DOI: 10.12398/j.issn.2096-7217.2026.01.006

基于改进YOLOv8n的苹果采摘机器人识别算法研究

门晓龙[1,2,3], 何义川[1,2,3*], 汤智辉[2,3], 刘湛[1,2], 潘思祺[1,3], 姚欢杰[4]

(1. 塔里木大学机械电气化工程学院,新疆阿拉尔,843300; 2. 南疆特色农林产物利用与装备兵团重点实验室,新疆阿拉尔,843300; 3. 新疆维吾尔自治区教育厅普通高等学校现代农业工程重点实验室,新疆阿拉尔,843300; 4. 阿拉尔市疆田果将香梨种植农民专业合作社,新疆阿拉尔,843300)

摘要:针对果园内枝叶遮挡、光照不均等复杂环境因素导致采摘机器人对苹果的检测精度低、鲁棒性差等问题。研究提出一种改进YOLOv8n网络轻量化模型的苹果检测算法,旨在提升复杂场景下果实检测的准确性与可靠性。首先,在主干网络末端加入多尺度位置注意力机制(multi-scale position-channel attention,MPCA),通过对不同尺度下苹果位置信息的精准捕捉,增强模型对目标位置特征的感知能力,有效降低复杂果园环境对检测结果的干扰;其次,在颈部网络中引入"FocusFeature"特征融合模块,借助多尺度深度可分离卷积实现信息融合与特征增强,提升模型的检测能力。最后,将损失函数CIoU替换为WIoUv3,依据预测框与真实框的匹配质量,动态调整样本在损失计算中的权重,有效提升了模型的定位精度。试验表明,改进模型在复杂果园环境下精确率达92.2%、召回率86.4%,*mAP*@0.5与*mAP*@[0.5: 0.95]分别为94.4%、72.4%,相较原始算法提升3.4%、6%、3.9%、2.8%。消融试验显示,WIoUv3、MPCA、FocusFeature特征融合模块分别使*mAP*@0.5提升0.3%、0.9%、1.1%,多模块协同实现性能叠加。模型大小为6.69 MB,较原始模型仅增加0.74 MB,未显著增加计算负载。与YOLOv5n、YOLOv7-tiny、YOLOv11n等主流算法相比,改进模型在精确率上分别提高3.9%、7.6%、5.7%,*mAP*@[0.5: 0.95]指标分别提升2.5%、9.9%、13%,在检测精度与泛化能力方面展现出显著优势。可见,提出的改进YOLOv8n算法通过多模块协同作用,显著提升了复杂果园环境下苹果检测的精度与鲁棒性,为采摘机器人的实际应用提供了有效技术方案,对推动苹果采收智能化发展具有重要意义。

关键词:YOLOv8n;采摘机器人;目标检测;轻量化;模型改进;复杂果园环境

中图分类号:S225.93;TP391.4 **文献标识码:**A **文章编号:**2096-7217(2026)01-0052-11

门晓龙, 何义川, 汤智辉, 刘湛, 潘思祺, 姚欢杰. 基于改进YOLOv8n的苹果采摘机器人识别算法研究[J]. 智能化农业装备学报(中英文), 2026, 7(1): 52-62

MEN Xiaolong, HE Yichuan, TANG Zhihui, LIU Zhan, PAN Siqi, YAO Huanjie. Research on apple picking robot recognition algorithm based on improved YOLOv8n [J]. Journal of Intelligent Agricultural Mechanization, 2026, 7(1): 52-62

0 引言

我国是全球最大的苹果生产国,产量占世界总产量的50%以上[1]。苹果产业在我国水果经济发展中起着不可或缺的作用,但目前我国苹果生产其采收方式仍以人工采摘为主,生产效率低、成本高[2]。随着我国农业现代化的加速推进,苹果采收方式正从传统人工采摘向智能化机器采摘转型。采摘机器人作为在农业领域的核心装备,通过精准的识别与定位技术,不仅提升了采摘作业效率,并减少了人工成本,为破解"用工难、用工贵"问题提供了技术解决方案,促进了整个苹果产业的可持续发展,具有较大的战略意义和市场前景[3]。

目标检测算法是采摘机器人实现采摘作业的核心技术支撑,对推动果园智能化发展具有关键作用。然而,依赖手工特征提取的传统目标检测算法在面对阴影、遮挡、重叠等复杂果园环境时,存在泛化能力差、检测性能下降等问题,导致对苹果的识别和检测效果仍存在显著不足,随着机器视觉技术的飞速发展,基于卷积神经网络的深度学习技术不断完善,其相较于传统

收稿日期:2025-10-17 修回日期:2025-12-15

基金项目:新疆生产建设兵团科技计划(2024BA005)

第一作者:门晓龙,男,1995年生,内蒙古呼和浩特人,硕士研究生;研究方向为南疆智能农业机械装备。E-mail: 2879955207@qq.com

*通信作者:何义川,男,1984年生,甘肃秦安人,硕士,教授,博导;研究方向为南疆智能农业机械装备。E-mail: hyc2003@163.com

目标检测算法在实时性与准确性上优势显著。因此，越来越多的专家学者开始对水果识别领域的深度学习算法展开研究。例如：程浈浈等[4]采用基于YOLOv3的目标检测算法，通过修正锚点大小优化苹果在不同尺度上的变化，为苹果采摘机器人的研发提供技术支撑。武星等[5]提出了一种基于改进YOLOv3的轻量化苹果检测模型，通过构建同构残差级联特征和采用可分离深度卷积的方式提高了苹果的检测速度和准确率。岳琳茜等[6]提出了一种基于改进YOLOv4苹果检测模型的方法，通过在模型中加入注意力模块及可变形卷积，从而提升复杂背景下对苹果的检测性能。周桂红等[7]通过在YOLOv4网络模型中加入scSE注意力模块及深度可分离卷积层，研发出一种全景图像的苹果检测方法，为复杂果园环境下的果实计数提供了有效方案。刘子龙等[8]提出了一种基于改进YOLOv5的目标检测算法，通过在YOLOv5的基础上加入坐标注意力机制，加强了对小目标苹果检测的能力。龙燕等[9]提出一种融合Transformer与改进损失函数的YOLOv7模型，该模型对苹果具有更好的检测效果和鲁棒性。陈青等[10]提出了一种基于改进YOLOv7苹果姿态融合识别方法。通过改进多尺度特征融合网络以及引入CBAM注意力机制，为苹果无损高效采摘提供抓取方向参考。

