doi:10.6041/j.issn.1000-1298.2024.11.004

基于改进 YOLO v5s 的水稻害虫识别研究

王泰华[1,2]　郭亚州[1]　张家乐[1]　张晨阳[1]

(1.河南理工大学电气工程与自动化学院,焦作 454003;
2.河南省煤矿装备智能检测与控制重点实验室,焦作 454003)

摘要:水稻害虫识别时,受稻田环境影响易出现目标被遮挡、与背景颜色相似、多目标相邻等问题导致识别精度降低。为此本文提出了一种基于改进 YOLO v5s 的水稻害虫识别模型 YOLO v5s-Coordslimneck,通过替换主干网络中的普通卷积为 CoordConv,增强了模型对目标位置信息的获取能力;引入 CBAM 注意力机制,提升了模型对目标区域的关注度;采用 Slim-neck 减少了计算量并增强了特征处理能力;引入 Soft-NMS 算法优化相邻目标边框筛选,减少漏检。实验结果表明,改进后的 YOLO v5s 模型在水稻害虫数据集上的平均精度均值达到 94.3%,相比原模型提升 3.8 个百分点,比其他主流模型 YOLO v3、YOLO R-CSP、YOLO v7、YOLO v8s 提升 1.5、12.7、13.6、1.9 个百分点。消融实验进一步验证了改进模型中各个组件的有效性。热力图分析也体现了改进模型能够更好地学习害虫特征。综上,本文提出的改进 YOLO v5s 模型在提高水稻害虫检测精度方面取得了显著成效,为防控水稻虫害提供了一种精确的识别方法。

关键词:水稻;害虫识别;CBAM;Slim-neck;YOLO v5s

中图分类号:S435.16　　**文献标识码**:A　　**文章编号**:1000-1298(2024)11-0039-10　　**OSID**:

Rice Pest Identification Based on Improved YOLO v5s

WANG Taihua[1,2]　GUO Yazhou[1]　ZHANG Jiale[1]　ZHANG Chenyang[1]

(1. *School of Electrical Engineering and Automation*, *Henan Polytechnic University*, *Jiaozuo* 454003, *China*
2. *Henan Key Laboratory of Intelligent Detection and Control of Coal Mine Equipment*, *Jiaozuo* 454003, *China*)

Abstract: When identifying rice pests, issues such as targets being obscured, similarity to the background color, and proximity of multiple targets due to the rice field environment can lead to reduced identification accuracy. To address this, a rice pest identification method was proposed based on an improved YOLO v5s. The method enhanced the model's ability to capture target location information by replacing ordinary convolution in the backbone network with CoordConv. It introduced the CBAM attention mechanism to increase the model's focus on the target area. The Slim-neck architecture was adopted to enhance feature processing capabilities and reduce computational load. The introduction of the Soft-NMS algorithm optimized the selection of adjacent target bounding boxes, reducing missed detections. Experimental results showed that the improved YOLO v5s model achieved an mAP of 94.3% on the rice pest dataset, which was an increase of 3.8 percentage points over the original model and 1.5, 12.7, 13.6 and 1.9 percentage points higher than that of the other mainstream models such as YOLO v3, YOLO R-CSP, YOLO v7, and YOLO v8s, respectively. Ablation experiments further validated the effectiveness of each component in the improved model. Heat map analysis also demonstrated that the improved model can better learn pest features. In summary, the improved YOLO v5s model proposed achieved significant results in improving the accuracy of rice pest detection, providing a more precise identification method for the prevention and control of rice pests.

Key words: rice; pest identification; CBAM; Slim-neck; YOLO v5s

收稿日期:2024-01-09　修回日期:2024-05-21
基金项目:国家自然科学基金项目(61903126)和河南省科技攻关项目(212102210503)
作者简介:王泰华(1976—),男,副教授,主要从事工业过程控制研究,E-mail:9567551@qq.com

0 引言

水稻是我国重要的粮食作物之一[1],其种植面积广泛[2]。水稻产量受虫害影响较大,据全国农业技术推广服务中心统计,2023 年水稻虫害受灾面积达 5.333×10^8 hm^2,造成的经济损失占所有农作物虫害经济损失的 55%。因此高效识别水稻害虫做好虫害预测防控,对保障粮食安全意义重大。

传统水稻害虫为人工识别,存在较大局限,识别效率低。近年来,图像识别技术发展迅猛,该技术也逐渐应用到虫害识别领域。目前应用于虫害检测的图像识别方法主要有两大类:基于传统图像处理的检测方法和基于深度学习的检测方法。这些基于传统图像处理的方法都需要通过人工的方式提取图像边缘、颜色和纹理等特征,其识别结果取决于人工提取特征的优劣,因此泛化能力比较差。

近几年,计算机硬件算力的提升使得深度学习得到了快速的发展,其模型在众多领域都展现出优异的性能。相比于机器学习模型,深度学习模型的泛化能力更强,性能更稳定,且无需进行特征提取[3-9]。这些方法虽然弥补了机器学习模型上的不足,但是在水稻害虫检测方面仍然存在一些问题:①由于田间环境复杂,害虫很容易被其他物体遮挡导致重要特征缺失[10],造成误检和漏检。②害虫颜色与背景颜色相似,其特征与背景关联性强,造成误检和漏检。③害虫数量较多且分布不均匀也容易出现漏检。

针对以上问题,本文提出一种基于改进 YOLO v5s 的水稻害虫识别方法。通过将 YOLO v5s 主干中除第 1 层卷积之外的普通卷积替换为 CoordConv,使模型能够获得更多位置信息,以解决对多目标识别时的漏检问题;通过在主干网络末端引入 CBAM 注意力机制[11],以提升模型对目标区域的关注度,进而提高对遮挡物体识别效果以及背景颜色相似目标识别效果;通过在 YOLO v5s 的颈部采用 Slim - neck 架构,以减少模型计算量;采用 Soft - NMS 算法[12]优化相邻目标边框的筛选。

1 实验数据集

实验所采用的水稻害虫数据集来自安徽省农业科学院。该数据集中水稻害虫总共有 10 种,如图 1 所示,分别为稻螟蛉虫卵、稻螟蛉幼虫、稻螟蛉成虫和虫苞、水稻二化螟幼虫和成虫、水稻二化螟虫蛹、稻绿蝽成虫、黑色带斑稻绿蝽幼虫以及绿色带斑稻绿蝽幼虫。其中,稻螟蛉虫苞是稻螟蛉老熟幼虫化为成虫的前一阶段,该阶段与成虫阶段近似,因此在进行数据集分类时,将稻螟蛉虫苞与稻螟蛉成虫归为一类。

