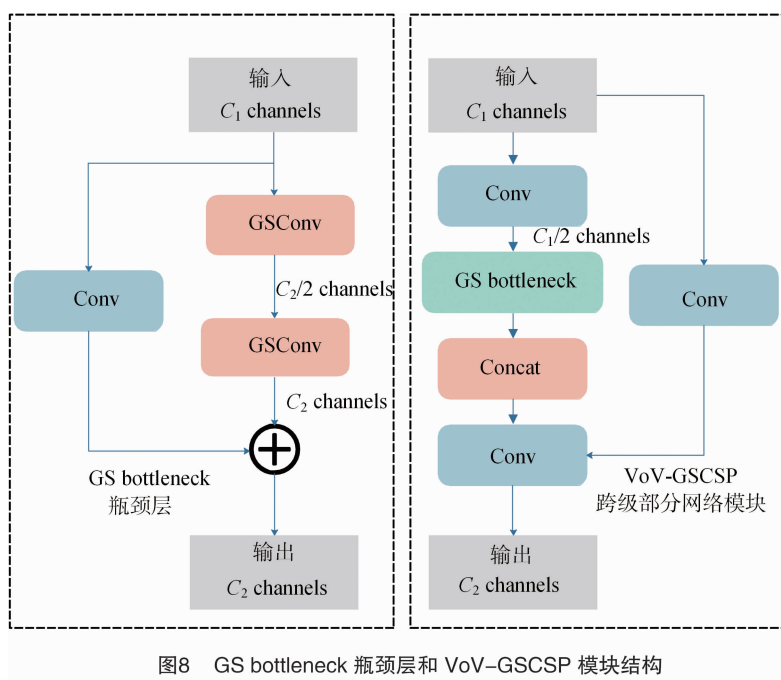
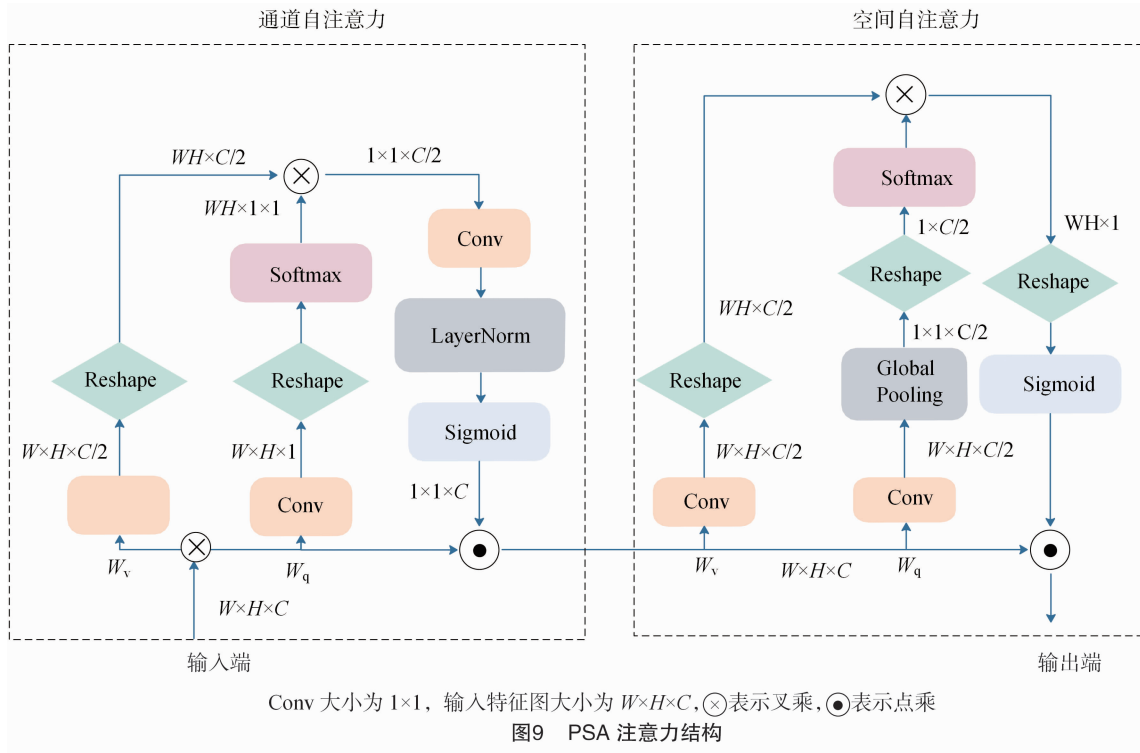


