

DOI: 10.13733/j.jcam.issn.2095-5553.2025.07.019

陈萍, 古丽巴哈尔·托乎提, 张国辉, 等. 基于改进 YOLOv7 的自然环境乌梅成熟度检测方法[J]. 中国农机化学报, 2025, 46(7): 124-130

Chen Ping, Gulbahar Tohti, Zhang Guohui, et al. Maturity detection method of black plum in the natural environment based on improved YOLOv7 [J]. Journal of Chinese Agricultural Mechanization, 2025, 46(7): 124-130

基于改进 YOLOv7 的自然环境乌梅成熟度检测方法^{*}

陈萍¹, 古丽巴哈尔·托乎提^{1, 2}, 张国辉¹, 买买提明·艾尼¹

(1. 新疆大学机械工程学院, 乌鲁木齐市, 830017; 2. 西安交通大学, 西安市, 710049)

摘要:不同成熟度的乌梅具有不同的药理作用, 为判断果园中大部分果实的成熟度, 基于 YOLOv7 目标检测算法进行一系列改进。在 YOLOv7 模型的 Backbone 中添加 BiFormer 模块以提高网络的特征表达能力; 设计果实成熟度精分模块以提高果实成熟度检测的正确率。研究表明, 改进的 YOLOv7—1 模型平均精度均值 *mAP* 达到 0.805, 比改进的 YOLOv7—2 模型、Faster R—CNN 模型、YOLOv3 模型、Mask R—CNN 模型、YOLOv5s 模型、YOLOv5l 模型、YOLOv7 模型和 YOLOv8 模型分别高 4.8、12.4、0.9、0.7、12.6、1.7、5.8 和 12.3 个百分点。改进的 YOLOv7—1 模型可以提高乌梅成熟度识别的准确性。

关键词: 乌梅; 成熟度检测; 深度学习; 自然环境

中图分类号: S567; TP399 **文献标识码:** A **文章编号:** 2095-5553 (2025) 07-0124-07

Maturity detection method of black plum in the natural environment based on improved YOLOv7

Chen Ping¹, Gulbahar Tohti^{1, 2}, Zhang Guohui¹, Mamtimin Geni¹

(1. College of Mechanical Engineering, Xinjiang University, Urumqi, 830017, China;

2. Xi'an Jiaotong University, Xi'an, 710049, China)

Abstract: Black plum with different maturity has different pharmacological effects. To judge the maturity of most fruits in the orchard, a series of improvements are carried out based on the YOLOv7 target detection algorithm. The Vision Transformer with Bi-Level Routing Attention (BiFormer) module is added to the Backbone of YOLOv7 model to improve the feature expression ability of the network. The fruit maturity refinement module is designed to improve the correct rate of fruit maturity detection. The study shows that the improved YOLOv7—1 model has a Mean Average Precision (*mAP*) of 0.805, and is higher by 4.8, 12.4, 0.9, 0.7, 12.6, 1.7, 5.8 and 12.3 percentage points, respectively, compared with the improved YOLOv7—2 model, Faster R—CNN model, YOLOv3 model, Mask R—CNN model, YOLOv5s model, YOLOv5l model, YOLOv7 model, and YOLOv8 model. The improved YOLOv7—1 model can improve the accuracy of identifying the maturity of black plum.

Keywords: black plum; maturity detection; deep learning; natural environment

0 引言

乌梅, 别名酸梅、黑莓, 蔷薇科植物梅的果实, 具有食用价值、药用价值、生态价值^[1]。我国是乌梅的原产地, 不同成熟度的乌梅作用大不相同^[2]。成熟的乌梅分布不均会加大采摘难度, 也会影响果农的收益, 因此快速准确地识别乌梅, 提供乌梅不同成熟度分布信息,

成为合理安排劳动力以及适时采摘的必要条件。

乌梅的成熟度主要依靠肉眼判断果实表皮颜色加以区分。传统机器视觉主要通过果实表皮的颜色特征进行提取分割, 建立果实的成熟度模型。刘宇飞等^[3]基于机器视觉对番茄表面红色面积检测, 进行颜色分级, 根据颜色特征的关系来判断番茄的表面成熟度等级。这个分级方法简易, 减少运算时间, 但有一定的局

