

周涛,王骥,麦仁贵.基于改进YOLOv8的实时菠萝成熟度目标检测方法[J].华中农业大学学报,2024,43(5):10-20.
DOI:10.13300/j.cnki.hnlkxb.2024.05.002

基于改进YOLOv8的实时菠萝成熟度目标检测方法

周涛¹,王骥^{1,2},麦仁贵³

1. 广东海洋大学电子与信息工程学院,湛江 524088;

2. 广东省智慧海洋传感网及其设备工程技术研究中心,湛江 524088;

3. 广东海洋大学数学与计算机学院,湛江 524088

摘要 为提高不同成熟度种植区域的机械采摘菠萝准确率,保证菠萝品质,提出了基于改进YOLOv8的实时菠萝成熟度目标检测方法。针对自然环境下菠萝机械采摘中存在目标小、数量密集和光线遮挡等问题,改进模型把原始YOLOv8模型中主干部分、颈部部分的公共卷积替换成深度可分离卷积(depthwise separable convolution, DSCConv),精简模型参数;在融合特征前增加了卷积注意力机制模块(convolutional block attention module, CBAM),使特征融合更关注重要的特征,提升目标检测的准确率;使用EIoU损失函数替换YOLOv8网络原损失函数CIoU,加快网络收敛速度。结果显示,改进模型对菠萝成熟度检测的平均精度均值为97.33%,与Faster R-CNN、YOLOv4、YOLOv5、YOLOv7对比发现,平均精度均值分别提升5.53、7.91、4.38、4.66百分点;在保证检测精度的前提下,算法模型参数量仅为 16.8×10^6 。结果表明,改进模型提高了菠萝成熟度识别的精度和推理速度,具有更强的鲁棒性。

关键词 菠萝果实;成熟度;YOLOv8;目标检测

中图分类号 TP391 **文献标识码** A **文章编号** 1000-2421(2024)05-0010-11

菠萝果实富含膳食纤维、矿物质等营养成分和多种生物活性化合物,深受广大消费者喜爱。粤西是我国菠萝的主产地,菠萝产业产量大,经济效益好。仅2021年粤西菠萝种植产量达109.92万t,占全省菠萝产量的90.09%,特别是位于徐闻曲界镇境内的“菠萝的海”是粤西菠萝种植的核心区,近733.33 hm²连片种植的菠萝成为《国家地理杂志》推荐的3A级景区^[1]。

菠萝的品质和贮存与采摘时刻的成熟度相关。由于菠萝为非跃变性果实,过早采摘会导致果实质量下降、口感不佳,而过晚采摘则容易引发果实腐烂,因此确定菠萝的不同成熟度是必不可少的,菠萝成熟度的精准识别可以降低误采概率^[2]。目前,菠萝果实采摘多以传统的人工采摘为主,成本高且采摘品质无法保证,近年来出现了机器自动化采摘方式^[3],而自然环境下菠萝成熟度目标检测是菠萝自动化采摘研究的重要内容^[4]。

对于自然场景下的水果成熟度目标检测研究,

近年来已取得了一定进展,国内外学者尝试采用机器视觉及光谱分析方法实现苹果^[5]、葡萄^[6]、香蕉^[7]、芒果^[8]等果实成熟度的检测。近几年,深度学习理论逐渐被许多学者用于水果成熟度检测的研究。熊俊涛等^[9]利用轻量化YOLOv5-Lite对木瓜的成熟度进行分类识别,实现了对不同拍摄距离、不同遮挡情况以及不同光照下的果实总体检测。陈仁凡等^[10]提出一种轻量化YOLO-ODM草莓成熟度检测方法,实现了在温室环境下草莓果实的快速识别。马瑞峻等^[11]采用PITAYA-YOLOv5对3种不同成熟阶段的火龙果成熟度进行了判别。苗荣慧等^[12]提出基于改进YOLOv7的自然环境下不同成熟度樱桃番茄果实的识别方法,结果显示该方法对绿熟期、转色期、初熟期、中熟期和完熟期果实识别精度较高。上述研究主要针对果实颜色和形状特征进行识别与分割。尽管现有的成熟度研究方法在单色水果成熟度识别方面具有可行性,但处于转色期的菠萝果实具有混合颜色,使得成熟度的检测具有挑战性,主要体现在:遮挡障碍物体积、形状与颜色导致菠萝目标特

收稿日期:2023-12-07

基金项目:广东省普通高校重点领域新一代信息技术专项(2020ZDZX3008)

周涛,E-mail:949339627@qq.com

通信作者:王骥,E-mail:13902576499@163.com

征不足或模糊,对网络模型的特征提取产生负面影响大;菠萝目标数量过多导致单个目标尺度小、分辨率低,造成模型漏检;在密集场景下,菠萝特征受到噪声干扰,使网络模型难以准确分类和定位单个菠萝特征,易发生误检。基于上述原因,李斌等^[13]提出了一种基于单目视觉的检测田间菠萝果实的识别方法,迎光条件下识别正确率为85%,但在背光或阴天的情况下识别正确率仅为40%。张星等^[14]采用改进YOLOv3网络,利用K-means聚类算法对构建的果实数据集进行聚类分析,识别率达到95%。Cuong等^[15]针对适宜期菠萝自动化采摘提出了一种基于改进Tiny YOLOv4的菠萝成熟度识别算法。李阳德等^[16]通过将YOLOv4骨干网络替换成轻量级网络MobileNetV3,提出了轻量级的MobileNetV3-YOLOv4生长期菠萝成熟度分析方法,结果显示该方法对于黄熟期菠萝和青熟期菠萝的检测平均精度分别达到87.62%和94.21%。现有的菠萝成熟度检测模型在一定程度上克服了当前深度学习模型只能识别无遮挡的单一果实或形态较好的相邻果实的局限性,显著提高了识别精度和效率,但密集场景下遮挡引起的漏检误检问题以及尺度变化情况下检测速率低下问题仍然存在。因此,本研究提出了基于改进YOLOv8的实时菠萝成熟度目标检测方法,通过探

