基于YOLOv5n-ChtaFSD模型的飞机草精准检测

袁俊杰[1] 龙 阳[1] 李 岩[2] 陈 文[1] 颜伟胜[3] 黎科江[2]
赵文锋[3] 桑 文[2*]

（1. 中华人民共和国湛江海关，广东 湛江 524022；2. 华南农业大学植物保护学院，广州 510642；
3. 华南农业大学人工智能与低空技术学院，广州 510642）

摘要：为解决飞机草*Chromolaena odorata*识别效率低、成本高的问题，基于YOLOv5n模型构建了一种改进的目标检测模型YOLOv5n-ChtaFSD，该模型在颈部网络引入重参数化泛化特征金字塔网络结构，该结构通过消除冗余的上采样操作来提升精度，降低推理耗时；该模型在颈部网络嵌入简单注意力机制，该机制根据特征重要性实现自适应筛选，从而增强模型对小目标的感知能力，缓解因特征稀疏导致的漏检问题；该模型引入动态检测头，该检测头通过动态注意力机制与统一尺度建模提升多尺度特征融合效果与小目标定位精度。结果显示：YOLOv5n-ChtaFSD模型的精确率、召回率与平均精度均值分别为90.7%、81.6%和85.9%，较原模型分别提升4.5百分点、4.7百分点和4.0百分点，明显优于YOLOv5n、YOLOv5s、YOLOv8n、YOLOv8s、YOLOv11n和YOLOv11s模型。YOLOv5n-ChtaFSD模型的参数量为2.4 M，浮点运算量为5.4 GFLOPs，模型体积为4.9 MB，表明该模型具有更优的检测性能，同时其参数量较小，有利于在移动机器人等边缘设备中部署应用。

关键词：飞机草；目标检测；YOLOv5模型；特征融合；改进

Precision detection of Siam weed *Chromolaena odorata* based on a YOLOv5n-ChtaFSD model

Yuan Junjie[1] Long Yang[1] Li Yan[2] Chen Wen[1] Yan Weisheng[3] Li Kejiang[2]
Zhao Wenfeng[3] Sang Wen[2*]

（1. Zhanjiang Customs of the People's Republic of China, Zhanjiang 524022, Guangdong Province, China;
2. College of Plant Protection, South China Agricultural University, Guangzhou 510642, Guangdong Province, China;
3. College of Artificial Intelligence and Low-Altitude Technology, South China Agricultural University,
Guangzhou 510642, Guangdong Province, China）

Abstract: To address the problems of low efficiency and high cost in the current identification of Siam weed *Chromolaena odorata*, an improved object detection model YOLOv5n-ChtaFSD is constructed based on YOLOv5n. In the neck of the model, an efficient re-parameterizable generalized feature pyramid network (Efficient RepGFPN) is introduced, which not only improves detection accuracy but also reduces inference time by eliminating redundant upsampling operations. Meanwhile, the model embeds a simple parameter-free attention module (SimAM), which adaptively selects features based on feature importance, thereby enhancing the model's small-target perception capability and mitigating missed detection caused by sparse features. Furthermore, the model adopts a dynamic head (DyHead) that improves the effect of multi-scale feature fusion and small-target localization accuracy through dynamic

基金项目：海关总署科研项目(2025HK110)，湛江海关科研项目(2024ZJHK009)
* 通信作者（Author for correspondence），E-mail：sangwen@scau.edu.cn
收稿日期：2025-12-06

attention mechanism and unified scale modeling. The results show that the precision, recall rate and mean average precision of the YOLOv5n-ChtaFSD model reach 90.7%, 81.6% and 85.9%, which are 4.5%, 4.7% and 4.0% higher than those of the original model, respectively, and the proposed model is significantly superior to YOLOv5n, YOLOv5s, YOLOv8n, YOLOv8s, YOLOv11n and YOLOv11s models. With 2.4 M parameters, 5.4 GFLOPs of computational complexity and a model size of 4.9 MB, YOLOv5n-ChtaFSD achieves superior detection performance in complex environments with fewer parameters, making it suitable for deployment on edge devices such as mobile robots.

Key words: *Chromolaena odorata*; object detection; YOLOv5 model; feature fusion; improvement

飞机草 *Chromolaena odorata* 是一种全球性入侵杂草，原产于中美洲，自20世纪20年代传入东南亚后迅速扩散，现已广泛分布于亚洲、非洲、大洋洲和西太平洋群岛的热带及亚热带地区，并于2001年被国际自然保护联盟列为全球最具威胁的100种外来入侵物种之一（陈臻和白天才，2018）。该植物繁殖力强，环境适应范围广，已对中国广西、广东、云南、贵州及港澳台等地的生态系统构成严重威胁（吴迎福等，2017），因此防控该恶性杂草迫在眉睫。

目前，飞机草的识别主要依赖于传统形态分类及分子生物学方法（Buddhachat et al.，2020）。尽管这些方法有一定的科学性，但普遍存在操作复杂、成本高、依赖专业人员等局限，难以满足快速、实地的监测需求。随着计算机视觉技术的快速发展，基于图像的目标检测方法在低成本、高效率方面展现出较大的优势，为入侵植物的智能识别提供了新的技术路径。在农业领域，基于卷积神经网络的目标检测模型已取得显著进展。目标检测算法主要包括两阶段与单阶段算法（余咏等，2024）。两阶段模型以区域卷积神经网络（region-based convolutional neural network，R-CNN）为主，后续逐渐出现快速R-CNN（Faster R-CNN）（Chen et al.，2022）、掩膜R-CNN（Mask R-CNN）（Ahn et al.，2023）和级联R-CNN（Cascade R-CNN）（Lee & Gao，2023）等检测精度较高的模型，但这些模型的计算复杂，推理速度慢；而YOLO（You Only Look Once）（Hsu & Lin，2021）、单次多框检测器（single shot multibox detector，SSD）（Qian et al.，2023）、视网膜网络（retina network，RetinaNet）（Zhu et al.，2021）等单阶段模型在保持较高精度的同时，具有更快的处理速度，更适合实时检测任务。近年来，国内外学者围绕农业场景下的模型优化开展了多项研究。例如，在茄子果实病虫害检测（郭臻和李玥等，2024）、农田害虫识别（徐悦等，2025）与水田杂草防控（Ju et al.，2025）等方面取得了较大的进展，深度学习在植物保护方面展现出了广泛的应用前景。但这些研究多集中于常见作物与杂草（Hu et al.，2024），未见针对飞机草的高效检测方法。