收稿日期: 2023 年 12 月 30 日 修回日期: 2024 年 3 月 29 日

^{*} 基金项目: 国家自然科学基金地区基金项目(12162031); 西安交通大学机械制造系统工程国家重点实验室(sklms2022022)

第一作者: 陈萍, 女, 1998 年生, 重庆人, 硕士研究生; 研究方向为机器视觉和智能装备。E-mail: chenping19980802@163.com

通讯作者: 古丽巴哈尔·托乎提, 女, 1972 年生, 乌鲁木齐人, 副教授; 研究方向为智能机器人、计算机图像处理等。E-mail: gulbahart@163.com

限性。周文静等^[4]将红提葡萄果穗用KNN模型从背景中分割出来,再运用圆形Hough变换提取出葡萄果实,提取果实的H颜色分量值,设置特定的颜色阈值对单个果实成熟度进行判定,经过统计后再判断整个果串的成熟度等级。上述机器学习进行果实检测仍存在检测精度较低、检测时间较长等问题。

随着深度学习R-CNN、VGG、Faster R-CNN、YOLO算法、Mask R-CNN的发展,并在各种视觉任务中有较好的检测效果,一些学者开始利用深度学习对水果成熟度进行识别分类。王立舒等^[5]提出一种改进YOLOv4-Tiny模型的蓝莓成熟度识别方法,蓝莓成熟度在自然环境中检测精度较高。但局限于未成熟、半成熟和成熟状态,不能对蓝莓整个成熟过程进行全面把控。李竹等^[6]利用VGG16神经网络,提取果皮颜色特征建立模型,对蓝莓成熟度的预测较为精细,提高了对蓝莓果实成熟度的预测。王勇等^[7]提出一种基于改进YOLOv5s模型的不同成熟度苹果目标检测方法,能够准确定位不同成熟度苹果的特征区域。苗荣慧等^[8]提出一种改进的轻量化YOLOv7模型的樱桃番茄果实成熟度检测方法,提高自然环境中相邻成熟度果实和被遮挡果实的检测精度,可见深度学习在果实成熟度检测研究中优势显著。其中Bochkovskiy团队^[9]提出YOLOv7模型,具有检测精度高、速度快、支持多种类型目标等优点,在检测性能上优于其他模型。因此,选取YOLOv7模型为基础模型用于自然环境乌梅成熟度的检测。

在自然环境中,大多数检测算法不能同时满足精度高、速度快、计算量小的要求。对于平衡好检测速度、检测精度和模型计算复杂度的乌梅成熟度检测方法的研究还较少。乌梅具有以下特性:乌梅形状较小而密集,果实之间往往存在遮挡问题。乌梅同一花序所结果实成熟度不一致,且相邻成熟度特征差异不明显。针对以上问题,本文提出一种改进YOLOv7的目标检测方法。在YOLO7的Backbone中加入BiFormer注意力模块^[10, 11],并设计果实成熟度精分模块^[12],加强模型对全局信息的捕捉能力,提高对乌梅相邻成熟度的区分能力,提高自然环境中乌梅成熟度检测效果,以确定果园成熟果实的分布区域,为乌梅的采摘提供参考依据。

1 材料与方法

1.1 数据采集

数据集是由智能手机在自然环境下收集的,采集在不同成熟时间段的乌梅图像共1106张。采集地位于新疆维吾尔自治区喀什地区莎车县米夏镇塔瓦克斯村家庭乌梅种植果园,该种植果园涵盖自然环境下乌梅树的多种复杂情况的需求。采集时间为2023年5月23—

26日、2023年7月15—18日、2023年8月2—5日、2023年9月13—16日,包含早、中、晚时间段,有晴天、阴天;有顺光、逆光、背光拍摄方式,包含自然环境下光线明暗不同、乌梅堆叠程度不同、树叶遮挡程度不同、成熟度不同,数据集贴近自然状态下的乌梅。