实验所采用的原数据集总共有 2 180 幅图像,这对于模型的训练是远远不够的。为了使模型能够充分地学习到各类害虫的特征,采用翻转、平移、旋转、改变亮度、加噪声、擦除等数据增强方式对原数据集中的样本进行扩充,扩充后样本为 6 540 幅。数据增强效果如图 2 所示。

数据集所使用的标注软件为 labelImg,该软件标注后的数据格式为 Pascal VOC[13],在所有的数据标注完成后,将数据格式转化为 YOLO 格式。数据集被划分为训练集、验证集、测试集,比例为 7∶1∶2。划分数据集后,各类别样本在这 3 个数据集中所对应的数量如表 1 所示。

2 水稻害虫检测方法

2.1 YOLO v5s 网络结构

YOLO v5s 的网络结构主要有主干网络、颈部网络、检测头。主干网络由分类网络构成,分类网络为 CSPDarknet53[14],主要是对 Darknet53 的每个残差模块进行 CSP 连接得来,而使用 CSP 连接可以减少其参数量和计算量;颈部网络由 FPN[15] + PAN[16]结构与 CSP 模块构成,FPN + PAN 结构主要用于对主干网络提取的特征进行处理[17],通过上采样和下采样的方式,将不同尺度的特征图进行融合,以此增强网络对特征的表达能力;而颈部中的 CSP 模块通过 CSP 连接的方式进一步降低了模型的参数量和计算量;检测头由 3 个输出头构成,每个输出头针对 3 个不同尺度的特征图进行预测,尺度分别为80 像素 × 80 像素、40 像素 × 40 像素以及 20 像素 × 20 像素,因此输出头可以对大、中、小 3 个目标进行检测。YOLO v5s 网络结构如图 3 所示。

2.2 网络结构

为了解决水稻害虫识别时因遮挡、数量多、与背景颜色相似导致误检、漏检的问题,本文在 YOLO v5s 网络基础上进行了改进:①将 YOLO v5s 主干网络中除第 10 层外的普通卷积全部替换为 CoordConv。②在主干网络中的 C3 模块与 SPPF 模块之间加入 CBAM 注意力机制。③将 YOLO v5s 的颈部架构修改为 Slim - neck 架构。④将网络中的 NMS 算法更换为 Soft - NMS 算法。改进后网络结构如图 4 所示。

2.2.1 CoordConv

CoordConv 主要由 2 部分构成:对输入的特征图进行 Concat 操作为其增加 2 个坐标通道,其中一个通道表示每个像素的 x 坐标,另一个通道表示每个

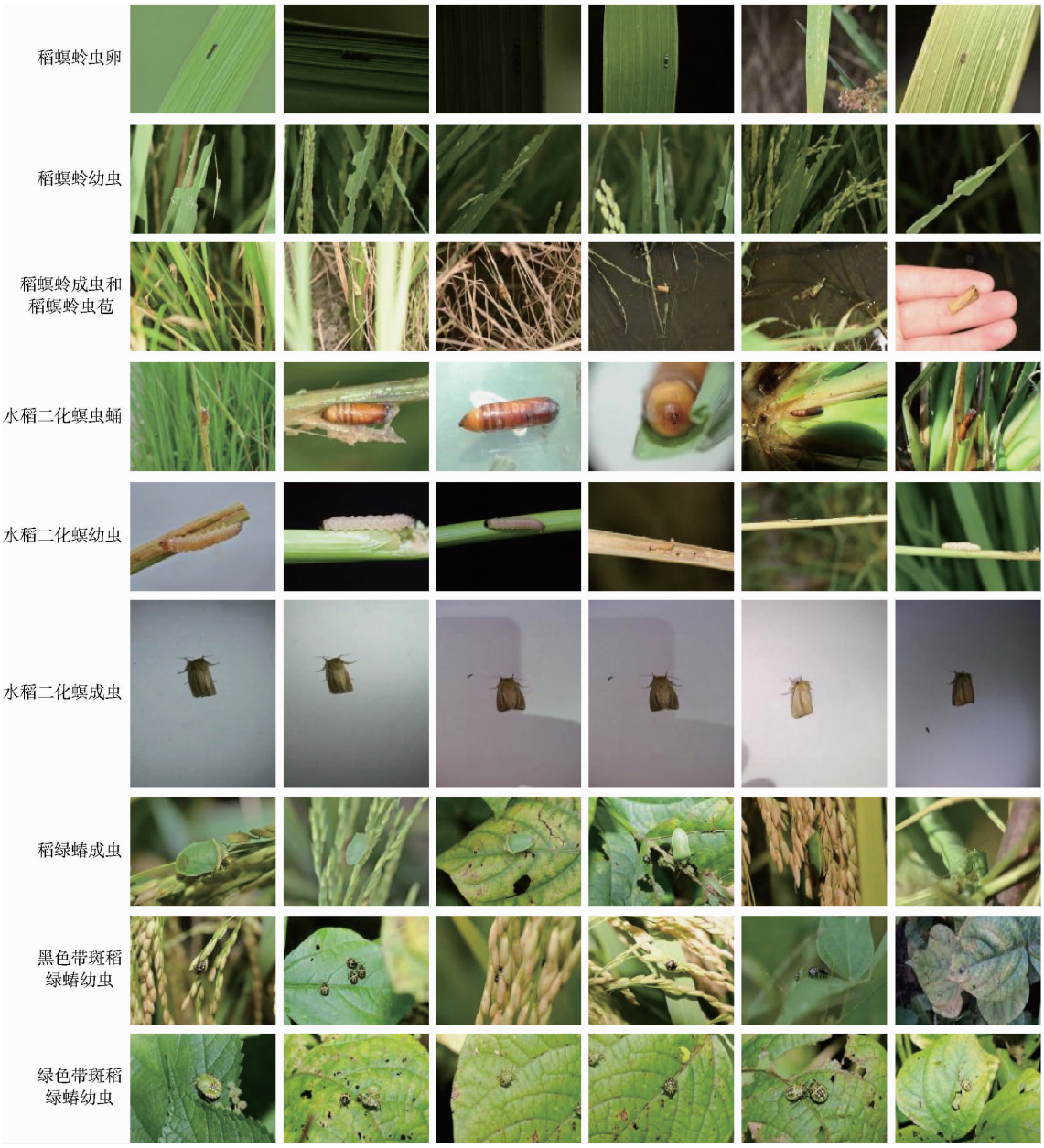

图 1　水稻害虫图像示例

Fig. 1　Images of rice pests



图 2　数据增强效果图

Fig. 2　Data enhancement effect diagram

像素的 y 坐标；普通卷积。CoordConv 在进行卷积计算时，先对输入图像添加两个通道，使得图像中的每个像素都带有位置信息，之后再对其进行普通卷积操作。CoordConv 所添加的位置信息可以使其在对图像进行局部运算时，增加图像局部信息与整体信息的关联性，从而更精准地确定目标位置，并且 CoordConv 还具备传统卷积的平移不变性，使得图像在进行卷积运算时，其不同位置的像素可以共享统一的卷积核参数，进而更有效地学习图像的本质特征，CoordConv 结构如图 5 所示。