为实现复杂环境下飞机草的快速、高效检测，本研究提出一种基于YOLOv5n模型改进的检测模型——YOLOv5n-ChtaFSD，该模型在颈部网络中引入重参数化泛化特征金字塔网络（efficient re-parameterizable generalized feature pyramid network，Efficient RepGFPN）来提升特征融合效率，在颈部网络嵌入无参注意力机制（simple, parameter-free attention module，SimAM）来增强对小目标特征的感知能力，采用动态检测头优化多尺度定位性能，以期使模型在保持较低参数量的前提下显著提升检测精度，从而为飞机草的实时监测与防控提供技术支撑。

1 材料与方法

1.1 材料

数据集构建：飞机草数据集采集于广东省湛江市森林公园周边区域路边野生植株，在自然光照条件下，使用iPhone 13 Pro Max手机进行拍摄。原始图像以HEIC格式保存，分辨率为4 032×3 024像素，共采集906张原始图像。为增强数据多样性并提升模型在复杂环境下的泛化能力，拍摄过程中采用了顺光、逆光等不同光照条件，并涵盖了多种拍摄角度与距离（图1）。

试验平台和参数配置：本研究在Windows11操作系统环境下完成，模型使用Pytorch1.3，GPU为NVIDIA GeForce RTX 4090，Python版本为3.8，CPU为AMD EPYC 7C13 64核心处理器。模型训练时，Epochs设置为150，Batch设置为32，Imgsz设置为640，Workers设置为32，Optimizer设置为SGD。

1.2 方法

1.2.1 数据集处理

使用LabelImg工具对图像中的飞机草目标进行标注，生成对应的txt格式标注文件。标注完成

后，将数据集按照8∶1∶1的比例随机划分为训练集、验证集和测试集，分别用于模型训练、参数调优与性能评估。为缓解因训练样本不足可能引起的过拟合问题，本研究对训练集图像进行了旋转、裁剪、缩放、亮度调整及去噪等数据增强处理，以提升训练数据的多样性并增强模型的鲁棒性（李善军等，2024；赵文锋等，2024）。

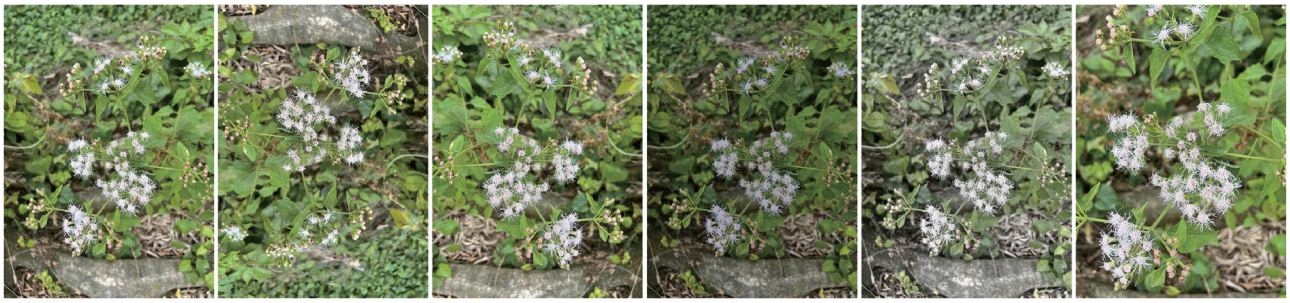


图1　飞机草数据集样本

Fig. 1　Samples of the *Chromolaena odorata* dataset

1.2.2　YOLOv5n基线模型

YOLOv5是YOLO系列的第5代目标检测模型，由Ultralytics团队于2020年发布。它具有YOLO系列的实时检测优势，即通过优化网络结构、引入自适应训练机制和高效数据增强显著提升了模型的速度和精度。根据网络宽度与深度不同，YOLOv5系列可划分为YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l和YOLOv5x五个版本，模型复杂度依次递增。YOLOv5n是由Ultralytics团队开发的高效单阶段目标检测模型，主要包括输入层、主干网络、颈部网络和检测头（Redmon et al.，2015），该模型融合了跨阶段局部网络、特征金字塔网络、Mish激活函数和焦点损失函数等多种策略，在保持结构相对简洁的同时，有效提升了模型的检测性能和鲁棒性。YOLOv5n模型具有极低的参数量与良好的检测鲁棒性，能在资源受限的边缘设备上实现对目标的实时精准识别，满足田间场景对延迟与功耗的严格要求，因此本研究选择其作为基线模型。

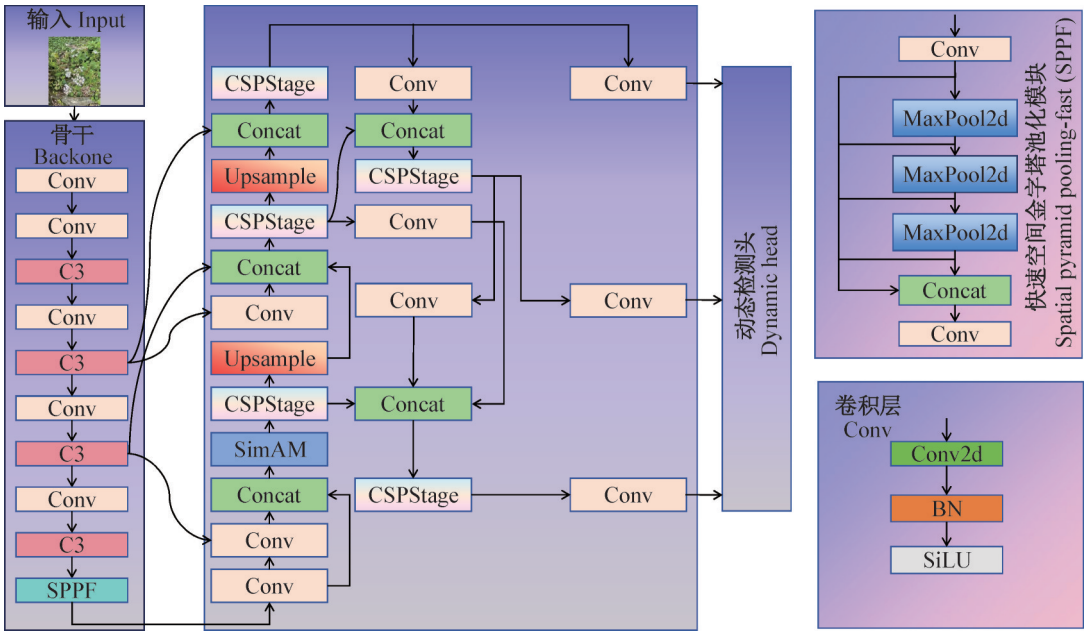
1.2.3　YOLOv5n模型的改进

针对飞机草检测中YOLOv5n模型的多尺度特征融合能力不足、准确率偏低及细节信息丢失等问题，本研究从以下3个方面进行改进：首先，在颈部网络中引入Efficient RepGFPN（Xu et al.，2022）；其次，嵌入轻量级、无参数的SimAM；最后，引入动态检测头框架。改进后的模型为YOLOv5n-ChtaFSD（图2）。