1.2 乌梅成熟度区分

成熟度区分有很多类型,果皮颜色是衡量果实成熟度的重要指标之一。在乌梅成熟过程中,果实里的各种激素的转换,颜色会相应地发生变化。从最初的绿色,明显逐步转变为红色,最后呈蓝紫色。因此,根据果皮颜色预测乌梅成熟度是一种可行的方式。为了更高效地识别乌梅在自然环境中的不同成熟度,乌梅果实成熟度主要以颜色为判断依据,根据感官评价法、表皮比较法,以及人工采摘专家的经验,将乌梅成熟度大致分成4个类别:绿熟期、半转色期、转色期、成熟期,如图1所示。

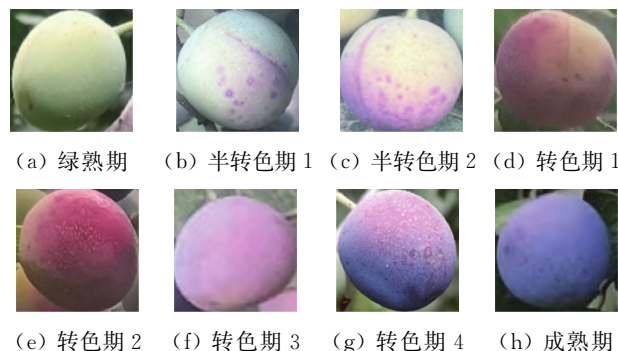


图1 乌梅成熟度分类

Fig. 1 Classification of black plum maturity

绿熟期乌梅的表皮颜色完全是绿色,如图1(a)所示。半转色期乌梅表皮有轻微蓝色或黄色,具体性状表现为乌梅果实表皮的绿色区域面积占比超过3/4,如图1(b)和图1(c)所示。转色期果实表皮颜色逐渐发黄和变红几乎没有绿色,具体性状表现为果实表面的黄色区域和红色区域面积占比大于1/2,如图1(d)~图(g)所示。图1(h)是成熟期乌梅,这个阶段的具体性状表现为乌梅果实表皮颜色完全是蓝色。

1.3 乌梅成熟度检测模型

YOLOv7网络模型比当前最优模型计算量少,拥有更快的推理速度和更高的检测精度。为提高YOLOv7模型目标检测算法性能,对其进行改进。

将BiFormer注意力模块分别放在YOLOv7模型的Backbone的前端(改进的YOLOv7-1模型的具体检测识别过程如图2所示)和后端(改进的YOLOv7-2模型的具体检测识别过程如图3所示)训练,添加果实成熟度精分模块,并进行结果分析。

