

基于 YOLOv5 的白芍除草机目标检测模型设计与试验

冯雅¹ 柳佳俊² 蒋杰² 方园³ 刘洋洋^{2*}

1. 绍兴职业技术学院信息工程学院 浙江绍兴 312000

2. 安徽农业大学工学院 安徽合肥 230036

3. 安徽农业大学园艺学院 安徽合肥 230036

摘要: 目的: 针对白芍中药田中杂草和草药的精准识别问题。方法: 本文提出一种基于 YOLOv5 的白芍除草机目标检测模型, 可以实现对白芍和两种杂草的精准识别检测, 提高了除草效果、精度与实时性, 并通过试验验证了该目标检测模型的有效性。实验结果: 基于 yolov5n 目标检测模型的 mAP 为 87.5%, 精准率为 92.8%, 召回率为 82.1%, 模型预测速度为 10.1ms, 目标检测的置信度能达到接近 85% 到 90% 左右。结论: 证明本文提出目标检测模型对白芍以及中草药行间杂草的识别检测精度较高, 满足白芍除草机的工作需要, 为中药材智能植保技术研发提供了参考。

关键词: 深度学习; YOLOv5; 白芍; 除草机; 目标检测

引言

白芍具有极大的医用价值, 但杂草严重影响产量与质量。机械除草不仅效率高, 还不会造成污染, 是中药材植保的最佳选择^[1]。但目前除草机械作业过程容易对农作物造成损害, 严重制约中药业发展, 因此亟需研发适用中药材除草机的目标检测算法。Wang 等^[2]结合 YOLOv5 中的注意力机制构建了 YOLO-CBAM 模型。权龙哲等^[3]提出基于 YOLOv4 卷积神经网络目标检测模型, 修改后 YOLOv4 网络的 mAP 为 0.828, 优于其他 3 类模型。姜红花等^[4]提出基于 Mask R-CNN 的杂草检测方法, 当 IoU 为 0.5 时, mAP 为 0.785、单样本耗时为 285ms。李金阳等^[5]提出了杂草检测模型 YOLOv7-FWeed, 采用 F-ReLU 作为卷积模块的激活函数, 增加了 MaxPool 多头自注意力 (M-MHSA) 模块, 以提高杂草的识别精度。上述的目标检测模型在一定程度上满足了实际使用场景所需的精确度高、检测速度快、成本较低等需求, 但是未有针对于白芍除草机的目标检测模型。

因此本研究提出一种基于 YOLOv5 的白芍除草机目标检测模型, 并选择出最优的边框损失。该模型包含对白芍和杂草的识别检测两部分, 提高了除草精度和实时性, 并通过试验验证了模型的有效性。

1 材料和方法

1.1 图像采集和处理

本研究选用白芍田中常见的狗尾草和喜旱莲子草两种

杂草为例, 采集白芍开花期图像 2000 张和两种杂草的清晰图片各 800 张, 共计 3600 张图片来建立白芍田作物数据集。通过 LabelImg 软件对图像进行手动标记 (如图 1-1 所示), 再将标记文件按照 8:1:1 的比例分为训练集, 验证集和测试集, 如表 1 所示。



图 1 LabelImg 标注

表 1 数据集分类数量表

数据集名称	杂草种类	数量	合计
训练集	喜旱莲子草	1600	2880
	狗尾草	640	
	白芍	640	
验证集	喜旱莲子草	200	360
	狗尾草	80	
	白芍	80	
测试集	喜旱莲子草	200	360
	狗尾草	80	
	白芍	80	

1.2 图像的灰度处理

在 RGB 模型中，灰度化的优势在于相较于彩色图像，灰度图像占内存较小，运行速度更快，可在视觉上增加对比，突出目标区域。常见的图像灰度处理方法有最大值法、平均值法和加权平均法。最大值法是直接取分量中最大的分量数值，如式（1）所示：

$$R = G = B = \max (R, G, B) \quad (1)$$

平均值法是取分量中数值的均值，如式（2）所示：

$$R = G = B = (R + G + B) / 3 \quad (2)$$

加权平均法是根据重要性将三个分量以不同的权重进行加权平均，如式（3）所示：

$$\text{Gray} (i, j) = 0.299 * R (i, j) + 0.578 * G (i, j) + 0.114 * B (i, j) \quad (3)$$