索复杂自然环境下不同成熟度菠萝果实的有效识别方法,旨在实现模型识别精度、推理速度和鲁棒性的相对平衡,为快速精确识别目标提供新的研究思路。

1 材料与方法

1.1 图像采集

菠萝果实图像数据集采集地位于广东省湛江市徐闻县曲界镇“菠萝的海”。于2023年2月1日、20日和同年3月20日,分别使用佳能相机(EOS 5D Mark IV 5D4)、无人机(DJI MAVIC 3)等设备采集自然环境下不同种植区域及不同成熟度的菠萝果实图像。采集时间包括上午、中午和下午,以获取不同光照条件下的图像,共获取2 520张菠萝果实图像。采集图像类型包括顺光图像、逆光图像、遮挡图像、无遮挡图像、中远距离图像等。

1.2 果实成熟度等级划分

依据国家农业行业标准NY/T 450—2001^[17],菠萝成熟度可划分为4种类型,即未成熟、转色、中熟和成熟(表1)。根据菠萝实际采摘情况及果皮颜色变化,4种成熟度类型合并为3种成熟度类型,将转色、中熟合并为近生理成熟。由表2可知,不同成熟度的菠萝直接影响着其后续加工、贮藏以及运输方式。

表1 菠萝4级成熟度

Table 1 The 4th degree of ripeness of pineapple

级别 Level	成熟度 Ripeness	描述 Description	采摘情况 Picking situation
1	未成熟 Unripe	果眼饱满,全果以绿色为主 Full fruit eyes, the entire fruit predominantly green	不宜采摘贮藏 Not suitable for picking and storage
2	转色 Color-changing	果眼饱满,果底部开始出现桔黄色 Full fruit eyes, orange-yellow appearing at the base	可采摘贮藏或远距离运输 Suitable for picking and storage or long-distance transportation
3	中熟 Mid-ripening	从果实下部四分之一处为桔黄色发展到果实的一半为桔黄色 Orange-yellow color develops from the bottom quarter of the fruit to the halfway point	可采摘,短距离运输 Suitable for picking, short-distance transportation
4	成熟 Overripe	从果实的一半为桔黄色发展到整个果实均为桔黄色 Orange-yellow color develops from the halfway point to the entire fruit	可采摘,当日销售 Suitable for picking, same-day sale

1.3 数据集搭建

本研究使用了自建的菠萝果实图像数据集,使用了Labelimg数据标注工具对每张图片进行标注,总计包含2 520张图片。为了增加模型的鲁棒性,数据集涵盖了多旋转角度和多样光照条件下的特征图像。研究中构建的数据集分为3个部分,包括训练集、验证集和测试集,图1是部分样本图像的示意图。

为尽可能减小图片失真,将所有图片调整为640

像素×640像素大小,并保持原有的宽高比例。此外,为增强模型的泛化能力和鲁棒性,本研究采用数据增强技术,包括原始模型自带的随机旋转、缩放、裁剪和颜色变换等方法,以扩充数据集并减少过拟合的风险。菠萝果实图像数据集合计2 520张,包括训练集样本图像2 016张、验证集样本图像252张、测试集样本图像252张,其中含未成熟、近生理成熟、成熟菠萝3类样本,分别为3 100、927、1 145个。数据集的基本信息如表3所示。

表 2 菠萝 3 级成熟度
Table 2 The 3rd degree of ripeness of pineapple



级别 Level	成熟度 Maturity	图像 Image	描述 Description	采摘情况 Picking situation
1	未成熟 Unripe		果实大小定型, 果面为绿色 Fruit size stabilized, fruit surface is green	不可采摘贮藏 Not suitable for picking and storage
2	近生理成熟 Ripe		果面呈现黄绿色 Fruit surface presents a yellow-green color	可采摘, 远近距离运输 Suitable for picking, transportation over short and long distances
3	成熟 Overripe		果面呈现出完全黄色的外观 Fruit surface displays a fully yellow appearance	可采摘, 当日销售 Suitable for picking, same-day sale



图 1 不同拍摄条件下的菠萝图像
Fig. 1 Pineapple images under different shooting conditions

1.4 改进 YOLOv8 的实时菠萝成熟度检测模型构建

改进 YOLOv8 算法网络结构如图 2 所示, 由骨干网络、颈部网络和无锚检测头组成。主干网络用于从输入图像中提取特征映射, 颈部网络和头部网络用于预测特征映射中物体的边界框和标签^[18]。

自然复杂环境下快速准确地对菠萝成熟度进行检测、区分近生理成熟期和成熟期的关键点在于对菠萝表面黄色区域的识别和量化。综合考虑色彩特

表 3 菠萝成熟度数据集基本信息
Table 3 Basic information of pineapple maturity dataset

数据集 Data set	图像样本数量 Number of image samples	未成熟 Unripe	近生理成熟 Ripe	成熟 Overripe
训练集 Training set	2 016	2 482	741	946
验证集 Validation set	252	328	93	117
测试集 Testing set	252	317	101	118
合计 Total	2 520	3 100	927	1 145

征和特定阈值下黄色区域的面积, 当菠萝表面的黄色区域占整体面积的 90% 以上时, YOLOv8 模型将其识别为成熟菠萝。

传统的卷积操作由于其固定的卷积核和有限的感受野(receptive field), 只能获取局部信息, 因此在检测小目标方面存在一定的局限性, 特别是对多尺度遮挡的菠萝容易导致漏检。本研究采用深度可分离卷积(DSCConv)替代原始网络的主干部分和颈部部分的常规卷积操作, 旨在减小模型尺寸并提升推理速度。本研究在 DSCBS 模块卷积之前、上采样阶段后以及下采样阶段的每个 DSC2f 模块后引入卷积注意力机制模块(CBAM), 即在特征融合前插入该模块, 特征注意力增强, 提高菠萝识别和定位的准确度。

传统卷积对每个通道的输入特征映射应用卷积核, 通过对每个通道的结果进行叠加来生成输出特征映射, 这一过程在所有输入通道上重复, 生成多个输出特征映射。相反, 深度可分离卷积可分解为 2 个步骤: ①深度卷积层对输入特征图的每个通道使用