2.2.2　CBAM 注意力机制

CBAM 注意力机制是一种结合通道注意力机制与空间注意力机制而成的模块[18]，其结构如图 6 所示。

表 1 害虫数据集信息

Tab. 1 Pest dataset information

害虫种类	训练集样本数量	验证集样本数量	测试集样本数量	标签
稻螟蛉虫卵	585	89	166	DMP1
稻螟蛉幼虫	1 025	79	293	DMP2
水稻二化螟幼虫	400	134	115	DDMP1
水稻二化螟虫蛹	468	62	130	DDMP2
水稻二化螟成虫	621	107	166	DDMP3
绿色带斑稻绿蝽幼虫	665	109	216	DLP2
黑色带斑稻绿蝽幼虫	652	70	172	DLP1
稻绿蝽成虫	535	61	151	DLP3
稻螟蛉虫苞和稻螟蛉成虫	565	81	142	DMP3

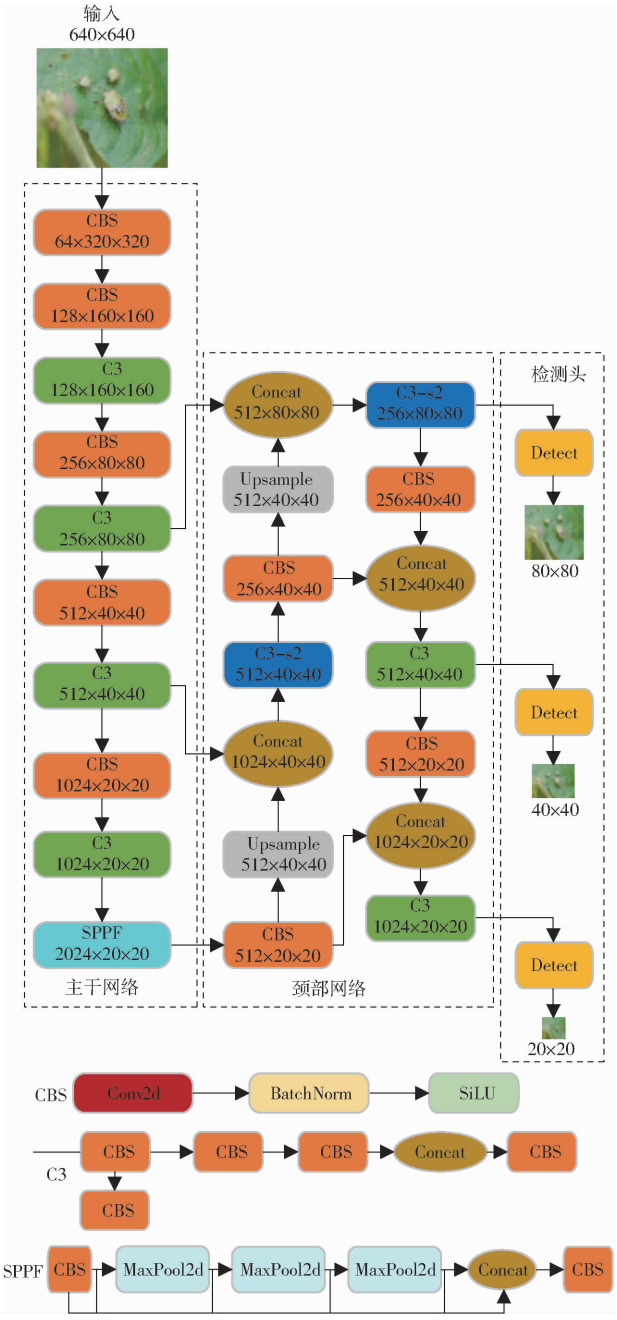

图 3 YOLO v5s 结构图

Fig. 3 YOLO v5s structure diagram

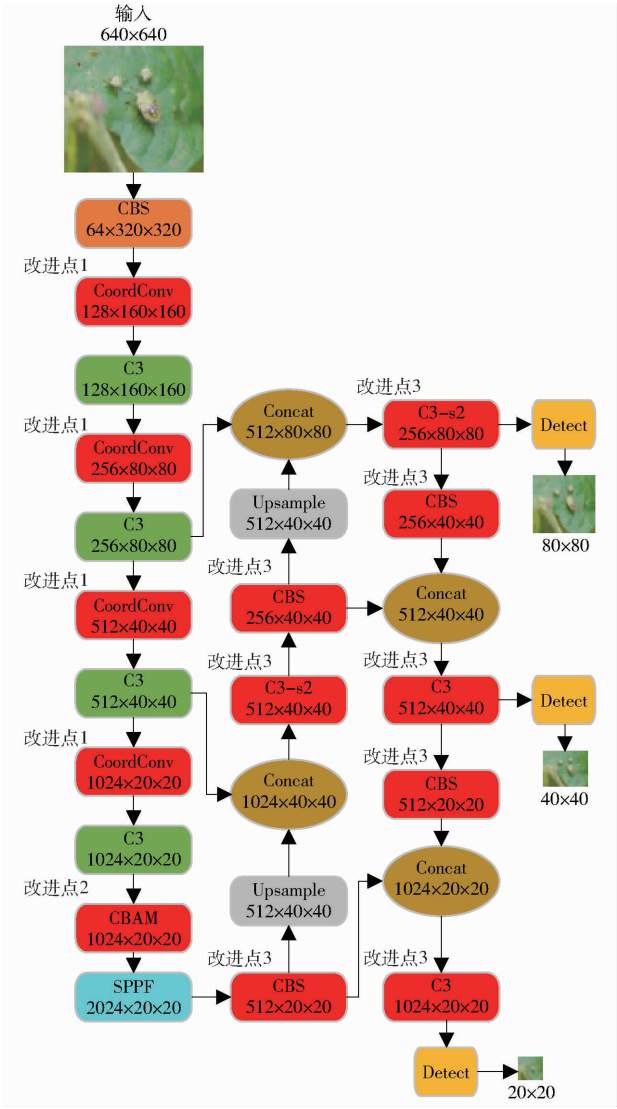

图 4 改进模型结构图

Fig. 4 Improved model structure diagram

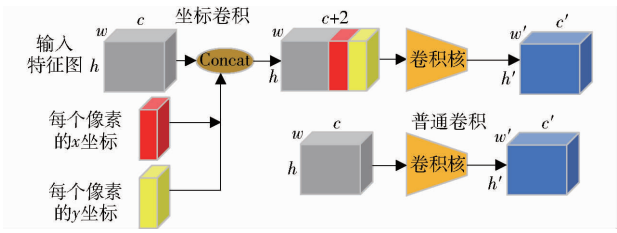

图 5 CoordConv 结构图

Fig. 5 CoordConv structure diagram

通道注意力机制能够提升模型对目标轮廓特征的关注度。方法是将输入的特征图同时进行全局最大池化和全局平均池化，得到两个相同维度的 1×1 特征图，之后将这两个特征图送入一个两层全连接结构，再通过激活函数 sigmoid 输出一个维度与输入特征图通道数相同的一维向量，并与输入特征图进行相乘。其计算公式为

$$MC(F)=\sigma(W_1(W_0(F^c_{avg}))+W_1(W_0(F^c_{max}))) \tag{1}$$

式中 $MC(F)$——通道注意力输出权值

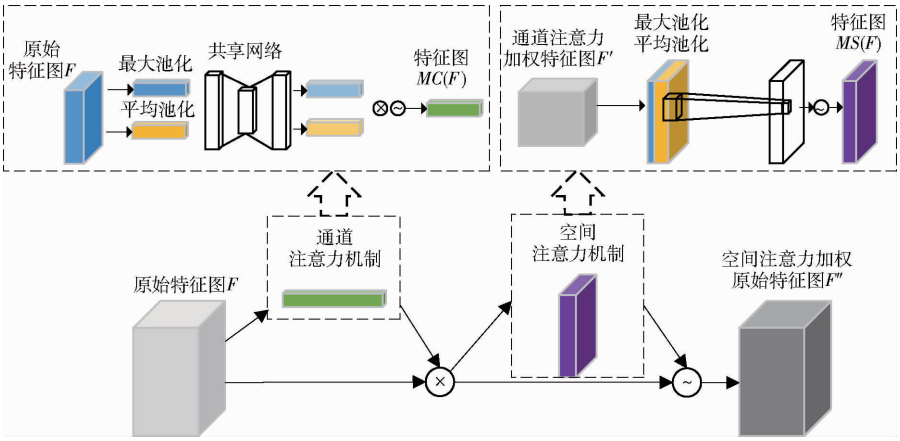

图 6　CBAM 注意力机制结构图

Fig. 6　CBAM attention mechanism structure diagram

σ——激活函数 sigmoid

