

胡玲艳, 孙浩, 徐国辉, 等. 基于机器视觉的温室蓝莓花期蜜蜂授粉监测[J]. 华中农业大学学报, 2023, 42(3): 105-114.  
DOI: 10.13300/j.cnki.hnlkxb.2023.03.013

## 基于机器视觉的温室蓝莓花期蜜蜂授粉监测

胡玲艳<sup>1</sup>, 孙浩<sup>1</sup>, 徐国辉<sup>1</sup>, 郭睿雅<sup>2</sup>, 郭占俊<sup>2</sup>, 陈鹏宇<sup>1</sup>, 裴悦琨<sup>1</sup>, 汪祖民<sup>1</sup>

1. 大连大学信息工程学院, 大连 116622; 2. 大连市现代农业生产发展服务中心, 大连 116000

**摘要** 为评估日光温室蓝莓开花窗口期内授粉蜜蜂投放量的合理性, 基于机器视觉对蓝莓的蜜蜂授粉次数进行统计。针对检测环境复杂、目标尺度小、易被遮挡等问题, 对数据集进行改进泊松融合数据增强处理; 并优化设计YOLOv5模型结构, 通过引入GAM注意力机制和Transformer模块, 增强模型特征提取能力, 特征金字塔网络采用BiFPN结构及CARAFE模块补充上下文信息; 使用EIoU损失函数和Soft NMS边界框筛选算法, 提高边界框的定位精度, 解决目标遮挡漏检问题。结果显示, 改进后网络的平均精度均值达到96.6%, 较原网络提高3.5个百分点, 在GPU上对单张蓝莓授粉图像的平均检测时间为11.4 ms。研究表明, 本研究建立的模型的识别准确度、检测速度及鲁棒性满足对蓝莓的蜜蜂授粉次数的实时监测。

**关键词** 机器视觉; 改进泊松融合; YOLOv5; 注意力机制; Transformer; CARAFE

**中图分类号** TP391 **文献标识码** A **文章编号** 1000-2421(2023)03-0105-10

蓝莓花期授粉是温室蓝莓种植中至关重要的一个环节, 借助蜜蜂授粉不仅均匀而且效率高, 还可提高蓝莓的坐果率和果实品质, 解决温室蓝莓在相对隔离条件下的授粉难题。蜜蜂的投放量直接影响授粉效果。目前蓝莓花期授粉次数统计大多数依靠人工, 效率低且精度不高。因此, 采用先进的信息技术手段, 记录每朵花上蜜蜂的访花行为, 实现日光温室蓝莓开花窗口期下的蜜蜂授粉次数统计, 具有重要的现实意义。

现有目标检测算法主要包括基于图像特征的目标检测和基于深度学习的目标检测方法<sup>[1]</sup>。传统的机器学习在昆虫检测中分为图像预处理、特征提取以及识别分类3个步骤<sup>[2]</sup>, 图像预处理和特征提取部分主要使用图像处理技术, 使用滤波、去噪以及图像增强技术对图像进行预处理, 再提取昆虫的轮廓、纹理等特征<sup>[3-4]</sup>, 提取的特征通过人工设计进行特征融合用来训练分类器, 最终完成昆虫的分类任务。李文勇<sup>[5]</sup>采用图像采集系统与图像预处理方法和基于分割点定位的粘连害虫分割算法, 构建一套在线果园虫害监测系统和设备, 结果表明对多种害虫的识别率均在90%以上。Qin等<sup>[6]</sup>提出光谱残差显著性检测方法用于储藏粮食害虫的边缘检测, 实现害虫

位置的准确检测。白云飞等<sup>[7]</sup>使用改进的形态学二次重建方法结合模糊推理的方法识别蜂王, 在2种不同环境条件下该方法的识别成功率分别达到88.8%和84.6%。

近年来随着人工智能技术的发展, 以卷积神经网络为主干特征提取网络的目标检测算法在图像识别领域得到广泛应用<sup>[8]</sup>, 如Lei等<sup>[9]</sup>使用引入Swin Transformer为主干网络的改进YOLOv5进行水下生物探测, 改进后的模型优于一般的目标检测模型, 在复杂的水下场景中具有更强的鲁棒性。张小花等<sup>[10]</sup>建立了基于EfficientDet-D1的草莓快速检测及分类方法, 该方法具有EfficientNet网络中快速归一化特征加权融合特点, 在成熟草莓与未成熟草莓的检测性能均优于其他目标检测算法。李瑞<sup>[11]</sup>针对田间聚集性小目标害虫识别精度低、鲁棒性差的问题, 提出一种基于Coarse-to-Fine网络的小目标害虫识别与计数方法, 所提出的检测网络模型准确率与鲁棒性均得到有效提升。

为确保蓝莓授粉样本采集的可操作性与可验证性, 本研究设计一套端到端数据采集设备, 远程采集蓝莓温室大棚内蜜蜂授粉样本图片, 并结合改进的YOLOv5进行检测识别, 通过设计主干网络、颈部网

收稿日期: 2022-09-12

基金项目: 国家自然科学基金青年基金项目(61601076); 大连市科技创新基金项目(2020JJ26SN058, 2021JJ13SN78)

胡玲艳, E-mail: hulingyan@dlu.edu.cn

通信作者: 汪祖民, E-mail: wangzumin@dlu.edu.cn

络结构,从而融合各个尺度的特征,增强网络的特征表示能力,进一步结合注意力机制、Transformer等模块,将全局信息和关键特征进行有效融合,旨在实现模型准确率、检测速度以及鲁棒性的综合平衡,为快速精确识别蓝莓蜜蜂授粉目标提供新的研究思路。

## 1 材料与方 法

### 1.1 数据采 集

本研究的数据来源于大连市金州区与庄河市蓝莓大棚种植基地,针对普通相机采集数据耗时耗力的问题,设计一套数据采集设备,该设备主要包括3部分:(1)iDS-2DC4223IW-/GLT(S5)型号摄像头,用于采集蓝莓授粉样本图像;(2)各类生长环境参数传感器,用于采集大棚内的各项环境参数数据;(3)USR-G805插卡式4GWIFI版工业路由器、电源模块和铁制支架部分,用于将数据传输至云端服务器。该数据采集设备实时收集大棚内蓝莓生长环境参数和蓝莓授粉样本图片,通过5G通信模块上传到云端服务器。并使用EOS 550D数码相机进行人工图像采集,进一步对数据集进行扩充,提高数据集丰富性。试验选取2个大棚,共放置3套数据采集设备,设置摄像头抓拍周期为10 s,图像分辨率为1 920像素×1 080像素,图像为JPG格式,数据采集设备采集图像1 929张,EOS 550D数码相机拍摄图像240张,图像分辨率为6 000像素×4 000像素,图像为JPG格式。经过人工数据清洗,删除不包含蓝莓授粉样本的图像,保留2 100张有效图像,使用LabelImg软件对图像进行人工标注,构建蓝莓授粉样本图像数据集,数据采集设备以及数据集的部分样本如图1所示。



A: 数据采集设备 Data acquisition equipment; B, C: 金州区样本 Sample of Jinzhou District; D: 庄河市样本 Sample of Zhuanghe.