图2 三种灰度处理结果

由图2可知，加权平均法转换的灰度图亮度最佳，所以本研究采用加权平均法来处理杂草图像。

1.3 目标检测模型设计

YOLO 系列算法拥有十分快速的推理速度及较高的检测精度，其中 YOLOv5 在保持网络轻量化的同时，对杂物遮挡和小目标等问题上较其他算法。所以本实验选择 YOLOv5 网络模型为本研究基本的目标检测模型，整体架构如图3所示。

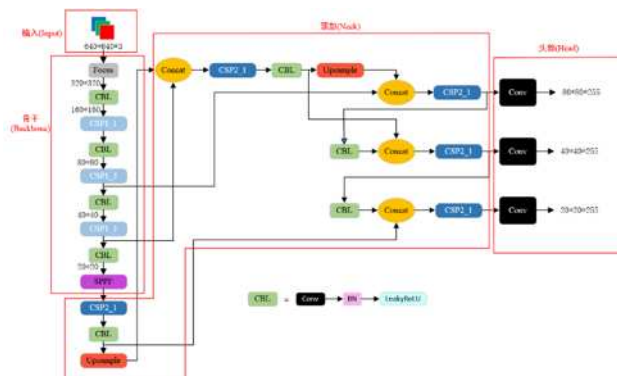


图3 YOLOv5 网络模型结构

YOLOv5 模型通过马赛克 (Mosaic) 数据增强、自适应锚框计算和自适应图片缩放处理后的图像。马赛克数据增强处理丰富了数据集，随机缩放增加了许多小尺寸的白芍图像，使得网络模型的鲁棒性更好，同时减少 GPU 显存的需要。在目标检测的推理过程中会产生很多检测框，NMS 的主要作用就是通过计算检测框的 IoU 值与设定阈值进行对比，不断筛选 IoU 值最高的检测框，最终选择最优检测框。由网络模型预测出的边界框与样本真实边界框之间存在交集和

并集，交集与并集之比为交并比 (Intersection over Union, IoU)，如公式（4）所示。

$$\text{IoU} = \frac{A \cap B}{A \cup B} \quad (4)$$

式中，A 为真实框；B 为预测框，如图9所示。

使用 IoU 计算边框损失，边框损失越小说明预测框和真实框的交集越大，即目标检测的效果越好。YOLOv5 的边框损失函数为 CIoU，减少了预测框、目标框的重叠面积和中心点距离对边框损失的影响，如图4所示。由于 CIoU 对长宽比的描述是相对值，存在一定的模糊，由公式（5）（6）可知，若预测框的宽与高之比等于真实框的宽与高之比，则 CIoU 等于 DIoU。

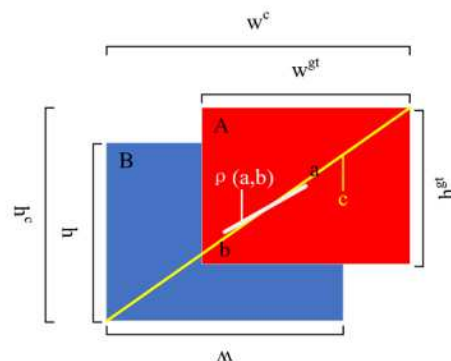


图4 CIoU 图解

本文将采用 CIoU、EIoU、SIoU、Alpha-IoU 以及 C-Alpha IoU 与 YOLOv5 n、YOLOv5 s、YOLOv5 m 进行组合，得到不同目标检测模型，通过分析模型训练数据集的结果，从而选

择最优边框损失函数，进一步增强目标检测的效果。相应计算公式如下。

$$L_{C_{IoU}} = 1 - IoU + \frac{\rho^2(a,b)}{c^2} + \beta v \quad (5)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2, \begin{cases} \beta = 0, \text{当IoU} < 0.5 \text{时;} \\ \beta = \frac{v}{(1 - IoU) + v}, \text{当IoU} \geq 0.5 \text{时.} \end{cases} \quad (6)$$

式中，a、b分别为真实框和预测框的中心点；c为预测框和真实框之间最小外接矩形的对角线距离；w、h、wgt、hgt分别为预测框和真实框的宽与高；ρ(a,b)为欧式距。

$$L_{EIoU} = 1 - IoU + \frac{\rho^2(a,b)}{c^2} + \frac{\rho^2(w, w^{gt})}{(w^c)^2} + \frac{\rho^2(h, h^{gt})}{(h^c)^2} \quad (7)$$