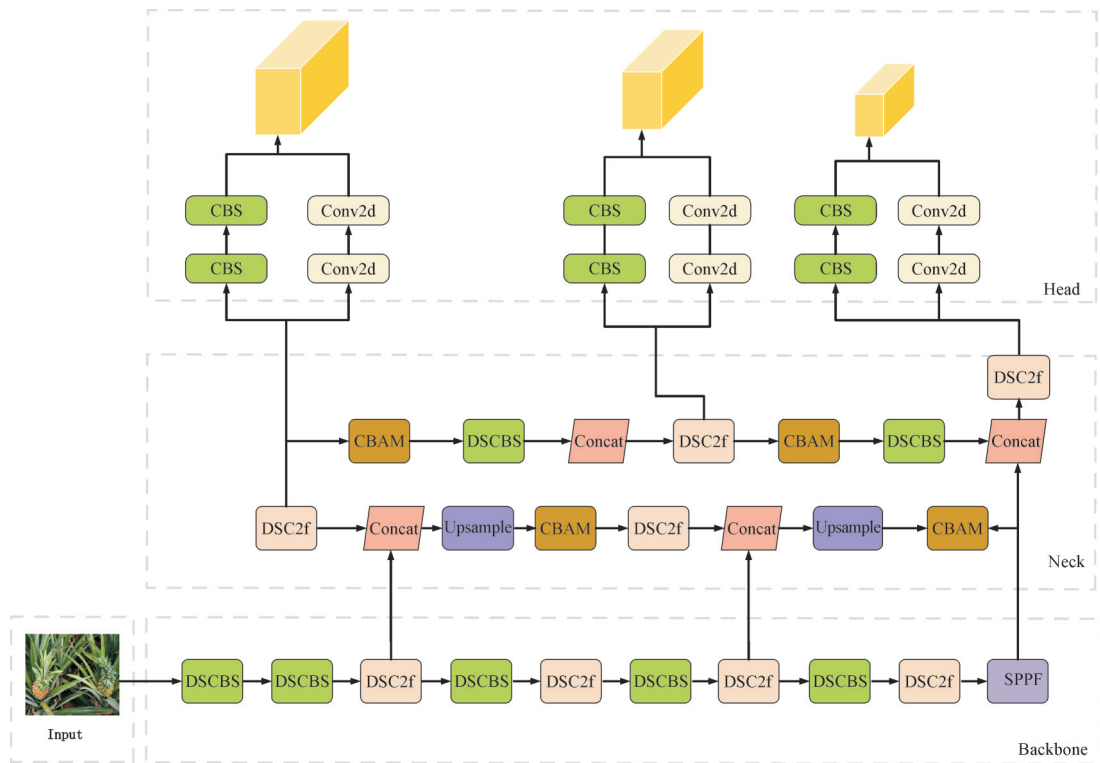


图 2 改进的 YOLOv8 模型

Fig. 2 Improved YOLOv8 model

单独的卷积核进行单独卷积操作后输出相应的特征图;②逐点卷积层前一阶段输出的特征图作为输入,每个卷积核只针对单个位置进行卷积操作,但需要处理输入特征图的所有通道。普通卷积由大小为 $3 \times 3 \times C$ 的卷积核实现,深度可分离卷积由大小分别为 $3 \times 3 \times 1$ 和 $1 \times 1 \times C$ 的 2 个卷积核组成,如图 3 所示。

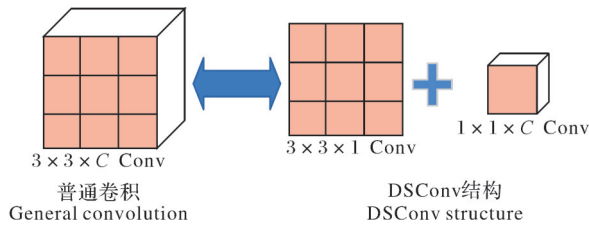


图 3 深度可分离卷积

Fig. 3 Depthwise separable convolution

深度可分离卷积与传统卷积的比值如公式(1)所示,相比普通卷积,深度可分离卷积的参数量更少。

$$R = \frac{1}{N} + \frac{1}{k^2} \tag{1}$$

式(1)中, R 为参数数量的比值, N 是输入通道的数量, k 代表卷积核的大小。

在特征提取过程中,YOLOv8 模型对不同特征采取相同的加权方式,导致目标的漏检和误检,不利于对菠萝小尺度目标信息的提取,因此本研究引入 CBAM 注意力机制,通过通道注意力模块和空间注意力模块有效地增加整个特征图中遮挡或小目标的权重,使信息更容易被网络学习。CBAM 注意力过程如图 4 所示。

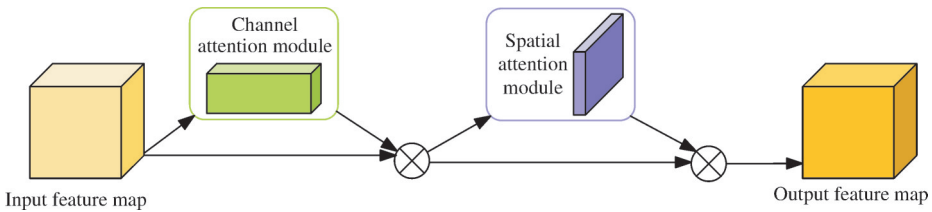


图 4 CBAM 注意力过程图

Fig. 4 CBAM attention process diagram

将提取的菠萝图像特征标记为 F ,利用特征之间的通道关系生成通道注意力特征图 $M_c(F)$,再与 F 逐点相乘形成新的特征 F' ,增强与菠萝果实目标区域相关的特征。然后,利用特征之间的内部空间关系生成空间注意力特征图 $M_s(F')$,与 F' 逐点相乘得到 F'' ,从通道和特征之间的空间关系上增强了菠萝目标区域特征的权重,计算公式如式(2)所示:

$$\begin{aligned} F' &= M_c(F) \otimes F \\ F'' &= M_s(F') \otimes F' \end{aligned} \quad (2)$$

式(2)中, \otimes 代表逐点相乘。

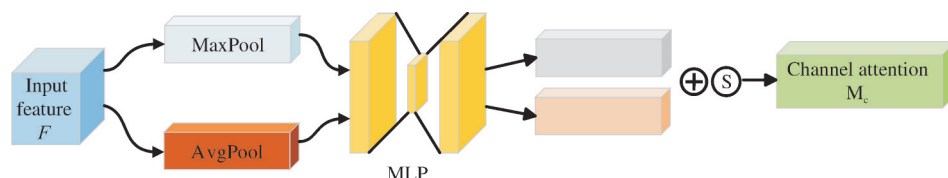


图5 通道注意力模块结构图

Fig. 5 Channel attention module structure diagram

$$M_c(F) = \sigma \left(\begin{aligned} &W_1(W_0(F_{avg}^c)) \\ &+ W_1(W_0(F_{max}^c)) \end{aligned} \right) \quad (3)$$

