

# 自然环境下基于改进 YOLOv7 的梨花识别方法

张秀花, 魏华杰, 孔德刚, 刘尚坤, 黄 征, 王洪森

(1. 河北农业大学机电工程学院, 保定 071001; 2. 河北省智慧农业装备技术创新中心, 保定 071001)

**摘 要:** 针对自然环境下梨花易被遮挡、背景杂乱、光照条件与目标距离不断变化等特点导致梨花识别难和精度不高的问题, 该研究提出了一种基于改进 YOLOv7 模型的梨花识别算法。该算法首先加入 P2 小目标层, 增加了特征提取与模型多尺度融合能力, 使被遮挡的梨花目标更好地被捕获; 其次, 在输出检测端末尾加入 CBAM (convolutional block attention module) 注意力机制模块, 提高模型的上下文理解能力, 提升 YOLOv7 在各种场景下 (不同光照条件、复杂背景等) 的表现; 最后, 将 CIoU (complete intersection over union) 损失函数优化为 NWD (normalized weighted distance) 损失函数, 针对不同形状的目标进行精确的边界框回归, 提高模型对复杂背景梨花目标与远距离梨花目标的检测精度。试验结果表明: 改进模型与原模型相比, 精确率、召回率、mAP 和 F1-score 分别提高了 2.1、1.2、1.9 和 0.6 个百分点, 达到了 99.4%、99.6%、96.4% 和 89.8%; 与其他主流算法相比, 各评价指标均有优势。研究结果可为梨园自然环境下梨花精准识别提供支撑。

**关键词:** 梨花; 图像识别; YOLOv7; 自然环境; 小目标层; CBAM; 损失函数

doi: 10.11975/j.issn.1002-6819.202408209

中图分类号: S24; TP391.4

文献标志码: A

文章编号: 1002-6819(2025)-02-0224-09

张秀花, 魏华杰, 孔德刚, 等. 自然环境下基于改进 YOLOv7 的梨花识别方法[J]. 农业工程学报, 2025, 41(2): 224-232.

doi: 10.11975/j.issn.1002-6819.202408209 <http://www.tcsae.org>

ZHANG Xiuhua, WEI Huajie, KONG Degang, et al. Recognizing pear blossom in the natural environment using improved YOLOv7[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(2): 224-232. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202408209 <http://www.tcsae.org>

## 0 引 言

随着农业向智能化和精准化转型, 花朵识别技术的应用愈发突出, 花朵识别对于花量估计<sup>[1]</sup>、花期判断<sup>[2]</sup>与智能机械化疏花<sup>[3]</sup>等具有重要意义。但在自然环境下, 现有识别方法易受遮挡、背景、光照与距离等因素影响, 梨花精准识别有待深入研究。

目前, 花朵目标识别已有很多研究<sup>[4-6]</sup>, 夏烨等<sup>[7]</sup>使用一种融合 Ghost 卷积块与 Bifpn 改进的 YOLOv5s 算法, 对棚架梨园中梨花与花苞进行识别与分类, 检测速度较快但精度不高, 花朵的检测召回率与精确率分别为 85.3% 与 89.4%; 周桂红等<sup>[8]</sup>基于改进的 YOLOv5s 对梨花进行识别并进行了密度分析, 在原模型中添加了小目标层和 CBAM 注意力机制, 改进模型相对于原模型精确率、召回率等均有所提升, 且改进模型在多个品种上表现出较好的适应性, 但其对远近距离、不同背景与不同光照等不同梨花识别场景没有进行分类与识别试验, 其精确率与召回率分别为 91.62% 与 83.05%, 仍有提高空间; 余贤海等<sup>[9]</sup>改进了 YOLOv5s 算法, 添加 CBAM 注意力模块并使用加权框融合方法检测番茄花朵, 对不同花期以及不同花蕊姿态朝向的番茄花朵建立识别模型, 两种模型总体识别准确率较高; 王小荣等<sup>[10]</sup>在 YOLOv7

模型的基础上, 引入 Swin Transformer 注意力机制并对损失函数进行了改进, 使其能够在遮挡、光照以及密集干扰因素等情况下对红花进行有效识别并提高了检测速度, 但其算法的准确率还有待加强; LI 等<sup>[11]</sup>将 YOLOv5 应用于猕猴桃花朵检测, 采集多种类猕猴桃花朵数据集并进行迁移学习, 提高模型检测精度, 该模型可准确识别不同类别的猕猴桃花朵及获取其花朵分布; WU 等<sup>[12]</sup>使用通道剪枝技术对 YOLOv4 模型进行剪枝, 简化了苹果花检测模型并提高了检测效率, 实现了自然环境下苹果花快速准确检测, 但仅对不同光照条件苹果花进行了分析; SHANG 等<sup>[13]</sup>提出了一种基于轻量级的 YOLOv5s 算法的苹果花检测方法, 减少了模型占用空间大小, 提高了检测速度, 适用于不同花朵颜色、不同天气条件与不同光照条件下的苹果花快速实时检测, 但模型的准确率、召回率不高, 分别为 88.4% 与 86.1%。

上述研究表明, 采用深度学习模型可有效识别花朵, 但梨花与其他品种有所不同, 一般花朵开放先于叶片生长, 个别品种花叶同期生长, 致使识别时的环境更加复杂, 因此梨花精准识别有其自身的特点和难度。本研究基于 YOLOv7 模型进行改进, 以解决自然环境下梨花识别过程中的易被遮挡、背景杂乱、光照条件与目标距离不断变化的难题, 有效提升识别效果。

## 1 材料与方法

### 1.1 图像采集

本研究根据实际识别过程中识别装置与目标远近距

收稿日期: 2024-08-28 修订日期: 2024-12-02

基金项目: 河北省现代农业产业技术体系建设专项资金项目 (HBCT2024170207)

作者简介: 张秀花, 教授, 研究方向为现代果园智能化装备研发。

Email: zhang72xh@163.com

离变化的特点,将拍摄装置与梨花之间的距离划分为近距离(20~40 cm)、远距离(80~100 cm)两种方式,拍摄目标分别为单簇梨花、多簇梨花与远距离一片梨花。现代梨园一般为南北行栽植,实际识别数据采集作业时,识别装置在行间行驶,梨花与相机相对方位不断变化,选取典型的8个方向(东、东北、北、西北、西、西南、南、东南)进行梨花图片拍摄,贴合实际作业过程中梨花与相机的相对方位关系,如图1所示,其中包括正面、侧面、背面、顺光、逆光、遮挡等情况,还拍摄不同背景(梨花、地面、天空与各种背景互相结合)梨花图片,还原梨园自然环境下梨花真实场景。