图1 数据采集设备以及部分数据集的样本示例

Fig.1 Part samples of the datasets and data acquisition equipment

### 1.2 改进泊松融合数据增强方法

针对固定拍摄设备采集不到足够数量的蜜蜂访花行为图片,线上样本采集维度受限等问题,本研究提出一种蓝莓授粉样本图像的改进泊松融合数据增强方法,改进泊松融合数据增强方法结构如图2所示,主要包括3个步骤:(1)选取部分图像裁剪出蜜蜂图像,建立蜜蜂数据集,使用颜色迁移算法<sup>[12]</sup>对蜜蜂图像进行处理,实现蜜蜂图像与蓝莓花朵图像色调的统一。(2)随机选取若干颜色迁移后的蜜蜂图像分别进行高斯模糊、运动模糊处理,模拟摄像头抓拍过程中产生的对焦模糊、运动模糊现象;随机选取若干蓝莓花朵图片进行对比度增强处理,模拟大棚内的强光、暗光时的亮度状态。(3)使用改进的泊松融合算法将蜜蜂图像融入到蓝莓花朵图像上,生成新的数据集。

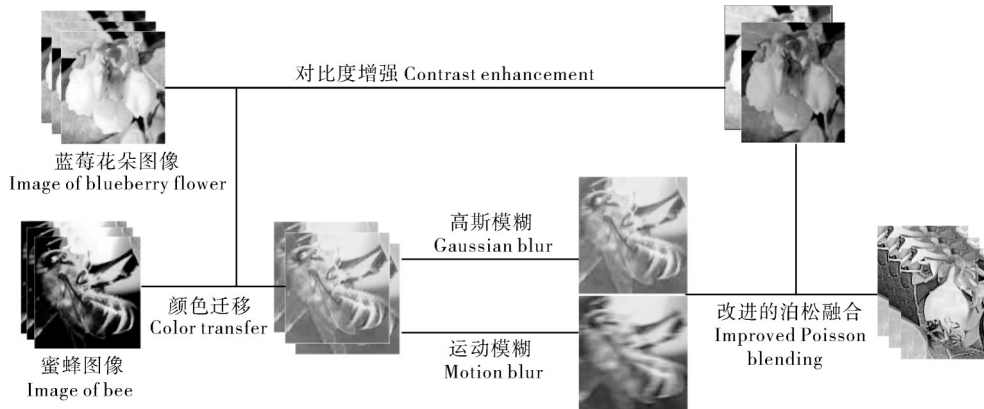


图2 改进泊松融合数据增强方法结构

Fig.2 The structure of improved Poisson blending data enhancement method

1)改进的泊松融合算法。泊松融合算法<sup>[13]</sup>作为一种图像融合方法,能够将目标图像较好地融合到原图像中,泊松融合算法的计算公式如下:

$$\min_f \iint_{\Omega} |\nabla f - v|^2 \text{ and } f|_{\partial\Omega} = f^*|_{\partial\Omega} \quad (1)$$

式(1), $f$ 表示融合后的结果图像, $f^*$ 表示蓝莓花朵图像, $v$ 表示蜜蜂图像的梯度, $\nabla f$ 表示 $f$ 的一阶梯度, $\Omega$ 表示要融合的区域, $\partial\Omega$ 代表融合区域的边缘部分,即蜜蜂花朵图像边缘不变的情况下,求融合部分的图像 $C$ ,使得 $C$ 在融合部分的梯度与蜜蜂花朵图像在融合部分的梯度最为接近。

2)泊松融合算法具体步骤。①通过差分方法,计算蜜蜂图像与蓝莓花朵图像的梯度场。②蜜蜂图像梯度场做Mask操作,提取待融合区域的梯度场。

③蜜蜂待融合区域与蓝莓花朵图像的梯度场进行融合,得到合成图像的梯度场。④对合成图像的梯度场求偏导,计算合成图像的散度。⑤通过散度构建系数矩阵求解约束方程。

传统的泊松融合算法能较好地将蜜蜂图像融入到蜜蜂花朵图像中,但存在融合区域边界消失的问题,通过改进步骤②,首先获取蜜蜂Mask图像,对Mask图像进行闭运算处理得到Triamp图<sup>[14]</sup>,Triamp图经KNN算法处理得到Alpha图<sup>[15]</sup>,使用Alpha图对蜜蜂图像梯度场做Mask操作,获得更多的融合边界语义信息,改进后的泊松融合算法步骤如图3所示。