F^c_{avg}——在空间上平均池化后特征映射

F^c_{max}——在空间上最大池化后特征映射

W_0——第 1 个全连接层权重矩阵

W_1——第 2 个全连接层权重矩阵

空间注意力机制能够提升模型对目标位置特征的关注度。方法是将输入的特征图同时进行最大池化和平均池化，得到两个通道为 1、尺度与输入特征图大小相同的特征图，之后将这两个特征图进行 Concat 操作送入普通卷积中，再通过 sigmoid 函数得到一个与输入特征图大小相同且通道为 1 的权重特征图，并与输入特征图相乘。其计算公式为

$$MS(F)=\sigma(f^{7\times7}[F^s_{avg};F^s_{max}]) \tag{2}$$

式中　$MS(F)$——空间注意力输出权值

$f^{7\times7}$——尺寸为 7×7 的卷积核

F^s_{avg}——在通道上平均池化后特征映射

F^s_{max}——在通道上最大池化后特征映射

CBAM 注意力机制由上述通道注意力机制与空间注意力机制串接而成。识别害虫时，模型在这两种注意力机制的共同作用下能够对与背景颜色相似的害虫进行精准定位和识别，而在主干的末端引入 CBAM 注意力机制可以增强深层网络的感受野，使深层网络能够获取更多的上下文信息，从而也提高了被遮挡害虫识别精度。其计算公式为

$$F'=MC(F)F \tag{3}$$

$$F''=MS(F)F' \tag{4}$$

式中　F——输入特征图

F'——经过通道注意力加权得到的特征图

F''——经过空间注意力加权得到的特征图

2.2.3　Slim－neck 结构

Slim－neck 结构是把颈部中的所有普通卷积替换为 GSConv 后，再引入 VoV－GCSP 模块[19]，其结构如图 7 所示。

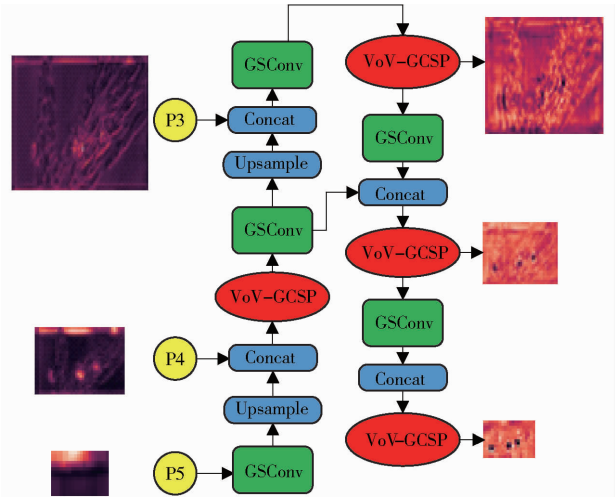

图 7　Slim－neck 结构图

Fig. 7　Slim－neck structure diagram

GSConv 是将输入特征图的通道平分成两部分，其中一部分通道特征图采用深度可分离卷积，另一部分通道特征图与深度可分离卷积的输出结果进行 Concat 操作，最后再对其采用 Shuffle 操作，将通道打乱，重新排序，使 GSConv 在降低计算成本的同时能够输出与普通卷积接近的特征图，从而保持精度不变。GSConv 结构如图 8 所示。