上述研究针对YOLO系列目标检测算法展开改进，提升了其对苹果的检测能力。然而，这些算法依旧存在一些不足，比如模型结构较为复杂，检测速度较慢，并且在实际部署时面临一定困难，相比之下，YOLOv8在模型架构、检测头设计和训练策略等方面都进行了优化，检测精度和推理速度均得到提升，对小目标检测能力及鲁棒性更强，且YOLOv8n轻量化版本符合采摘机器人端侧部署需求，故本文以YOLOv8n为基线模型进行改进，来实现对苹果的自动检测。

1　材料选取

为建立符合苹果检测算法训练要求的自然场景特征数据集，以新疆阿拉尔市苹果种植园为拍摄基地，全方位、多角度采集苹果自然场景特征。由于新疆南疆地区海拔高、紫外线强、昼夜温差大，对苹果表皮纹理和着色度会产生一定影响，有利于丰富数据集样本并提高模型的泛化能力。数据采集时，共拍摄5 549张苹果自然场景生长环境照片，并采用随机方式分别拍摄不同时间段、不同光照强度、不同拍摄距离、不同方向以及遮挡状态下的自然场景（如图1所示），其中，枝干遮挡949张，枝叶遮挡875张，果实重叠876张，强光光照747张，弱光光照736张，远景（4～6 m）场景446张，正常场景920张，为维持果实形态真实性，保留完整细节特征，本研究将保留所有图像的原始拍摄格式（JPG）与原生像素尺寸，旨在为模型提供更贴合真实检测场景的训练样本。苹果数据集按照8∶1∶1分割，得到训练集4 439张，验证集555张，测试集555张。



图1　自然场景苹果拍照图

Figure 1　Apple photos taken in natural scenes

试验环境硬件、软件配置及模型参数如表1所示。

表1　试验环境与模型参数配置

Table 1　Experimental environment and model parameter configuration

类别	配置项	型号/参数
硬件配置	操作系统	Windows11
	CPU	Intel(R)Core(TM)i7-14650HX
	GPU	NVIDIA GeForce RTX4060(显存为16 GB)
	核显	Intel(R)UHD Graphics
软件环境	编程语言	Python 3.8.20
	深度学习框架	Pytorch 2.0.1
	开发工具	PyCharm 2022.1.3
	CUDA版本	12.7
模型基础参数	输入图像尺寸/（像素×像素）	640×640
	训练轮次/轮	100
	批次大小/样本	32
	初始学习率	0.01
	优化器	AdamW

本研究使用开源图像标注工具Labellmg完成苹果数据集标注，此工具支持标注目标检测、图像分类等计算机视觉任务。标注规则是果实框优先使用最小外接矩形框，当苹果超过1/3面积在图像边界外时，因果实特征不完整难以支撑模型有效学习，故不作为标注对象。标注后的结果以YOLO格式存储为txt文件，苹果实例统一为“apple”标签进行存储，以保证数据集可用且格式规范，如图2所示。

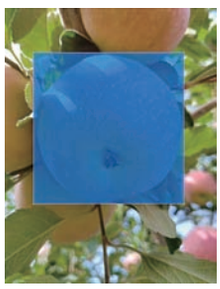


图2 苹果标注示意图

Figure 2 Diagram of apple labeling

2 YOLOv8n模型改进

2.1 YOLOv8网络结构

YOLOv8是Ultralytics公司发布的YOLO系列主流模型,自2023年推出后在目标检测领域广泛应用[11]。在检测精度和速度上都展现出了卓越的性能,目前已广泛应用于交通运输[12]、工业制造[13]、农业[14]、医疗[15]等各个领域。YOLOv8(图3)可分为:输入模块(Input)、主干网络(Backbone)、颈部网络(Neck)和头部网络(Head)4部分[16]。其中,输入模块负责接收并对输入数据进行缩放,色调调整,归一化等预处理及Mosaic增强、混合增强等数据增强操作,为检测模型提供标准化数据输入;主干网络采用改进的CSPDarknet结构作为骨干网络,通过融合C2f与SPPF模块,实现模型高效的特征提取与多尺度目标适应。C2f模块是在YOLOv7的CSP模块基础上改进而来,其内部的残差瓶颈结构与卷积层设计,有效保障了数据信息在网络深层传递与表达[17];SPPF模块延续自YOLOv5,该模块由多个池化层和卷积层组成,能够将不同尺度的特征信息进行高效融合,显著增强模型对不同大小目标的检测能力[18];颈部网络部分延续使用了FPN+PAN的模式,特征金字塔网络(FPN)通过自上而下的多层次特征来识别不同尺寸的目标,路径聚合(PAN)网络通过自下而上的路径聚合,加强底层到高层的特征信息流动。通过将两者融合,实现了特征的双向高效流动,显著增强模型了对目标的检测能力与鲁棒性。头部网络部分采用解耦头结构,将目标检测任务分为分类头和回归头,使得模型能够更加专注于各自目标的优化,进而提升分类和定位的准确性,再结合动态锚点机制和优化损失函数,最终输出高精度的目标预测结果[19]。