图8 GS bottleneck 瓶颈层和 VoV-GSCSP 模块结构

**1.3.5 PSA 注意力** 在目标检测任务中, 准确地提取个体的空间位置信息是至关重要的。对于包含害虫个体的数据图像, 稻叶、农田等自然背景占比较大, 在卷积等过程中这些无用背景同样参与计算, 产生较多的冗余信息, 对害虫目标的识别造成干扰。为提高空间信息的权重, 降低背景噪声的影响, 本研究在主干提取网络的 SPPF 层前嵌入了极化自注意力 (polarized self-attention, PSA) 机制模块以解决该问题<sup>[28]</sup>。

PSA 极化自注意力机制来源于光学摄影中的极化滤波思想, 即将某一方向上的特征完全折叠, 同时在其正交方向上保证较低的信息压缩损失。如图 9 所示, PSA 模块由通道自注意力和空间自注意力 2 个部分构成, 计算方法见公式 (8)、(9)。

在通道自注意力部分中, 输入的特征图会首先经过 2 次卷积操作转换成通道被完全压缩的特征  $Q$  ( $W \times H \times 1$ ) 以及通道对半压缩的特征  $V$  ( $W \times H \times C/2$ ), 其次对  $Q$  进行降维转置, 并通过 Softmax 函数



对先前压缩的空间特征信息赋予权重,对  $V$  进一步执行降维操作,将通过上述操作后得到的  $Q$  ( $WH \times 1 \times 1$ ) 和  $V$  ( $WH \times C/2$ ) 相乘,并将得到的结果进行卷积和归一化处理,然后通过 Sigmoid 函数计算即可获得对应的通道权重值。

$$Ch(X) = f_{sg}[W(\sigma_1(W_v(X))) \times f_{sm}(\sigma_2(W_q(X)))]. \quad (8)$$

$$f_{sm}(X) = \sum_{j=1}^{N_p} \frac{e^{x_j}}{\sum_{m=1}^{N_p} e^{x_m}} x_j. \quad (9)$$

其中,  $Ch(X)$  代表通道自注意力;  $f_{sg}$  代表 Sigmoid 函数;  $f_{sm}$  代表 softmax 函数;  $W_v$ 、 $W_q$ 、 $W$  代表卷积操作;  $\sigma_1$ 、 $\sigma_2$ 、 $\sigma_3$  代表降/升维操作。

空间自注意力部分会接收来自通道自注意力部分的权重值,同样经过 2 次卷积,得到通道数减半的  $Q$  和  $V$ ,  $Q$  经过全局池化后,在空间维度上信息被完全压缩,经过同自注意力模块中相同的操作后,与被降维后的  $V$  进行相乘、升维、Sigmoid 函数处理,得到对应的空间权重值,将得到的通道和空间权重值进行点乘加权,即可得到特征输出,计算方法见公式(10)、(11)。

$$Sp(X) = f_{sg}[\sigma_3(f_{sm}(\sigma_1(f_{gp}(w_q(X)))) \times \sigma_2(W_v(X))]; \quad (10)$$

$$f_{gp} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X(:, i, j). \quad (11)$$

其中,  $f_{gp}$  代表全局池化操作。

#### 1.4 试验平台及参数设置

软件试验环境:64 位 Windows 10 操作系统,使用 python 3.8 版本进行开发,开发软件为 Pycharm2023 专业版,深度学习框架为 Pytorch 2.0.1, cuda 版本为 11.8; 硬件配置:CPU 为 Intel Core i5 - 13400f, GPU 使用 NVIDIA GeForce RTX2060s, GPU 显存为 8 GB。所有试验均在相同环境下进行,试验时间为 2024 年 1—3 月,试验地点为湖北省荆州市长江大学计算机科学学院。模型训练过程中的超参数设置如表 2 所示。

表 2 改进害虫识别模型训练超参数设置

迭代次数 (轮)	批量大小	优化器	初始 学习率	动量因子	权重衰减
200	16	SGD	0.01	0.937	0.000 5

#### 1.5 评价指标

为检验本研究提出模型的优越性,采用目标检测领域中公认的精确率 (precision,  $P$ )、召回率 (recall,  $R$ )、平均精度均值 (mean average precision,  $mAP$ )、浮点计算量 (GFLOPs) 作为评价指标进行优越性评价,计算公式分别为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \times 100\%; \quad (12)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\%; \quad (13)$$

$$AP = \int_0^1 P(R) dR; \quad (14)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n P_{AP_i} \quad (15)$$