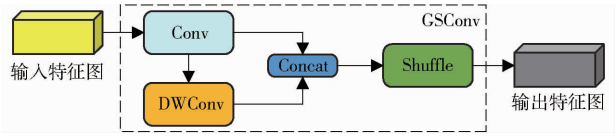

图 8　GSConv 结构图

Fig. 8　GSConv structure diagram

VoV－GCSP 模块是将输入特征图先经过 1×1 普通卷积进行降维，然后将结果同时输入进两条支路，其中一条支路为两层 GSConv，另一条支路是与两层 GSConv 的结果做 Concat 操作，最后经过 1×1 卷积进行输出。该模块能够充分提取浅层网络与深层网络的特征信息，并进一步融合，从而使输出的特

征图同时具备丰富的位置信息和语义信息。VoV - GCSP 模块结构如图 9 所示。总之,Slim - neck 结构在减少模型参数量与计算量的同时,能够进一步提升模型精度。

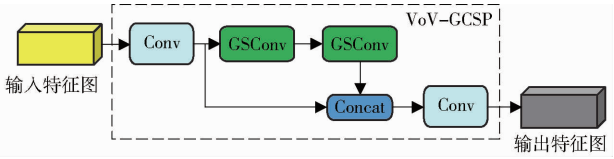

图 9 VoV - GCSP 结构图

Fig. 9 VoV - GCSP structure diagram

2.2.4 Soft - NMS

YOLO v5s 算法在对图像进行预测时,是对最终输出图的每个像素进行预测,一个像素会预测出多个边框,这会导致同一个目标出现多个边框,为了抑制冗余边框,采用了 NMS[20] 算法。NMS 算法通过对每个类别下所有预测框的分类得分进行排序,然后以最大得分对应的预测框为基准,计算剩下所有预测框与基准预测框的 IOU 值,之后将高于阈值的 IOU 所对应预测框得分变为 0,即删除该预测框,最后保留低于阈值的 IOU 所对应的预测框。该算法的筛选方式是直接删除 IOU 高于阈值的重叠预测框,而当检测两个相邻目标时,预测框很有可能发生重叠,若其 IOU 高于阈值时,两个目标的预测框会被删除一个,从而造成漏检。为了提升相邻目标的检测精度,对模型采用 Soft - NMS 算法。该算法会对高于阈值的 IOU 所对应的预测框得分进行惩罚,以避免将其置零,从而确保这些预测框能够保留在得分序列中,避免相邻目标出现漏检。

3 结果与分析

3.1 实验环境及评价指标

操作系统为 Windows 10 专业版,CPU 型号为 Intel(R) Core(TM) i5 - 13400F@ 2.50 Hz,GPU 型号为 RTX3060,运行内存为 32 GB,1TB 固态硬盘,编程语言为 Python 3.8.17,深度学习框架为 Pytorch 1.12.0,GPU 加速库为 CUDA 11.6、CUDNN 8.9.4。训练过程中的模型参数设置:分辨率为 640 像素 × 640 像素,批量大小为 16,迭代轮数为 300,优化器为 SGD,初始学习率为 0.01。

选用准确率、召回率、每个类别的 AP、mAP 以及浮点运算量作为评估模型性能的指标。

3.2 结果评估

3.2.1 不同模型检测结果对比

选用 YOLO v3[21]、YOLO R - CSP、YOLO v7、YOLO v8s、YOLO v5s 以及改进模型在同一个测试集下进行实验,并利用准确率、召回率、平均精度均值、参数量、浮点运算量对各个模型检测结果进行对比。不同模型实验结果如表 2 所示。

表 2 不同模型的实验结果

Tab. 2 Experimental results of different models

模型	准确率/%	召回率/%	平均精度均值/%	参数量	浮点运算量	平均检测时间/ms
YOLO v3	93.7	84.8	92.8	9 322 798	2.32×10^{10}	15.4
YOLO R - CSP	85.0	73.4	81.6	9 125 932	2.06×10^{10}	14.5
YOLO v7	86.7	73.9	80.7	9 156 668	2.61×10^{10}	13.9
YOLO v8s	94.8	85.1	92.4	3 007 403	8.10×10^{9}	15.6
YOLO v5s	94.6	86.0	90.5	7 034 398	1.58×10^{10}	15.6
改进模型	96.1	87.1	94.3	7 269 440	1.48×10^{10}	18.6

从表 2 可以看出,改进后模型准确率、召回率以及平均精度均值分别为 96.1%、87.1%、94.3%。相比于原模型准确率提升 1.5 个百分点,召回率提升 1.1 个百分点,平均精度均值提升 3.8 个百分点;而与 YOLO v3、YOLO R - CSP、YOLO v7、YOLO v8s 对比,各个指标均有提升,准确率上升 2.4、11.1、9.4、1.3 个百分点,召回率上升 2.3、13.7、13.2、2 个百分点,平均精度均值上升 1.5、12.7、13.6、1.9 个百分点。在模型的复杂度方面,改进模型的浮点运算量相比于原算法降低 6.3%,但是参数量和平均检测时间却上升 3.3%、3 ms。从模型的整体性能上看,改进后的模型在水稻害虫检测方面仍具有非常大的优势,能够对害虫进行更为精确的识别。

不同模型对每一种害虫的识别效果也有优劣之分,为了综合分析模型在水稻害虫数据集上所表现出来的性能,选用各类害虫的平均精度对不同模型进行分析,结果如表 3 所示。

从表 3 可以看出,改进后模型对不同种类害虫平均精度为 98.1%、85.9%、94.2%、99.5%、99.5%、94.8%、89.2%、91.8%、95.7%,其中水稻二化螟虫蛹和水稻二化螟成虫的平均精度最高,均为 99.5%,而黑色带斑稻绿蝽幼虫最低,仅有 89.2%。与其他模型相比,稻螟蛉幼虫最高提升了 14 个百分点,最低提升 2.7 个百分点;水稻二化螟幼虫最高提升 14.9 个百分点,最低提升 2.6 个百分点;水稻二化螟虫蛹最高提升 19.4 个百分点;水稻二化螟成虫最高提升 10.5 个百分点;绿色带斑稻绿蝽幼虫最高提升 9.2 个百分点,稻绿蝽成虫最高提升 13.6 个百分点,最低提升 5.8 个百分点;稻螟蛉成虫最高提升 42.1 个百分点,最低提升 0.2 个百分点;稻螟蛉虫卵平均精度相比于原模型和 YOLO v8s