1.2.4　在颈部网络中引入Efficient-RepGFPN

原YOLOv5n模型采用特征金字塔网络和路径聚合网络进行特征融合，虽能在一定程度上兼顾语义强度与定位精度，但其信息流动仍存在单向性限制。为应对飞机草所处背景复杂、形态多样所导致的特征提取困难，本研究在颈部网络中引入Efficient-RepGFPN。该结构通过优化跨尺度连接与特征复用机制增强多尺度特征的融合效率，从而提升模型在复杂环境下对飞机草关键特征的提取能力。

与YOLOv5模型所采用的路径聚合特征金字塔网络相比，Efficient RepGFPN主要在以下4个方面进行了优化：第一，通过引入广义特征金字塔网络和重参数化结构，实现对不同尺度特征图的动态通道调整。高层特征图采用较少通道以降低计算开销，而低层特征图保留更多通道以增强细节捕捉能力，在有限算力条件下优化了计算资源的分配。而YOLOv5模型中路径聚合特征金字塔网络采用固定通道数配置，这可能限制了高层语义信息或低层空间特征的表达能力。第二，Efficient RepGFPN移除了广义特征金字塔网络中冗余的上采样操作Quenn-Fusion部分，在保持高精度的同时有效减少了推理时间。第三，训练阶段采用多分支并行的跨阶段局部网络结构，该结构融合了3×3卷积、1×1卷积及残差连接等操作，通过丰富的特征变换增强模型的非线性表达能力；在推理阶段，多个分支被重参数化为单一3×3卷积核，既保留了跨阶段局部结构在特征交互方面的优势，又提升了前向推理效率。第四，引入两阶段蒸馏机制（分别基于强数据增强域和无马赛克域），并结合高层次聚合网络的层聚合设计，将大模型的知识有效迁移至轻量化模



C3：跨阶段局部三卷积模块；CSPStage：跨阶段局部网络；Concat：特征拼接；Upsample：上采样；SimAM：简单无参数注意力；MaxPool2d：二维最大池化层；BN：批归一化层；SiLU：SiLU激活函数；Conv2d：二维卷积。C3: Cross stage partial 3-conv module; CSPStage: cross stage partial stage; Concat: feature concatenation; Upsample: upsample; SimAM: simple attention module; MaxPool2d: 2D max pooling layer; BN: batch normalization layer; SiLU: SiLU activation function; Conv2d: 2D convolution.

图2 YOLOv5n-ChtaFSD模型的框架图

Fig. 2 Framework of YOLOv5n-ChtaFSD model

型中；该策略不仅增强了小模型对复杂场景的适应性，还通过多尺度特征交互进一步提升了模型的鲁棒性。

1.2.5 在颈部网络中嵌入SimAM

SimAM是一种无需引入额外参数的轻量级注意力机制（Yang et al.，2021；Ding et al.，2022）。在飞机草图像中，相邻像素间通常表现出较高的特征相似性，而距离较远的像素间相似性则显著降低。SimAM基于这一视觉特性，通过计算特征图中每个像素与其邻域之间的神经能量差异，自适应地生成具有明确神经科学解释的3D注意力权重。与传统的通道注意力（1D权重）或空间注意力（2D权重）机制不同，SimAM能同时沿通道和空间维度对特征进行协同校准，形成完整的3D注意力权重。这种3D权重结构能更全面地捕捉特征间的依赖关系，其表达能力优于传统的1D或2D注意力机制。

SimAM模块定义了能量函数 $e_t(\boldsymbol{\omega}_t,\boldsymbol{b}_t,\boldsymbol{y},\boldsymbol{x}_i)=(\boldsymbol{y}_t-t)^2+\frac{1}{M-1}\sum_{i=1}^{M-1}(y_o-\boldsymbol{x}_i)^2$，式中，$t$代表当前神经元的特征数值；$\boldsymbol{x}_i$代表同一特征通道内其余神经元的特征数值；$\boldsymbol{\omega}_t$和$\boldsymbol{b}_t$分别为线性变换操作的权重和偏置；$y_o$为负样本目标值；$\boldsymbol{y}_t$为对第$t$个目标的预测得分；$\boldsymbol{y}$为模型输出，$M$为样本总数。为确定各神经元的重要性程度，对能量函数进行优化，得到最终能量函数，即 $e_t(\boldsymbol{\omega}_t,\boldsymbol{b}_t,\boldsymbol{y},\boldsymbol{x}_i)=\frac{1}{M-1}\sum_{i=1}^{M-1}(-1-(\boldsymbol{\omega}_t\boldsymbol{x}_i+\boldsymbol{b}_t))^2+(1-(\boldsymbol{\omega}_t t+\boldsymbol{b}_t))^2+\lambda\boldsymbol{\omega}_t^2$。由最终能量函数直接求解得到最优的$\boldsymbol{\omega}_t$和$\boldsymbol{b}_t$，即 $\boldsymbol{\omega}_t=-\frac{2(t-\mu_t)}{(t-\mu_t)^2+2\sigma_t^2+2\lambda}$，$\boldsymbol{b}_t=-\frac{1}{2}(t+\mu_t)\boldsymbol{\omega}_t$，式中，$\sigma_t$是特征值的标准差；$\mu_t$为特征图在某一局部区域或整个空间维度上的均值；$\lambda$为设定的超参数。最终的注意力权重则由能量函数的最小值来确定。整个SimAM的计算流程过程可表示为 $\hat{\boldsymbol{X}}=\mathrm{sigmoid}\left(\frac{1}{\boldsymbol{E}}\right)\odot\boldsymbol{X}$，式中$\boldsymbol{E}$为所有神经元的能量汇总矩阵，$\hat{\boldsymbol{X}}$和$\boldsymbol{X}$分别代表输出与输入特征图。通过SimAM模块可得到最终的输出特征图。

1.2.6 引入动态检测头

在目标检测中，YOLOv5模型检测头采用静态结构和多尺度预测的设计，主要负责预测目标的边界框、分类得分及置信度，但该结构在应对小目标、遮挡及多尺度目标共存的复杂场景时表现有限。为此，本研究引入了动态检测头（Dai et al.，2021），通