式(3)中, σ 表示sigmoid激活函数, F_{avg}^c 和 F_{max}^c 表示平均池化和最大池化后的特征, W_0 和 W_1 表示多层感知中的两层权重。

空间注意力模块侧重于局部信息,通过池化对信息进行过滤,然后对过滤后的信息进行卷积提取重要信息,如图6所示,将 F' 作为空间注意力模块的输入,进行最大池化和平均池化,通过Concat运算进行叠加,然后通过卷积运算和sigmoid激活函数得到空间注意力特征图 M_s ,计算公式如式(4)所示:

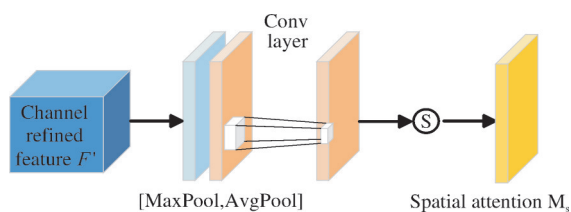


图6 空间注意力模块结构图

Fig. 6 Spatial attention module structure diagram

$$M_s(F) = \sigma \left[f^{5 \times 5} \left([F_{avg}^s; F_{max}^s] \right) \right] \quad (4)$$

式中, σ 表示sigmoid激活函数, F_{avg}^s 和 F_{max}^s 表示平均池化特征和最大池化特征, $f^{5 \times 5}$ 表示 5×5 卷积。

1.5 损失函数

对于原YOLOv8模型,其分类损失为Varifocal Loss (VFL),其回归损失为CIoU Loss+DFL(distribution focal loss)的形式,为了更好地提高识别准确率,本研究考虑改进损失函数,将CIoU Loss替换成EIoU Loss^[19]。

分别采用最大池化和平均池化提取通道注意力模块信息,然后进行过滤、激活和归一化,提高通道信息的提取能力。通道注意力模块的结构如图5所示,首先,输入特征 F ,使用最大池化和平均池化获取每个特征通道的全局信息,然后将特征图输入到MLP(multilayer perceptron)多层感知器中并将池化后的信息进行加权求和,通过sigmoid激活函数对结果进行处理,最终得到通道注意力特征图 M_c ,其中包含对输入特征图 F 不同通道关注程度的信息,计算公式如式(3)所示:

VFL提出了非对称的加权操作,计算公式如式(5)所示:

$$VFL(p, q) = \begin{cases} -q \left(q \log(p) + (1-q) \log(1-p) \right) \\ -\alpha p^q \log(1-p) \end{cases} \quad (5)$$

式(5)中, q 为预测框和标签框的交并比, p 代表概率。

YOLOv8使用的交并比(intersection over union, IoU),损失函数为CIoU,它考虑了预测框和真实框之间的重叠面积、中心点距离以及宽高比^[20]。然而,对于宽高比的描述仍然存在一定的模糊性,计算公式如式(6)~(8)所示:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (6)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (7)$$

$$\alpha = \frac{v}{(1 - IoU + v)} \quad (8)$$

式(6)~(8)中,IoU为交并比, ρ 表示2个矩形框之间的欧式距离, b 和 b^{gt} 分别代表预测框和标签框, c 代表2个矩形框闭合区域的对角线距离, α 为权重系数, v 用于测量预测框和标签框相对比例的一致性, w^{gt} 和 h^{gt} 为标签框的宽度和高度, w 和 h 分别代表预测框的高度和宽度。

EIoU在CIoU的基础上将长宽比分解为长宽的差异值,以加速预测框的回归速度,使回归过程更专注于高质量的锚框,进而提高预测框的回归精度。此外,EIoU优化了边框回归中的样本不平衡问题,

样本不平衡问题是指一幅图像中回归误差小的高质量锚框数量远远小于误差大的低质量锚框数量样本,样本质量差会产生过大的梯度,影响训练过程,计算公式如式(9)所示:

$$\begin{aligned} L_{\text{IoU}} &= L_{\text{IoU}} + L_{\text{dis}} + L_{\text{asp}} \\ &= 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} \\ &\quad + \frac{\rho^2(w, w^{\text{gt}})}{c_w^2} + \frac{\rho^2(h, h^{\text{gt}})}{c_h^2} \end{aligned} \quad (9)$$

式(9)中, ρ 表示2个矩形框之间的欧式距离, b 和 b^{gt} 分别代表预测框和标签框, w 和 h 分别代表预测框的高度和宽度, c 代表2个矩形框闭合区域的对角线距离, w^{gt} 和 h^{gt} 为标签框的宽度和高度, c_w 和 c_h 是覆盖预测框和标签框的最小外部矩形的宽度和高度。

DFL以交叉熵形式优化了与标签 y 最接近的左右2个位置的概率,从而让网络更快地聚焦于目标位置附近的分布情况,计算公式如式(10)所示:

$$\text{DFL}(S_i, S_{i+1}) = - \left(\begin{aligned} &(y_{i+1} - y) \log(S_i) \\ &+ (y - y_i) \log(S_{i+1}) \end{aligned} \right) \quad (10)$$

式(10)中, S 是输出分布, y_i 和 y_{i+1} 是浮点值 y 的左右整数值。

1.6 模型评价指标

本研究模型采用精确度(precision, P)、召回率(recall, R)、平均精度均值(mean of average precision, P_{mA})作为评价指标。具体来说, P 表示预测的算法面积与实际检测面积的比值, R 表示准确预测的类别占所需类别总数的比例, P_{mA} 基于 Precision-Recall 曲线得出,能够全面评估算法的准确性和鲁棒性, P_{mA} 值越高,说明预测精度越高。此外,本研究还利用推理时间和网络参数体积作为性能指标,较短的推理时间意味着更好的实时性能,而较小的模型大小则代表较小的内存使用。其中, T_p 表示真阳性, F_p 表示假阳性, N 表示样本总数, F_N 表示漏报的真阳性数量,本研究中检测到的菠萝总类别数, P_{iA} 表示每个类别平均精度的平均值, Q 表示菠萝总类别数。计算公式如式(11)~(14)所示:

$$P_{iA} = \frac{T_p}{T_p + F_p} \quad (11)$$

$$P_{\text{mA}} = \frac{\sum_{i=1}^Q P_{iA}}{Q} \times 100\% \quad (12)$$

$$P = \frac{T_p}{T_p + F_p} \quad (13)$$

$$R = \frac{T_p}{T_p + F_N} \quad (14)$$

2 结果与分析

2.1 环境及参数调整

本研究中使用的操作系统是 Windows 10, PyTorch 作为深度学习模型的框架(表4)。在训练阶段,采用随机梯度下降(SGD)进行优化,初始学习率为0.01,动量因子为0.937,权重衰减系数为0.000 5,输入图像分辨率设置为640像素×640像素,批尺寸为16,交并比为0.5,使用改进YOLOv8在数据集上训练得到的权重来进行迁移学习并训练300个轮次(Epoch)。