本研究中,  $N_{TP}$  (true positives) 代表正确识别出害虫的数量;  $N_{FP}$  (false positives) 代表错误识别的害虫数量;  $N_{FN}$  (false negatives) 表示未检测出的害虫数量;  $AP$  是 P-R 曲线对应的积分, 代表平均精度;  $n$  为类别数量, 在本研究中  $n=6$ 。

## 2 结果与分析

### 2.1 消融试验结果

为验证本研究改进的有效性, 在相同的试验环境和参数设定下, 基于 YOLO v8n 进行消融试验, 分析各部分改进的有效性。记前文主干改进为试验 A、颈部改进为试验 B、添加注意力模块为试验 C, a~f 依次表示的本研究中的 6 类害虫: 稻褐飞虱、黑尾叶蝉、稻纵卷叶螟、稻蜡、二化螟、稻水蝇。

由表 3 可知, (1) 在主干网络中引入 SD\_Conv 后, 模型对 6 类害虫的识别精度分别提高 6.4、2.3、

2.4、1.6、3.3、4.0 百分点, 说明加入 SPD 模块的特征提取网络尽可能地保留了相对细小的特征信息, 对于图像数据中表现出小目标特征的害虫具有更好的特征提取能力; (2) 在颈部引入 GSCConv 和 VoV-GSCSP 模块后, 虽然在  $mAP_{0.5}$  上仅提升 1.0 百分点, 但计算量降低 14.6%, 满足在算力较低的设备上部署的要求, 同时证明在 YOLO v8n 中引入 Slim-Neck 范式具有可行性; (3) 加入 PSA 注意力模块后, 模型的平均检测精度提升 2.8 百分点, 说明加入注意力模块后的模型能够更加关注个体的空间信息, 有效降低背景噪声的影响, 提高模型的识别能力。此外, 模型对于稻水蝇和稻褐飞虱的识别精度偏低, 核验数据集后发现, 这 2 类害虫在不同生长时期的外观有一定差异, 且存在部分生长时期害虫特征不易区分的问题, 导致检测结果偏低, 后期会针对不同生长时期的害虫进行进一步细分, 使得模型有更好的检测效果。

表 3 消融试验结果

模型	$AP_{0.5} (%)$						$mAP_{0.5} (%)$	计算量 (GFLOPs)	参数量 (M)
	a	b	c	d	e	f			
YOLO v8n	73.2	89.1	91.5	89.7	79.3	72.2	82.5	8.9	3.15
A	79.6	91.4	93.9	91.3	82.6	76.2	85.8	8.3	2.93
B	74.1	90.7	92.2	90.0	80.3	73.7	83.5	7.6	2.85
C	78.6	90.9	93.1	92.2	80.6	76.4	85.3	9.0	3.15
YOLO v8-SDPS	80.6	92.5	94.2	92.4	82.7	77.1	86.6	7.5	2.62

由表 3 和图 10 可知, 基于上述 3 种改进的 YOLO v8-SDPS 相较于基准模型计算量有所降低,  $mAP_{0.5}$  提升 4.1 百分点, 准确率和召回率均明显提升, 说明 YOLO v8-SDPS 充分融合了各模块改进之后的优势, 实现了模型检测性能与计算量的平衡, 符合实际生产过程中的需求。