利用改进的YOLOv7-1模型和改进的

YOLOv7—2 模型检测乌梅在 4 个不同成熟阶段的表皮颜色,建立乌梅成熟度预测模型。模型在包含乌梅

4 个成熟度的图像数据集上进行训练和测试,训练得到最优权值模型。

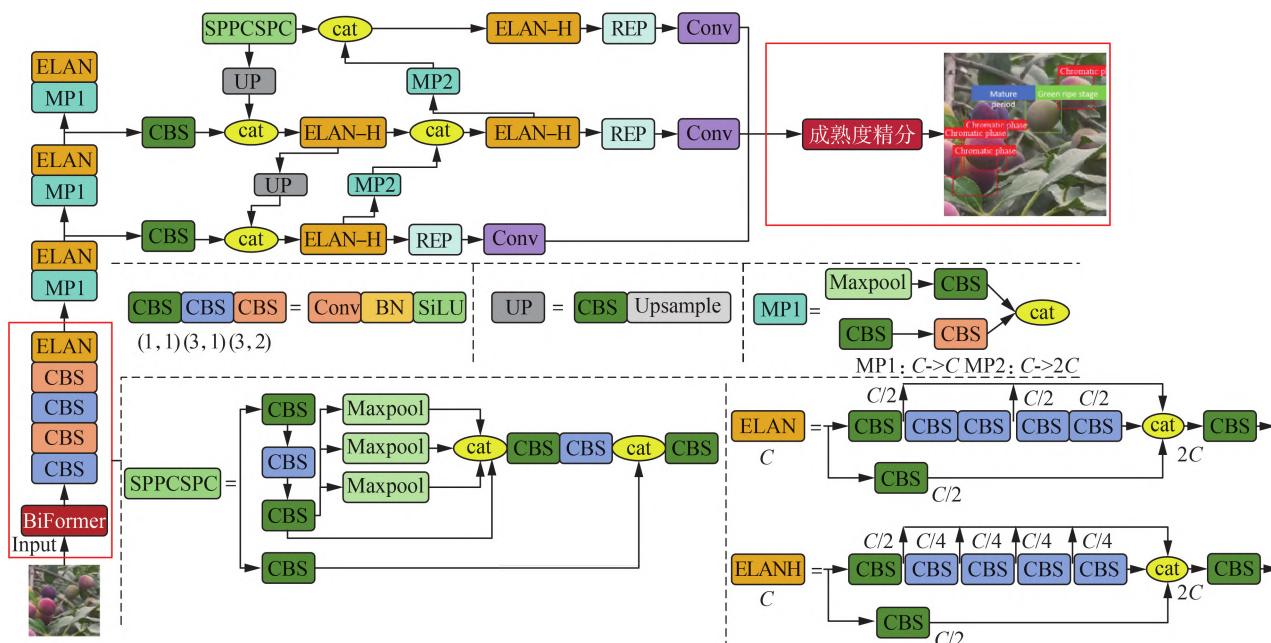


图 2 改进的 YOLOv7—1 模型的乌梅成熟度检测及识别过程

Fig. 2 Detection and recognition process of black plum maturity of the improved YOLOv7—1 model