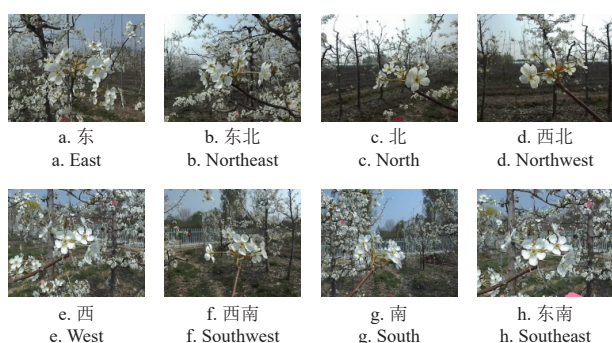


图1 不同拍照方向下的梨花图片

Fig.1 Pear blossom pictures in different directions

拍摄时间为2024年4月8—12日,正处于梨花盛开时期,花苞均已绽放。拍摄地点为河北农业大学教学试验农场与雄安新区容城县高宇农业产业园。使用手持可移动设备进行拍摄,拍摄图片以3648×2736像素分辨率保存为.jpg文件。拍摄目标为九年生鸭梨梨花、九年生黄金梨花与五年生新梨七号梨花,各品种花朵如图2所示。鸭梨花开单层,花瓣4~6片,花朵体型较大;新梨七号花开单层,花瓣4~6片,花朵体型较小,分布密集;黄金花开单层或多层,花瓣5~11片,盛花期伴有叶片生长,花朵体型较大。



图2 不同品种梨花图像

Fig.2 Pear blossom images of different varieties

剔除模糊与无法使用照片后,剩余鸭梨梨花照片3240张,黄金梨花照片1582张,新梨七号梨花照片2184张。由于鸭梨样本数量较多,且拍摄背景元素丰富,使用鸭梨作为主要研究对象进行自然环境下梨花识别,其余两个品种作为对照,以检验该改进算法的效果。

## 1.2 数据集构建

对采集到的原始梨花图像使用标注工具 Labellmg 进行人工标注,标签名称为 pearb,使用最小外接边框最大程度排除无关像素干扰。将采集到的3240张鸭梨梨花

图像进行数据标注并生成.xml格式文件,文件中含有标注目标检测框的中心坐标及其高度与宽度,最后是标注目标标签类别,将数据集按照比例为7:2:1分配为训练集、验证集和测试集,每个数据集都含有标注过的梨花标签信息。

## 1.3 试验平台

本研究采用 Intel core i7-12700KF 的 Win11 操作系统,主频为3.6 GHz,运行内存为32 G,搭载 Nvidia GeForce RTX 3070 显卡。试验在 Pycharm 平台上进行操作,配置 PyTorch 深度学习框架进行环境搭建,运用 Python 语言进行算法编写,所有试验均在相同环境下进行。

## 2 梨花识别算法及改进

### 2.1 YOLOv7 网络结构

本研究所采用的 YOLOv7 算法框架主要由输入端(Input)、骨干网络(Backbone)、融合网络(Neck)和输出检测头(Head)4部分组成<sup>[14]</sup>。其中 Input 模块包括自适应的动态锚框计算、自适应的图像调整和 Mosaic 数据增强功能<sup>[15]</sup>; Backbone 主要负责特征提取,从原始输入图像中提取出多层次的特征信息,包括 CBS 模块、ELAN 网络和 MP 模块, CBS 模块由 CONV 卷积层、Batch Normlization 批量归一化与 SiLU 激活函数组成, ELAN 通过利用多层次的特征信息提高模型的表现, MP 模块包括 CBS 模块和 MaxPOOL 模块; Neck 主要负责融合特征, YOLOv7 使用了典型的 FPN<sup>[16]</sup>+PAN<sup>[17]</sup> 的结构, FPN (feature pyramid network) 通过一条自上而下的路径聚合,实现高层语义信息指导低层位置信息, PAN (path aggregation network) 在 FPN 的基础上额外增加了一条自下而上的路径,实现低层位置信息与高层语义信息的进一步融合; Head 是目标检测网络的最后部分,主要负责将从 Backbone 和 Neck 获得的特征图转化为最终的检测结果,包括类别、边界框坐标和置信度等。

### 2.2 模型改进

为了解决原始 YOLOv7 网络模型在自然环境下梨花识别精度不高的问题,本文提出了改进 YOLOv7 检测模型。在模型中首先加入 P2 小目标层,增加了特征提取与模型多尺度融合能力,使被遮挡的梨花目标更好地被捕获;其次,在输出检测端末尾加入 CBAM 注意力机制模块,提高模型的上下文理解能力,提升 YOLOv7 在各种场景下(不同光照条件、复杂背景等)的表现;最后,将 CIOU 损失函数优化为 NWD 损失函数,针对不同形状的目标进行精确的边界框回归,提高模型对复杂背景梨花目标与远距离梨花目标的检测精度。本文提出的改进 YOLOv7 模型结构如图3所示。