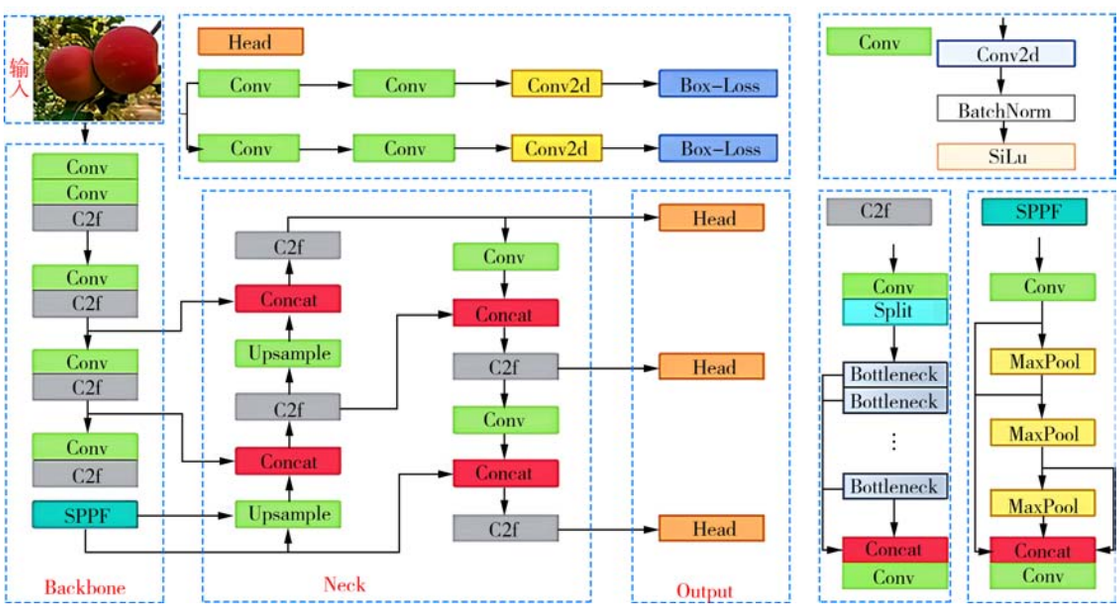


图3 YOLOv8网络架构图

Figure 3 YOLOv8 network architecture diagram

2.2 改进YOLOv8n网络结构

为提高模型在复杂果园环境下对苹果的检测精度,本文以YOLOv8n为基础框架进行改进,首先在主干网络末端加入多尺度位置注意力机制(multi-scale position-channel attention,MPCA)注意力机制,然后在颈部网络引入FocusFeature特征融合模块,最后将损失函数替换为WIoUv3。改进后的YOLOv8n网络结构如图4所示。

2.3 主干网络末端引入MPCA注意力机制模块

为增强果园复杂环境下苹果果实特征在空间维度与通道维度上的表达能力,设计引入了一种融合空间方向感知与通道注意力的特征增强机制。该机制以输入特征图为基础,分别从X方向(水平)和Y方向(垂直)对空间上下文信息进行建模,同时结合通道维度上的全局信息,实现对果实关键区域的精确聚焦。

苹果在图像中通常呈现近似圆形或椭圆形,其空

间分布具有中心对称性[20]。传统空间注意力机制（如CBAM）[21]对全局区域平等加权，难以聚焦果实中心区域。而 MPCA 通过水平（X 方向）和垂直（Y 方向）池化，可分别提取果实的横向轮廓与纵向中轴线特征，契合果实的形态学特征。其中，水平池化能抑制果实上下边缘的枝叶干扰，垂直池化可增强果实左右两侧的边界响应，两者结合可精准定位果实几何中心，如图 5 所示。

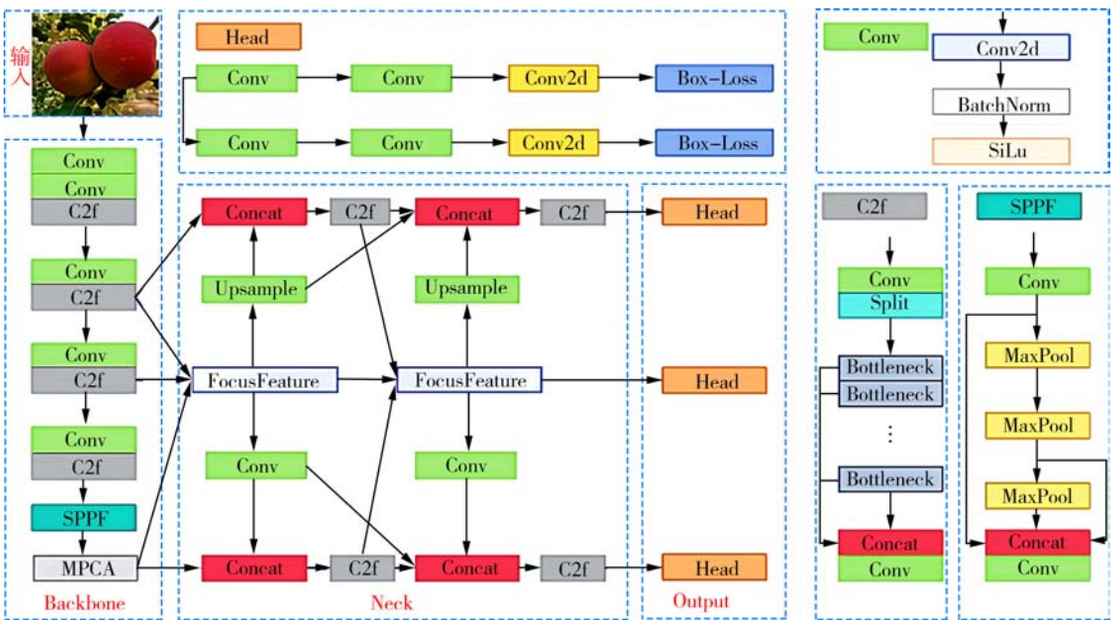


图 4　改进后的 YOLOv8n 网络架构图

Figure 4　The improved YOLOv8n network architecture diagram

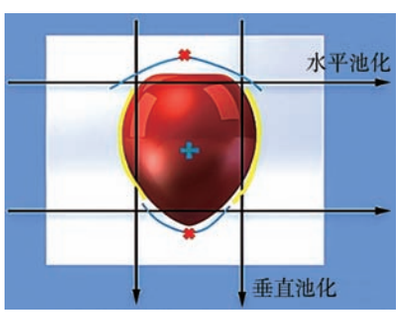


图 5　水平—垂直池化增强果实特征提取示意图

Figure 5　Schematic diagram of fruit feature extraction enhanced by horizontal-vertical pooling

具体运行上，首先对输入特征图分别在水平和垂直方向上进行平均池化，具体计算方式如式（1）－式（2）所示。

$$F_{\text{h-avg}}=\frac{1}{W}\sum_{j=1}^{W}F_j \tag{1}$$

$$F_{\text{v-avg}}=\frac{1}{H}\sum_{i=1}^{H}F_i \tag{2}$$

式中：$F_{\text{h-avg}}$——水平方向平均池化后的特征图；

$F_{\text{v-avg}}$——垂直方向平均池化后的特征图；

W——输入特征图在水平方向上的尺寸；

H——输入特征图在垂直方向上的尺寸；

F_j——取特征图中第 j 列的所有元素；

F_i——取特征图中第 i 行的所有元素。

通过对输入特征图进行水平及垂直方向平均池化后，可提取出输入特征图的一维方向特征。随后，Y 方向特征通过维度转置后与 X 方向特征进行串联，以促进水平和垂直方向间的信息交互，经卷积操作后（Conv1），特征图被拆分（Split）为 X 和 Y 两支注意力通道，分别用于增强对应方向的原始特征信息。为精细化建模空间维度（$H\times W$）的依赖关系，对卷积输出采用二次卷积（Conv2）扩大感受野，以挖掘跨尺度细微空间关联，通过 Sigmoid 激活函数生成归一化全局权重，再经通道拆分解耦为 X 方向权重 X_{weight} 与 Y 方向权重 Y_{weight}。这两个方向权重分别与前述经卷积操作拆分后的 X、Y 两支注意力通道进行逐元素乘法操作，使对应方向特征得到增强，之后再次融合，形成增强后的空间方向特征。该过程通过动态调整各空间维度的响应强度，使空间权重系数更加精细，实现对特征空间分布的自适应优化，有效抑制冗余信息并突出关键位置特征。