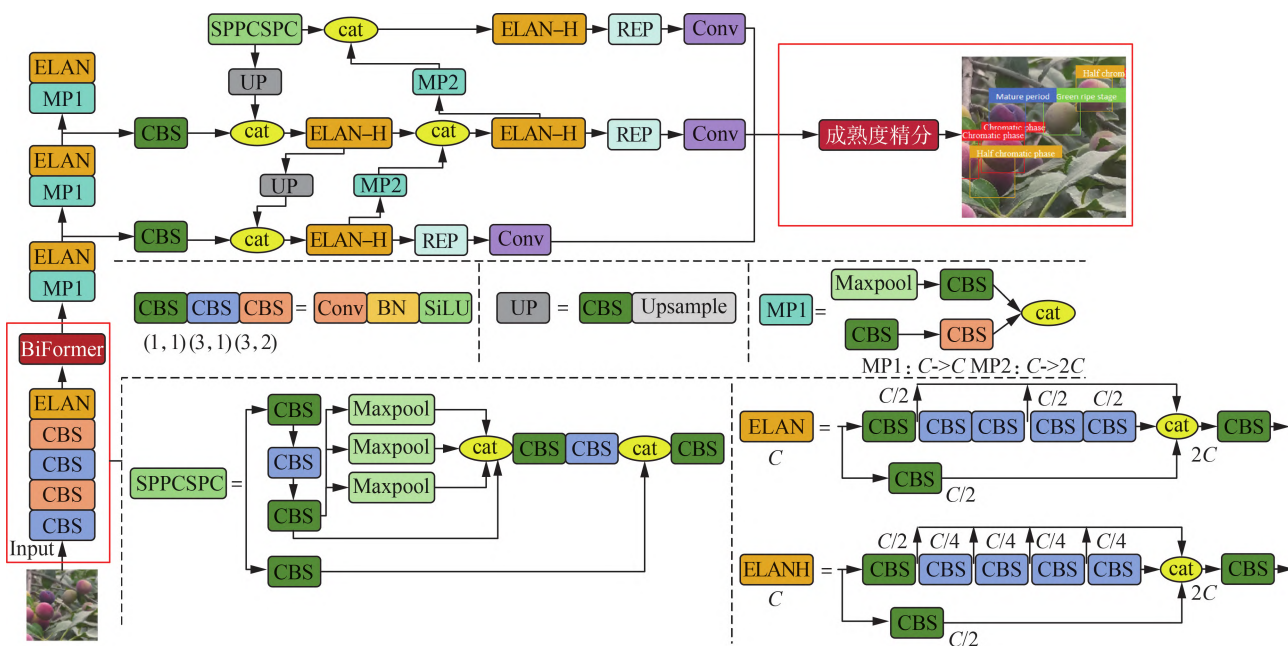


图 3 改进的 YOLOv7—2 模型的乌梅成熟度检测及识别过程

Fig. 3 Detection and recognition process of black plum maturity of the improved YOLOv7—2 model

1.3.1 BiFormer 注意力模块

BiFormer 注意力模块是基于 Vision Transformer (ViT 模型),引入一种新的双层路由注意力机制 (Bi-Level Routing Attention),这是一种新的通用 ViT 模型。BiFormer 注意力模块是一种动态的、查询感知的稀疏注意力机制,可以帮助网络捕捉长距离上下文依赖。在 YOLOv7 的 Backbone 中加入 BiFormer 注意力模块,提高网络对全局语义信息的捕捉能力。

ViT 模型是基于 Transformer 结构的深度学习模型,对图像处理有很好的分类能力。ViT 的核心流程包括图像分块处理 (make patches)、图像块嵌入 (patch embedding) 与位置编码、Transformer 编码器和 MLP 分类处理等 4 个主要部分。ViT 会将整幅图像拆分成小图像块,然后把把这些小图像块的线性嵌入序列作为 Transformer 的输入送入网络,然后使用监督学习的方式进行图像分类的训练。ViT 模型的整体

结构如图 4 所示。

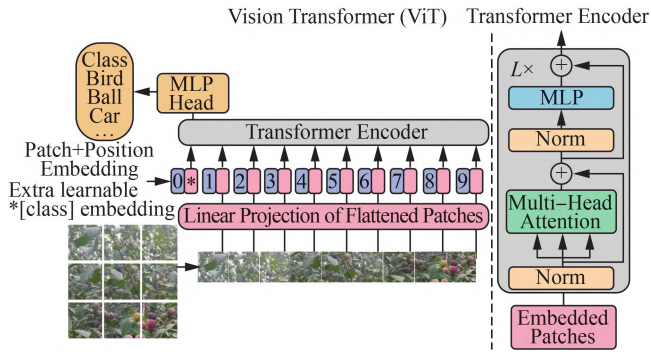


图 4 ViT 模型

Fig. 4 ViT model

BiFormer 注意力模块的整体结构如图 5 所示。BiFormer 注意力模块遵循大多数的 ViT 模型架构设计,也是采用四级金字塔结构,即下采样 32 倍。BiFormer 注意力模块在第 1 阶段使用重叠块嵌入,在第 2~第 4 阶段使用块合并模块来降低输入空间分辨率,同时增加通道数,然后采用连续的 BiFormer 块做特征变换。在每个块的开始均使用深度卷积来隐式编码相对位置信息。随后依次应用 BRA 模块和扩展率为 2 层的多层感知机 MLP 模块,分别用于交叉位置关系建模和每个位置嵌入。