式中，wc、hc分别为预测框和真实框之间最小外接矩形的宽与高。

$$L_{S_{IoU}} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (8)$$

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma t}), \rho_x = \left(\frac{a_x - b_x}{c_w} \right)^2, \rho_y = \left(\frac{a_y - b_y}{c_h} \right)^2 \quad (9)$$

$$\gamma = 2 - \Lambda, \Lambda = 1 - 2 \times \sin^2(\arcsin(x) - \frac{\pi}{4}) \quad (10)$$

$$x = \frac{c_h^*}{\rho(a,b)} = \sin(\pi), c_h^* = \max(a_y, b_y) - \min(a_y, b_y) \quad (11)$$

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^0, \omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})}, \omega_h = \frac{|h, h^{gt}|}{\max(h, h^{gt})} \quad (12)$$

式中，ax、ay、bx、by分别为真实框和预测框的中心点的横、纵坐标；cw、ch分别为预测框和真实框之间最小外接矩形的宽与高。

$$L_{Alpha-IoU} = 1 - IoU^\alpha \quad (13)$$

$$L_{C-Alpha IoU} = 1 - IoU^\alpha + \left[\frac{\rho^2(a,b)}{c^2} \right]^\alpha + (\beta v)^\alpha \quad (14)$$

式中，α=3。

分类损失和目标损失均采用 BCEwithLogitsLoss，即对输出的置信度进行 sigmoid 函数激活运算，然后再进行交叉熵损失函数（Binary Cross Entropy Loss, BCELoss）运算，其计算公式如下：

$$L_{class} = -\frac{1}{N} \sum_{i=1}^N \{y_i \log C(x_i) + (1 - y_i) \log [1 - C(x_i)]\}, C(x_i) = \frac{1}{1 + e^{-x_i}} \quad (15)$$

式中，N表示训练时的样本总数；xi表示第i次输入的分类预测值；yi表示第i次输入的分类真实值（0或1）；

C(xi)表示第i次输入的分类预测值经过 sigmoid 函数激活后得到的概率。

$$L_{object} = -\frac{1}{N} \sum_{i=1}^N \{q_i \log O(p_i) + (1 - q_i) \log [1 - O(p_i)]\}, O(p_i) = \frac{1}{1 + e^{-p_i}} \quad (16)$$

式中，pi表示第i次输入的目标预测值；qi表示第i次输入的目标真实值（0或1）；O(pi)表示第i次输入的目标预测值经过 sigmoid 函数激活后得到的概率。

1.3.1 模型训练参数设置

本研究针对 yolov5s、yolov5m 和 yolov5n 对于白芍和杂草的检测进行训练，参数如表 2 所示。

表 2 超参数配置参数

参数名称	原始默认 coco 参数
初始学习率 lr0	0.01
余弦退火算法 lrf	0.2
学习率动量 momentum	0.937
权重衰减系数 weight_decay	0.0005
预热学习 epoch warmup_epochs	3
预热学习率动量 warmup_momentum	0.8
预热学习率 warmup_bias_lr	0.1
giou 损失系数 box	0.05
分类损失系数 cls	0.5
有无物体损失系数 obj	1
bceloss 中性样本重量 obj_pw	1
IoU 训练阈值 iou_t	0.2
多个锚点阈值 anchor_t	4

1.3.2 模型评价指标

本研究中使用准确率，召回率，平均精度均值作为目标检测评价指标，通过混淆矩阵计算评价指标，其中 TP (True Positive) 指的是模型将多少原本是正类成果预测成正类的数量，FN (False negative) 是模型将正类预测为负类的数量，FP (False Positive) 是模型将负类预测成正类的数量，最后的 TN (True negative) 指的是模型将负类预测成负类的数量。混淆矩阵如表 3 所示。

表 3 模型预测混淆矩阵

序号	真实值与预测值关系	字母表示
1	真实值是正 & 预测值是正	TP
2	真实值是正 & 预测值是负	FN
3	真实值是负 & 预测值是正	FP
4	真实值是负 & 预测值是负	TN