表 3　不同模型对各个类别识别平均精度

Tab. 3　Average accuracy of different models for each category　%

模型	害虫种类								
	稻螟蛉虫卵	稻螟蛉幼虫	水稻二化螟幼虫	水稻二化螟虫蛹	水稻二化螟成虫	绿色带斑稻绿蝽幼虫	黑色带斑稻绿蝽幼虫	稻绿蝽成虫	稻螟蛉成虫
YOLO v3	97.2	76.2	86.9	99.3	99.5	93.1	84.3	86.0	94.5
YOLO R－CSP	93.0	76.1	79.3	80.1	89.0	85.6	80.2	79.5	71.5
YOLO v7	95.4	71.9	80.0	86.0	97.6	91.5	71.8	78.2	53.6
YOLO v8s	98.8	83.2	91.6	99.5	99.5	93.0	87.9	82.9	95.5
YOLO v5s	98.5	78.1	90.2	99.5	99.5	88.0	84.2	85.9	90.3
改进模型	98.1	85.9	94.2	99.5	99.5	94.8	89.2	91.8	95.7

的稻螟蛉虫卵平均精度略有下降，分别为 0.4、0.7 个百分点，但是与其他模型的稻螟蛉虫卵平均精度相比，仍有较大的提升。

从整体上看，改进模型对于各类害虫的平均精度均有提升，分析其原因，害虫识别受农田环境的影响，存在遮挡、数量多、与背景颜色相似等问题，从而造成识别精度下降。本文对 YOLO v5s 模型的改进点是针对这些问题提出的，其 CBAM 注意力机制提升了模型对害虫的关注度，使得被遮挡的害虫更容易识别，同时颈部所采用的 Slim－neck 结构也增强了模型对害虫特征的表达能力，使其能够更好地识别出数量多、与背景颜色相似的害虫。

各个模型对水稻害虫数据集检测效果如图 10 所示。显然，改进模型对各类害虫的识别效果提升明显，能够准确地对各类害虫进行定位与识别；YOLO v3 没有识别出图像 2、图像 9 的稻螟蛉成虫和稻绿蝽成虫；YOLO R－CSP 和 YOLO v7 对害虫的识别效果最差，仅仅识别出了图像 1、图像 4、图像 7 的害虫，而图像 8 中的黑色带斑稻绿蝽幼虫存在漏检情况；YOLO v8s 相对于 YOLO R－CSP 和 YOLO v7，检测效果提升了一些，但是仍然没有识别出图像 9 中的稻绿蝽成虫；原模型在这 9 幅图像的检测中，没有识别出稻绿蝽成虫，并且图像 1 和图像 8 中的稻螟蛉幼虫、黑色带斑稻绿蝽幼虫均存在漏检情况。

3.2.2　消融实验

使用 5 组不同的模型进行消融实验，分别为 YOLO v5s、YOLO v5s＋CBAM、YOLO v5s＋CBAM＋CoordConv、YOLO v5s＋CBAM＋CoordConv＋Slim－neck、YOLO v5s＋CBAM＋CoordConv＋Slim－neck＋Soft－NMS，并利用平均精度均值与各类害虫的平均精度对这 5 组模型的实验结果进行分析，实验结果如表 4 所示。

通过表 4 可以看出，YOLO v5s＋CBAM、YOLO v5s＋CBAM＋CoordConv、YOLO v5s＋CBAM＋CoordConv＋Slim－neck、YOLO v5s＋CBAM＋CoordConv＋Slim－neck＋Soft－NMS 模型相比于原模型，平均精度均值提升 0.6、1.1、3.3、3.8 个百分点。其中，YOLO v5s＋CBAM 对绿色带斑稻绿蝽幼虫、黑色带斑稻绿蝽幼虫、稻绿蝽成虫、稻螟蛉成虫的平均精度与原模型对比，分别提升 3.2、2.8、2、1.1 个百分点，说明 CBAM 注意力机制在密切关注目标的同时，使模型获得了更多的全局上下文信息，因此提升了与背景颜色相似目标和被遮挡目标的识别精度；YOLO v5s＋CBAM＋CoordConv 对稻螟蛉幼虫、绿色带斑稻绿蝽幼虫、黑色带斑稻绿蝽幼虫、稻绿蝽成虫、稻螟蛉成虫的平均精度相较于原模型分别提升 2、3.6、0.6、2.2、3 个百分点，说明 CoordConv 能够获取更多目标的位置信息，减少漏检情况的发生；YOLO v5s＋CBAM＋CoordConv＋Slim－neck 对稻螟蛉虫卵、稻螟蛉幼虫、水稻二化螟幼虫、绿色带斑稻绿蝽幼虫、黑色带斑稻绿蝽幼虫、稻绿蝽成虫、稻螟蛉成虫的平均精度相较于原模型分别提升 0.1、8、1.5、7.4、3.1、3.7、6.2 个百分点，说明 Slim－neck 结构能够有效融合浅层网络与深层网络的特征信息，并进一步提高网络对特征的表达能力；YOLO v5s＋CBAM＋CoordConv＋Slim－neck＋Soft－NMS 对稻螟蛉幼虫、水稻二化螟幼虫、绿色带斑稻绿蝽幼虫、黑色带斑稻绿蝽幼虫、稻绿蝽成虫、稻螟蛉成虫的平均精度相较于原模型分别提升 7.8、4、6.8、5、5.9、5.4 个百分点，说明 Soft－NMS 能够有效减少相邻目标边框被删除的情况，从而使识别精度提升。因此，改进模型能够更精确地识别出被遮挡目标、与背景颜色相似目标、多目标以及相邻目标。

3.2.3　热力图解释

Gard－CAM[22] 通过对输出特征图的对应权重进行加权求和，得到了能够表达网络对水稻害虫特征学习程度的热力图。在热力图中，红色区域表示模型对特征学习的最大程度，黄色区域次之，蓝色区域最小，其中蓝色区域认为是冗余信息。为了进一步验证改进模型在水稻害虫识别中的有效性，从测试集中随机选择 3 幅图像，对比改进模型和原模型生成的热力图，如图 11 所示。