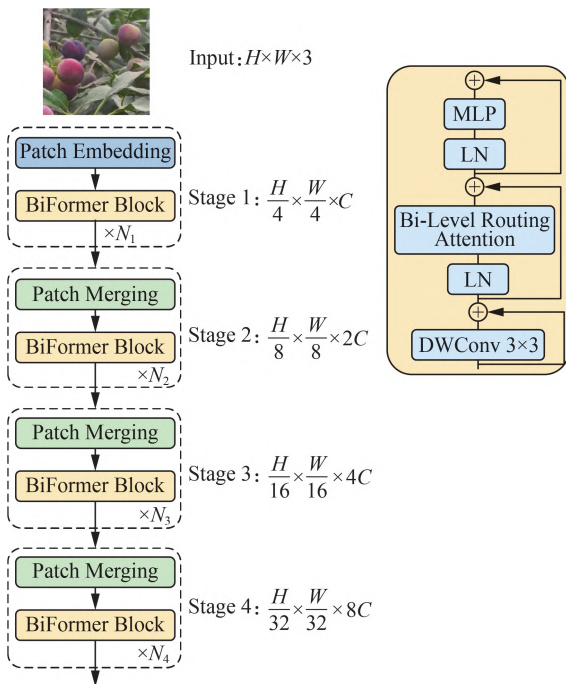


图 5 BiFormer 注意力模块的整体结构

Fig. 5 Overall architecture of the BiFormer attention model

1.3.2 HSV 颜色分割

运用 HSV 颜色分割方法,从图像中提取出改进的 YOLOv7 模型的检测框部分。将框内图像进行提取,对提取后的果实图像颜色特征进行再提取、计算与分析。流程如图 6 所示。

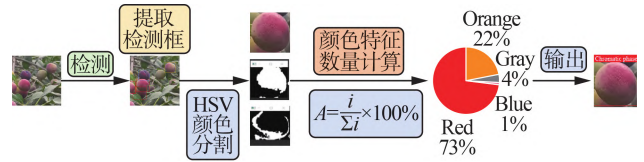


图 6 HSV 颜色分割及输出

Fig. 6 HSV color segmentation and output

从图像中提取出改进的 YOLOv7 模型的检测框部分,并将框内的 RGB 图像转换为 HSV 图像。HSV 与 RGB 相比,能够更直观地表达颜色的明暗程度、鲜艳程度以及色调,且亮度对色彩的影响较小,常用于分割指定颜色的目标。其转换如式(1)~式(3)所示。

$$H = \arccos \left\{ \frac{[(R-G) + (R-B)/2]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\} \quad (1)$$

$$S = 1 - \frac{3}{R+G+B} [\min(R, G, B)] \quad (2)$$

$$V = \frac{1}{3(R+G+B)} \quad (3)$$

式中: H ——Hue 颜色的色相;

S ——Saturation 饱和度;

V ——Value 色明度;

R ——Red 红色;

G ——Green 绿色;

B ——Blue 蓝色。

遍历整个分割出来的检测框图像的像素点,对每一个像素点进行 HSV 提取;再对该像素点颜色进行判定,判定为某种颜色;直到全部像素点判定完成,就可以知道该颜色的数目。占比计算和输出,判断检测框内的果实为某一成熟期的乌梅。像素占比 A 计算如式(4)所示。

$$A = \frac{i}{\sum i} \times 100\% \quad (4)$$

式中: i ——分割部分该颜色像素个数;

$\sum i$ ——改进的 YOLOv7 模型检测框内分割部分的所有颜色的总像素个数。

2 深度学习模型实验

2.1 数据集处理

为增加数据集丰富度,使改进的 YOLOv7—1 模型和改进的 YOLOv7—2 模型能够适应更多的场景,使用旋转、缩放、水平翻转和剪切等方法来增强训练数据集。数据增强技术为训练过程创建了 5 530 张图像。按照 7:1.5:1.5 比例随机分为训练集、验证集、测试集。

2.2 实验配置环境及训练参数

训练与测试在一台配置为 i5-12490F 3.00 GHZ 的 CPU、8 GB 的 NVIDIA GeForce RTX 3070 的 GPU 的计算机上运行,采用 python 3.8 作为编程语

言、Pytorch 深度学习框架进行训练。采用改进的 YOLOv7—1 模型和改进的 YOLOv7—2 模型为主要框架与预训练权重,设置输入图像尺寸为 1 280 像素×1 280 像素,以 4 张图像为一个批处理量,最大迭代次数为 50 次的模型参数,进行网络训练。