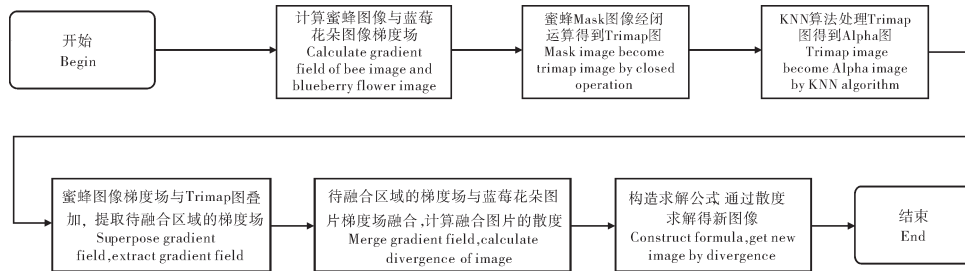


图3 改进泊松融合算法步骤图

Fig.3 Step of improved Poisson blending algorithm

最终得到融合后的蓝莓授粉样本图片如图4所示,其中红框部分为融合区域。

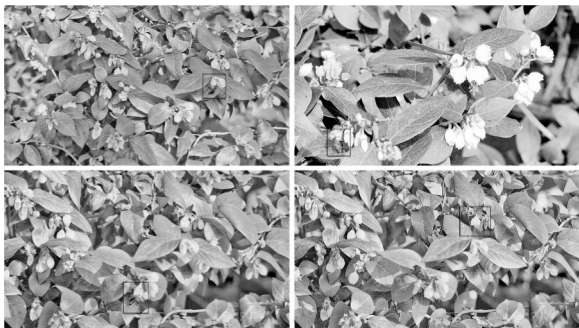


图4 融合后的蓝莓授粉样本

Fig.4 Samples of blueberry pollination after fusion

### 1.3 YOLOv5网络改进设计

YOLOv5网络由4个主要部分组成:输入端、Backbone、Neck、Prediction Head。图像输入网络后,Backbone主干网络在不同的图像粒度上聚集并形成图像特征,Neck颈部网络进行图像多尺度特征融合并传输到Prediction Head,Prediction Head对图像特征进行预测,生成边界框和预测类别。其中,Backbone采用CSPDarknet结构进行特征提取,Neck采用FPN+PAN<sup>[16]</sup>结构实现多尺度特征融合。

1)主干网络的优化。①Bottleneck结构的优化蜜蜂尺寸较小或远距离拍摄导致蜜蜂在图像中的占比较小,并且由于池化层的存在,会使蜜蜂信息量进一步减少,导致深层特征对小目标的表达能力较弱。本研究通过改进C3模块中的Bottleneck结构,在Bottleneck结构中加入GAM注意力模块形成GAMBottleneck结构,引入GAMBottleneck结构的C3模块称之为GAMC3模块,在减少信息扩散的情况下能放大全局跨维度的交互特征,GAMC3结构如图5所示。

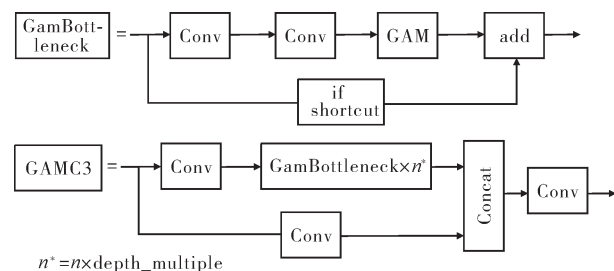


图5 GAMC3模块结构图

Fig.5 GAMC3 module structure

GAM注意力<sup>[17]</sup>模块采用CBAM注意力中的顺序通道-空间注意力机制,对模块内部的通道与空间注



注意力子模块进行重新设计。GAM注意力模块结构如图6所示,输入特征图首先经过通道注意力子模块处理,通道注意力子模块使用1个2层的MLP多层感知器放大跨维通道-空间依赖性,特征图进一步由空间注意力子模块处理,使用2个 $7\times 7$ 卷积层进行

空间信息融合,同时删除最大池化层,在增强特征信息保留的同时有效减少网络的计算复杂度和参数量。GAM注意力模块在减少信息扩散的情况下放大全局交互特征,提高各个特征在通道和空间上的联系,实现对蜜蜂关键特征的精准提取。

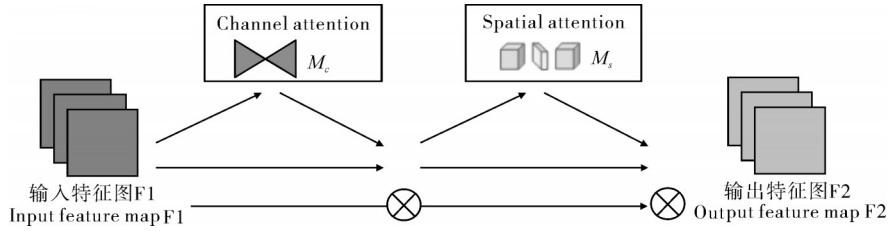


图6 GAM注意力模块结构

Fig. 6 Structure of GAM Attention

②MHSA模块的引入。本研究引入BoTNet<sup>[18]</sup>中的Multi-Head Self-Attention替换Bottleneck结构内的 $3\times 3$ 卷积层构建MSBottleneck模块,引入MS-Bottleneck结构的C3模块称之为MSC3模块,本研究将Multi-Head Self-Attention缩写为MHSA模块。卷积运算可以有效地捕获信息,但这种操作具有局部性,为了捕获全局信息,将MSC3模块添加至主干网络最后1层,有效提取图像基本特征和高层语义信息,实现全局特征信息的整合,MSC3结构图如图7所示。

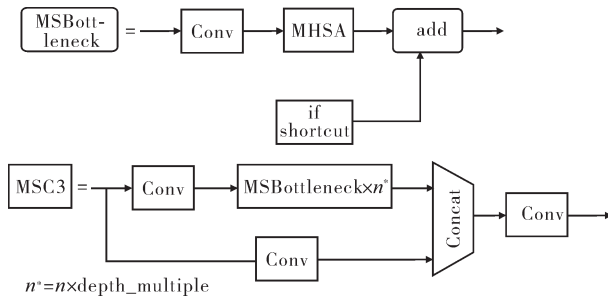


图7 MSC3模块结构图

Fig. 7 MSC3 module structure

2)颈部网络的优化设计。为了进一步加强对小目标关键特征的提取,本研究采用权重分配的BiFPN<sup>[19]</sup>结构替换FPN+PAN结构,BiFPN在FPN+PAN的基础上进行改进,对不同特征的重要程度进行评估,获取权重值,对各个尺度乘以相应的权重,强化网络对小目标关键特征的表达。

FPN+PAN结构中采用卷积方式进行上采样,对于特征图每个位置采用相同的上采样核,不能很好地捕捉到特征图的语义信息,另外引入了大量参数和计算量,本研究采用CARAFE<sup>[20]</sup>上采样模块,其上采样核与特征图的语义信息相关,并具有较大的感受野,能更好地捕获小目标的特征信息,其次模