过动态注意力机制与统一检测头框架实现自适应特征融合。动态检测头的核心创新在于融合了动态通路注意力机制，将尺度、空间与任务感知统一于一个协同注意力框架中，这种设计显著提升了模型在复杂环境下的泛化能力与检测精度，更适于在机器人等嵌入式设备中部署。

动态检测头对输入特征张量$\boldsymbol{F}\in\boldsymbol{R}^{L\times S\times C}$进行自注意力变换，即$\boldsymbol{W}(\boldsymbol{F})=\pi(\boldsymbol{F})\cdot\boldsymbol{F}$，式中$L$、$S$、$C$分别表示序列长度维度、空间位置和通道维数，$\boldsymbol{F}$为输入多尺度特征图；$\boldsymbol{\pi}(\boldsymbol{F})$是基于输入特征动态生成的注意力权重函数，可根据尺度感知注意力模板、空间感知注意力模板和任务感知注意力模块自适应调整特征响应，从而增强模型对多尺度目标特别是小目标的判别与定位能力。尺度感知注意力模块主要根据不同尺度的特征在语义上的重要性进行动态融合，可表达为$\boldsymbol{\pi}_L(\boldsymbol{F})\cdot\boldsymbol{F}=\sigma\left(f\left(\frac{1}{SC}\sum_{S,C}\boldsymbol{F}\right)\right)\cdot\boldsymbol{F}$，式中，$f(x)$为由1×1卷积层近似表示的线性函数；$\sigma(x)=\max\left(0,\ \min\left(1,\frac{x+1}{2}\right)\right)$为硬Sigmoid函数。空间感知注意力模块通过可变形卷积使注意力学习稀疏化，再在相同空间位置上跨层级聚合特征，对特征图中的每个位置进行动态加权，即$\boldsymbol{\pi}_S(\boldsymbol{F})\cdot\boldsymbol{F}=\frac{1}{\mathrm{L}}\sum_{l=1}^{L}\sum_{k=1}^{K}w_{l,k}\cdot\boldsymbol{F}(l;p_k+\Delta p_k;c)\cdot\Delta m_k$，式中，$\boldsymbol{\pi}_{\mathrm{S}}(\boldsymbol{F})$为空间注意力权重图，是输入特征$\boldsymbol{F}$的函数，维度与$\boldsymbol{F}$完全一致（$C\times H_l\times W_l$，其中$H_l$、$W_l$为第$l$层特征的高和宽）；$l$是序列长度维度的索引变量；$k$是稀疏采样位置数目的索引变量；$K$是稀疏采样位置的数目；$p_k+\Delta p_k$通过自学习空间偏移$\Delta p_k$移位的位置以聚焦于区分区域；而$\Delta m_k$是位置$p_k$处的自学习重要性标量。为了实现联合学习和对不同的对象表征进行泛化，任务感知注意力模块动态地打开和关闭功能通道，来支持不同的任务，即$\pi_C(\boldsymbol{F})\cdot\boldsymbol{F}=\max(\alpha^1(\boldsymbol{F})\cdot\boldsymbol{F}_c+\beta^1(\boldsymbol{F}),\ \alpha^2(\boldsymbol{F})\cdot\boldsymbol{F}_c+\beta^2(\boldsymbol{F}))$，$\boldsymbol{F}_c$是第$c$个通道的特征切片；$\alpha^1(\boldsymbol{F})\cdot\boldsymbol{F}_c+\beta^1(\boldsymbol{F})$是对第$c$个通道进行第1次线性变换；$\alpha^2(\boldsymbol{F})\cdot\boldsymbol{F}_c+\beta^2(\boldsymbol{F})$是对第$c$个通道进行第2次线性变换。整体公式表示为通道注意力$\boldsymbol{\pi}_C(\boldsymbol{F})$通过对输入特征$\boldsymbol{F}$的每个通道$\boldsymbol{F}_c$施加两组自适应线性变换，并取最大值实现非线性增强，最终与原始特征逐元素相乘，得到通道注意力增强的特征图。

最后，动态检测头模块通过上述3个注意力模块的顺序叠加，动态优化输入特征，最终得到优化后的输出，即$\mathrm{W}(\boldsymbol{F})=\boldsymbol{\pi}_C(\boldsymbol{\pi}_S(\boldsymbol{\pi}_L(\boldsymbol{F})\cdot\boldsymbol{F})\cdot\boldsymbol{F})\cdot\boldsymbol{F}$。

1.2.7　YOLOv5n-ChtaFSD模型的稳定性和鲁棒性测试

在保证试验平台和Batch、Imgsz及Workers等试验参数不变的情况下，YOLOv5n-ChtaFSD模型分别进行100个Epoch和150个Epoch试验，得到边界框损失率、目标置信度损失率、分类损失率、精确率、召回率、交并比阈值为0.50时的平均精度均值、交并比阈值从0.50至0.95且步长0.05的平均精度均值，从而验证模型的稳定性和鲁棒性。

1.2.8　消融试验及结果评价

为验证各改进模块对模型性能的独立贡献与协同增益，本研究以基准模型YOLOv5n（试验1）为基础，采用单变量控制法设计消融试验。首先，分别单独引入Efficient-RepGFPN（试验2）、单独嵌入SimAM（试验3）和单独引入动态检测头（试验4），采用精确率、召回率、交并比阈值为0.50时的平均精度均值、浮点运算量、参数量以及模型体积这6个指标对各模块的性能进行评价；其次，为验证模块间的协同作用，在上述试验基础上，依次进行两两模块组合试验（分别为试验5~试验7），同样采用上述6个指标对各模块的性能进行评价；最终，同时引入Efficient-RepGFPN、嵌入SimAM和引入动态检测头（试验8），仍然采用上述6个指标对各模块的性能进行评价。浮点运算量即模型完成一次前向推理所需的计算量；参数量即模型中可学习参数的总数量；模型体积即模型文件占用的存储空间。浮点运算量、参数量以及模型体积这3个指标越小，则试验结果越好；而其他3个指标越大，则试验结果越好（魏冰冰等，2025）。

1.2.9　对比试验及结果评价

本研究选取当前主流的YOLO系列模型YOLOv5n、YOLOv5s、YOLOv8n、YOLOv8s（Hussain，2023）、YOLOv11n和YOLOv11s（Khanam & Hussain，2024）作为基准模型进行对比试验，将这些模型在1.2.6的环境参数下与改进后的YOLOv5n-ChtaFSD模型进行比较。采用精确率、召回率、交并比阈值为0.50时的平均精度均值和模型体积这4个指标对结果进行评价。