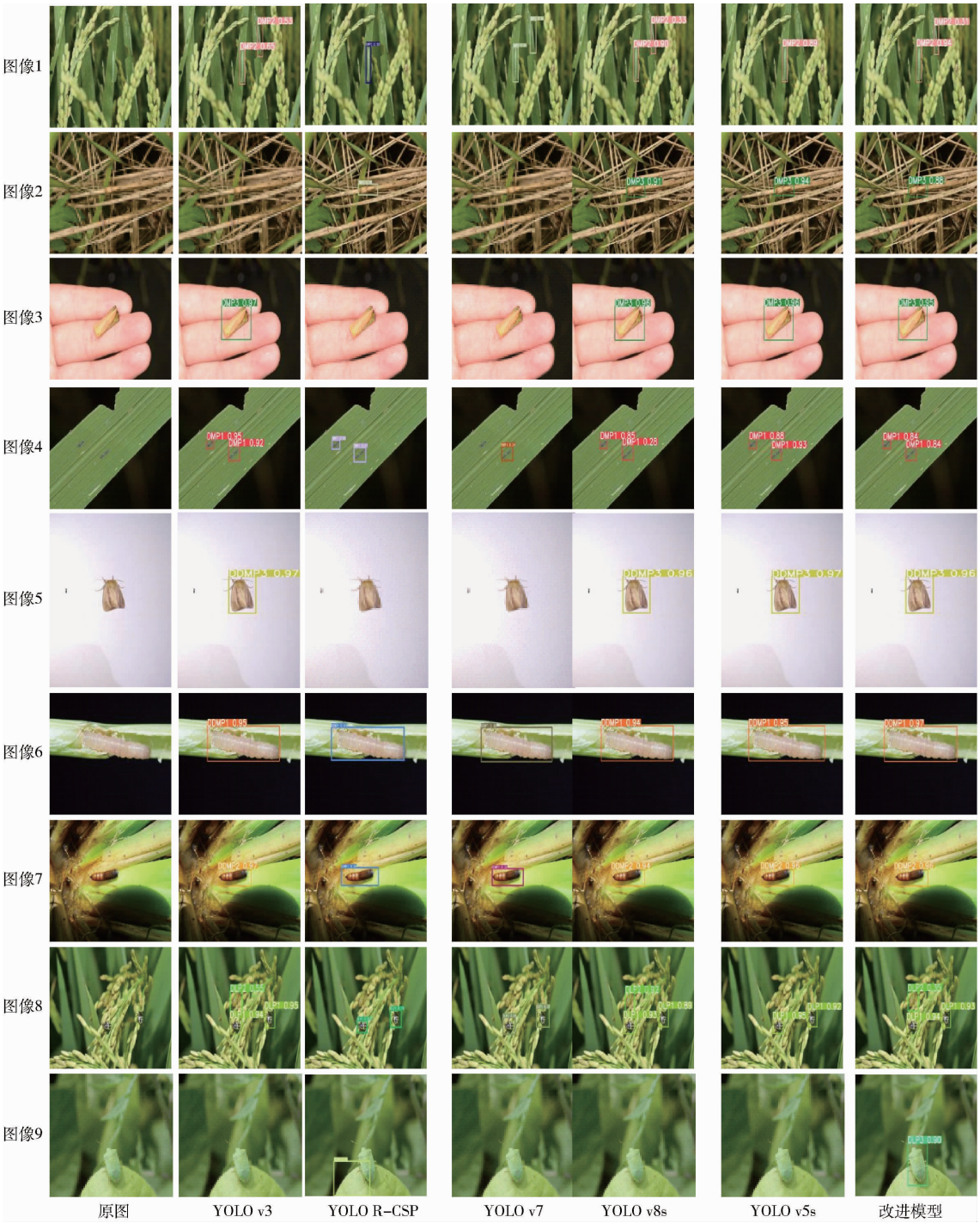

图 10 不同模型检测效果

Fig. 10 Detection performance of different models

表 4 消融实验结果

Tab. 4 Results of ablation experiments %

模型	平均精度均值	平均精度								
		稻螟蛉虫卵	稻螟蛉幼虫	水稻二化螟幼虫	水稻二化螟虫蛹	水稻二化螟成虫	绿色带斑稻绿蝽幼虫	黑色带斑稻绿蝽幼虫	稻绿蝽成虫	稻螟蛉成虫
YOLO v5s	90.5	98.5	78.1	90.2	99.5	99.5	88.0	84.2	85.9	90.3
YOLO v5s + CBAM	91.1	96.9	76.1	90.2	99.5	99.5	91.2	87.0	87.9	91.4
YOLO v5s + CBAM + CoordConv	91.6	97.4	80.1	89.9	99.4	99.4	91.6	84.8	88.1	93.3
YOLO v5s + CBAM + CoordConv + Slim − neck	93.8	98.6	86.1	91.7	99.5	99.5	95.4	87.3	89.6	96.5
YOLO v5s + CBAM + CoordConv + Slim − neck + Soft − NMS	94.3	98.1	85.9	94.2	99.5	99.5	94.8	89.2	91.8	95.7

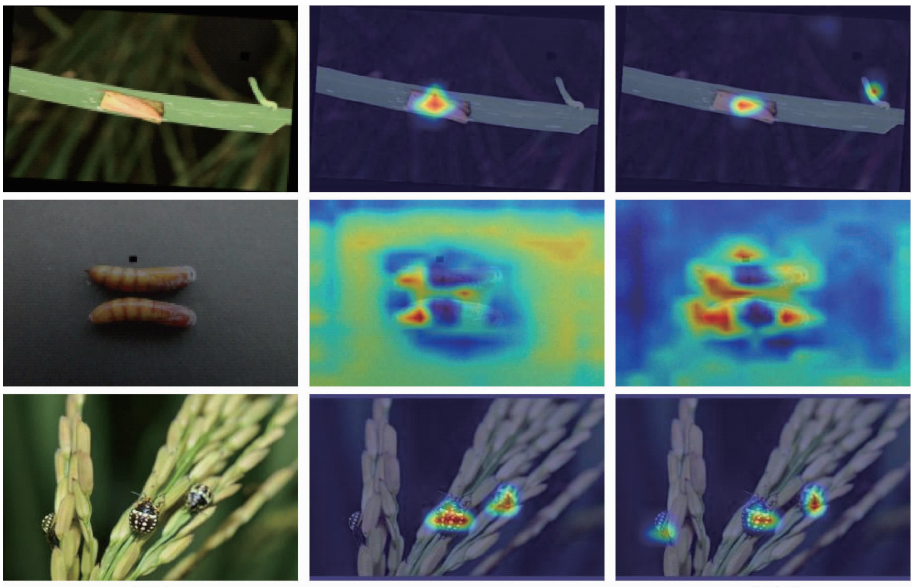

图 11　热力图效果

Fig. 11　Heat map effect

由图 11 可知，改进模型对水稻害虫的特征学习比较充足，能够较好地识别出目标，而原模型所学习到的害虫特征不够充分，容易漏检。通过热力图对比也印证了改进模型在水稻害虫识别上表现更优。