表4 试验环境配置

Table 4 Experimental environment configuration

类别 Type	配置 Configuration
中央处理器 Central processing unit	Intel®core™i9-9900CPU @3.10 GHz 32 GB
图形处理器 Graphics processing unit	GeForce RTX 2070 Super 8 GB
系统环境 System environment	Windows10
框架 Framework	PyTorch 1.11.0
编程语言 Programming language	Python 3.8

2.2 模型训练过程

在训练过程中记录模型的损失函数、精确度 P 、召回率 R 以及 $P-R$ 曲线。由图7可以看出,模型训练30轮之前其损失函数下降速度快,损失曲线平滑;当训练轮数大于30、小于180时,损失函数稳定下降,并伴随小范围的波动;模型经过300轮训练后损失值完成收敛为0.023。最终模型精确度为97.38%,召回率为95.57%,平均精度均值为97.33%。

2.3 与YOLOv8模型的比较

本研究通过深度可分离卷积替换公共卷积,提高了小目标和遮挡目标的识别效果,如图8所示,使用深度可分离卷积的模型能够有效识别小目标菠萝以及叶片遮挡较为严重的目标。

本研究采用改进YOLOv8模型在测试集进行性能测试。检测结果如表5所示,改进YOLOv8模型在未成熟、近生理成熟及成熟果实识别试验中精确度、召回率、平均精度较YOLOv8模型有明显提高,平均检测时间为5.4 ms,相较于原始YOLOv8模型平均精度均值提升了5.51百分点,平均检测时间减少了0.1 ms。由于改进的网络中引入了注意力机制使模型更加关注重要的区域,表现为较高的精确度和召回率。改进YOLOv8网络与原始网络的检测效果如图9所示,改进模型可以更加准确地辨识目标并

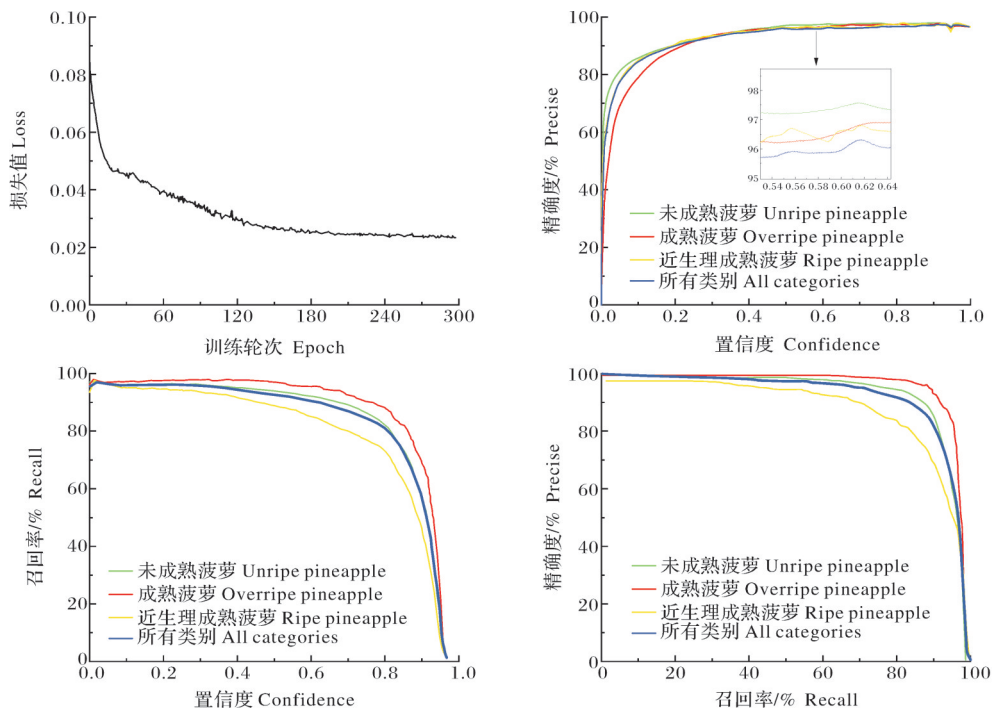


图7 训练过程图

Fig. 7 Training process diagram



图8 卷积替换效果对比

Fig. 8 Comparison of convolutional substitution effect

表 5 改进 YOLOv8 模型与原始模型的试验结果对比

Table 5 Comparison of experimental results between the improved YOLOv8 model and the original model						
网络模型 Model	类别 Type	精确度/% P	召回率/% R	平均精度/% P_A	平均精度均值/% P_{mA}	平均检测时间/ms Average detection time
YOLOv8	a	93.10	91.90	92.50	91.82	5.5
	b	91.52	89.48	90.50		
	c	92.76	92.16	92.46		
改进 YOLOv8 Improved YOLOv8	a	97.30	97.12	97.21	97.33	5.4
	b	98.87	93.55	95.73		
	c	99.24	96.22	99.09		

注 Note:a:未成熟 Unripe;b:近生理期成熟 Ripe;c:成熟 Overripe. 下同。The same as below.