与此同时，对输入特征图进行全局平均池化（global average pooling，GAP），计算方式如式（3）所示。

$$F_{\text{GAP}}=\frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}F_{i,j} \tag{3}$$

式中：$H\times W$——特征图的空间维度大小；

$F_{i,j}$——输入特征图 F 在第 i 行、第 j 列的所有通道值。

池化后的特征F_{GAP}再进行卷积操作,为避免直接处理导致特征信息缺乏交互性与关联性,对经Sigmoid激活的权重在第二个通道维度上计算均值,使权重维度与特征图一致,从而利用水平与垂直方向信息动态调整权重。调整后的权重与对应特征相乘,实现多支路特征的交互融合与协同优化,增强模型对特征的综合表征能力。

所有增强分支在模块末尾进行融合,与原始输入特征图相乘输出,实现空间与通道信息的协同强化。该机制在不显著增加参数量的前提下,能够有效提升模型对空间位置信息与关键目标区域的感知能力,如图6所示。

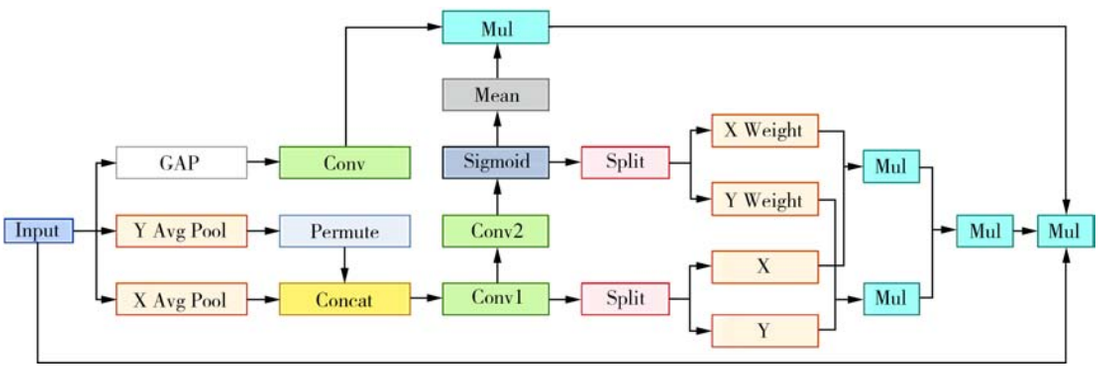


图6 MPCA多尺度位置注意力机制流程图

Figure 6 Flowchart of the MPCA multi-scale position attention mechanism

2.4 颈部网络架构中嵌入FocusFeature特征融合模块

为增强模型对多尺度损伤特征的建模能力,本文设计了名为"FocusFeature"的特征融合增强模块。该模块的核心理念是通过不同尺度的深度可分离卷积(DWConv)[22]提取多尺度上下文信息,并实现融合表达,从而提升特征图对目标区域的关注能力。

具体而言,FocusFeature特征融合模块将来自主干网络的3层特征图作为输入,包括浅层F_{P3}、中间层F_{P4}与深层F_{P5}。

为了使3层特征图能够有效融合,保障模型训练的稳定性,对浅层进行下采样(ADown),深层进行上采样(AUp),中间层不作处理。3层特征图统一尺寸后进行卷积处理,最后在通道维度进行拼接Concat,生成融合特征F_{concat}。随后,融合特征F_{concat}分别经过5×5、7×7、9×9和11×11的深度可分离卷积,以提取不同感受野下的上下文信息,深度可分离卷积计算方式如式(4)所示。

$$F_{k\times k}=DSConv_{k\times k}F_{concat} \tag{4}$$

式中:$DSConv_{k\times k}F_{concat}$——对特征值$F_{concat}$执行卷积核大小为$k\times k$的深度可分离卷积操作。

若采用单一尺寸的小卷积核[23],模型在处理大尺寸果实图像时,难以捕捉其周围背景环境、相邻果实或枝叶的相对位置等上下文语义信息,这是由于小尺寸卷积核的感受野局限于局部区域,无法有效获取大目标的全局结构信息。

而本研究中采用的多尺度核(5×5、7×7、9×9、11×11),通过融合不同感受野特征,在应对不同尺寸果实检测时表现出明显优势。对于小果实,较小尺度的卷积核(如5×5)能够保留如梗洼等精细细节;对于大果实,较大尺度的卷积核(如9×9、11×11)则可有效捕捉其全局轮廓信息。多尺度卷积完成后,与通过恒等映射(Identity)保留的原始融合特征F_{concat}构成残差路径并进行首次融合(Add1),生成特征图F_{Add1},计算方式如式(5)所示。

$$F_{Add1}=\sum_{k\in\{5\times5,7\times7,9\times9,11\times11\}}DSConv_{k\times k}(F_{concat})+F_{concat} \tag{5}$$

这一设计通过保留原始特征,既能避免多尺度卷积过程中的信息损耗,丰富融合后的特征,又能利用"原始特征+多尺度变换特征"的残差学习机制,缓解深层网络的梯度消失问题,提升模型训练的稳定性。融合后的特征先经1×1卷积压缩通道数,减少计算复杂度,再通过ReLU激活函数引入非线性变换,增强特征表达能力。为进一步丰富特征信息,在已处理和增强的特征基础上,将其与原始融合特征F_{concat}进行第二次融合(Add2),使模型能够综合利用先前处理得到的强表达特征以及最初的基础信息,形成最终输出F_{output},计算方式如式(6)所示。

$$F_{output}=\mathrm{ReLU}\left[Conv_{1\times1}(F_{Add1})\right]+F_{concat} \tag{6}$$

通过上述操作,该模块在保持参数量轻量化的同时,实现多尺度语义增强与细节保持的统一,有效增强了网络对空间细节与上下文环境的感知能力,FocusFeature特征融合模块结构如图7所示。