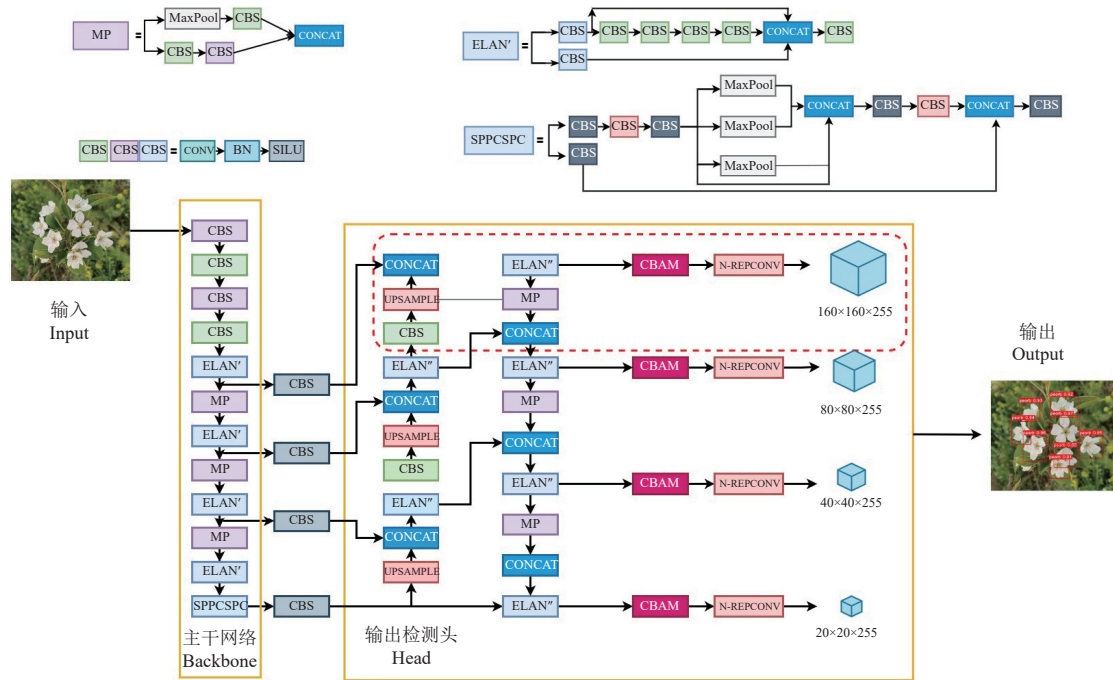
#### 2.2.1 增加小目标层

由于小目标的尺寸较小, YOLOv7 对小目标的探测能力较差,处于复杂环境中时,常规的模型识别精度较低。原始 YOLOv7 算法中仅存在3个检测层,分别输出特征图为 8×8×255、16×16×255 与 32×32×255 像素的目



标。在本文中,远距离梨花目标较小,有些距离较远的梨花尺寸甚至小于 10 个像素,目标在经过 YOLOv7 系

列处理后,所输出特征图尺寸更小,而原 P3 检测层对较小的特征图检测能力较差。



注: CBAM 为添加了 CBAM 注意力机制的卷积块;虚线圈出范围为所添加小目标层;N-RECONV 为优化了 NWD 损失函数的输出卷积块。

Note: CBAM is a convolutional block with added CBAM attention mechanism; The dotted circle indicates the range of the added small target layer; N-RECONV optimizes the output convolution block of the NWD loss function.

图 3 改进 YOLOv7 模型结构图  
Fig.3 Improved YOLOv7 model structure diagram

为改善模型对于较小梨花的检测能力,本研究增加了小目标检测层 P2 以更快速的检测小目标,当输入图像为 640×640 像素时,最终通过 P2 检测层输出的特征图为 160×160 像素,与其他检测层输出的特征图相比具有更小的感受野,并且增加了一组初始锚框,初始锚框由聚类算法自动生成。这个检测层虽然会提高计算时的数据量与内存使用,但增加了特征提取与模型多尺度融合能力,使改进模型在梨花部分被遮挡的情况下识别出梨花的特征,尤其是在目标有不同部分被遮挡而整体轮廓仍然可见时有更好效果,同时对小目标梨花的检测能力有很大的提高。

### 2.2.2 CBAM 注意力机制

注意力机制是模仿人脑的算法,可以把已有的计算能力集中于重要部分<sup>[18]</sup>。在真实环境与复杂的光照条件下拍摄,会产生背景全为梨花或背景全为天空的情况,背景颜色与目标梨花颜色非常相似,存在背景干扰与边缘干扰。此外,远距离密集的梨花分布使得识别目标更小且遮挡更加严重,增加了检测的难度。原 YOLOv7 模型对密集梨花与不同距离下梨花经常出现漏检与错检情况。

本研究使用 WOO 等<sup>[19]</sup>提出的 CBAM 注意力机制通过捕捉梨花图像边缘部位的特征以及空间与通道之间的联系,借助 CBAM 注意力机制减少边缘干扰与背景干扰。CBAM 详细结构如图 4 所示。

经过 CBAM 处理后,新的特征图将通过通道模块和空间模块的注意力加权,从而增强特征在通道和空间维

度上的关联性,通道注意力模块与空间注意力模块的处理过程如式(1)、(2)所示:

$$F^* = M_C(F) \otimes F \quad (1)$$

$$F^{**} = M_S(F^*) \otimes F^* \quad (2)$$

式中  $F$  为输入特征图;  $F^*$  为通道注意力输出;  $F^{**}$  为空间注意力输出;  $\otimes$  为加权;  $M_C$ 、 $M_S$  为通道和空间注意力。

引入 CBAM 注意力机制可以提高模型的上下文理解能力,在不显著增加模型占用空间的前提下,提升 YOLOv7 在各种场景下(不同光照条件、复杂背景等)的表现,提高检测性能。

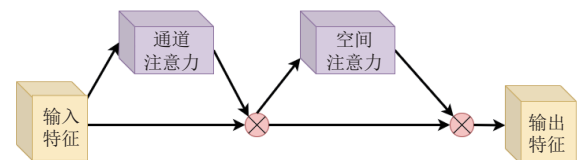


图 4 CBAM 结构  
Fig.4 CBAM structure

### 2.2.3 NWD 损失函数优化

本文针对的是自然环境下现代梨园梨花目标的检测问题, YOLOv7 所使用的 CIoU 损失函数难以应对梨花小目标检测带来特征信息不足的限制性,且梨花目标为花瓣形状,而 YOLOv7 的边界框为标准矩形,因此,梨花的边界框内会携带一些背景的信息,而自然环境下背景杂乱,梨花边界与背景相似,对损失函数进行改进优

化, 引入 NWD 损失函数<sup>[20]</sup>, 对生成和真实边界框分布的相似性进行评估, 可以针对不同形状的目标进行更精确的边界框回归, 更精准的反应梨花目标的几何特征, 减少漏检率。

NWD 函数使用一种新的计算边界框与边界框之间的相似度计算方法来代替传统 IoU 计算方法。使用二维高斯分布对边界框建模<sup>[21]</sup>, 并使用 Optimal Transport (OT) 理论中的 Wasserstein distance 进行归一化处理<sup>[22]</sup>, 该过程分为两步进行:

1) 对边界框进行二维高斯建模

水平边界框  $K = (c_x, c_y, w, h)$ , 其中  $(c_x, c_y)$ 、 $w$  和  $h$  分别表示中心坐标、边界框的宽度和边界框的高度。其内切椭圆的方程如式 (3) 所示:

$$\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} = 1 \quad (3)$$

式中  $(\mu_x, \mu_y)$  为椭圆的中心坐标;  $\sigma_x$  为沿  $x$  轴的半轴长度;  $\sigma_y$  为沿  $y$  轴的半轴长度。