获得更高的置信度,同时避免原始模型中出现的漏检现象。

2.4 消融试验性能对比

为评估改进后算法的有效性,本研究使用原始

网络与改进的 YOLOv8 网络在菠萝数据集上进行消融试验。由表 6 可知,在同一数据集的消融试验中,加入 DSConv 模块对比原模型 YOLOv8,其精确度、召回率和平均精度均值分别下降 0.80、2.69、1.66 百

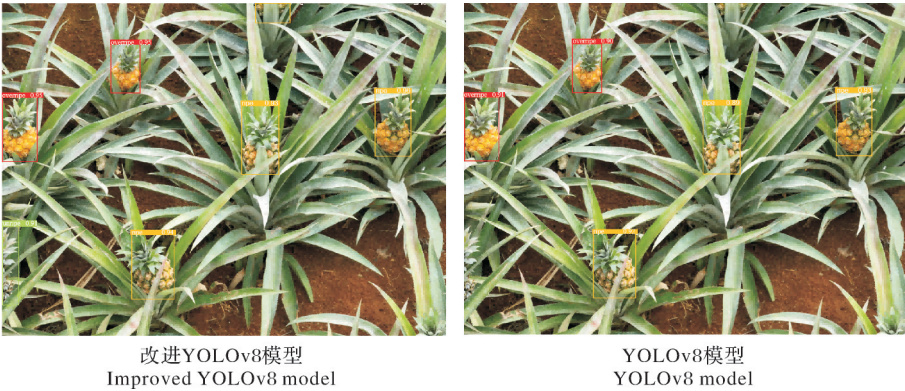


图 9 改进前后检测网络识别效果

Fig. 9 Identification effect of the improved pre- and post- detection network

分点,每秒帧数增加 17.4 帧,模型参数量由 22.6×10^6 减少到 6.0×10^6 。特征融合前加入 CBAM 模块后,其精确度和平均精度均值分别上升 1.6、1.25 百分点,召回率下降 0.37 百分点,每秒帧数下降 2.4 帧,参数量增加 1.1×10^6 。相较于原始 YOLOv8 模型,改进后模型的精确度、召回率和平均精度均值分别提升了 5.27、3.98、5.51 百分点,在参数量整体减少 6.2×10^6 的情况下,每秒帧数下降 13.7 帧。

表 6 消融试验结果

Table 6 Results of ablation experiments

模型 Model	精确 度/% P	召回 率/% R	平均精度 均值/% P_{mA}	帧率/ (帧/s) FPS	参数量 Parameter count
YOLOv8	92.10	91.59	91.82	149.1	22.6×10^6
YOLOv8+DSConv	91.30	88.90	90.16	166.5	6.0×10^6
YOLOv8+CBAM	93.70	91.22	93.07	146.7	23.7×10^6
YOLOv8+ DSConv+CBAM	97.38	95.57	97.33	135.4	16.8×10^6

2.5 模型对比试验

为了比较本研究提出的改进算法的效率,选择双阶段经典模型 Faster R-CNN 和单阶段经典模型 YOLOv4、YOLOv5、YOLOv7 以及 YOLOv8 进行对比试验。由表 7 可知,利用改进的 YOLOv8 模型进行目标检测,对未成熟期、近生理成熟期及成熟期的菠萝检测精确度分别达到 97.30%、98.87% 和 99.24%,召回率分别为 97.12%、93.55% 和 96.22%,平均精度分别为 97.21%、95.73% 和 99.09%,均高于其他基础模型。同时,改进后的 YOLOv8 模型相对于 YOLOv7 模型参数量增加 3.8×10^6 ,识别速度增加 27.4 帧/s,对比原 YOLOv8 模型,识别速度接近,模型参数量减少 5.8×10^6 ,平均精度均值增加 5.51 百分点。

2.6 模型检测效果分析

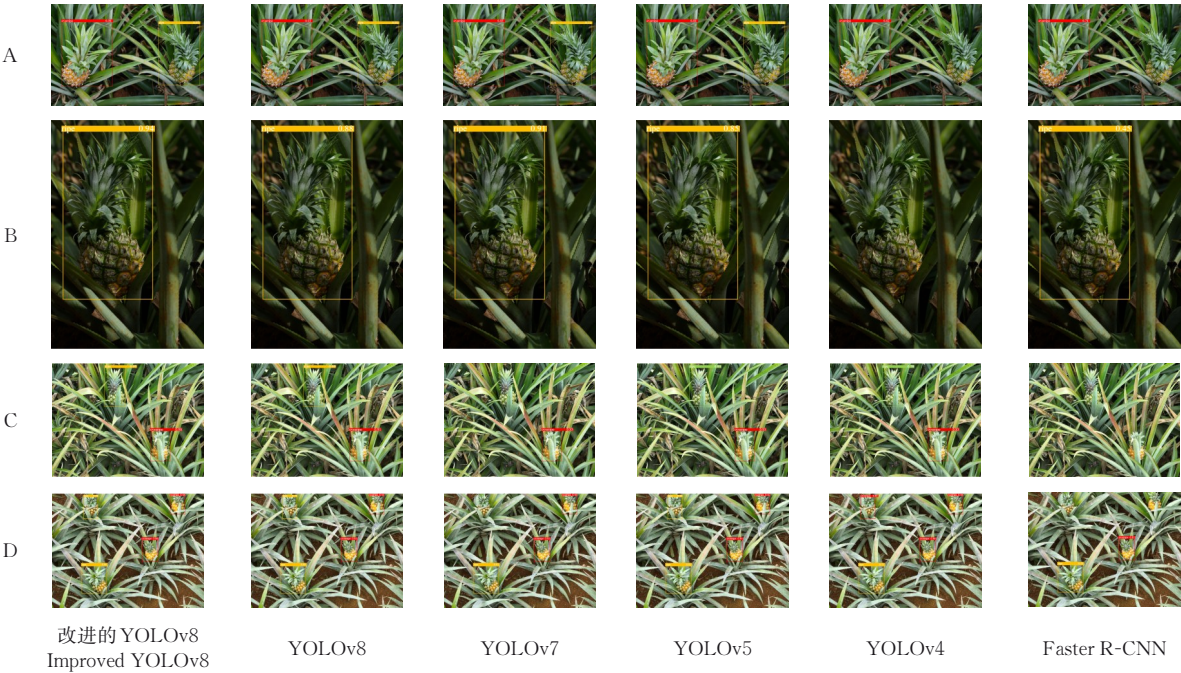
本研究检测了不同目标检测模型在复杂环境下针对菠萝成熟度的检测效果,由图 10 可知,改进 YOLOv8 模型在自然光条件下可准确检测出菠萝的 3 种成熟度,而其他模型均存在漏检现象;在逆光条件下,改进 YOLOv8 模型测试效果最佳,YOLOv4 模型存在漏检现象;当菠萝处于遮挡情况下,YOLOv7 和 Faster R-CNN 出现漏检现象,YOLOv5 和 YOLOv4 出现误检,改进 YOLOv8 模型一定程度上避免了漏检、误检的情况。在中远距离的条件下,YOLOv7、Faster R-CNN 出现漏检现象,YOLOv4 存在漏检现象且检测置信度较低,改进 YOLOv8 模型的漏检情况有明显的改善,能够准确地检测出图中目标。综上所述,改进 YOLOv8 模型鲁棒性强,能够实现在多种复杂环境下对菠萝成熟度的精确检测。