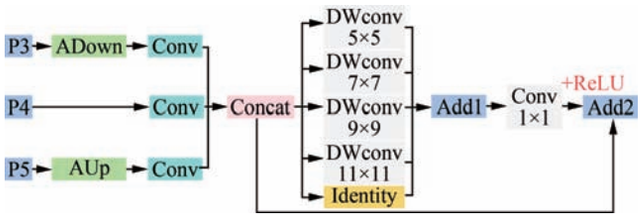


图 7　FocusFeature 特征融合模块架构图

Figure 7　Architecture diagram of the FocusFeature feature fusion module

2.5　采用 WIoUv3 损失函数替换原 CIoU 损失函数

在 YOLOv8 网络结构中，采用 CIoU[24] 作为边界框回归损失来度量预测框与真实框之间的差异，CIoU 在 *IoU*[25] 的基础上引入了中心点距离和长宽比的惩罚项来计算损失，能够一定程度提升边界框回归的精度，有效优化了苹果检测效果，其损失计算如式(7)－式(10)所示。

$$L_{CIoU}=1-IoU+\frac{\rho^2(b,b_{gt})}{c^2}+\alpha\vartheta \tag{7}$$

$$IoU=\frac{\left|B\cap B_{gt}\right|}{\left|B\cup B_{gt}\right|} \tag{8}$$

$$\alpha=\frac{\vartheta}{(1-IoU)+\vartheta} \tag{9}$$

$$\vartheta=\frac{4}{\pi^2}\left[\tan^{-1}\left(\frac{w_{gt}}{h_{gt}}\right)-\tan^{-1}\left(\frac{w}{h}\right)\right]^2 \tag{10}$$

式中：*IoU*——交并比，即预测框和真实框交集面积与并集面积的比值；

b——预测框中心点坐标；

b_{gt}——真实框中心点坐标；

$\rho^2(b,b_{gt})$——预测框和真实框中心点之间的欧几里得距离的平方；

c^2——预测框和真实框的最小外接矩形对角线长度的平方；

α——一个权重系数，用于调整长宽比约束的影响程度；

ϑ——衡量预测框和真实框长宽比差异的一个度量；

B——预测框；

B_{gt}——真实框；

w_{gt}——真实框的宽；

h_{gt}——真实框的高；

w——预测框的宽；

h——预测框的高。

然而，CIoU 仍然存在一定局限，由于在训练过程中采用固定的权重分配机制，未能根据样本质量进行动态权重调整，特别是在复杂果园环境下的小目标检测场景中，CIoU 存在对复杂遮挡、重叠及光照变化等场景适应性不足，小目标检测精度欠佳，因此，本文将 CIoU 损失函数替换为 WIoU 系列中具有动态非单调聚焦机制且效果更为均衡的 WIoUv3 损失函数。WIoU 是一种改进型边界框回归损失函数，旨在进一步提升模型在复杂场景下的定位精度与训练稳定性[26-27]。其核心理念是根据预测框与真实框的匹配质量，动态调整样本在损失计算中的权重贡献，实现高质量样本和低质量样本之间的平衡。其中，WIoUv3 是在 WIoUv1 的基础上优化而来，WIoUv1 损失计算式如式(11)－式(12)所示。

$$L_{WIoUv1}=R_{WIoUv1}\cdot L_{IoU} \tag{11}$$

$$R_{WIoUv1}=\exp\left[\frac{(x-x_{gt})^2+(y-y_{gt})^2}{w_{gt}^2+h_{gt}^2}\right] \tag{12}$$

式中：R_{WIoUv1}——预测框与真实框的位置偏差因子；

L_{IoU}——预测框与真实框的交并比损失；

(x,y)——预测框中心坐标；

(x_{gt},y_{gt})——真实框的中心坐标；

$w_{gt}^2+h_{gt}^2$——一个归一化因子，确保距离惩罚不受目标尺寸影响。

但 WIoUv1 仅依赖单样本位置偏差调整权重，未考虑批次内样本的相对质量。WIoUv3 在 WIoUv1 的基础上，新增了基于样本质量的动态权重机制，损失计算方式如式(13)－式(15)所示。

$$L_{WIoUv3}=r\cdot L_{WIoUv1} \tag{13}$$

$$r=\frac{\beta}{\delta\varepsilon^{\beta-\delta}} \tag{14}$$

$$\beta=\frac{L_{IoU}^{*}}{L_{IoU}^{-}} \tag{15}$$

式中：r——梯度增益；

β——离群度；

ε,δ——超参数；

L_{IoU}^{*}——具有动量的指数移动值；

L_{IoU}^{-}——具有动量的指数移动平均值。

通过引入离群度 β 量化样本质量、梯度增益 r 动态调整权重，WIoUv3 进一步实现了高低质量预框之间的平衡，不仅显著提高了模型定位的准确性，还增强了对微小目标、复杂背景等难检测场景的鲁棒性，提升了模型的整体性能。

3　试验结果与分析

3.1　评估方法

本研究以精确率(P)、召回率(R)、平均精度(AP)及平均精度均值(mAP)[28-29]作为评价指标，来更加直

观地反映YOLOv8n模型的改进性能。评价指标计算方式如式(16)—式(19)所示。

$$P=\frac{TP}{TP+FP} \tag{16}$$

$$R=\frac{TP}{TP+FN} \tag{17}$$

$$AP=\int_0^1 P(h)\mathrm{d}h \tag{18}$$

$$mAP=\frac{1}{n}\sum_{i=1}^{n}\int_0^1 P(h)\mathrm{d}h \tag{19}$$

式中:TP——被模型正确识别为正样本的数量;

FP——被模型错误识别为正样本的数量;

FN——被模型错误识别为负样本的数量;

$P(h)$——精确率随召回率变化的函数关系;

n——样本的类别总数。

3.2 消融试验

为了验证改进算法的有效性,本研究进行8组消融试验,结果如表2所示。试验1为YOLOv8n基础试验,试验2中引入WIoUv3损失函数优化边界框回归,召回率提升1.3%至81.7%,mAP@0.5、mAP@[0.5:0.95]达90.8%、70.7%,显著提升了小目标检测能力。试验3在主干网络末端加入MPCA注意力机制,精确率提升到89.9%,mAP@0.5提升到91.4%,有效增强了复杂场景下的特征表达能力。试验4在颈部网络引入FocusFeature特征融合模块,精确率进一步提升至90.5%,mAP@0.5达91.6%,有效提升多尺度特征利用效率,优化模型检测性能。试验5同时加入WIoUv3损失函数与MPCA注意力机制,模型大小为6.58 MB,保障模型轻量化同时优化检测指标。试验6同时加入MPCA注意力机制和FocusFeature特征融合模块,在模型大小微增的情况下,其余性能指标都显著增长。试验7同时加入WIoUv3损失函数和FocusFeature特征融合模块,在模型参数减少的情况下,其余性能指标进一步增长。试验8融合所有优化模块,通过多模块协同作用实现了检测性能的全面提升,模型大小仅增加12.4%,精确度提升至92.2%、召回率为86.4%、mAP@0.5达94.4%。