当  $(X-\mu)^T \Sigma^{-1} (X-\mu) = 1$  时, 二维高斯分布的密度函数如式 (4) 所示:

$$f(X|\mu, \Sigma) = \frac{\exp(-\frac{1}{2}(X-\mu)^T \Sigma^{-1} (X-\mu))}{2\pi|\Sigma|^{\frac{1}{2}}} \quad (4)$$

式中  $X$  为高斯分布的坐标  $(x, y)$ ;  $\mu$  为均值向量;  $\Sigma$  为协方差矩阵。

水平边界框可以服从  $N(\mu, \Sigma)$  的二维高斯分布, 而  $\mu$  与  $\Sigma$  的值可以由式 (5)、(6) 给定:

$$\mu = \begin{bmatrix} C_x \\ C_y \end{bmatrix} \quad (5)$$

$$\Sigma = \begin{bmatrix} \frac{w^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix} \quad (6)$$

2) 使用 Wasserstein 计算真实边界框与预测边界框之间的相似度。

对于真实边界框  $K_1 = (c_{x1}, c_{y1}, w_1, h_1)$  和预测边界框  $K_2 = (c_{x2}, c_{y2}, w_2, h_2)$ , 其高斯分布可以表示为  $N_1$  和  $N_2$ 。Wasserstein 距离用于计算这两个高斯分布之间的相似度, 如式 (7) 所示:

$$W_2^2(N_1, N_2) = \left\| \begin{bmatrix} c_{x1}, c_{y1}, \frac{w_1}{2}, \frac{h_1}{2} \end{bmatrix}^T - \begin{bmatrix} c_{x2}, c_{y2}, \frac{w_2}{2}, \frac{h_2}{2} \end{bmatrix}^T \right\|_2^2 \quad (7)$$

由于  $W_2^2(N_1, N_2)$  是距离度量, 需要通过指数归一化来计算其相似度度量, 如式 (8) 所示:

$$NWD(N_1, N_2) = \exp\left(-\frac{\sqrt{W_2^2(N_1, N_2)}}{S}\right) \quad (8)$$

式中  $S$  为与数据集密切相关的常数, 根据训练数据集所有真实样本分布与生成样本分布之间的最大可能的

Wasserstein 距离作为  $S$ 。

而以 NWD 指标作为损失函数, 具体表达如式 (9) 所示:

$$L_{NWD} = 1 - NWD(N_1, N_2) \quad (9)$$

式中  $L_{NWD}$  为以 NWD 函数对传统 IoU 进行优化后的损失函数。

### 2.3 模型评价指标与训练参数

为了评价模型性能, 使用精确率  $P$  (precision)、召回率  $R$  (recall)、F1-score、平均精度均值 mAP (mean average precision) 和模型大小 (model size) 进行评价。

训练网络时设置学习率为 0.01, 动量因子系数为 0.937, 权重衰减参数为 0.0005。网络训练周期为 300, 输入图像大小为 3648×2736 像素, 在对改进 YOLOv7 模型及其他模型进行训练时, 超参数均保持一致, 并且均不加载预训练模型。

## 3 结果与分析

### 3.1 梨花图像检测结果

近距离检测结果如图 5 所示, 在逆光情况下, YOLOv7 模型和改进 YOLOv7 模型均可精确识别近距离梨花数量与位置信息, 但当存在枝叶遮挡以及梨花互相遮挡时, YOLOv7 的检测框会出现漏检现象, 黄色方框所标注的梨花被枝干或梨花遮挡, 出现漏检, 而改进 YOLOv7 模型可全部检测出来。

同时, 如图 5b 原 YOLOv7 模型检测结果所示, 6 朵花从左到右检测结果的置信度依次为 0.91、0.95、0.82、0.83、0.84 和 0.82; 如图 5b 改进 YOLOv7 模型检测结果所示, 相同花朵从左到右检测结果的置信度依次为 0.96、0.98、0.92、0.91、0.86 和 0.85。检测结果表明, 改进 YOLOv7 模型与原 YOLOv7 模型相比, 改进 YOLOv7 模型在近距离梨花检测准确度与置信度方面均有更好表现。

复杂背景检测结果如图 6 所示, 近距离梨花在大地、天空、梨花与天空与梨花背景下 YOLOv7 模型与改进模型均可准确识别, 在大地与梨花背景相接处, YOLOv7 产生了漏检, 而改进 YOLOv7 可检测出所有的梨花; 且改进 YOLOv7 模型比在不同背景的梨花识别结果中具有较高置信度, 如在大地背景中, YOLOv7 模型检测置信度从左到右依次为 0.93、0.90、0.92、0.97、0.93、0.91、0.82 和 0.92, 改进 YOLOv7 模型检测相同花朵置信度从左到右依次为 0.95、0.93、0.98、0.98、0.95、0.98、0.94 和 0.93。检测结果表明, 改进 YOLOv7 同样表现出更好的检测效果与更高的置信度。表 1 为 YOLOv7 模型和改进 YOLOv7 模型应用于大地背景、天空背景和梨花背景测试集的具体效果。

如表 1 所示, 改进 YOLOv7 的精确率、召回率、mAP 在大地、天空和梨花 3 种背景下均比 YOLOv7 模型高, 在测试大地数据集时, 改进 YOLOv7 的 mAP 较 YOLOv7 高 1.8 个百分点; 在测试天空数据集时, 改进 YOLOv7 的 mAP 较 YOLOv7 高 1.4 个百分点; 在测试梨



花数据集时,改进 YOLOv7 的 mAP 较 YOLOv7 高 1.5 个百分点。



注: 黄框为漏检梨花, 红框为识别到的梨花, 下同。  
Note: The yellow box represents missed pear blossoms, and the red box represents recognized pear blossoms, the same below.