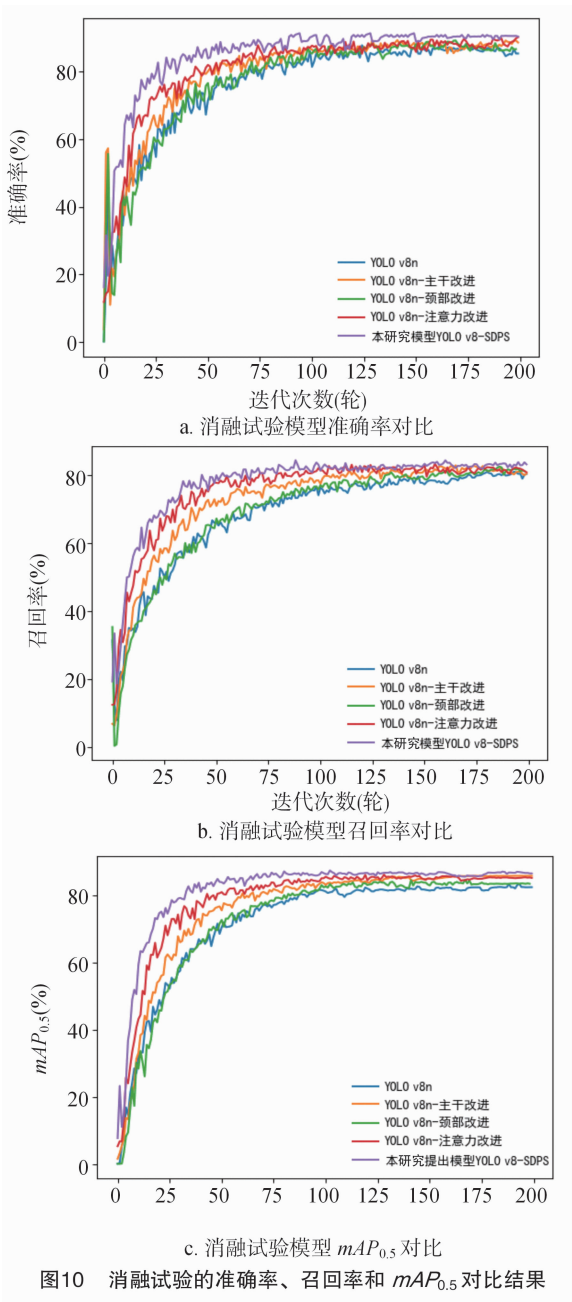
### 2.2 对比试验

**2.2.1 注意力对比** 在模型中加入注意力模块是研究者们常用的提升模型性能的方法, 为验证本研究模型添加 PSA 注意力模块的有效性, 将 PSA 与 SE、CA、ECA、CBAM 等经典注意力机制进行对比, 注意力添加位置保持一致, 对比结果如图 11、表 4 所示, CA 模块对于模型精度略有提升 (提升 1.0 百分点), CBAM 模块对于模型提升较为明显 (提升 2.3 百分点), 但低于 PSA, 且模型收敛速度较慢, ECA 模块加快了模型的收敛速度, 但对于精度的提

升效果甚微, SE 模块对模型并未起到提升精度的作用。本研究采用的 PSA 注意力模块不仅对于模型精度提升明显 (提升 2.8 百分点), 并且模型在迭代 60 轮时接近收敛, 均优于其他注意力模块, 证明了本研究引入 PSA 模块的优越性。

**2.2.2 不同模型对比** 在本研究建立的数据集上, 对比分析 YOLO v8-SDPS 与当前主流的目标检测算法, 结果如表 5 所示。由表 5 可知, 双阶段检测模型 Faster R-CNN 的检测效果最差且运算量最大。YOLO v8-SDPS 相比于改进前的基准模型 YOLO v8n 降低了一定的计算量, 精确率、召回率、 $mAP_{0.5}$  均不同程度地优于其他网络模型, 平均精度较 YOLO 系列中的 YOLO X、YOLO v5s、YOLO v7-Tiny、YOLO v8n 分别提高了 6.4、3.9、5.2、4.1 百分点, 而 FPS 虽然比改进前有所降低, 但与 YOLO v7-Tiny 相当, 仍满足实时性检测需求, 相较于其他模型





展现出更为突出的性能。

2.3 目标检测结果及分析

本研究对基准模型和改进后的模型在尺度不一、部分遮挡、背景相似、目标微小、强光照等不同条件下的检测效果进行可视化对比分析,以验证改进后模型的性能效果,对比结果如图 12 所示。由图 12 可以看出,原始模型存在一定的漏检、误检情况,将图 12 - b 图左下角部分遮挡的稻褐飞虱未被识别到,将图 12 - e 图的背景错误识别为稻蜡,将图 12 - h 图背景中的叶片识别成黑尾叶蝉等,说明原始模型在特征提取阶段丢失较多特征信息。而改