块轻量化,可以有效减少网络的计算复杂度和参数量,进而提高网络训练速度。

3)Prediction Heads的优化。为了进一步提高模型对中小目标的检测能力,捕获全局信息和丰富的上下文信息。本研究使用Transformer Prediction Heads<sup>[21]</sup>中的Transformer encoder模块与Prediction Heads中C3模块相结合构建C3TR模块,改进后的Prediction Heads是由低层、高分辨率的特征图生成,对微小物体的检测性能得到了很大的提高。由于网络架构的第18层与第21层的输出特征图负责预测小目标型物体,故对网络架构的第18层与第21层的Prediction Heads进行优化。通过上述对网络模型结构的优化设计,改进的网络结构可以更好地适应复杂环境下的蜜蜂授粉样本检测,同时提高蜜蜂授粉样本检测的准确率,改进的YOLOv5网络结构如图8所示。

4)损失函数和NMS的改进。YOLOv5网络中使用GIoU作为边界框的回归损失函数,如式(2)所示,其中IoU表示预测框与真实框的交并比, $A$ 表示真实框, $B$ 表示预测框, $C$ 为 $A$ 与 $B$ 并集区域的最小外接矩阵。

$$\text{GIoU}_{\text{loss}} = \text{IoU} - \frac{|C - (A \cup B)|}{|C|} \quad (2)$$

本研究采用EIoU<sup>[22]</sup>损失函数作为边界框的回归损失函数。EIoU基于GIoU进行完善,不仅考虑了中心点距离和长宽比,还考虑了真实框与预测框的宽度和高度的真实差异。其公式定义如式(3)所示。

$$L_{\text{EIoU}} = L_{\text{IoU}} + L_{\text{dis}} + L_{\text{asp}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{C^2} + \frac{\rho^2(\omega, \omega^{\text{gt}})}{C_{\omega}^2} + \frac{\rho^2(h, h^{\text{gt}})}{C_h^2} \quad (3)$$

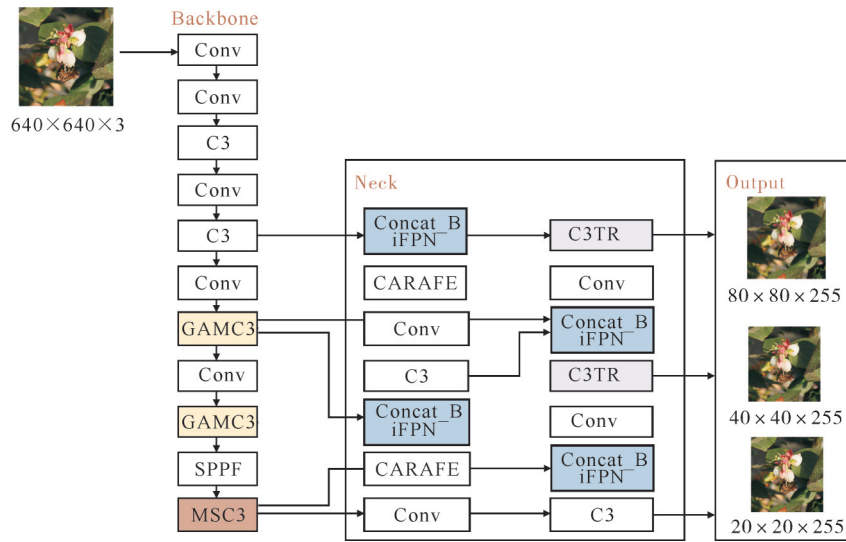


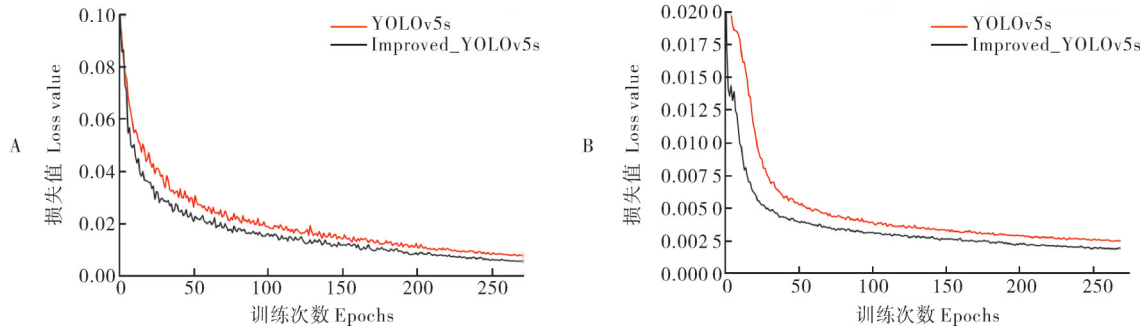
图 8 改进的 YOLOv5 网络结构

Fig.8 Structure of improved YOLOv5

其中,  $C_w$  和  $C_h$  是真实框与预测框的最小外接矩阵的宽度和高度,  $b$  和  $b'$  表示预测框和真实框的中心点,  $\rho$  表示欧氏距离,  $C$  表示真实框与预测框的最小外接矩阵的对角线长度。相比于其他目标框回归损失函数, EIoU 使模型回归框定位更加准确, 并加速了模型的收敛性, 损失值的变化曲线如图 9 所示, 从图 9 可以看出, 采用 EIoU 的网络模型收敛效果更好, 模型经过 100 次迭代后得到收敛。

NMS 算法利用得分高的边框抑制得分低且重叠

度高的边框从而去除冗余的检测框, 从而获得正确的检测结果, 但当 2 个目标距离较近且存在部分重叠时, 置信度较小的目标漏检的可能性较大。针对此问题, 使用 Soft NMS<sup>[23]</sup> 对预测框进行筛选处理, 针对上述 NMS 中 IoU 大于阈值的边框, 不是直接令其得分为 0, 而是采用高斯置信度降低策略降低这些边框的得分, 进而保留得分较少的预测框成为正确检测框的机会, 有效改善蜜蜂被遮挡而漏检的情况。



A: 定位损失 Localization loss; B: 置信度损失 Confidence loss.