图 5 近距离梨花图像检测结果  
Fig.5 Detection results of close-up pear blossom images

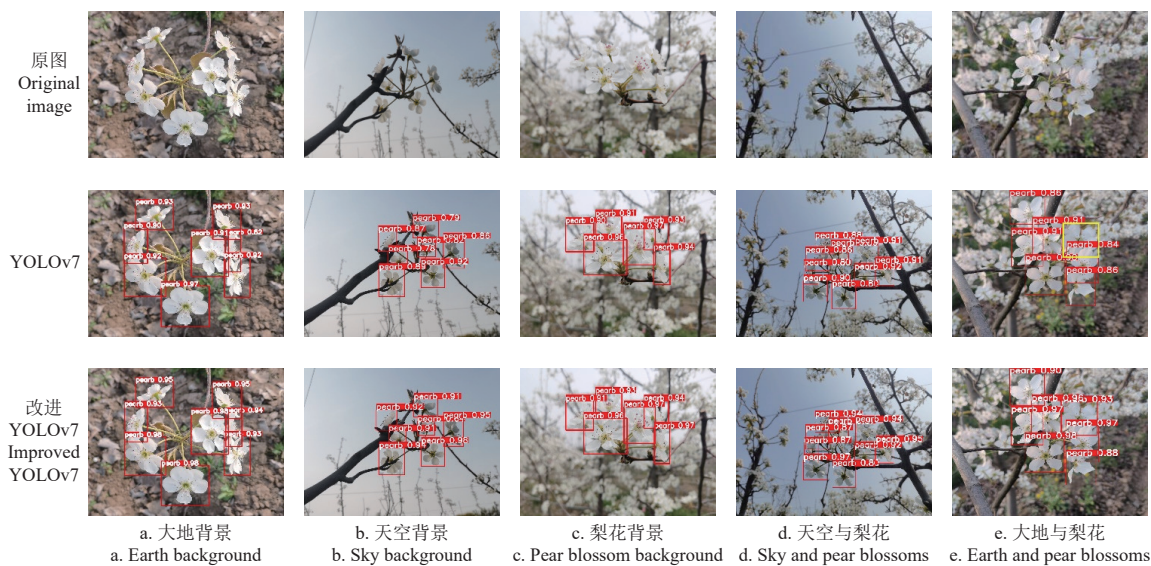


图 6 复杂背景梨花图像检测结果  
Fig.6 Detection results of pear blossom images with complex backgrounds

表 1 YOLOv7 和改进 YOLOv7 在 3 种背景下的检测结果 Table 1 Detection results of YOLOv7 and improved YOLOv7 under three different backgrounds				
背景 Background	模型 Model	精确率 Precision <i>P</i>	召回率 Recall <i>R</i>	平均精度均值 Mean average precision mAP
大地 Ground	YOLOv7	93.4	92.6	96.4
	改进 YOLOv7	94.7	95.4	98.2
天空 Sky	YOLOv7	93.5	92.7	96.3
	改进 YOLOv7	96.7	94.1	97.7
梨花 Pear blossom	YOLOv7	94.5	93.8	95.7
	改进 YOLOv7	96.4	98.7	97.2

远距离检测结果如图 7 所示, 由于梨花目标较小且

分布密集, 存在花朵互相遮挡与树枝遮挡, YOLOv7 模型在顺光、密集分布情况下产生了较多漏检, 而改进 YOLOv7 模型能检测出漏检的梨花。

由上述结果可知, 改进 YOLOv7 模型具有更高的置信度与检测准确度, 并能对不同遮挡情况、不同背景、不同光照条件和不同距离下的梨花目标进行精准检测, 证明该模型可有效的检测自然环境下盛花期梨花且检测精度较高。

### 3.2 多模块消融试验结果

为验证改进 YOLOv7 模型对梨花检测效果的提升,

通过消融试验逐步在 YOLOv7 原始模型中加入相应模块来对比和评价每个模块的性能，试验结果如表 2 所示。

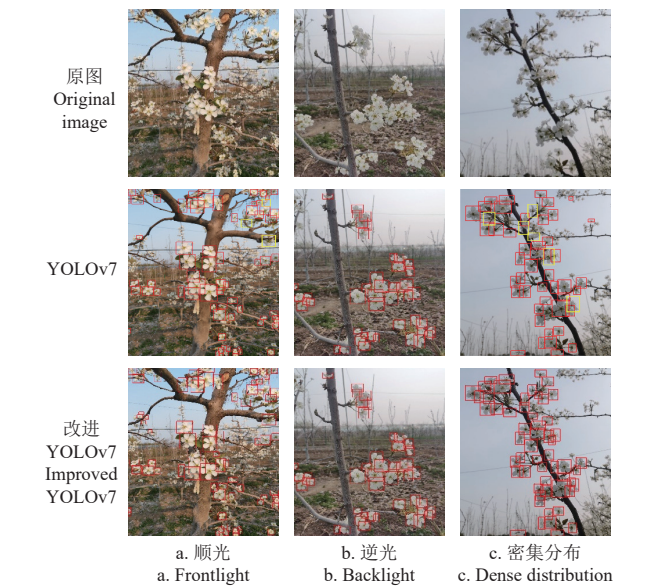


图 7 远距离梨花图像检测结果  
Fig.7 Detection results of distant pear blossom images

表 2 YOLOv7 消融试验结果					
Table 2 YOLOv7 ablation experiment results					
Small-object layer	CBAM	NWD	P	R	mAP
×	×	×	97.3	98.4	94.5
√	×	×	96.9	99.2	95.5
√	√	×	98.1	97.4	95.6
√	×	√	99.4	99.2	95.9
√	√	√	99.4	99.6	96.4

注：√为在 YOLOv7 中添加了对应模块。  
Note: √ indicates the addition of corresponding modules in YOLOv7.

表 2 中，加入小目标层后，召回率与 mAP 分别提升 0.8 与 1.0 个百分点，但精确率降低了 0.4 个百分点，说明模型对小目标梨花样本的检测中表现良好，加入了 CBAM 注意力机制，将算力集中在了重要特征的提取，精确率和 mAP 分别增长了 0.8 和 1.1 个百分点，使用 NWD 优化损失函数后，精确率、mAP 均有提升，但召回率由于已经处于较高水平，并未提升，说明使用相似度计算 IoU 方法提高了模型对梨花数据的泛化能力与检测精度，而改进 YOLOv7 在精确率、召回率和 mAP 上达到了 99.4%、99.6% 和 96.4%，相对于原 YOLOv7 分别增加 2.1、1.2 和 1.9 个百分点。在梨花目标检测的精度方面取得了较好的提升。

3.3 梨花图像检测热力图分析

本文采用基于梯度的网络可视化方法 Grad-CAM (gradient-weighted class activation mapping) [23] 生成目标检测热力图，如图 8 所示。