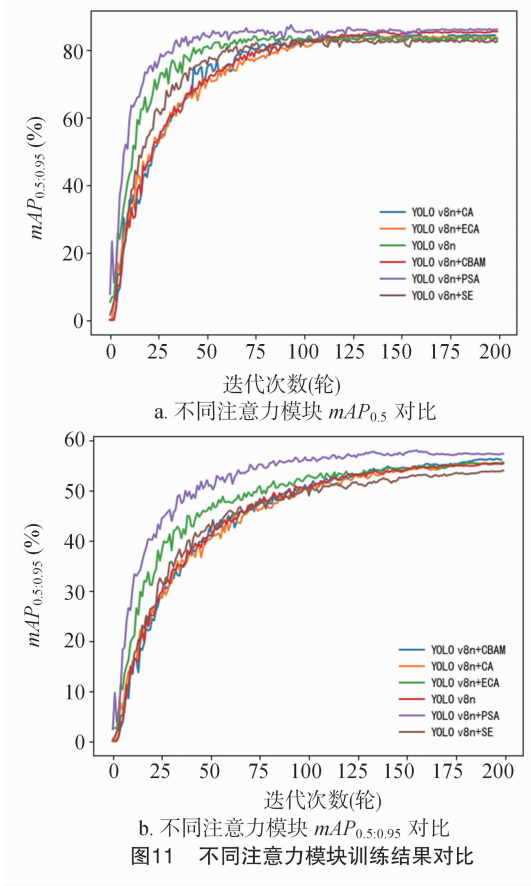


表 4 注意力机制试验结果对比

模型	精确率 (%)	召回率 (%)	$mAP_{0.5}$ (%)	$mAP_{0.5:0.95}$ (%)
YOLO v8n	83.2	80.6	82.5	54.1
YOLO v8n + SE	81.9	77.2	81.9	52.3
YOLO v8n + CA	85.7	81.3	83.5	54.4
YOLO v8n + ECA	84.3	80.9	82.7	54.5
YOLO v8n + CBAM	86.2	82.5	84.8	55.6
YOLO v8n + PSA	87.3	82.9	85.3	56.3

表 5 不同模型对比结果

模型	精确率 (%)	召回率 (%)	$mAP_{0.5}$ (%)	帧率 (帧/s)	计算量 (GFLOPs)
Faster R - CNN	69.5	67.9	64.6	19	203.5
CenterNet	81.7	82.6	77.5	41	70.2
YOLO X	82.3	78.7	80.2	96	27.2
YOLO v5s	83.6	79.3	82.7	113	16.3
YOLO v7 - Tiny	84.1	78.6	81.4	101	13.4
YOLO v8n	83.2	80.6	82.5	124	8.9
YOLO v8 - SDPS	88.7	84.3	86.6	102	7.5