图 9 损失值变化曲线

Fig.9 Loss value curve

### 1.4 试验环境与参数配置

本试验的模型训练的试验环境为: Windows10 操作系统、Intel Xeon W-2235 处理器、Quadro RTX 4000 显卡、编程语言为 Python 3.8、深度学习模型采用 Pytorch1.10, 采用 CUDA10.2 支持 GPU 加速。使用 YOLOv5s.pt 预训练权重进行迁移学习, 将数据集根据 4:1 的比例划分为训练集和验证集, 采用 SGD 随机梯度下降进行网络参数的梯度优化, 对模型参

数进行初始化。设置训练 epoch 为 300, 初始学习率为 0.01, batch size 为 4, 动量为 0.937, 前 3 轮使用 Warm up 稳定前期训练参数, 设定图片输入尺寸为 640 像素 × 640 像素。

### 1.5 模型评价指标

为了评估改进的 YOLOv5s 网络的检测性能, 本试验采用精确率 (precision,  $P$ )、召回率 (recall,  $R$ )、平均精度均值 ( $P_{m\Delta}$ )、蜜蜂授粉样本检测率、检测速度

5个评价指标衡量改进的网络模型的检测性能,其中准确率、召回率的公式定义如式(4)~(5):

$$P = \frac{T_p}{F_p + T_p} \times 100\% \quad (4)$$

$$R = \frac{T_p}{F_n + T_p} \times 100\% \quad (5)$$

其中,  $T_p$ (true positive): 正确检验框, 表示已标定的预测框与标签框匹配。  $F_p$ (false positive): 误检框, 表示将背景预测成为物体的框。  $F_n$ (false negative): 漏检框, 表示原本需要检测到物体的框, 没有检测出来。

精确率( $P$ )为模型预测的所有目标中, 预测正确的比例, 有利于突出结果的相关性。查全率又称召回率( $R$ ), 指所有的正样本目标中, 预测正确的目标比例。为了结合这2个指标, 引入了平均精度( $P_A$ )来测量检测精度,  $P_{mA}$ 为各类别  $P_A$  的平均值, 如式(6)。本研究为单类别目标检测, 即  $P_{mA}$  等于  $P_A$ 。

$$P_{mA} = P_A = \int_0^1 P(R) d(R) \times 100\% \quad (6)$$

本研究进一步引入蓝莓授粉样本检测率与检测速度作为模型的评价指标, 使用训练好的模型对选取的200张数据集图片进行检测, 其中蜜蜂授粉样本图像179张, 蜜蜂图像21张, 实际蜜蜂授粉样本数量189个, 统计识别正确的检测框数量、识别错误的检测框数量, 计算得到样本检测率与误检率, 衡量模型的实际检测效果。检测速度为模型在GPU上对单张图像测试时所需的推理时间。

## 2 结果与分析

### 2.1 与不同YOLOv5模型的比较

YOLOv5模型可根据深度和宽度分为4种不同的模型: YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。YOLOv5s作为最轻的模型, 包含最少的参数, 便于在实际应用场景中部署。其他网络模型都是在YOLOv5s的基础上不断加深、加宽, 理论上随着网络深度和宽度的增加, 网络特征提取和特征融合的能力也在不断增强, 但检测能力越强, 模型参数数量和模型大小也呈线性增加, 消耗了大量的计算资源。为了减少计算资源的消耗, 本研究选择YOLOv5s作为改进模型, 在不过多增加网络参数的前提下, 实现对YOLOv5其他模型的性能超越, 保证模型的轻量化以及高性能。由表1可知, 随着网络深度和宽度的增加, YOLOv5m、YOLOv5l 在YOLOv5s基础上平均精度均值分别提升2.2、3.0个百分点,

YOLOv5x更适用于极小目标的检测, 因此, 在本研究数据集上的表现不佳, 相较检测性能最好的YOLOv5l模型, 本研究所采用的改进YOLOv5s的平均精度均值提升0.5个百分点, 模型参数约为其20%, 模型更轻量化, 样本检测率相比其他模型更高、误检率更高, 检测性能较优。

表1 不同YOLOv5模型检测性能对比

Table 1 Contrast experiment of different YOLOv5 models

模型 Models	检测率/% Detection rate	误检率/% False detection rate	模型参数 Number of parameters	平均精度 均值/% $P_{mA}$
YOLOv5s	87.5	14	$7.2 \times 10^6$	93.1
YOLOv5m	88.7	4	$20.8 \times 10^6$	95.3
YOLOv5l	92.5	9	$46.3 \times 10^6$	96.1
YOLOv5x	88.7	14	$86.2 \times 10^6$	94.8
改进的YOLOv5s Improved-YOLOv5s	95.5	4	$8.7 \times 10^6$	96.6

### 2.2 消融实验

本研究通过消融实验直观地观察不同的改进方法对YOLOv5s网络模型检测性能的优化作用, 结果如表2所示。首先对主干网络进行优化, 在CSP结构中加入GAM注意力, 利用更多浅层信息来检测小目标同时增强检测效果, 同时在主干网络最后一层加入MHSA模块, 捕获全局的信息, 主干网络的优化使模型平均精度均值提升了2.1个百分点, 对主干网络最后一层特征图进行可视化, 原主干网络所提取的有效特征较为分散、不完整, 优化后的主干网络提取的有效特征相对集中与完整, 特征可视化对比图如图10所示。进一步对颈部网络进行设计, 将原先的FPN+PAN结构替换为BiFPN结构, 并采用CARAFE上采样模块替代Upsample, 使模型有效平均精度均值提升0.3个百分点, 同时在Prediction Heads

表2 消融实验结果

Table 2 Ablation experiment result

主干网络的优化 Optimization of Backbone	颈部网络的设计 Design of neck	Prediction Heads的优化 Optimization of Prediction Heads	损失函数、NMS的改进 Improvement of loss function and NMS	平均精度 均值/% $P_{mA}$
				93.1
✓				95.2(+2.1)
✓	✓			95.5(+0.3)
✓	✓	✓		96.4(+0.9)
✓	✓	✓	✓	96.6(+0.2)



中引入 Transformer 模块,提高小目标的检测效果,有效提升平均精度均值 0.9 个百分点,最后对损失函数和 NMS 进行改进,采用 EIoU 替代 GIoU,使模型回归框定位更加准确,采用 Soft NMS 替代原 NMS 算法,改善目标被遮挡的问题,最终改进的