4　结论

（1）提出了一种基于改进 YOLO v5s 的害虫识别方法。将 YOLO v5s 主干网络中除第 1 层之外的普通卷积替换为 CoordConv，使主干网络在提取目标特征时，能够获得更多目标位置信息；在主干网络末端添加 CBAM 注意力机制，既扩大了模型感受野，又提升了目标关注度；对颈部网络采用 Slim－neck 结构，减少模型计算量，增强模型对特征的处理能力；引入 Soft－NMS 算法优化了相邻目标边框筛选，进一步提高了模型对害虫的识别精度。实验结果表明，改进模型平均精度均值为 94.3%，相比于原算法提升 3.8 个百分点，且与主流模型 YOLO v3、YOLO R－CSP、YOLO v7、YOLO v8s 相比，平均精度均值分别提升 1.5、12.7、13.6、1.9 个百分点。

（2）YOLO v8s 的平均检测时间与复杂度是其 2 倍的 YOLO v5s 持平，显然 YOLO v5s 在进一步缩短检测时间上更具发展空间。改进后 YOLO v5s 识别精度取得了较好的效果，但在模型复杂度方面仍存不足，改进后模型参数量比原模型增加 3.3%，模型平均检测时间比原模型多 3 ms。

参　考　文　献

[1]　曾文伟，罗华，石小江，等. 我国常规稻发展现状与优势分析[J]. 作物研究，2023，37(2)：167－169.
ZENG Wenwei, LUO Hua, SHI Xiaojiang, et al. Analysis of the development status and advantages of conventional rice in China[J]. Crop Research, 2023, 37(2): 167－169. (in Chinese)

[2]　吴大秀. 我国水稻病虫害防治现状及发展策略[J]. 种子科技，2023，41(20)：120－122.
WU Daxiu. Current situation and development strategies of rice disease and pest control in China[J]. Seed Technology, 2023, 41(20): 120－122. (in Chinese)

[3]　崔金荣，魏文钊，赵敏. 基于改进 MobileNetV3 的水稻病害识别模型[J]. 农业机械学报，2023，54(11)：217－224，276.
CUI Jinrong, WEI Wenzhao, ZHAO Min. Rice disease identification model based on improved MobileNetV3[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 217－224, 276. (in Chinese)

[4]　赵辉，黄镖，王红君，等. 基于改进 YOLO v7 的农田复杂环境下害虫识别算法研究[J]. 农业机械学报，2023，54(10)：246－254.
ZHAO Hui, HUANG Biao, WANG Hongjun, et al. Research on pest identification algorithm in complex farmland environment based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 246－254. (in Chinese)

[5]　鲍文霞，吴德钊，胡根生，等. 基于轻量型残差网络的自然场景水稻害虫识别[J]. 农业工程学报，2021，37(16)：145－152.
BAO Wenxia, WU Dezhao, HU Gensheng, et al. Identification of rice pests in natural scenes based on lightweight residual networks[J]. Transactions of the CSAE, 2021, 37(16): 145－152. (in Chinese)

[6] 张楠楠,张晓,白铁成,等. 基于 CBAM - YOLO v7 的自然环境下棉叶病虫害识别方法[J]. 农业机械学报,2023,54(增刊 1):239 - 244.
ZHANG Nannan, ZHANG Xiao, BAI Tiecheng, et al. Identification method of cotton leaf diseases and pests in natural environment based on CBAM - YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(Supp. 1):239 - 244. (in Chinese)

[7] 林文树,张金生,何乃磊. 基于改进 YOLO v4 的落叶松毛虫侵害树木实时检测研究[J]. 农业机械学报,2023,54(4):304 - 312,393.
LIN Wenshu, ZHANG Jinsheng, HE Nailei. Real time detection of larch caterpillar infestation on trees based on improved YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(4):304 - 312,393. (in Chinese)

[8] 刘双喜,刘思涛,屈慧星,等. 基于 MS - YOLO v7 多尺度稻飞虱识别分类[J]. 农业机械学报,2023,54(增刊 1):212 - 221.
LIU Shuangxi, LIU Sitao, QU Huixing, et al. Multi scale rice planthopper identification and classification based on MS - YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(Supp. 1):212 - 221. (in Chinese)

[9] 肖德琴,黄一桂,张远琴,等. 基于改进 Faster R - CNN 的田间黄板害虫检测算法[J]. 农业机械学报,2021,52(6):242 - 251.
XIAO Deqin,HUANG Yigui,ZHANG Yuanqin,et al. Pest detection algorithm of yellow plate in field based on improved faster R - CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(6): 242 - 251. (in Chinese)

[10] YU Z, HUANG H,CHEN W, et al. YOLO - FaceV2: a scale and occlusion aware face detector[J]. arXiv Preprint, arXiv: 2208. 02019, 2022.

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

[12] BODLA N, SINGH B, CHELLAPPA R, et al. Soft - NMS: improving object detection with one line of code[C]//2017 IEEE International Conference on Computer Vision(ICCV), 2017: 5562 - 5570.

[13] EVERYINGHAM M, GOOL L V, CHRISTOPHE K L, et al. The pascal visual object classes (VOC) challenge[J]. International Journal of Computer Vision, 2010, 88:303 - 308.

[14] WANG C,LIAO H M, WU Y, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 390 - 391.

[15] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 2117 - 2125.

[16] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759 - 8768.

[17] TAN M, PANG R, LE Q V. Efficientdet: scalable and efficient object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781 - 10790.

[18] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904 - 1916.

[19] LI H, LI J, WEI H,et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles [J]. arXiv Preprint, arXiv: 2206.02424,2022.

[20] BODLA N, SINGH B, CHELLAPPA R, et al. Soft - NMS—improving object detection with one line of code[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 5561 - 5569.

[21] REDMON J, FARHADI A. Yolov3: an incremental improvement[J]. arXiv Preprint, arXiv: 1804.02767,2018.

[22] SELVARAJU R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 618 - 626.