准确率 (Precision)，是在所有检测完成的目标中被检测正确的比例，如式 (17) 所示。

$$Precision = \frac{TP}{TP + FP} \quad (17)$$

召回率 (recall)，是从全部正样本中被正确确定为正例的类别的可能性，如式 (18) 所示。

$$Recall = \frac{TP}{TP + FN} \quad (18)$$

平均精度均值 (mAP) 是指阈值等于 0.5 时, 样本中所有类别的平均精度的平均值, 反应了模型的性能优劣程度, 如式 (19) 所示。

$$mAP = \frac{\sum_{i=1}^n \int_0^1 P(R) dR}{n} \quad (19)$$

2 实验结果和讨论

2.1 模型实验环境与方案

本研究模型实验设备的 CPU 为 AMD Ryzen 7 4800H with

Radeon Graphics (8 核 2.90GHz), GPU 为 NVIDIA GeForce RTX 2060 (6GB), 操作系统为 Windows11, 运行环境为 Python 3.8.16、Pytorch 1.13.1、CUDA 11.7。参数设置如下: 超参数文件采用 hyp.scratch-low.yaml, 其中初始学习率为 0.01、学习率动量为 0.937、权重衰减系数为 0.0005, 图片输入尺寸为 640 × 640。模型的批量大小设置为 16, 当在最近 300 轮实验中模型效果没有得以提升时终止实验。

2.2 数据集的实验与结果分析

2.2.1 白芍数据集的实验结果

不同模型进行数据集的实验, 其边框损失如图 5 所示。

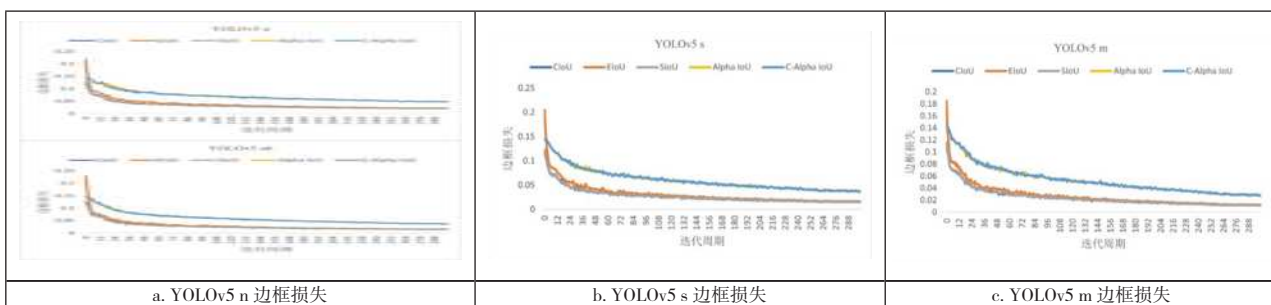


图 5 三组模型的边框损失对比

通过图 5 可以清楚地看出, 边框损失在第 60 轮迭代之后基本趋于平缓。

表 4 模型实验结果

模型	IoU 类型	mAP	P	R	模型大小 (MB)
YOLOv5 n	CIoU	0.797	0.802	0.75	3.71
	EIoU	0.848	0.897	0.774	
	SIoU	0.847	0.839	0.786	
	AlphaIoU	0.871	0.928	0.786	
	C-AlphaIoU	0.875	0.916	0.781	
YOLOv5 s	CIoU	0.826	0.782	0.769	13.7
	EIoU	0.848	0.847	0.727	
	SIoU	0.871	0.841	0.821	
	AlphaIoU	0.832	0.911	0.734	
	C-AlphaIoU	0.856	0.872	0.75	
YOLOv5 m	CIoU	0.845	0.911	0.727	40.2
	EIoU	0.842	0.863	0.774	
	SIoU	0.859	0.911	0.774	
	AlphaIoU	0.842	0.874	0.738	
	C-AlphaIoU	0.855	0.876	0.762	

由表 4 可知, mAP 值、精准率和召回率的最大值分别是 0.875、0.928 和 0.821。其中对于 YOLOv5 n 采用 C-Alpha IoU 可以获得最高的 mAP 值 0.875, 第二高的精准率 0.916 和第二高的召回率 0.781; 对于 YOLOv5 s 采用 SIoU 可以获得第二高的 mAP 值 0.871, 和最高的召回率 0.821, 但是准确率较低, 仅为 0.841。YOLOv5 m 模型 mAP 值、精准率和召回率都较低。基于以上分析, 本研究 YOLOv5n 网络模型的边框损失采用 C-Alpha IoU。