3 讨 论

为了解决现有菠萝成熟度检测模型无法准确识别不同环境下的菠萝果实成熟度,本研究对 YOLOv8 模型进行了改进,采用深度可分离卷积(DSConv)替换 YOLOv8 的公共卷积、引入 CBAM 注意力机制以及使用 EIoU 损失函数。改进模型对菠萝成熟度检测 P_{mA} 为 97.33%,与 Faster R-CNN、YOLOv4、YOLOv5、YOLOv7 对比发现, P_{mA} 分别提升 5.53、7.91、4.38、4.66 百分点;在保证检测精度的前提下,算法模型参数量仅为 16.8×10^6 。这一结果表明所提出的改进模型在精确度、召回率和平均精度均值方面均明显优于其他模型,相对于其他模型在中远距离和遮挡情况下实现了更优异的目标检测性能。同时,本次研究证明了使用 YOLOv8 对菠萝成熟度检测的可行性,实验数据相比一些基础模型有

表 7 不同网络模型的测试对比
Table 7 Test comparison of different network models

模型 Model	类别 Type	精确度/% P	召回率/% R	平均精度/% P_A	平均精度均值/% P_{mA}	帧率/(帧/s) Frame per second	参数量 Parameter count
Faster R-CNN	a	93.10	91.50	92.30	91.80	10.0	151.3×10^6
	b	89.30	86.90	88.10			
	c	96.50	93.50	95.00			
YOLOv4	a	91.75	89.25	90.50	89.42	63.0	25.6×10^6
	b	91.50	86.50	89.00			
	c	91.20	86.32	88.76			
YOLOv5	a	93.20	92.30	92.75	92.95	99.0	15.8×10^6
	b	90.50	90.00	90.25			
	c	96.17	95.53	95.85			
YOLOv7	a	95.69	94.29	94.99	92.67	108.0	13.0×10^6
	b	91.49	90.15	90.82			
	c	94.30	91.30	92.80			
YOLOv8	a	93.10	91.90	92.50	91.82	149.1	22.6×10^6
	b	91.52	89.48	90.50			
	c	92.76	92.16	92.46			
改进的 YOLOv8 Improved YOLOv8	a	97.30	97.12	97.21	97.33	135.4	16.8×10^6
	b	98.87	93.55	95.73			
	c	99.24	96.22	99.09			



A:顺光条件Frontlighting condition;B:逆光 Backlighting condition;C:遮挡Occlusion condition;D:中远距离 Long-distance condition.

图 10 不同目标检测网络在复杂环境下针对菠萝成熟度的检测效果

Fig. 10 The effect of different target detection networks on pineapple ripeness in complex environment

很大的改进,与李阳德等^[16]的基于轻量级 Mobile-NetV3-YOLOv4 的生长期菠萝成熟度分析(1 580 张样本图,数据集划分为青熟期和黄熟期 2 个类别)相

比,本研究样本数量(2 520 张)更多,进一步将数据集划分成未成熟、近生理期成熟和成熟 3 个类别。因此,本研究模型的数据集划分更加科学,具有更强的

泛化能力,能够满足机械采摘菠萝识别监测的需要。因此,未来研究工作将围绕与应用嵌入式设备与智能采摘机器人相结合,实现更高效、高质量的采摘。

本研究虽然在算法上进行了优化,并取得了显著的效果,但并未结合农业机器人进行实际应用,考虑到农业作业环境中不确定的因素,对于本研究是一个很大的挑战。后续的工作中,需考虑环境因素的影响,针对环境不确定因素进行算法优化,实现搭载农业机器人针对菠萝的自动化采摘进行算法辅助,以提高菠萝采摘的速度并节省人力成本。