进后的模型对于小目标边缘特征的提取更为充分,同时能有效降低背景噪声的影响,使模型更关注个

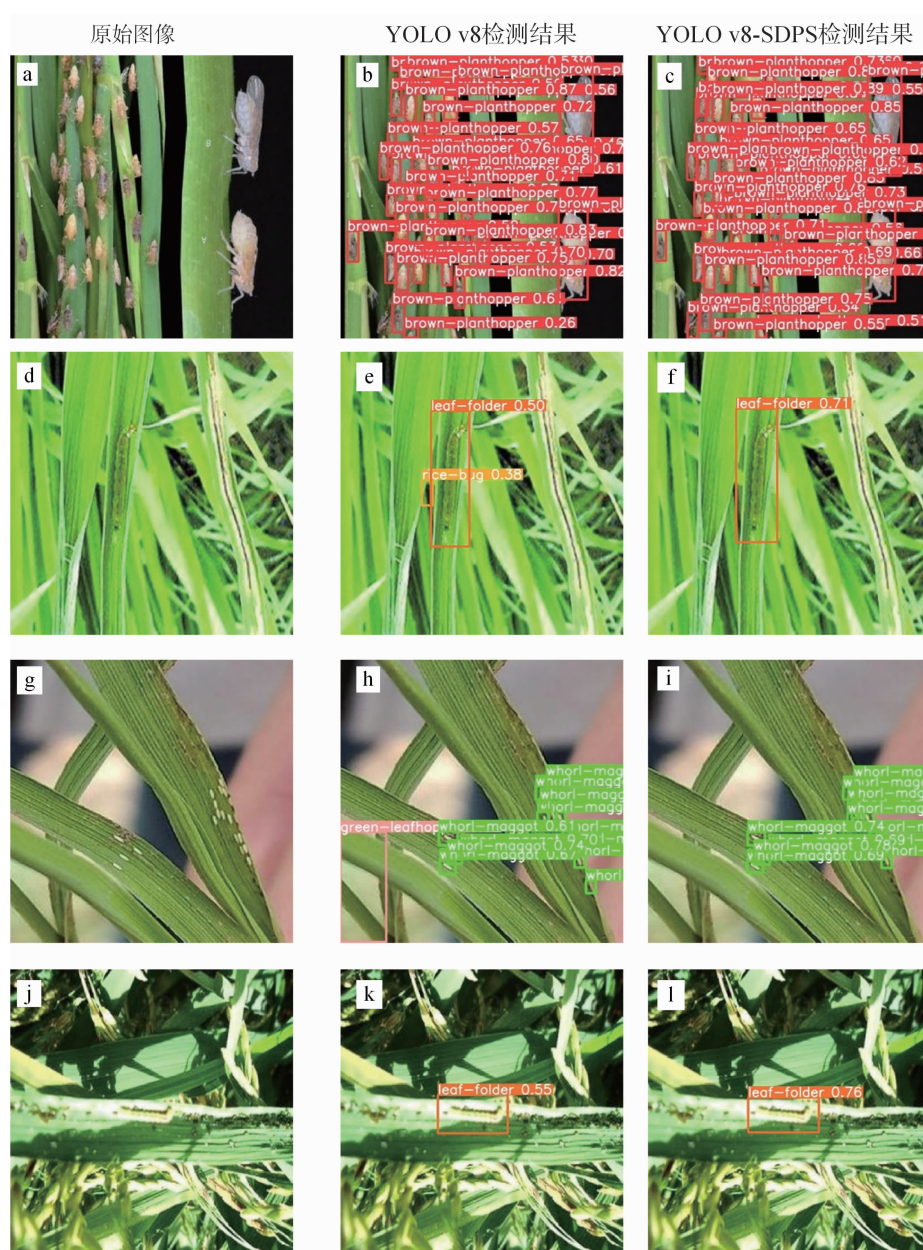


图12 模型在尺度不一、部分遮挡、背景相似、微小目标、强光照条件下检测效果对比

体特征,在不同环境下依然能保持优于基准模型的检测性能。

### 3 结论与讨论

针对现有水稻害虫识别方法准确率较低、计算较为复杂以及缺乏检测功能等问题,本研究提出一种基于 YOLO v8 的改进模型并命名为 YOLO v8 - SDPS,对 6 种常见且危害较大的水稻害虫进行研究。该模型通过引入 SD\_Conv 模块替代标准卷积、引入基于 Slim - Neck 范式的 GSConv 和 VoV - GSCSP 模块、添加 PSA 自注意力,降低参数量,同时

提高模型的检测性能。通过对比试验结果可知,本研究提出的改进后的模型在自建水稻害虫数据集上的平均识别精度提升 4.1 百分点,参数量、计算量分别降低了 16.8%、15.7%,满足部署在低能耗设备上的要求。这表明本研究取得较好的效果,证明改进的有效性。本研究中的 YOLO v8 - SDPS 模型不仅可以对水稻害虫进行识别分类,而且可以对发生虫害的水稻进行检测定位。在实际生产中,可以帮助相关从业者及时、准确定位稻田虫害发生的位置,从而采取相应的防治措施。该模型亦可为农业领域中相关农作物的害虫识别防治提供一定的

参考。

本研究针对现有的水稻害虫识别问题提出的 YOLO v8 - SDPS 模型取得了良好的试验效果,但仍存在部分害虫在不同生长周期因外观特征相似而识别精度较低的问题。未来的研究将重点关注相似外观特征的害虫识别问题,通过扩大数据集、细化标签标注等方式,持续优化模型在多种复杂情况下的识别能力,为农业工作者提供准确的害虫信息,促进智慧农业的发展。