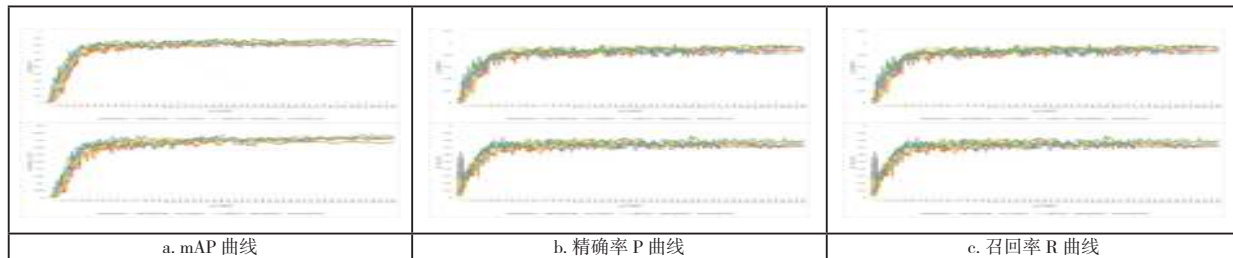


图 6 采用 C-Alpha IoU 的指标参数曲线

如图6所示,通过对比采用C-Alpha IoU的YOLOv5 n、YOLOv5 n6、YOLOv5 s、YOLOv5 s6、YOLOv5 m和YOLOv5 m6这6个模型的mAP、精确率P和召回率R的曲线,可以发现相应指标均已趋于稳定,进一步证明采用C-Alpha IoU的YOLOv5s网络模型在白芍目标检测中有效。

2.2.2 杂草数据集实验结果

通过对建立的杂草数据集进行实验,采用yolov5n权重模型进行实验,本次实验迭代次数100次、batch-size为8,可以在tensorboard中进行查询。

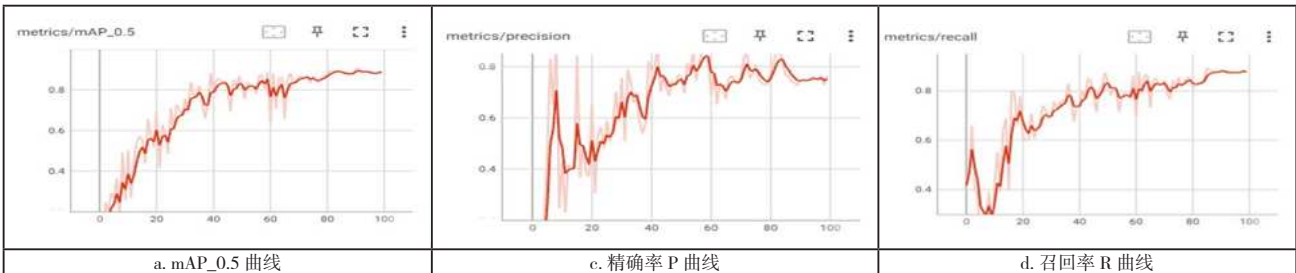


图7 YOLOv5n mAP、精确率、召回率实验结果

从图7可知,mAP=0.5时曲线极大值为91.7%;精确度最高可达88.93%;召回率为86.93%。

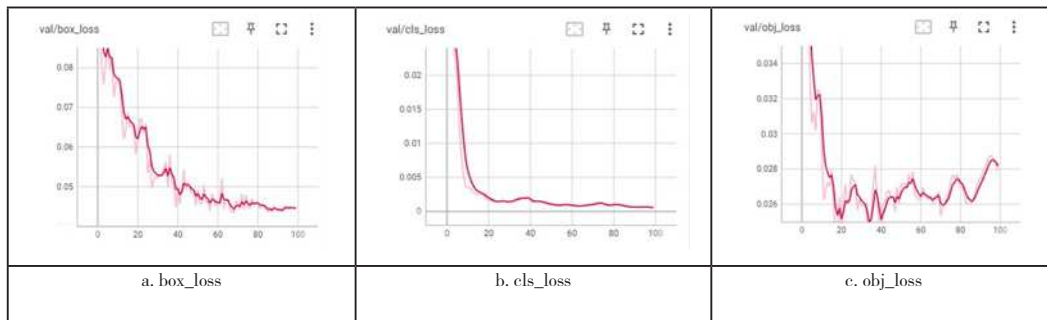


图8 YOLOv5s box_loss、cls_loss、obj_loss 实验结果

由图8中可知,box_loss值接近于0,说明预测框与标定框之间的误差较小,即目标框与真实框接近。cls_loss值接近于0,说明目标检测的分类信息较准确。obj_loss接近于0,说明计算网络置信度值较高,检测到目标的概率较大。图中三类曲线都平缓倾斜下降,没有出现过拟合现象,说明曲线拟合较为准确。但由于杂草随环境变化很大,因此后期