1.2.10　YOLOv5n-ChtaFSD模型特征的可视化

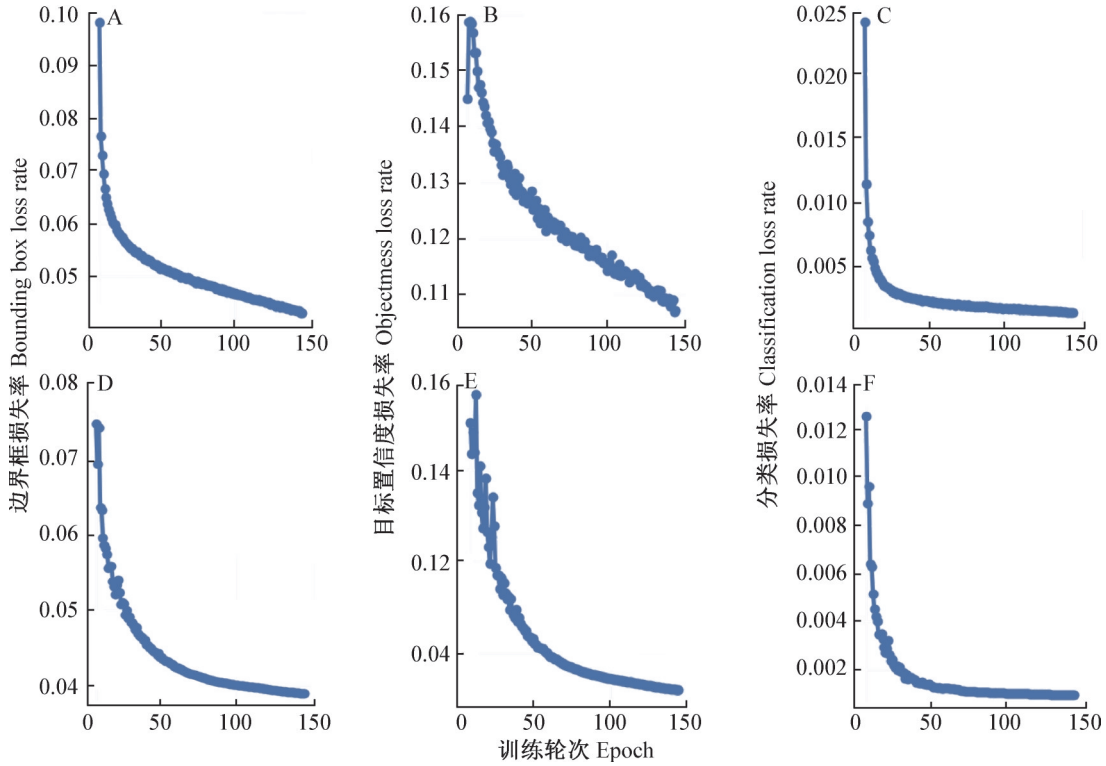
热力图可视化是一种数据可视化技术，通过颜色的变化来表示数值的大小、密度或强度。热力图通过颜色的渐变来编码数据的值（Selvaraju et al.，

2020)。通常用蓝色、绿色等冷色调表示低数值或低密度区域,用黄色、红色等暖色调表示高数值或高密度区域,中间数值则对应过渡色。从测试集中随机选取3幅图像,采用XGradCAM算法验证改进后模型YOLOv5n-ChtaFSD的有效性,参数设置如下:梯度权重归一化,特征图选取模型最后一层卷积层输出,置信度阈值设为0.2,热力图比例系数设为0.02。XGradCAM算法通过梯度反向传播定位模型关注目标区域,量化模型对目标特征的响应强度,进而清晰呈现模型的特征学习与定位聚焦效果;分别利用YOLOv5n模型与YOLOv5n-ChtaFSD模型作为基础模型,输入选取的测试图像,通过XGradCAM算法生成对应目标区域的热力图,对比两种模型热力图的目标区域聚焦度、特征响应强度及背景冗余度。

2 结果与分析

2.1 YOLOv5n-ChtaFSD模型的稳定性与鲁棒性

YOLOv5n-ChtaFSD模型经过100个Epoch训练后,测试集和训练集的边界框损失率、目标置信度损失率和分类损失率均达到收敛状态(图3);YOLOv5n-ChtaFSD模型的精确率、召回率、交并比阈值为0.50时的平均精度均值、交并比阈值从0.50至0.95且步长0.05的平均精度均值参数均较高(图4),表明改进后的模型稳定性与鲁棒性较好,既保证模型充分收敛,又避免过度训练的资源浪费。



A~C:测试集;D~F:验证集。A–C: Test set; D–F: validation set.

图3 训练时YOLOv5n-ChtaFSD模型的边界框损失率(A、D)、目标置信度损失率(B、E)和分类损失率(C、F)

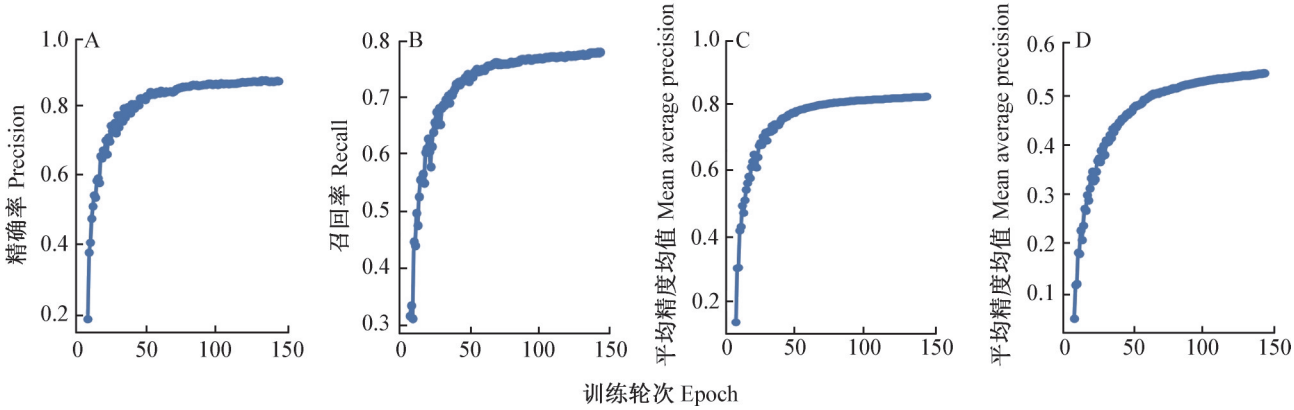
Fig. 3 Loss rates of bounding box (A, D), objectness (B, E), classification (C, F) for YOLOv5n-ChtaFSD model during training