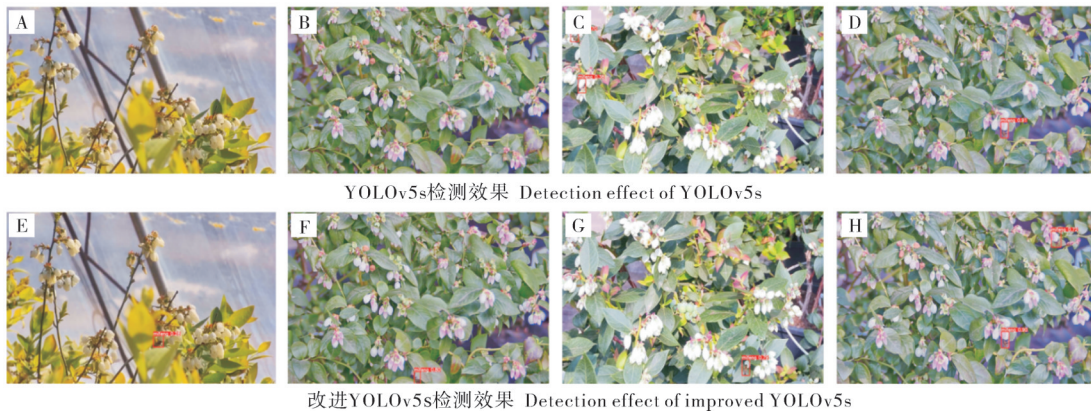
YOLOv5s 模型平均精度均值达到 96.6%,相较于原 YOLOv5s 模型提升 3.5 个百分点,改进前后的检测效果对比如图 11 所示,有效提高了复杂环境下被遮挡目标的检测能力,同时进一步降低了误检率。



A:原图 Original figure; B:原主干网络 Original Backbone; C:优化后的主干网络 Optimized Backbone.

图 10 特征可视化对比图

Fig.10 Comparison of feature visualization



改进YOLOv5s检测效果 Detection effect of improved YOLOv5s

A, E: 第一组遮挡目标检测效果对比 Comparison of detection effect of occluded targets in group 1; B, F: 第二组遮挡目标检测效果对比 Comparison of detection effect of occluded targets in group 2; C, G: 第三组遮挡目标检测效果对比 Comparison of detection effect of occluded targets in group 3; D, H: 误检情况下检测效果对比 Comparison of detection effect under false detection condition.

图 11 改进前后的检测效果对比

Fig.11 Comparison of detection effect of before and after improvement

为了进一步验证本研究改进算法的鲁棒性,对图像分别进行运动模糊、伽马变化处理,分析模型在弱光、遮挡模糊情况下的检测效果,具体检测效果如图 12 所示。随着图片模糊程度与对比度增加,原模型检测准确率有所降低,出现漏检、误检情况。而对于改进的 YOLOv5s 模型,在存在干扰的环境中,仍然具备良好的检测性能,准确率虽有所降低但在可接受范围内,证明该模型具有较好的鲁棒性。

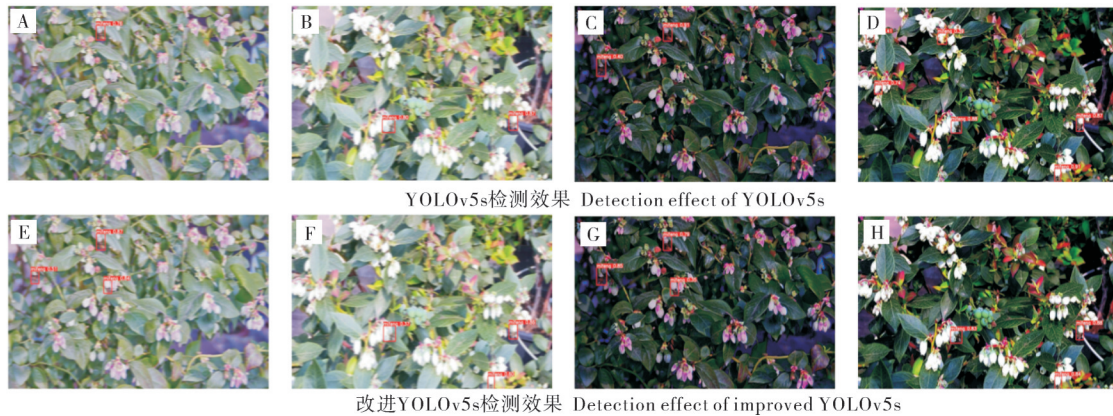
### 2.3 与不同目标检测算法的比较

将改进的 YOLOv5s 算法与 YOLOv4、Faster RCNN、SSD-MobilenetV3、YOLOX-S、EfficientDet-D1 等主流的目标检测网络部署于同一服务器端并选用相同的试验环境与数据增强算法进行对比试验(表 3)所示。由表 3 可知,与其他模型相比,改进

YOLOv5s 模型的平均精度均值最高,较 SSD-MobilenetV3、Faster RCNN 分别提升 13.3、9.2 个百分点,比 YOLOv4、YOLOX-S、EfficientDet-D1 分别提高 7.5、2.9 和 4.2 个百分点。在检测速度方面,改进 YOLOv5s 检测速度远快于二阶段检测模型 Faster RCNN,对比一阶段检测模型 YOLOv4、YOLOX-S,改进 YOLOv5s 检测速度更具优势,与 2 种轻量级网络 EfficientDet-D1 和 SSD-MobilenetV3 相比,三者识别速度相当,但改进 YOLOv5s 的平均精度均值更高,检测性能更好,表明该方法在复杂的种植环境检测具有明显的应用优势。

### 2.4 与不同数据增强算法的比较

将本研究的数据增强方法分别与图像 hsv 增强(方案 1)、图像随机角度变化增强(方案 2)、图像 mo-



A,E:第一组运动模糊后的检测效果对比 Comparison of detection effect after motion blur in group 1; B,F:第二组运动模糊后的检测效果对比 Comparison of detection effect after motion blur in group 2; C,G:第一组弱光条件下的检测效果对比 Comparison of detection effect under low light condition in group 1; D,H:第二组弱光条件下的检测效果对比 Comparison of detection effect under low light condition in group 2.