#### 参考文献:

- [1] 褚世海,李儒海,黄启超,等. 湖北省水稻田农药使用现状调查[J]. 中国植保导刊,2022,42(12):65-68.
- [2] 钟朝军,戴长庚,毛向华. 水稻主要虫害绿色防控技术研究[J]. 耕作与栽培,2022,42(6):123-124.
- [3] 蒋心璐,陈天恩,王 聪,等. 农业害虫检测的深度学习算法综述[J]. 计算机工程与应用,2023,59(6):30-44.
- [4] Zou K L, Ge L Z, Zhou H, et al. Broccoli seedling pest damage degree evaluation based on machine learning combined with color and shape features[J]. Information Processing in Agriculture, 2021, 8(4):505-514.
- [5] 田 冉,陈梅香,董大明,等. 红外传感器与机器视觉融合的果树害虫识别及计数方法[J]. 农业工程学报,2016,32(20):195-201.
- [6] Rustia D J A, Lin C E, Chung J Y, et al. Application of an image and environmental sensor network for automated greenhouse insect pest monitoring[J]. Journal of Asia - Pacific Entomology, 2020, 23(1):17-28.
- [7] 王春桃,梁炜健,郭庆文,等. 农业害虫智能视觉检测研究综述[J]. 中国农机化学报,2023,44(7):207-213.
- [8] 温艳兰,陈友鹏,王克强,等. 基于机器视觉的病虫草害检测综述[J]. 中国粮油学报,2022,37(10):271-279.
- [9] Domingues T, Brandão T, Ferreira J C. Machine learning for detection and prediction of crop diseases and pests: a comprehensive survey[J]. Agriculture, 2022, 12(9):1350.
- [10] Kasinathan T, Uyyala S R. Machine learning ensemble with image processing for pest identification and classification in field crops[J]. Neural Computing and Applications, 2021, 33(13):7491-7504.
- [11] 吴子炜,夏 芳,陆林峰,等. 基于改进 YOLO v5 的水稻主要害虫识别方法[J]. 江苏农业科学,2023,51(21):218-224.
- [12] 温艳兰,陈友鹏,王克强,等. 基于迁移学习和改进残差网络的复杂背景下害虫图像识别[J]. 江苏农业科学,2023,51(8):171-177.
- [13] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6):84-90.
- [14] Xiao J Q, Zhou Z Y. Research progress of RNN language model [C]//2020 IEEE International Conference on Artificial Intelligence and Computer Applications(ICAICA). Dalian, China: IEEE, 2020: 1285-1288.
- [15] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [EB/OL]. (2017-06-12)[2024-04-01]. <https://arxiv.org/abs/1706.03762>.
- [16] 慕君林,马 博,王云飞,等. 基于深度学习的农作物病虫害检测算法综述[J]. 农业机械学报,2023,54(增刊2):301-313.
- [17] 余 颢,吴 伶,单鲁泉. 基于 SSD 网络模型改进的水稻害虫识别方法[J]. 郑州大学学报(理学版),2020,52(3):49-54.
- [18] 范春全,何彬彬. 基于迁移学习的水稻病虫害识别[J]. 中国农业信息,2020,32(2):36-44.
- [19] Wu X P, Zhan C, Lai Y K, et al. IP102: a large - scale benchmark dataset for insect pest recognition [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach, CA, USA: IEEE, 2019: 8779-8788.
- [20] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real - time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas, NV, USA: IEEE, 2016: 779-788.
- [21] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Honolulu, HI, USA: IEEE, 2017: 6517-6525.
- [22] Redmon J, Farhadi A. YOLO v3: an incremental improvement[EB/OL]. (2018-04-08)[2023-12-16]. <https://arxiv.org/abs/1804.02767v1>.
- [23] Thuan D. Evolution of Yolo algorithm and YOLO v5: the State - of - the - Art object detection algorithm[J/OL]. Oulu University of Applied Sciences, 2021: 1-61 (2021-03-04)[2023-04-27]. <https://urn.fi/URN:NBN:fi:amk-202103042892>.
- [24] Fu J, Liu J, Tian H J, et al. Dual attention network for scene segmentation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach, CA, USA: IEEE, 2019: 3141-3149.
- [25] Sunkara R, Luo T. No more strided convolutions or pooling: a new CNN building block for low - resolution images and Small objects [EB/OL]. (2022-08-07)[2023-11-13]. <https://doi.org/10.48550/arXiv.2208.03641>.
- [26] 毛远宏,贺占庄,刘露露. 目标跟踪中基于深度可分离卷积的剪枝方法[J]. 西安交通大学学报,2021,55(1):52-59.
- [27] Li H L, Li J, Wei H B, et al. Slim - neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles [EB/OL]. (2022-06-06)[2023-11-01]. <https://arxiv.org/abs/2206.02424v3>.
- [28] Liu H J, Liu F Q, Fan X Y, et al. Polarized self - attention: towards high - quality pixel - wise mapping[J]. Neurocomputing, 2022, 506:158-167.