2.2 消融试验结果

仅引入Efficient RepGFPN后(试验2),参数量与模型体积略有增加,但精确率、召回率与平均精度均值分别为88.9%、78.9%和83.9%,较试验1分别提升了2.7百分点、2.0百分点和2.0百分点,表明该模块有效增强了多尺度特征融合能力;仅嵌入SimAM后(试验3),平均精度均值较试验提升了0.3百分点,表明该模块对关键区域的特征强化有积极作用;仅引入动态检测头后(试验4),精确率、召回率以及平均精度均值分别为88.9%、80.3%和85.0%,尤其是召回率与平均精度均值较试验1和试验2提升明显,分别提升了3.4百分点和3.1百分点,表明该模块在尺度、空间与任务感知方面的自适应融合能力有助于减少漏检并提升定位精度(表1)。

引入Efficient RepGFPN和嵌入SimAM后(试验5),虽精确率较试验2有小幅度下降,但召回率和平均精度均值相应提升,而精确率、召回率与平均精度均值较试验3分别提升了0.8百分点、2.0百分点

和1.8百分点，表现出更强的泛化性能；引入Efficient RepGFPN和动态检测头后（试验6），浮点运算量、参数量和模型体积虽较试验2和试验4有增加，但精确率、召回率与平均精度均值也均有很大程度的提升，表明二者具备良好的协同增强效应；嵌入SimAM和引入动态检测头后（试验7），参数量与模型体积虽较试验2有所增大，但精确率、召回率与平均精度均值均获得进一步提升；同时引入GFPN、嵌入SimAM和引入动态检测头后（试验8），精确率、召回率与平均精度均值分别为90.7%、81.6%与85.9%，较试验1提升4.5百分点、4.7百分点和4.0百分点，模型体积为4.9 MB，浮点运算量为5.4 GFLOPs，参数量为2.4 M，均处于合理范围（表1），表明本研究所提出的改进策略在增强模型表达能力的同时，未引起显著的冗余负担。



A~D：精确率、召回率、交并比（intersection over union，IoU）阈值为0.50时的平均精度均值、IoU阈值从0.50至0.95且步长0.05的平均精度均值。A–D: Precision, recall, mean average precision at IoU threshold of 0.50, mean average precision across IoU thresholds from 0.50 to 0.95 with step 0.05.

图4　训练时YOLOv5n-ChtaFSD模型的评价指标

Fig. 4　Evaluation metrics of YOLOv5n-ChtaFSD model during training

表1　消融试验结果

Table 1　Results of ablation experiments

组数 Group	Efficient RepGFPN	SimAM	动态检测头 Dynamic head	精确率 Precision/%	召回率 Recall/%	平均精度均值 Mean average precision/%	浮点运算量 Floating point operation/GFLOP	参数量 Parameter/M	模型体积 Volume/MB
试验1 Experiment 1	×	×	×	86.2	76.9	81.9	4.1	1.7	3.6
试验2 Experiment 2	√	×	×	88.9	78.9	83.9	5.1	2.3	4.7
试验3 Experiment 3	×	√	×	87.4	77.3	82.2	4.1	1.7	3.6
试验4 Experiment 4	×	×	√	88.9	80.3	85.0	4.4	1.8	3.8
试验5 Experiment 5	√	√	×	88.2	79.3	84.0	5.1	2.3	4.8
试验6 Experiment 6	√	×	√	89.1	81.1	85.3	5.4	2.4	4.9
试验7 Experiment 7	×	√	√	89.0	80.5	84.8	4.4	1.8	3.8
试验8 Experiment 8	√	√	√	90.7	81.6	85.9	5.4	2.4	4.9

√：使用此方法；×：不使用此方法。√: This method is adopted; ×: this method is not adopted. operation.

2.3　YOLOv5n-ChtaFSD模型与其他模型的对比结果

与YOLOv5n、YOLOv8n、YOLOv8s、YOLOv11n和YOLOv11s模型相比，YOLOv5n-ChtaFSD模型的精确率分别提高了4.5百分点、10.8百分点、2.7百分点、8.7百分点与3.9百分点。与YOLOv5s模型相比，YOLOv5n-ChtaFSD模型的精确率略低0.8百分

点，但模型体积降低8.8 MB，更适用于轻量化部署场景。与YOLOv5n、YOLOv8n、YOLOv11n模型相比，YOLOv5n-ChtaFSD模型的平均精度分别提高了4.0百分点、7.9百分点和6.1百分点。此外，与YOLOv5s、YOLOv8n、YOLOv8s、YOLOv11n和YOLOv11s模型相比，YOLOv5n-ChtaFSD模型的体积分别减少8.8、1.0、16.6、0.3与13.4 MB，进一步体现其在模型效率方面的优势（表2）。

表2 YOLOv5n-ChtaFSD与其他YOLO模型的对比结果

Table 2 Comparison results between YOLOv5n-ChtaFSD and other YOLO models

模型 Model	精确率 Precision/%	召回率 Recall/%	平均精度均值 Mean average precision/%	模型体积 Volume/MB
YOLOv5n	86.2	76.9	81.9	3.6
YOLOv5s	91.5	83.5	88.4	13.7
YOLOv8n	79.9	73.0	78.0	5.9
YOLOv8s	88.0	80.2	86.2	21.5
YOLOv11n	82.0	74.5	79.8	5.2
YOLOv11s	86.8	79.3	84.8	18.3
YOLOv5n-ChtaFSD	90.7	81.6	85.9	4.9

2.4 YOLOv5n-ChtaFSD模型特征的可视化

在飞机草目标区域，YOLOv5n-ChtaFSD模型的热力图表现出较强的特征响应，暖色区域更集中，亮度更高（图5），表明该模型对飞机草关键判别性特征具有更强的捕捉和响应能力；而YOLOv5n模型的热力图响应较分散，冷色调区域占主导（图5），表明该模型特征学习不够深入，关键信息提取能力有限，从而影响了识别的稳定性与精度。