表2 消融试验

Table 2 Ablation experiment

序号	WIoUv3	MPCA	FocusFeature	P/%	R/%	mAP@0.5/%	mAP@[0.5:0.95]/%	*Parameters*/M	计算量/GFLOPs
1				88.8	80.4	90.5	69.6	3.01	8.1
2	√			85.8	81.7	90.8	70.7	3.01	8.1
3		√		89.9	81.8	91.4	70.0	3.46	9.1
4			√	90.5	80.5	91.6	69.7	3.39	9.9
5	√	√		86.7	82.4	91.5	70.8	3.46	9.1
6		√	√	88.3	84.2	92.3	70.7	3.84	10.9
7	√		√	90.1	85.6	93.1	71.0	3.39	9.9
8	√	√	√	92.2	86.4	94.4	72.4	3.84	10.9

3.3 对比试验

为评估改进模型的检测性能,本研究进行了对比试验,选取YOLOv5n、YOLOv7-tiny、YOLOv8n、YOLOv11n以及本研究改进算法共5种轻量化模型在相同数据集及试验环境下进行比较试验,试验结果如表3所示。

从数据来看,在精确率与召回率方面,改进模型分别以92.2%和86.4%高于YOLOv5n、YOLOv7-tiny、YOLOv8n、YOLOv11n,检测准确性与目标捕捉能力更优;平均精度mAP@0.5达94.4%,表明基础检测精度更全面;多阈值平均精度mAP@[0.5:0.95]为72.4%,凸显复杂场景下的泛化能力优势。综上所述,改进模型综合性能显著优于对比算法,在苹果检测等实际应用中兼具高精度与实用性。

表3 对比试验

Table 3 Comparative test

比较算法	P/%	R/%	mAP@0.5/%	mAP@[0.5:0.95]/%
YOLOv5n	88.3	81.7	91.3	69.9
YOLOv7-tiny	84.6	76.1	87.9	62.5
YOLOv8n	88.8	80.4	90.8	70.7
YOLOv11n	86.5	72.1	83.3	59.4
本研究改进算法	92.2	86.4	94.4	72.4

为直观评估改进模型性能,本文展示了其在果园复杂环境下与其他4种模型的检测效果。如图8所示,从可视化结果可见,在枝叶遮挡、光照不均、果实密集等复杂场景下,改进模型仍能通过边界框精准锁定果园中的

苹果目标，并稳定输出最高精度检测结果，充分验证了其在真实果园环境中的抗干扰能力与实用价值。

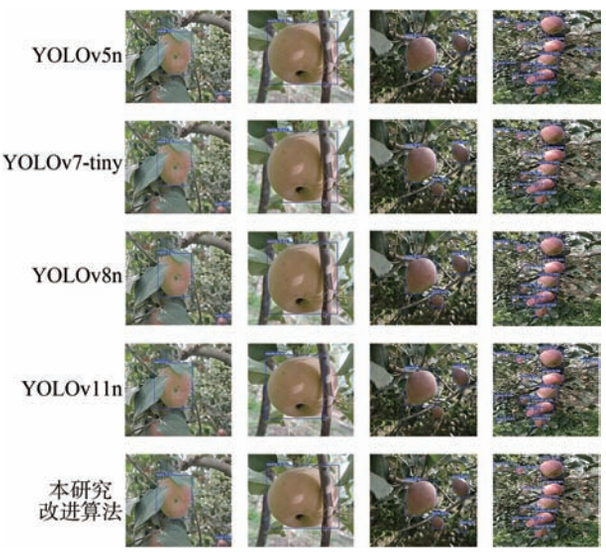


图8　不同检测模型的检测效果图

Figure 8　Shows the detection effect diagrams of different detection models

4　结论

本研究提出一种基于改进YOLOv8n模型的复杂果园环境苹果抗干扰检测方法，可精准实现枝叶遮挡、光照不均等复杂场景下苹果目标的检测，为采摘机器人实际应用提供技术支撑，主要结论如下：

1）针对复杂果园环境下传统检测模型对苹果检测精度低、鲁棒性差的问题，以YOLOv8n为基线模型，在主干网络末端引入MPCA多尺度位置注意力机制。该机制通过水平—垂直池化精准捕捉不同尺度下苹果位置信息，增强模型对目标位置特征的感知能力，有效降低枝叶遮挡、背景干扰对检测结果的影响，为复杂场景下苹果特征提取提供保障。

2）针对颈部网络多尺度特征融合效率不足的问题，设计并引入FocusFeature特征融合模块。该模块基于多尺度深度可分离卷积，对主干网络输出的P3、P4、P5三层特征图进行尺寸统一与多感受野特征提取，通过残差融合保留原始特征信息，缓解深层网络梯度消失问题，显著提升模型对不同大小苹果目标的检测能力，尤其优化小目标苹果的特征表达。

3）针对原始YOLOv8n采用的CIoU损失函数无法动态调整样本权重、定位精度有限的问题，将损失函数替换为WIoUv3。该损失函数依据预测框与真实框的匹配质量，通过动态权重调整实现高低质量样本的平衡，有效提升模型定位精度与复杂场景适应性。

改进后的YOLOv8n模型综合性能显著优于原始模型及主流对比算法：在复杂果园环境下，精确率达92.2%、召回率86.4%，*mAP*@0.5与*mAP*@[0.5：0.95]分别为94.4%、72.4%，相较原始YOLOv8n分别提升3.4%、6%、3.9%、2.8%；与YOLOv5n、YOLOv7-tiny、YOLOv11n相比，精确率分别提高3.9%、7.6%、5.7%，*mAP*@[0.5：0.95]分别提升2.5%、9.9%、13%。同时，模型大小为6.69 MB，较原始模型仅增加0.74 MB，未显著增加计算负载，在保障检测精度与鲁棒性的同时，满足实际应用中的轻量化需求，对推动苹果采收智能化发展具有重要意义，也为其他复杂场景下的目标检测研究提供参考。

参　考　文　献

[1] 张静，张昊，聂尚卿，等. 基于改进YOLOv8的苹果采摘机器人视觉算法[J]. 液压气动与密封，2025，45(3)：36－43.
ZHANG Jing，ZHANG Hao，NIE Shangqing，et al. Vision algorithm for apple picking robot based on improved YOLOv8 [J]. Hydraulics Pneumatics & Seals，2025，45(3)：36－43.