出现收敛震荡现象。

2.3 模型的实时性实验

不同权重大小的YOLOv5网络模型,目标检测的精确率和检测时间均不同,模型目标检测精确率越高,则所需要的检测时间越长,检测时间如表5所示。

表5 模型检测时间

模型	IoU类型	预处理时间 (ms)	推理时间 (ms)	非极大值抑制时间 (ms)	总计 (ms)
YOLOv5 n	C-AlphaIoU	3.2	15.1	4.8	23.1
YOLOv5 s	C-AlphaIoU	3.2	23.1	4.3	30.6
YOLOv5 m	C-AlphaIoU	3	28.2	4.1	35.3

通过表5可以看出,在采用C-Alpha IoU的模型中YOLOv5 n、YOLOv5 s、YOLOv5 m时间分别为23.1ms、30.6ms、35.3ms,其中,YOLOv5 n模型检测时间最短,满足除草机对目标检测实时性的需求。

2.4 模型检测结果

本研究使用以C-Alpha IoU为边框损失的YOLOv5 n进行测试,输入测试集图片进行检测,平均每张图检测速度为10.1ms,目标检测的置信度平均能达到接近85%到90%左右,识别效果如图9所示,模型检测速度和准确度较高。

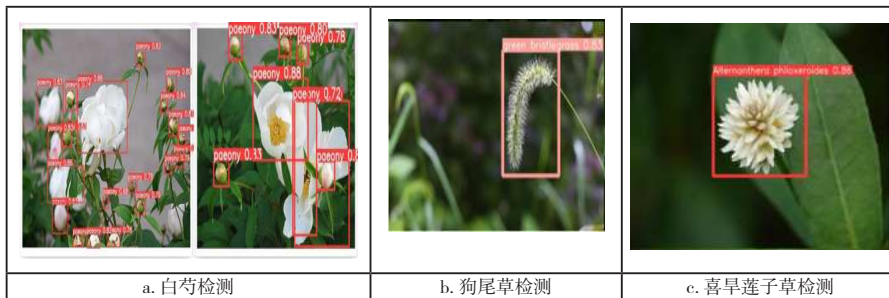


图 9 YOLOv5s 目标检测效果图

由图 9 可知，模型对测试集中白芍与杂草检测无遗漏，证明该模型对白芍、狗尾草和喜旱莲子草的目标检测实验效果较好，可以满足除草机械对目标检测的需求。

3 结论

为了提高中草药白芍除草机的除草效果、精度与实时性，本研究选择出最优的 YOLOv5 目标检测模型，并通过对数据集的训练结果，确定出边框损失采用 C-Alpha IoU 最合适。实验结果显示，模型的检测速度、精度及目标检测的置信度较高，证明本文提出目标检测模型对白芍以及杂草的识别检测满足白芍除草机的工作需要。本研究结果可以为实现白芍田杂草的机械化防控管理提供参考。

参考文献：

[1] 袁洪波, 赵努东, 程曼. 基于图像处理的田间杂草识别研究进展与展望 [J]. 农业机械学报, 2020, 51(S2):323-334.
[2] WANG Aichen, PENG Tao, Cao Huadong, et al. TIA-YOLOv5: An improved YOLOv5 network for real-time detection

of crop andweed in the field [J]. Frontiers in Plant Science, 2022.

[3] 权龙哲, 夏福霖, 姜伟, 等. 基于 YOLOv4 卷积神经网络的农田苗草识别研究 [J]. 东北农业大学学报, 2021, 52(7): 89-98.

[4] 姜红花, 张传银, 张昭, 等. 基于 Mask R-CNN 的玉米田间杂草检测方法 [J]. 农业机械学报, 2020, 51(6): 220-228.

[5] LI Jinyang, ZHANG Wei, ZHOU Hong, et al. Weed detection in soybean fields using improved YOLOv7 and evaluating herbicide reduction efficacy [J]. Frontiers in Plant Science, 2024, 14: 1284338.

作者简介：

冯雅（1992—），女，汉，浙江绍兴人，硕士研究生，绍兴职业技术学院，讲师，研究方向：中药材智能装备研究。

基金项目：

浙江省教育厅一般科研项目（Y202352930）。