图12 不同场景下的检测效果对比

Fig.12 Comparison of detection effect in different scenarios

表3 不同目标检测算法的比较

Table 3 Contrast experiment of different object detection algorithm

模型 Model	主干网络 Backbone	精确率/% P	召回率/% R	平均精度均值/% P <sub>mA</sub>	推理时间/ms Reason time
SSD	MobilenetV3	87.6	81.2	83.3	17.5
YOLOv4	CSPDarknet53	90.5	88.4	89.1	26.1
Faster RCNN	ResNet50	86.8	90.4	87.4	42.7
YOLOX-S	Darknet53	95.6	88.7	93.7	13.3
EfficientDet-D1	EfficientNet	94.1	92.4	92.8	14.2
改进的YOLOv5s Improved-YOLOv5s	Improved-CSPDarknet53	97.8	93.6	96.6	11.4

saic 增强(方案3)3种传统数据增强方法进行对比(表4),以上3种数据增强方法均采用YOLOv5s模型进行测试,采用同样的训练参数以及相同的验证集。由表4可知,方案1较未使用数据增强方法的P<sub>mA</sub>降低2.1个百分点,通过改变图像hsv通道生成新图像,本质上改变图像的颜色特征,提高了数据集的泛化性,但检测精度有所降低。方案2与方案3都是在原基础图像上进行扩充,精度有所上升,但增强得到的

表4 不同数据增强方法的检测效果

Table 4 Contrast experiment of different data enhancement method

方法 Method	精确率 Precision	召回率 Recall	平均精度均值 P <sub>mA</sub> %
未数据增强 Without data enhancement	94.4	85.0	90.3
方案1 Plan 1	97.0	80.0	88.2
方案2 Plan 2	92.5	92.3	93.1
方案3 Plan 3	94.9	92.5	92.5
本研究方法 The study	97.3	91.1	93.6

图像的特征信息与原图相比变化不大,模型训练时易过拟合。与其他3种方法相比,本研究所提方法的P<sub>mA</sub>较方案1、方案2及方案3分别提升5.4、0.5、1.1个百分点,与未使用数据增强方法相比提升3.3个百分点。考虑到蜜蜂访花行为相对具有规律性,本研究算法与其他增强方法相比,将蜜蜂与花朵背景进行改进后的泊松融合达到数据增强的目的,进一步提升了蓝莓授粉样本的检测精度与鲁棒性。

### 3 讨论

本研究针对蓝莓授粉样本检测任务,首先提出一种改进泊松融合数据增强方法,通过蜜蜂数据集与蓝莓花朵数据集以改进的泊松融合方式生成蜜蜂授粉样本数据集,有效解决了样本价值性不高、模型训练样本不足的问题,提高了模型小目标的检测精度与鲁棒性。其次针对待检测图像环境背景复杂、蜜蜂目标较小、被叶片遮挡等问题,提出1种基于改进YOLOv5s的蓝莓授粉样本检测方法,将注意力机制GAM与MHSA模块与原始YOLOv5s主干网络融合,提高模型网络对特征信息的提取能力;引入



BiFPN结构与Transformer模块,进一步强化模型网络对小目标关键特征的表达;采用EIoU损失函数计算预测框回归损失提高定位精度,并使用Soft NMS改善遮挡等复杂情况下漏检以及小目标物体识别效果差的问题。

本研究采用数据增强方法处理后的图像数据集,将网络模型的准确率提高3.3个百分点,与其他传统数据增强方法相比,进一步提高了网络的检测准确性和增强鲁棒性,说明改进泊松融合数据增强方法扩充的数据集更具有价值性,与神经网络的学习相配合可起到正向增强作用。所提出的改进YOLOv5s模型平均精度均值达到96.6%,对单张蓝莓授粉图片的检测时间为11.4 ms,能够提取到更关键更细致的目标特征,同时对小目标有较强的检测能力,在平均识别准确率以及检测速度上明显优于SSD<sup>[24]</sup>、Faster RCNN、YOLOX<sup>[25]</sup>、EfficientDet-D1等现有模型,在复杂的场景中具有更强的鲁棒性,并能完成蓝莓授粉次数统计功能。

在前期工作基础上,结合精准农业的发展蓝图以及现有的技术背景,本研究模型的计算复杂度还可进一步优化与提升,后续对模型进行轻量化处理,且在不损失准确性的情况下加速推理将是未来研究的重点,并进一步将模型部署在嵌入式设备上,最终完成真实环境下的检测任务,为蓝莓授粉统计终端应用提供有力的技术支撑。