3 讨论

本研究构建的YOLOv5n-ChtaFSD模型通过三模块协同优化实现了飞机草检测精度与轻量化的平衡。本研究的消融试验结果显示，仅引入Efficient RepGFPN后，平均精度均值提升2.0百分点，与Xu et al.（2022）在Damo-YOLO模型中提出该结构时的设计初衷一致。Damo-YOLO模型聚焦通用目标检测，而本研究将其用于入侵杂草检测场景时，特征融合效率提升更显著，因为飞机草背景复杂、形态多样的特性恰好匹配了该结构消除冗余上采样、强化跨尺度特征复用的核心优势，从而有效解决了杂草小目标特征弥散问题。本研究的消融试验结果显示，SimAM注意力机制的单独优化虽仅使平均精度均值提升了0.3百分点，但与Yang et al.（2021）提出该机制时的通用化应用不同。本研究将其嵌入颈部网络而非主干网络，既避免了计算冗余，又精准强化了多尺度融合后的有效特征，此设计与茄子病虫害检测中注意力机制针对嵌入特征融合阶段的策略相符（郭臻和李玥等，2024），且针对飞机草特征稀疏的特异性进一步降低了小目标漏检率。

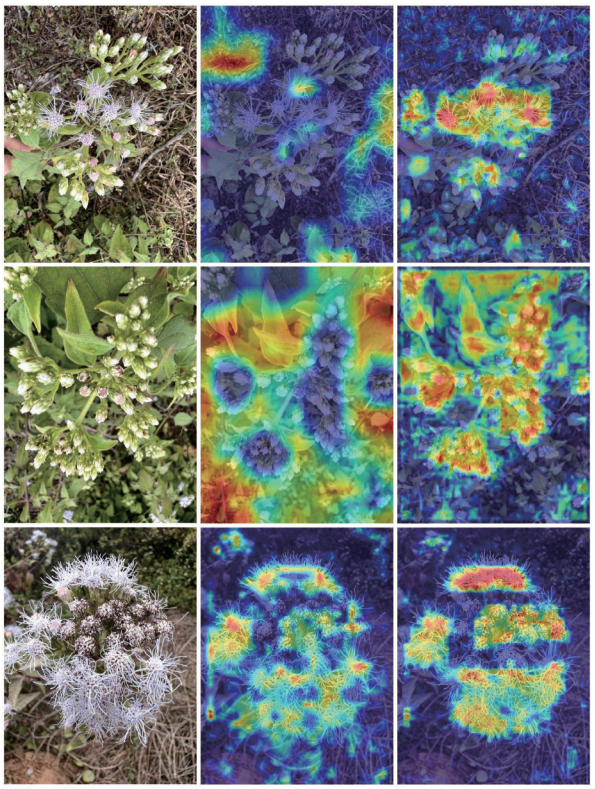


图5 不同场景下YOLOv5n与YOLOv5n-ChtaFSD模型的热力图

Fig. 5 Heatmaps of YOLOv5n and YOLOv5n-ChtaFSD models under different scenarios

本研究的消融试验结果显示，引入动态检测头后平均精度均值提升了3.1百分点，显著优于其在Dai et al.（2021）原始研究中的基础表现，究其原因可能是本研究通过Efficient RepGFPN特征融合–SimAM特征筛选–动态检测头定位的全链路协

同，弥补了单一动态检测头在小目标特征不足时的定位偏差，与Ju et al.（2025）针对水田杂草的多模块改进逻辑一致，但本研究更注重轻量化框架下的资源控制，未因模块叠加导致浮点运算量激增。本研究的对比试验结果显示YOLOv5n-ChtaFSD模型的精确率为90.7%，虽略低于YOLOv5s模型的精确率（91.5%），但模型体积仅4.9 MB，较YOLOv5s模型减少8.8 MB，突破了YOLO系列模型精度与体积正相关的普遍规律（Hussain，2023）；YOLOv5n-ChtaFSD模型在参数量接近YOLOv11n模型的情况下平均精度均值提升6.1百分点，这一优势源于YOLOv11n侧重通用目标检测优化，而YOLOv5n-ChtaFSD模型针对飞机草的稀疏分布、形态变异等特性进行了专项适配，印证了物种特异性模型优化优于通用模型直接迁移的结论（Hu et al.，2024）。

针对入侵植物飞机草形态细碎、背景杂乱、小目标与遮挡频发的特点，本研究构建轻量化检测模型YOLOv5n-ChtaFSD。彭镇坤和侯周国（2025）构建了基于轻量化YOLOv8n的玉米苗与杂草识别方法，技术路线与评价体系均与本研究一样，即以轻量化YOLO系列为基线，通过优化特征融合结构、增强注意力机制、改进检测头实现精度与效率的平衡，采用平均精度均值、参数量、浮点运算量等指标对模型进行评价。但玉米作物中杂草识别采用Ghost卷积与内容感知特征重组上采样实现模型压缩，面向玉米田间规整场景，开展二分类检测（彭镇坤和侯周国，2025）；而本研究未压缩主干网络，采用Efficient RepGFPN、SimAM无参注意力与动态检测头的组合策略，专注复杂野外环境下单类别飞机草检测，更适配入侵杂草野外监测需求，为复杂环境下轻量化检测提供了与已有农田杂草识别不同的技术路径。

本研究的飞机草数据集仅采集于广东省湛江市森林公园周边，虽进行了数据增强处理，但未覆盖中国西南、华南不同气候区飞机草的生长场景，这可能导致模型在新入侵区域的泛化能力不足；本研究未针对飞机草与其近缘种如紫茎泽兰*Eupatorium adenophorum*的相似性形态设计鉴别机制，因此现有模型难以区分这类易混淆物种；本研究仅完成离线性能验证，缺乏在移动机器人等边缘设备上的实地部署测试。未来将扩充跨区域、多季节、多生态系统的飞机草数据集，增加近缘种样本以提升模型鉴别能力；将多光谱图像技术与超绿特征处理融合，优化极端光照、阴雨等环境下的特征提取鲁棒性；通过模型量化、剪枝技术进一步降低推理耗时，推进在移动机器人平台的实地部署，最终构建图像快速检测–分子辅助验证–田间精准防控的一体化技术体系，为外来入侵杂草的治理提供更全面的支撑。

参 考 文 献（References）

Ahn YP, Kim KB, Park HJ. 2023. Comparison of the effect of interpolation on the mask R-CNN model. Journal of Information and Communication Convergence Engineering, 21(1): 17−23

Buddhachat K, Kongket B, Pandith H. 2020. Differentiation of Siam weed (*Chromolaena odorata* (L.) R. M. King & H. Rob.) and morphologically related species in Asteraceae by ITS barcode coupled with high resolution melting analysis (HRM). Plant Gene, 24: 100246

Chen MQ, Yu LJ, Zhi C, Sun RJ, Zhu SW, Gao ZY, Ke ZX, Zhu MQ, Zhang YM. 2022. Improved faster R-CNN for fabric defect detection based on Gabor filter with genetic algorithm optimization. Computers in Industry, 134: 103551