如图 8 中所示为部分梨花图像的检测热力图，图中红色越深说明其对目标的检测结果贡献较大，蓝色越深说明其对目标的检测结果贡献较少，从图 8b 与图 8c 可知，相较于原始模型，改进 YOLOv7 模型的检测热力图热力值更接近于实际梨花区域，着重提取梨花边缘特征。

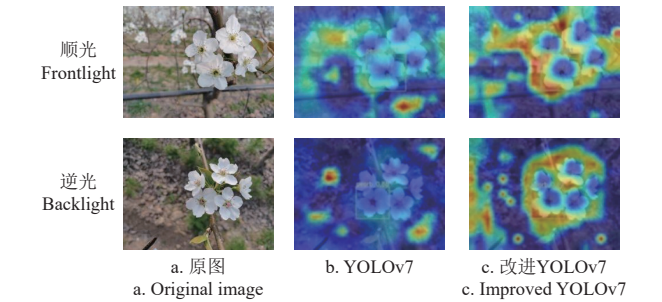


图 8 部分梨花图像检测热力图  
Fig.8 Partial pear blossom image detection heatmap

3.4 改进前后模型适用性性能试验

如表 3 所示，鸭梨梨花改进后模型精确率、召回率、mAP 和 F1-score 相较于改进前分别提升了 2.1、1.2、1.9 和 0.6 个百分点，新梨七号改进后精确率、召回率、F1-score 和 mAP 相较于改进前分别提升了 2.1、1.2、0.1 和 1.1 个百分点，黄金梨改进后精确率、召回率、F1-score 和 mAP 相较于改进前分别提升了 0.3、2.6、1.2 和 1.0 个百分点。试验结果说明，针对不同的品种的梨花识别，本文改进 YOLOv7 模型相较于原模型均有提升，证明了改进算法对不同品种梨花有较好的鲁棒性。

表 3 模型改进前后 3 种梨花检测结果					
Table 3 Three types of pear blossom detection results before					
品种	模型	P	R	F1-score	mAP
鸭梨	YOLOv7	97.3	98.4	89.2	94.5
Yali	改进 YOLOv7	99.4	99.6	89.8	96.4
新梨七号	YOLOv7	95.9	98.0	93.1	96.8
New pear No.7	改进 YOLOv7	98.0	99.2	93.2	97.9
黄金梨	YOLOv7	97.2	97.2	89.0	95.2
Gold pear	改进 YOLOv7	97.5	99.8	90.2	96.2

表 4 为 YOLOv7 模型和改进 YOLOv7 模型应用于近景梨花和远景梨花数据集的性能对比。由表 4 可知，改进 YOLOv7 的精确率、召回率与 mAP 在近景与远景两种场景下均比 YOLOv7 模型高，在训练近景数据集时，改进 YOLOv7 的 mAP 较 YOLOv7 高 3.9 个百分点；在训练远景数据集时，改进 YOLOv7 的 mAP 较 YOLOv7 高 3.7 个百分点。改进 YOLOv7 的精确率、召回率与 mAP 在顺光与逆光两种光照下均比 YOLOv7 模型高，在训练顺光数据集时，改进 YOLOv7 的 mAP 较 YOLOv7 高 4.4 个百分点；在训练逆光数据集时，改进 YOLOv7 的 mAP 较 YOLOv7 高 1.6 个百分点。

表 4 YOLOv7 和改进 YOLOv7 在不同条件下的性能指标

Table 4 Performance indicators of YOLOv7 and improved YOLOv7 under different conditions				
场景	模型	P	R	mAP
近景	YOLOv7	97.4	94.4	88.6
Close view	改进 YOLOv7	99.2	95.5	92.5
远景	YOLOv7	96.7	93.0	90.8
Distant view	改进 YOLOv7	99.4	94.1	94.5
顺光	YOLOv7	96.2	89.1	86.3
Frontlight	改进 YOLOv7	98.7	92.4	90.7
逆光	YOLOv7	99.4	96.2	91.0
Backlight	改进 YOLOv7	99.6	97.7	92.6



### 3.5 不同目标检测模型性能比较

为了评估改进 YOLOv7 网络对盛花期梨花的识别性能, 使用 SSD<sup>[24]</sup>、Faster R-CNN<sup>[25]</sup>、YOLOv3<sup>[26]</sup>、YOLOv4<sup>[27]</sup>、YOLOv5s、YOLOv7、YOLOv8、YOLOv9、YOLOv10 网络模型对鸭梨梨花数据集进行模型训练和性能评估。

由表 5 可知, 将改进模型与目前存在的主流检测算法进行对比试验, 改进后的 YOLOv7 模型在相同的硬件设备和环境下, mAP、召回率与 F1-score 均优于其他经典网络模型。改进 YOLOv7 同 SSD、Faster R-CNN 相比有明显优势。YOLOv5s 模型占用空间虽小, 为 13.7 MB, 但各项指标均低于改进后 YOLOv7。YOLOv8 虽取得了较高的精确率, 但 mAP、召回率与 F1-score 较低, 改进 YOLOv7 在精确率、召回率、mAP 与 F1-score 上达到了 99.4%、99.6%、96.4% 和 89.8%, 相对于原 YOLOv7 分别增加了 2.1、1.2、1.9 和 0.6 个百分点, 模型占用空间也与其相当, 改进 YOLOv7 模型与 YOLOv9、YOLOv10 相比, 精确率、召回率、mAP 与 F1-score 评价指标也均有优势, 证明了本文模型在梨花检测具有较高的准确性与良好的综合表现, 可以实现自然环境下盛花期梨花的准确识别。

表 5 不同目标检测模型性能

Table 5 Performance of different object detection models					
模型 Model	P/%	R/%	mAP/%	F1-score/%	模型大小 Model size/MB
SSD	74.3	91.6	78.5	61.6	90.6
Faster R-CNN	92.3	74.5	84.7	78.2	108.0
YOLOv3	83.1	90.1	88.9	76.2	235.0
YOLOv4	80.2	90.6	86.4	71.4	244.0
YOLOv5s	94.0	96.0	91.1	86.4	13.7
YOLOv7	97.3	98.4	94.5	89.2	71.3
YOLOv8l	92.1	93.2	89.3	83.4	84.5
YOLOv9	92.7	92.1	88.9	83.2	116.0
YOLOv10	95.7	92.4	87.4	82.4	52.2
Improved YOLOv7	99.4	99.6	96.4	89.8	87.6