[2] ZHANG Y，LI N，ZHANG L，et al. Parallel depthwise separable multiscale information aggregation network for real-time fruit detection in trellis apple orchard [J]. Computers and Electronics in Agriculture，2025，237：110503.

[3] ZHAO Z，HICKS Y，SUN X，et al. A novel self-supervised method for in-field occluded apple ripeness determination [J]. Computers and Electronics in Agriculture，2025，234：110246.

[4] 程浈浈，程一帆，缪百灵，等. 基于YOLOv3的不同场景下苹果目标检测方法[J]. 北方园艺，2024(22)：137－146.
CHENG Zhenzhen，CHENG Yifan，MIAO Bailing，et al. Apple target detection method in different scenarios based on YOLOv3 [J]. Northern Horticulture，2024 (22)：137－146.

[5] 武星，齐泽宇，王龙军，等. 基于轻量化YOLOv3卷积神经网络的苹果检测方法[J]. 农业机械学报，2020，51(8)：17－25.
WU Xing，QI Zeyu，WANG Longjun，et al. Apple detection method based on Light-YOLOv3 convolutional neural network [J] Transactions of the Chinese Society for Agricultural Machinery，2020，51(8)：17－25.

[6] 岳琳茜，李文宽，杨晓峰，等. 基于改进YOLOv4的苹果检测与果径估测方法[J]. 激光杂志，2022，43(2)：58－65.
YUE Linxi，LI Wenkuan，YANG Xiaofeng，et al. Apple detection and fruit diameter estimation method based onimproved YOLOv4 [J]. Laser Journal，2022，43(2)：

58－65.

[7] 周桂红，马帅，梁芳芳．基于改进YOLOv4模型的全景图像苹果识别[J]．农业工程学报，2022，38(21)：159－168.
ZHOU Guihong，MA Shuai，LIANG Fangfang. Recognition of the apple in panoramic images based on improved YOLOv4 model [J]. Transactions of the Chinese Society of Agricultural Engineering，2022，38 (21)：159－168.

[8] 刘子龙，张磊．自然环境下改进YOLOv5对小目标苹果的检测[J]．系统仿真学报，2025，37(8)：2124－2138.
LIU Zilong，ZHANG Lei. Detection of small apple targets based on improved YOLOv5 in natural environments [J]. Journal of System Simulation，2025，37(8)：2124－2138.

[9] 龙燕，杨智优，何梦菲．基于改进YOLOv7的疏果期苹果目标检测方法[J]．农业工程学报，2023，39(14)：191－199.
LONG Yan，YANG Zhiyou，HE Mengfei. Recognizing apple targets before thinning using improved YOLOv7 [J]. Transactions of the Chinese Society of Agricultural Engineering，2023，39(14)：191－199.

[10] 陈青，殷程凯，郭自良，等．基于改进YOLOv7的苹果生长状态及姿态识别[J]．农业工程学报，2024，40(6)：258－266.
CHEN Qing，YIN Chengkai，GUO Ziliang，et al. Apple growth status and posture recognition using improved YOLOv7 [J]. Transactions of the Chinese Society of Agricultural Engineering，2024，40(6)：258－266.

[11] JRONDI Z，MOUSSAID A，HADI M Y. Exploring End-to-End object detection with transformers versus YOLOv8 for enhanced citrus fruit detection within trees [J]. Systems and Soft Computing，2024，6：200103.

[12] 赵树恩，龚道元，田卓帅．基于改进YOLOv8模型的复杂交通场景目标检测算法[J]．重庆交通大学学报(自然科学版)，2025，44(12)：33－41.
ZHAO Shuen，GONG Daoyuan，TIAN Zhuoshuai. Algorithm for object detection in complex traffic scenes based on improved YOLOv8 model [J]. Journal of Chongqing Jiaotong University (Natural Science)，2025，44(12)：33－41.

[13] 宋奇恒，吴双，黄及远，等．基于YOLOv8改进算法的船体合拢焊缝外观缺陷检测系统设计[J]．制造业自动化，2025，47(5)：118－124.
SONG Qiheng，WU Shuang，HUANG Jiyuan，et al. Design of an appearance defect detection system for hull butt welds based on an improved YOLOv8 algorithm [J]. Manufacturing Automation，2025，47(5)：118－124.

[14] JIA X，HUA Z，SHI H，et al. A soybean pod accuracy detection and counting model based on improved YOLOv8 [J]. Agriculture，2025，15(6)：617.

[15] DEWI C，MANONGGA D，HENDRY，et al. Deep learning and YOLOv8 utilized in an accurate face mask detection system [J]. Big Data and Cognitive Computing，2024，8(1)：9.

[16] 罗友璐，潘勇浩，夏顺兴，等．基于改进YOLOv8的苹果叶病害轻量化检测算法[J]．智慧农业(中英文)，2024，6(5)：128－138.
LUO Youlu，PAN Yonghao，XIA Shunxing，et al. Lightweight apple leaf disease detection algorithm based on improved YOLOv8 [J] Smart Agriculture，2024，6(5)：128－138.

[17] 魏堂伟，张津诚，王晶，等．基于改进YOLOv7的茶叶嫩芽识别模型研究[J]．智能化农业装备学报(中英文)，2024，5(2)：42－50.
WEI Tangwei，ZHANG Jincheng，WANG Jing，et al. Study of tea buds recognition and detection based on improved YOLOv7 model [J]. Journal of Intelligent Agricultural Mechanization，2024，5(2)：42－50.

[18] 蔡家一，刘世伟，单龙祥，等．基于机器视觉与YOLO v5的裂纹蛋分拣机器人设计与试验[J]．智能化农业装备学报(中英文)，2025，6(1)：41－50.
CAI Jiayi，LIU Shiwei，SHAN Longxiang，et al. Design and experiment of cracked egg sorting robot based on machine vision and YOLO v5 [J]. Journal of Intelligent Agricultural Mechanization，2025，6(1)：41－50.

[19] SUN Y，ZHANG D，GUO X，et al. Lightweight algorithm for apple detection based on an improved YOLOv5 model [J]. Plants，2023，12(17)：3032.