参考文献 References

- [1] 方伟. 广东省菠萝产业现状及发展建议[J]. 中国果树, 2023(6): 123-126. FANG W. Present situation and development suggestion of pineapple industry in Guangdong Province[J]. China fruits, 2023(6): 123-126 (in Chinese with English abstract).
- [2] TANTINANTRAKUN A, SUKWANIT S, THOMPSON A K, et al. Nondestructive evaluation of SW-NIRS and NIR-HSI for predicting the maturity index of intact pineapples[J/OL]. Postharvest biology and technology, 2023, 195: 112141 [2023-12-07]. <https://doi.org/10.1016/j.postharvbio.2022.112141>.
- [3] 姜涛, 郭安福, 程学斌, 等. 菠萝自动采摘收集机结构设计与分析[J]. 工程设计学报, 2019, 26(5): 577-586. JIANG T, GUO A F, CHENG X B, et al. Structural design and analysis of pineapple automatic picking-collecting machine[J]. Chinese journal of engineering design, 2019, 26(5): 577-586 (in Chinese with English abstract).
- [4] 薛忠, 陈如约, 张秀梅. 菠萝机械化种植与收获研究现状[J]. 山西农业大学学报(自然科学版), 2021, 41(3): 110-120. XUE Z, CHEN R Y, ZHANG X M. Frontier of pineapple planting and harvesting mechanization in the world[J]. Journal of Shanxi Agricultural University (natural science edition), 2021, 41(3): 110-120 (in Chinese with English abstract).
- [5] 吴莎莎, 王振杰, 江梦薇, 等. 基于多成熟度光谱信息融合的阿森泰克苹果品质预测模型研究[J]. 食品工业科技, 2024, 45(7): 294-305. WU S S, WANG Z J, JIANG M W, et al. Prediction model of Aztec apples quality based on the fusion of multi-maturity spectral information[J]. Science and technology of food industry, 2024, 45(7): 294-305 (in Chinese with English abstract).
- [6] 杨圣慧, 郑永军, 刘星星, 等. 基于无人机近地面多光谱图像的蛇龙珠葡萄成熟度判别[J]. 光谱学与光谱分析, 2021, 41(10): 3220-3226. YANG S H, ZHENG Y J, LIU X X, et al. Cabernet gemischt maturity determination based on near-ground multispectral figures by using UAVs[J]. Spectroscopy and spectral analysis, 2021, 41(10): 3220-3226 (in Chinese with English abstract).
- [7] FU L H, YANG Z, WU F Y, et al. YOLO-banana: a lightweight neural network for rapid detection of banana bunches and stalks in the natural environment[J/OL]. Agronomy, 2022, 12(2): 391 [2023-12-07]. <https://doi.org/10.3390/agronomy12020391>.
- [8] IGNACIO J S, EISMA K N A, CAYA M V C. A YOLOv5-based deep learning model for *in-situ* detection and maturity grading of mango[C]//2022 6th International Conference on Communication and Information Systems (ICCIS). Chongqing, China. Chongqing: IEEE, 2022: 141-147.
- [9] 熊俊涛, 韩咏林, 王潇, 等. 基于YOLOv5-Lite的自然环境木瓜成熟度检测方法[J]. 农业机械学报, 2023, 54(6): 243-252. XIONG J T, HAN Y L, WANG X, et al. Method of maturity detection for papaya fruits in natural environment based on YOLOv5-lite[J]. Transactions of the CSAM, 2023, 54(6): 243-252 (in Chinese with English abstract).
- [10] 陈仁凡, 谢知, 林晨. 基于YOLO-ODM的温室草莓成熟度的快速检测[J]. 华中农业大学学报, 2023, 42(4): 262-269. CHEN R F, XIE Z, LIN C. YOLO-ODM based rapid detection of strawberry ripeness in greenhouse[J]. Journal of Huazhong Agricultural University, 2023, 42(4): 262-269 (in Chinese with English abstract).
- [11] 马瑞峻, 何浣冬, 陈瑜, 等. 基于改进YOLOv5的火龙果成熟度识别方法[J]. 沈阳农业大学学报, 2023, 54(2): 196-206. MA R J, HE H D, CHEN Y, et al. Recognition method for maturity of pitaya based on improved YOLOv5[J]. Journal of Shenyang Agricultural University, 2023, 54(2): 196-206 (in Chinese with English abstract).
- [12] 苗荣慧, 李志伟, 武锦龙. 基于改进YOLOv7的轻量化樱桃番茄成熟度检测方法[J]. 农业机械学报, 2023, 54(10): 225-233. MIAO R H, LI Z W, WU J L. Lightweight maturity detection of cherry tomato based on improved YOLO v7[J]. Transactions of the CSAM, 2023, 54(10): 225-233 (in Chinese with English abstract).
- [13] 李斌, 汪懋华, 李莉. 基于单目视觉的田间菠萝果实识别[J]. 农业工程学报, 2010, 26(10): 345-349. LI B, WANG M H, LI L. In-field pineapple recognition based on monocular vision[J]. Transactions of the CSAE, 2010, 26(10): 345-349 (in Chinese with English abstract).
- [14] 张星, 高巧明, 潘栋, 等. 基于改进YOLOv3的田间复杂环境下菠萝拾捡识别研究[J]. 中国农机化学报, 2021, 42(1): 201-206. ZHANG X, GAO Q M, PAN D, et al. Picking recognition research of pineapple in complex field environment based on improved YOLOv3[J]. Journal of Chinese agricultural mechanization, 2021, 42(1): 201-206 (in Chinese with English abstract).
- [15] CUONG N H H, TRINH T H, MEESAD P, et al. Improved YOLO object detection algorithm to detect ripe pineapple phase[J]. Journal of intelligent & fuzzy systems, 2022, 43(1): 1365-1381.
- [16] 李阳德, 马晓慧, 王骥. 基于轻量级MobileNetV3-YOLOv4

- 的生长期菠萝成熟度分析[J]. 智慧农业(中英文), 2023, 5(2): 35-44. LI Y D, MA X H, WANG J. Pineapple maturity analysis in natural environment based on MobileNet V3-YOLOv4[J]. Smart agriculture, 2023, 5(2): 35-44 (in Chinese with English abstract).
- [17] 中华人民共和国农业部. NY/T 450—2001, 菠萝[S]. 北京: 中国农业出版社, 2001. The Ministry of Agriculture of People's Republic of China. NY/T 450—2001, Pineapple [S]. Beijing: China Agriculture Press, 2001(in Chinese).
- [18] 张阳婷, 黄德启, 王东伟, 等. 基于深度学习的目标检测算法研究与应用综述[J]. 计算机工程与应用, 2023, 59(18): 1-13. ZHANG Y T, HUANG D Q, WANG D W, et al. Review on research and application of deep learning-based target detection algorithms [J]. Computer engineering and applications, 2023, 59(18): 1-13 (in Chinese with English abstract).
- [19] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IoU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146-157.
- [20] 杨永刚, 谢睿夫, 龚泽川. 改进YOLOv7-tiny的无人机目标检测算法[J]. 计算机工程与应用, 2024, 60(6): 121-129. YANG Y G, XIE R F, GONG Z C. Improved YOLOv7-tiny UAV target detection algorithm[J]. Computer engineering and applications, 2024, 60(6): 121-129 (in Chinese with English abstract).

Real-time object detection method of pineapple ripeness based on improved YOLOv8

ZHOU Tao¹, WANG Ji^{1,2}, MAI Rengui³

1. College of Electronic and Information Engineering, Guangdong Ocean University, Zhanjiang 524088, China;

2. Guangdong Intelligent Ocean Sensor Network and Equipment Engineering Technology Research Center, Zhanjiang 524088, China;

3. College of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang 524088, China

Abstract A real-time object detection method of pineapple ripeness based on improved YOLOv8 was proposed to improve the accuracy of mechanical harvesting of pineapples in planting areas with different ripeness and ensure the quality of pineapples. The improved model replaced the common convolutions in the backbone and neck parts of the original YOLOv8 model with depth-wise separable convolutions (DSConv) to streamline parameters of model to solve the problems of small object size, dense quantity, and light occlusion of picked mechanical pineapple picking in natural environments. Convolutional block attention mechanism (CBAM) module was introduced before feature fusion to prioritize important features and improve the accuracy of object detection. The original loss function CIOU of YOLOv8 network was replaced with the EIOU loss function to accelerate the speed of network convergence. The results showed that the mean of average precision (P_{mA}) of the improved model for detecting the pineapple ripeness was 97.33%. The P_{mA} of improved model was 5.53, 7.91, 4.38, and 4.66 percentage points higher than that of Faster R-CNN, YOLOv4, YOLOv5 and YOLOv7, respectively. The number of parameters of the algorithm model was only 16.8×10^6 on the premise of ensuring the accuracy of detection. It is indicated that the improved model improves the accuracy and inference speed of recognizing pineapple ripeness, and has stronger robustness.

Keywords pineapple fruit; ripeness; YOLOv8; object detection

(责任编辑: 葛晓霞)