## 4 结 论

针对梨花识别过程中, 因梨花易被遮挡、背景杂乱、光照条件差与目标距离变化等特点导致的梨花识别难和精度不高问题, 本文提出了改进 YOLOv7 检测模型, 主要结论如下:

1) 改进后的 YOLOv7 精确率、召回率和 mAP 分别为 99.4%、99.6% 和 96.4%, 相较于原 YOLOv7 分别提高 2.1、1.2 和 1.9 个百分点, 模型大小为 87.6 MB, 与原始 YOLOv7 网络模型相当。

2) 本文模型提升了不同距离、不同光照下和不同背景下梨花的检测精度, 减少了不同遮挡情况的梨花漏检情况, 在训练近景和远景数据集时, 改进 YOLOv7 的 mAP 较 YOLOv7 分别高 3.9 和 3.7 个百分点; 训练顺光和逆光数据集时, 改进 YOLOv7 的 mAP 较 YOLOv7 分别高 4.4 和 1.6 个百分点; 改进 YOLOv7 的 mAP 在大地、天空和梨花背景下比 YOLOv7 模型分别高 1.8、1.4 和 1.5 个百分点; 且通过对改进模型与原模型检测热力图结

果的检测分析, 证明改进模型能提取出更多的有效特征, 具有更强的鲁棒性, 提高了自然环境下梨花目标检测精度, 但在计算速度方面优势不突出, 后续可对模型轻量化与提高检测速度方面继续研究。

### [参 考 文 献]

- [1] 邓颖, 吴华瑞, 朱华吉. 基于实例分割的柑橘花朵识别及花量统计[J]. 农业工程学报, 2020, 36(7): 200-207.  
DENG Ying, WU Huarui, ZHU Huaji. Recognition and counting of citrus flowers based on instance segmentation[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(7): 200-207 (in Chinese with English abstract)
- [2] 高昂, 卢传兵, 任龙龙, 等. 基于改进 YOLOX-s 的苹果花生长状态检测方法验证分析[J]. 中国农机化学报, 2023, 44(8): 162-167.  
GAO Ang, LU Chuanbing, REN Longlong, et al. Detection method and validation analysis based on the improved YOLOX-s apple blossom growth state[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(8): 162-167. (in Chinese with English abstract)
- [3] 司永胜, 孔德浩, 王克俭, 等. 基于改进 YOLOv8-Seg 的苹果单枝条花序疏除方法[J]. 农业工程学报, 2024, 40(14): 100-108.  
SI Yongsheng, KONG Dehao, WANG Kejian, et al. Thinning apple inflorescence at single branch level using improved YOLOv8-Seg[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(14): 100-108 (in Chinese with English abstract)
- [4] 赵春江, 文朝武, 林森, 等. 基于级联卷积神经网络的番茄花期识别检测方法[J]. 农业工程学报, 2020, 36(24): 143-152.  
ZHAO Chunjiang, WEN Chaowu, LIN Sen, et al. Tomato florescence recognition and detection method based on cascaded neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(24): 143-152. (in Chinese with English abstract)
- [5] 尚钰莹, 张倩如, 宋怀波, 等. 基于 YOLOv5s 的深度学习在自然场景苹果花朵检测中的应用[J]. 农业工程学报, 2022, 38(9): 222-229.  
SHANG Yuying, ZHANG Qianru, SONG Huaibo, et al. Application of deep learning using YOLOv5s to apple flower detection in natural scenes[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(9): 222-229. (in Chinese with English abstract)
- [6] NGUYEN K T, DANG M, DOAN M T T, et al. Utilizing deep neural networks for chrysanthemum leaf and flower featurerecognition[J]. AgriEngineering, 2024, 6(2): 1133-1139.
- [7] 夏烨, 雷晓晖, 祁雁楠, 等. 基于改进 Ghost-YOLOv5s-BiFPN 算法检测梨树花序[J]. 智慧农业 (中英文), 2022, 4(3): 108-119.

- XIA Ye, LEI Gouhui, QI Yannan, et al. Detection of pear inflorescence based on improved Ghost-YOLOv5s-BiFPN algorithm[J]. *Smart Agriculture*, 2022, 4(3): 108-119. (in Chinese with English abstract)
- [8] 周桂红, 孙乐琳, 梁芳芳, 等. 基于改进密度峰值聚类算法的梨花密度分级[J]. *农业工程学报*, 2023, 39(1): 126-135.
- ZHOU Guihong, SUN Lelin, LIANG Fangfang, et al. Pear blossom density grading based on improved density peak clustering algorithm[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(1): 126-135. (in Chinese with English abstract)
- [9] 余贤海, 孔德义, 谢晓轩, 等. 基于深度学习的番茄授粉机器人目标识别与检测[J]. *农业工程学报*, 2022, 38(24): 129-137.
- YU Xianhai, KONG Deyi, XIE Xiaoxuan, et al. Deep learning-based target recognition and detection for tomato pollination robots[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(24): 129-137. (in Chinese with English abstract)
- [10] 王小荣, 许燕, 周建平, 等. 基于改进 YOLOv7 的复杂环境下红花采摘识别[J]. *农业工程学报*, 2023, 39(6): 169-176.
- WANG Xiaorong, XU Yan, ZHOU Jianping, et al. Safflower picking recognition in complex environments based on an improved YOLOv7[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(6): 169-176. (in Chinese with English abstract)
- [11] LI G, FU L, GAO C, et al. Multi-class detection of kiwifruit flower and its distribution identification in orchard based on YOLOv5l and Euclidean distance[J]. *Computers and Electronics in Agriculture*, 2022, 201, 107342.
- [12] WU D, LV S, JIANG M, et al. Using channel pruning -based YOLOv4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments [J]. *Computers and Electronics in Agriculture*, 2020, 178, 105742.
- [13] SHANG Y Y, XU X, JIAO Y, et al. Using lightweight deep learning algorithm for real-time detection of apple flowers in natural environments[J]. *Computers and Electronics in Agriculture*, 2023, 207, 107765.
- [14] WANG C, BOCHKOVSKIY A, LIAO H. YOLOv7: Trainable bag of freebies sets new state of the art for real-time object detectors[EB/OL]. (2020-07-06)[2022-05-10]<https://arxiv.org/abs/2207.02696>.
- [15] SONG Q, LI S, BAI Q, et al. Object detection method for grasping robot based on improved YOLOv5[J]. *Micromachines*, 2021, 12(11): 1273-1280.
- [16] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [17] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 99: 296-2969.
- [18] 李盼盼, 王朝立, 孙占全, 等. 基于注意力机制的多特征融合人脸活体检测[J]. *信息与控制*, 2021, 50(5): 631-640.
- LI Panpan, WANG Chaoli, SUN Zhanquan, et al. Face liveness detection based on multi-feature fusion with an attention mechanism[J]. *Information and Control*, 2021, 50(5): 631-640. (in Chinese with English abstract)
- [19] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//*Computer Vision – ECCV 2018*. Cham: Springer International Publishing, 2018: 3-19.
- [20] ZHOU X, JIANG L, GUAN X J, et al. Infrared small target detection algorithm with complex background based on YOLO-NWD[C]//*Proceedings of the 4th Inter-national Conference on Image Processing and Machine Vision*. Hong Kong, China: ACM, 2022: 6-12.
- [21] 张林鎡, 巴音塔娜, 曾庆松, 等. 基于 StyleGAN2-ADA 和改进 YOLOv7 的葡萄叶片早期病害检测方法[J]. *农业机械学报*, 2024, 55(1): 241-252.
- ZHANG Linxuan, BA Yintana, ZENG Qingsong, et al. Grape Disease Detection Method Based on StyleGAN2-ADA and Improved YOLO v7[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(1): 241-252. (in Chinese with English abstract)
- [22] WANG J, XU C, YANG W, et al. A normalized gaussian wasserstein distance for tiny object detection[EB/OL]. (2021-10-26)[2022-06-14]<https://arxiv.org/abs/2110.13389>.
- [23] SELVARAJU R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[J]. *International Conference on Computer Vision*, 2020, 128(2): 336-359.
- [24] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(6): 1137-1149.
- [25] WANG CY, MARK LIAO HY WU YH, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]//*Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Seattle: IEEE, 2020: 1571-1580.
- [26] RAHIM U F, MINENO H. Data augmentation method for strawberry flower detection in non-structured environment using convolutional object detection networks[J]. *Journal of Agricultural and Crop Research*, 2020, 8(11): 260-271.
- [27] JIAWWEI Z, GUANGZHAO T, CHANG Q, et al. Weed detection in potato fields based on improved YOLOv4: Optimal speed and accuracy of weed detection in potato fields[J]. *Electronics Newsweekly*, 2022, 11(22): 3709.