Chen Z, Bai TC. 2018. Study on the damage and control of exotic invasive plant *Eupatorium odoratum* to the ecosystem in Southwest Guizhou. The Farmers Consultant, (4): 117 (in Chinese) [陈臻, 白天才. 2018. 外来入侵植物飞机草对黔西南州生态系统的危害和防除研究. 农家参谋, (4): 117]

Dai XY, Chen YP, Xiao B, Chen DD, Liu MC, Yuan L, Zhang L. 2021. Dynamic head: unifying object detection heads with attentions// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, pp. 7369−7378

Ding XH, Zhang XY, Han JG, Ding GG. 2022. Scaling up your kernels to 31x31: revisiting large Kernel design in CNNs//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, pp.11954−11965

Guo Z, Li Y. 2024. HCI-YOLO: a detection model for eggplant fruit diseases and pests based on improved YOLOv8. Software Engineering, 27(12): 63−68 (in Chinese) [郭臻, 李玥. 2024. HCI-YOLO: 基于改进YOLOv8的茄子果实病虫害检测模型. 软件工程, 27(12): 63−68]

Hsu WY, Lin WY. 2021. Adaptive fusion of multi-scale YOLO for pedestrian detection. IEEE Access, 9: 110063−110073

Hu K, Wang ZY, Coleman G, Bender A, Yao TT, Zeng S, Song DZ, Schumann A, Walsh M. 2024. Deep learning techniques for in-crop weed recognition in large-scale grain production systems: a review. Precision Agriculture, 25(1): 10073

Hussain M. 2023. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. Machines, 11(7): 677

Ju JY, Ma XH, Yang CX, Qian FZ, Wang JF, Chen GQ. 2025. Lightweight detection method for weeds in complex paddy fields based on the improved YOLOv5. Smart Agricultural Technology, 12: 101418

Khanam R, Hussain M. 2024. YOLOv11: an overview of the key architectural enhancements. arXiv, 2410.17725

Lee SH, Gao G. 2023. A study on pine larva detection system using

swin transformer and cascade R-CNN hybrid model. Applied Sciences, 13(3): 1330

Li SJ, Liang QY, Yu YH, Chen YH, Fu HM, Zhang HY. 2024. Research on Asian citrus psyllid YOLO v8-MC recognition algorithm and insect remote monitoring system. Transactions of the Chinese Society for Agricultural Machinery, 55(6): 210−218 (in Chinese) [李善军, 梁千月, 余勇华, 陈耀晖, 付慧敏, 张宏宇. 2024. 柑橘木虱YOLO v8-MC识别算法与虫情远程监测系统研究. 农业机械学报, 55(6): 210−218]

Peng ZK, Hou ZG. 2025. Distinguishing between maize seedlings and weeds based on lightweight YOLOv8n. Hunan Agricultural Sciences, (11): 101−106 (in Chinese) [彭镇坤, 侯周国. 2025. 基于轻量化YOLOv8n的玉米苗与杂草识别. 湖南农业科学, (11): 101−106]

Qian HM, Wang HL, Feng S, Yan SY. 2023. FESSD: SSD target detection based on feature fusion and feature enhancement. Journal of Real-Time Image Processing, 20: 2

Redmon J, Divvala S, Girshick R, Farhadi A. 2015. You Only Look Once: unified, real-time object detection. arXiv, 1506.02640

Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. 2020. Grad-CAM: visual explanations from deep networks *via* gradient-based localization. International Journal of Computer Vision, 128(2): 336−359

Wei BB, Zhang Z, Yan K, Zhang GJ, Wang P, Huang QY, Niu CY. 2025. Intelligent monitoring of tephritid fruit fly pests using an improved YOLOv5s model. Journal of Plant Protection, 52(5): 1163−1171 (in Chinese) [魏冰冰, 张志, 闫科, 张桂健, 王鹏, 黄求应, 牛长缨. 2025. 基于改进YOLOv5s模型的实蝇类害虫智能监测. 植物保护学报, 52(5): 1163−1171]

Wu YF, Wang YR, Sun YB, Zhou YL, He LH, Huang Y, Chen Z. 2017. Risk analysis of harmful plant *Eupatorium odoratum* to Southwest Guizhou. Forest Science and Technology, (2): 46−48 (in Chinese) [吴迎福, 王亚蓉, 孙远彬, 周娅琳, 贺林晖, 黄勇, 陈臻. 2017. 有害植物飞机草对黔西南州危害的风险分析. 林业科技通讯, (2): 46−48]

Xu XZ, Jiang YQ, Chen WH, Huang YL, Zhang Y, Sun XY. 2022. Damo-YOLO: a report on real-time object detection design. arXiv, 2211.15444

Xu Y, Zhao H, Yue YJ. 2025. Detection and identification of crop pests based on improved YOLOv8n. Journal of Chinese Agricultural Mechanization, 46(11): 144−151, 160 (in Chinese) [徐悦, 赵辉, 岳有军. 2025. 基于改进YOLOv8n的农作物虫害检测与识别. 中国农机化学报, 46(11): 144−151, 160]

Yang LX, Zhang RY, Li LD, Xie XH. 2021. SimAM: a simple, parameter-free attention module for convolutional neural networks//Proceedings of the 38th International Conference on Machine Learning, PMLR139

Yu Y, Wu JP, He XX, Wei J, Gao XH. 2024. Arthropod object detection model based on improved YOLOv4-tiny. Computer Technology and Development, 34(1): 114−120 (in Chinese) [余咏, 吴建平, 何旭鑫, 韦杰, 高雪豪. 2024. 基于改进YOLOv4-tiny的节肢动物目标检测模型. 计算机技术与发展, 34(1): 114−120]

Zhao WF, Huang YJ, Zhong MY, Li ZY, Luo ZT, Huang JJ. 2024. Multi-task visual perception method in dragon orchards based on OrchardYOLOP. Transactions of the Chinese Society for Agricultural Machinery, 55(11): 160−170 (in Chinese) [赵文锋, 黄袁爵, 钟敏悦, 李振源, 罗梓涛, 黄家俊. 2024. 基于OrchardYOLOP的火龙果园多任务视觉感知方法. 农业机械学报, 55(11): 160−170]

Zhu MM, Hu GP, Zhou H, Wang SQ, Zhang YL, Yue SJ, Bai Y, Zang KX. 2021. Arbitrary-oriented ship detection based on RetinaNet for remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14: 6694−6706

（责任编辑：张俊芳）