[20] 王鑫．苹果采摘末端执行器结构设计与柔顺控制研究[D]．杨凌：西北农林科技大学，2023.
WANG Xin. Research on structural design and compliancecontrol of apple picking end-effector [D]. Yangling：Northwest A&F University，2023.

[21] ZHANG X，LI D，LIU X，et al. Research of segmentation recognition of small disease spots on apple leaves based on hybrid loss function and CBAM [J]. Frontiers in Plant Science，2023，14：1175027.

[22] WANG H，JU X，ZHU H，et al. SEFormer：A lightweight CNN-Transformer based on separable multiscale depthwise convolution and efficient self-attention for rotating machinery fault diagnosis [J]. Computers，Materials & Continua，2025，82(1).

[23] JIANG L，JIN Q，HUA F，et al. Numerical investigation of the effective receptive field and its relationship with convolutional kernels and layers in convolutional neural network [J]. Frontiers in Marine Science，2024，11：1492572.

[24] 周子渊，成苗，何莲，等．基于改进YOLOv8的小目标与细长目标检测模型[J]．计算机应用，2024，44(S2)：

286－295.
ZHOU Ziyuan, CHENG Miao, HE Lian, et al. Small and elongated object detection model based on improved YOLOv8 [J]. Journal of Computer Applications, 2024, 44(S2): 286－295.
[25] YU Q, HAN Y, HAN Y, et al. Enhancing YOLOv5 performance for small-scale corrosion detection in coastal environments using IoU-Based loss functions [J]. Journal of Marine Science and Engineering, 2024, 12(12): 2295.
[26] CHEN X, JIAO Z, LIU Y. Improved YOLOv8n based helmet wearing inspection method [J]. Scientific Reports, 2025, 15(1): 1945.
[27] 郝博，徐新岩，赵玉欣，等. 基于改进YOLOv8的铆接孔表面缺陷检测[J]. 东北大学学报(自然科学版)，2024，45(11)：1595－1603.
HAO Bo, XU Xinyan, ZHAO Yuxin, et al. Surface defect detection of riveting holes based on improved YOLOv8 [J]. Journal of Northeastern University (Natural Science), 2024, 45(11): 1595－1603.
[28] 李小祥，张洁，秦柯贝，等. 基于改进YOLOv8的轻量级复杂环境苹果叶片病害检测方法[J]. 南京农业大学学报，2025，48(3)：734－743.
LI Xiaoxiang, ZHANG Jie, QIN Kebei, et al. A lightweight method for apple leaf disease detection in complex environments based on improved YOLOv8 [J]. Journal of Nanjing Agricultural University, 2025, 48(3): 734－743.
[29] 李书贤. 基于多传感器阵列的水果分类及新鲜度检测方法研究[D]. 天津：天津职业技术师范大学，2024.
LI Shuxian. Research on fruit classification and freshness detection method based on multi-sensor array [D]. Tianjin: Tianjin University of Technology and Education, 2024.

Research on apple picking robot recognition algorithm based on improved YOLOv8n

MEN Xiaolong[1, 2, 3], HE Yichuan[1, 2, 3*], TANG Zhihui[2, 3], LIU Zhan[1, 2], PAN Siqi[1, 3], YAO Huanjie[4]

(1. *College of Mechanical and Electrical Engineering, Tarim University, Alar* 843300, *China*; 2. *Xinjiang Production and Construction Corps (XPCC) Key Laboratory of Utilization and Equipment of Special Agricultural and Forestry Products in Southern Xinjiang, Alar* 843300, *China*; 3. *Modern Agricultural Engineering Key Laboratory at Universities of Education Department of Xinjiang Uygur Autonomous Region, Alar* 843300, *China*; 4. *The Farmers' Professional Cooperative of Jiangtian guojiang of Alar City for Fragrant Pear Planting, Alar* 843300, *China*)

Abstract: To address the issues of low detection accuracy and poor robustness of apple picking robots caused by complex environmental factors such as branch and leaf occlusion and uneven light distribution in orchards, this study proposes an apple fruit detection algorithm that improves the lightweight model of the YOLOv8n network, aiming to enhance the accuracy and reliability of apple detection in complex scenarios. Firstly, an "MPCA" (multi-scale position-channel attention) is added at the end of the backbone network to precisely capture the position information of apples at different scales, thereby enhancing the model's perception of target position features and effectively reducing the interference of complex orchard environments on detection results. Secondly, a "FocusFeature" feature fusion module is introduced in the neck network, which uses multi-scale depth wise separable convolution to achieve information fusion and feature enhancement, improving the model's detection capability. Finally, the loss function CIoU is replaced with WIoUv3 (WiseIoU), dynamically adjusting the weight of samples in the loss calculation based on the matching quality of predicted and ground truth boxes, effectively enhancing the model's localization accuracy. Experiments show that the improved model achieves an accuracy of 92.2%, a recall rate of 86.4%, and *mAP*@0.5 and *mAP*@[0.5: 0.95] of 94.4% and 72.4% respectively in complex orchard environments, representing improvements of 3.4%, 6%, 3.9%, and 2.8% compared to the original algorithm. Ablation experiments reveal that WIoUv3, MPCA, and FocusFeature feature fusion module increase *mAP*@0.5 by 0.3%, 0.9%, and 1.1% respectively, with performance enhancement achieved through the synergy of multiple modules. The model size is 6.69 MB,

only 0.74 MB larger thanthe original model, without significantly increasing the computational load. Compared with mainstream algorithms such as YOLOv5n, YOLOv7-tiny, and YOLOv11n, the improved model increases the accuracy by 3.9%, 7.6%, and 5.7% respectively, and *mAP*@[0.5: 0.95] by 2.5%, 9.9%, and 13%, demonstrating significant advantages in detection accuracy and generalization ability. It is evident that the improved YOLOv8n algorithm proposed in this study significantly enhances the accuracy and robustness of apple detection in complex orchard environments through the collaborative effect of multiple modules, providing an effective technical solution for the practical application of picking robots and playing a crucial role in promoting the intelligent development of apple harvesting.

Keywords: YOLOv8n; picking robot; object detection; lightweight; model improvement; complex orchard environment

Foundation items: Science and Technology Program of Xinjiang Production and Construction Corps (XPCC) (2024BA005)

First Author: MEN Xiaolong, E-mail: 2879955207@qq.com

* **Correspondence Author:** HE Yichuan, E-mail: hyc2003@163.com