## Recognizing pear blossom in the natural environment using improved YOLOv7

ZHANG Xiuhua , WEI Huajie , KONG Degang , LIU Shangkun , HUANG Zheng , WANG Hongsen

(1. College of Mechanical and Electrical Engineering, Hebei Agricultural University, Baoding 071001, China; 2. Hebei Intelligent Agriculture Technology Innovation Center, Baoding 071001, China)

**Abstract:** This study aims to improve the accuracy of object detection on the pear blossoms under complex conditions, particularly for the easily obscured, complex backgrounds, varying lighting, and target distances in natural environments. An improved recognition algorithm was proposed for the pear blossom using the YOLOv7 model. Firstly, a P2 small-object layer was added to increase the capability of feature extraction and multi-scale fusion of the model, so that the improved model was selected to better capture the obscured targets. Secondly, a CBAM (convolutional block attention module) attention mechanism was introduced at the end of the output detection layer. CBAM was used to improve the context understanding of the improved model and the performance of YOLOv7 in various scenarios (different lighting conditions, and complex backgrounds). Lastly, the CIoU (complete intersection over union) loss function was optimized to the NWD (normalized weighted distance) loss function. NWD shared the accurate bounding box regression and then performed better for targets with different shapes. The detection accuracy of the model was improved for the complex background targets and distant targets. Additionally, a dataset was created to photograph the pear blossoms from different angles, backgrounds, lighting conditions, and distances during the peak blooming period. The data collection was achieved in a total of 3,240 photos of Ya pear blossoms, 1,582 photos of Golden Pear blossoms, and 2,184 photos of New pear No. 7 blossoms. Ya pear images were used to create a complex environment dataset for pear blossoms, due to the large number of Ya pear samples and their rich background elements. Images of the three pear varieties were used to create a dataset for different varieties. The improved YOLOv7 model was trained and tested using the complex environment dataset. The images with Ya pear blossoms were randomly selected under different lighting conditions, occlusions, backgrounds, and distances. A comparison was also made on the original and improved YOLOv7 model. The results showed that the improved YOLOv7 model shared better detection performance and higher confidence levels. Ablation experiments were conducted to validate the effectiveness of the three improvements. These improvements significantly enhanced the original YOLOv7 model, thus increasing the detection accuracy of pear blossom targets. The detection heatmaps of the improved and original models showed that the heatmap values of the improved YOLOv7 model were closer to the actual pear blossom regions. The focus was also observed to extract the pear blossom edge features. The improved YOLOv7 model was trained and tested using different datasets of pear blossom, indicating excellent adaptability and robustness. Comparative experiments with mainstream algorithms were conducted to evaluate the performance of the improved YOLOv7 model. The precision, recall, mAP, and F1-score of the improved YOLOv7 model increased by 2.1, 1.2, 1.9, and 0.6 percentage points, respectively, reaching 99.4%, 99.6%, 96.4%, and 89.8%, compared with the original. Furthermore, the improved YOLOv7 model also showed better performance in all evaluation metrics, compared with the Faster R-CNN, SSD, YOLOv3, YOLOv4, YOLOv5, YOLOv8, YOLOv9, and YOLOv10 models. A series of experiments were conducted to evaluate the applicability and performance of the improved model. The mAP of the improved YOLOv7 was 3.9 and 3.7 percentage points higher than that of YOLOv7, when training close and distant datasets, respectively; When training the forward and backward light datasets, the mAP of improved YOLOv7 was 4.4 and 1.6 percentage points higher than that of YOLOv7, respectively; The mAP of the improved YOLOv7 increased by 1.8, 1.4, and 1.5 percentage points, respectively, in the ground, sky, and pear blossom backgrounds, compared with the YOLOv7 model. The high accuracy of detection was achieved to recognize the pear blossom in complex environments with varying backgrounds, distances, occlusions, and lighting conditions. The findings can also provide the data support to accurately identify the pear blossoms under the natural environment in pear orchards.

**Keywords:** pear blossom; image recognition; YOLOv7; natural environment; small-object layer; CBAM; loss function.