## 参考文献 References

- [1] 周中奎. 基于机器学习的智能汽车目标检测与场景增强技术研究[D]. 重庆: 重庆邮电大学, 2020. ZHOU Z K. Research on machine learning based object detection and augmented reality technology for intelligent vehicle[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2020 (in Chinese with English abstract).
- [2] 吕宗旺, 金会芳, 甄彤, 等. 图像处理技术在粮食害虫识别中的应用进展[J]. 河南工业大学学报(自然科学版), 2021, 42(3): 128-137. LÜ Z W, JIN H F, ZHEN T, et al. Application development of image processing technologies in grain pests identification [J]. Journal of Henan University of Technology (natural science edition), 2021, 42(3): 128-137 (in Chinese with English abstract).
- [3] 王贵财, 张德贤, 李保利, 等. 粮虫视觉检测技术的现状与展望[J]. 中国粮油学报, 2014, 29(4): 124-128. WANG G C, ZHANG D X, LI B L, et al. Present situation and prospects of storage pests based on vision inspection technology[J]. Journal of the Chinese cereals and oils association, 2014, 29(4): 124-128 (in Chinese with English abstract).
- [4] 张卫芳, 郭敏. 基于图割理论的储粮害虫图像分割[J]. 科学技术与工程, 2010, 10(7): 1661-1664, 1679. ZHANG W F, GUO M. Stored grain insect image segmentation method based on graph cuts[J]. Science technology and engineering, 2010, 10(7): 1661-1664, 1679 (in Chinese with English abstract).
- [5] 李文勇. 基于机器视觉的果园性诱害虫在线识别与计数方法研究[D]. 北京: 中国农业大学, 2015. LI W Y. On-line identification and counting of sex-pheromones lured orchard pest based on machine vision[D]. Beijing: China Agricultural University, 2015 (in Chinese with English abstract).
- [6] QIN Y, WU Y L, WANG Q F, et al. Method for pests detecting in stored grain based on spectral residual saliency edge detection [J]. Grain & oil science and technology, 2019, 2(2): 33-38.
- [7] 白云飞, 张翔, 林建. 基于改进的形态学二次重建蜜蜂王模糊识别方法[J]. 计算机应用与软件, 2020, 37(11): 139-145. BAI Y F, ZHANG X, LIN J. An improved morphology-based secondary reconstruction method for queen bee recognition[J]. Computer applications and software, 2020, 37(11): 139-145 (in Chinese with English abstract).
- [8] 杨万里, 段凌凤, 杨万能. 基于深度学习的水稻表型特征提取和穗质量预测研究[J]. 华中农业大学学报, 2021, 40(1): 227-235. YANG W L, DUAN L F, YANG W N. Deep learning-based extraction of rice phenotypic characteristics and prediction of rice panicle weight [J]. Journal of Huazhong Agricultural University, 2021, 40(1): 227-235 (in Chinese with English abstract).
- [9] LEI F, TANG F F, LI S H. Underwater target detection algorithm based on improved YOLOv5[J/OL]. Journal of marine science and engineering, 2022, 10(3): 310 [2022-09-12]. <https://doi.org/10.3390/jmse10030310>.
- [10] 张小花, 李浩林, 李善军, 等. 基于EfficientDet-D1的草莓快速检测及分类[J]. 华中农业大学学报, 2022, 41(6): 262-269. ZHANG X H, LI H L, LI S J, et al. Rapid detection and classification of strawberries based on EfficientDet-D1 [J]. Journal of Huazhong Agricultural University, 2022, 41(6): 262-269 (in Chinese with English abstract).
- [11] 李瑞. 小目标害虫图像自动识别与计数研究[D]. 合肥: 中国科学技术大学, 2021. LI R. Research on automatic tiny pest recognition and counting[D]. Hefei: University of Science and Technology of China, 2021 (in Chinese with English abstract).
- [12] 尉迟姝毅. 基于反向映射的图像间颜色迁移算法仿真[J]. 计算机仿真, 2021, 38(1): 212-216. YUCHI S Y. Simulation of color transfer algorithms between images based on reverse mapping [J]. Computer simulation, 2021, 38(1): 212-216 (in Chinese with English abstract).
- [13] 刘佳, 刘孝保, 阴艳超, 等. 面向不均衡样本空间的工件表面缺陷检测方法[J]. 机械科学与技术, 2022, 41(5): 755-763. LIU J, LIU X B, YIN Y C, et al. Surface defect detection method of workpiece for unbalanced sample space [J]. Mechanical science and technology for aerospace engineering, 2022, 41(5): 755-763 (in Chinese with English abstract).
- [14] CHEN G W, LIU Y, WANG J, et al. PP-matting: high-accuracy natural image matting[DB/OL]. arXiv, 2022: 2204.09433 [2022-09-12]. <https://arxiv.org/abs/2204.09433>.
- [15] CHEN Q F, LI D, TANG C K. KNN matting[J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 35(9): 2175-2188.
- [16] ZHAO Y, SHI Y B, WANG Z L. The improved YOLOV5 algorithm and its application in small target detection[M]. [S.l.]: In-

- telligent Robotics and Applications. Cham; Springer International Publishing, 2022: 679-688.
- [17] LIU Y C, SHAO Z R, HOFFMANN N. Global attention mechanism: retain information to enhance channel-spatial interactions [DB/OL]. arXiv, 2021; 2112.05561 [2022-09-12]. <https://arxiv.org/abs/2112.05561>.
- [18] SRINIVAS A, LIN T Y, PARMAR N, et al. Bottleneck transformers for visual recognition [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. Nashville: IEEE, 2021: 16514-16524.
- [19] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. Seattle: IEEE, 2020: 10778-10787.
- [20] YING J, LI H, YANG H M, et al. GPK-YOLOv5s: content-aware reassembly of features and self attention for high altitude parabolic detection and tracking [J/OL]. MATEC web of conferences, 2022, 363: 01012 [2022-09-12]. <https://doi.org/10.1051/mateconf/202236301012>.
- [21] ZHU X K, LYU S C, WANG X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), October 11-17, 2021, Montreal, BC, Canada. Montreal: IEEE, 2021: 2778-2788.
- [22] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [DB/OL]. arXiv, 2021: 2101.08158 [2022-09-12]. <https://arxiv.org/abs/2101.08158>.
- [23] LIN J R, YANG C M, LU Y, et al. An improved soft-YOLOX for garbage quantity identification [J/OL]. Mathematics, 2022, 10(15): 2650 [2022-09-12]. <https://doi.org/10.3390/math10152650>.
- [24] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector [M]. Cham; Springer International Publishing, 2016: 21-37.
- [25] GE Z, LIU S, WANG F, et al. Yolox: Exceeding yolo series in 2021 [DB/OL]. arXiv, 2021; 2107.08430 [2022-09-26]. <https://arxiv.org/abs/2107.08430>.

## Machine vision-based monitoring honeybee pollination of blueberry in greenhouse

HU Lingyan<sup>1</sup>, SUN Hao<sup>1</sup>, XU Guohui<sup>1</sup>, GUO Ruiya<sup>2</sup>, GUO Zhanjun<sup>2</sup>,  
CHEN Pengyu<sup>1</sup>, PEI Yuekun<sup>1</sup>, WANG Zumin<sup>1</sup>

1. College of Information Engineering, Dalian University, Dalian 116622, China;

2. Dalian Modern Agricultural Production Development Service Center, Dalian 116000, China

**Abstract** The statistics of pollination times of honeybee was evaluated based on machine vision to evaluate the rationality of the dosage of pollinating honeybee during the window period of blueberry blooming in the solar greenhouse. The dataset was processed with the method of improved poisson blending data enhancement to solve problems that the detection environment is complex, the target scale is small, and it is easy to be covered. The structure of YOLOv5 was optimized. The detection precision of model was improved by introducing GAM attention mechanism and Transformer module. BiFPN and CARAFE were introduced in feature pyramid network to complement the contextual information. EIou\_loss and Soft NMS were used to enhance the positioning precision of bounding box and solve the problem of detecting target occlusion. The results showed that the mean average precision of the improved YOLOv5 was 96.6%, 3.5 percentage points higher than that of the original algorithm. The detection time of a single blueberry pollination image on the GPU was 11.4 ms.

**Keywords** machine vision; improved Poisson blending; YOLOv5; attention mechanism; Transformer; CARAFE

(责任编辑:陆文昌)