2.3 评价指标

在模型检测精度方面,选用精确率 P 、召回率 R 和平均精度 AP 作为评价指标。精确率表示识别正确的乌梅数在识别目标中所占比率,召回率表示在所有乌梅中被识别出来的比率。在模型检测性能方面,选取平均精度均值 mAP 、平均检测时间 ADT 和模型内存占用量作为评价指标。检测时间使用目标检测网络检测一张图所消耗的平均时间为标准,ms。 P 、 R 、 AP 、 mAP 计算如式(5)~式(8)所示。

$$P = \frac{TP}{TP + FP}$$
(5)

$$R = \frac{TP}{TP + FN}$$
(6)

$$AP = \int_0^1 P(R) dR$$
(7)

$$mAP = \frac{1}{M} \sum_{k=1}^M AP(k)$$
(8)

式中： TP ——被模型预测为正类的正样本；
 FP ——被模型预测为正类的负样本；
 FN ——被模型预测为负类的正样本；
 M ——类别总数；
 $AP(k)$ ——第 k 类 AP 值。

3 结果分析

3.1 检测效果及分析

使用预训练好的深度学习模型进行训练,将按照 7 : 1.5 : 1.5 比例进行训练,得到 4 个成熟阶段乌梅识别正确率。在绿熟期(2023 年 5 月)、半转色期(2023 年 7 月)、转色期(2023 年 8 月)、成熟期(2023 年 9 月)4 个成熟阶段分别摘取乌梅果实 300 个,用改进的 YOLOv7—1 模型进行检测,结果如表 1 所示。

表 1 改进的 YOLOv7—1 模型不同成熟度级别检测结果
Tab. 1 Detection results of different maturity levels of the improved YOLOv7—1 model

成熟阶段	总数 /个	果实个数/个				正确率 /%
		绿熟期	半转色期	转色期	成熟期	
绿熟期	300	276	19	5	0	98.66
半转色期	300	21	232	45	2	94.33
转色期	300	1	37	244	18	94.66
成熟期	300	0	7	14	279	98.33

3.2 果园识别效果

为测试算法的实用性,把改进的 YOLOv7—1 模型算法、改进的 YOLOv7—2 模型对果园中的乌梅进行检测与识别,检测输出图片中乌梅的总数,并分别显示果实 4 个成熟度阶段的占比。由此可以推算出果园大部分果实的成熟度,改进的 YOLOv7—1 模型、改进的 YOLOv7—2 模型的乌梅不同成熟度检测效果如图 7 和图 8 所示。



图 7 改进 YOLOv7—1 网络模型的乌梅不同成熟度检测效果
Fig. 7 Detection effect of the improved YOLOv7—1 network model on different maturity of black plum

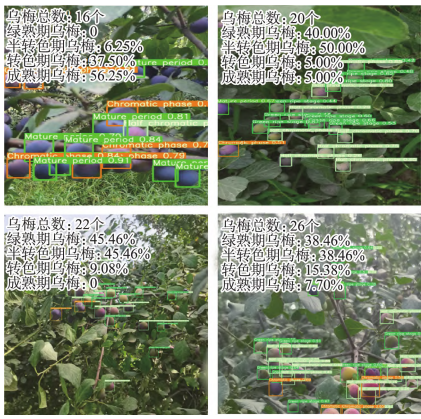


图 8 改进 YOLOv7—2 网络模型的乌梅不同成熟度检测效果
Fig. 8 Detection effect of the improved YOLOv7—2 network model on different maturity of black plum

综合对比可以看出,改进的 YOLOv7—1 模型相比于改进的 YOLOv7—2 模型目标检测算法具有更高的识别准确率。因此,本研究算法具有较强的鲁棒性,可以适应自然环境下的不同情况。

3.3 不同算法检测能力对比

为验证所选模型在自然复杂环境下对乌梅成熟度检测的准确性和快速性,选择改进的 YOLOv7—1 模型与 Faster R—CNN 模型、YOLOv3 模型、Mask R—CNN 模型、YOLOv5s 模型、YOLOv5l 模型、YOLOv7 模型、YOLOv8 模型和改进的 YOLOv7—2 模型对测试集的乌梅图像进行算法对比,结果如图 9 和图 10 所示。

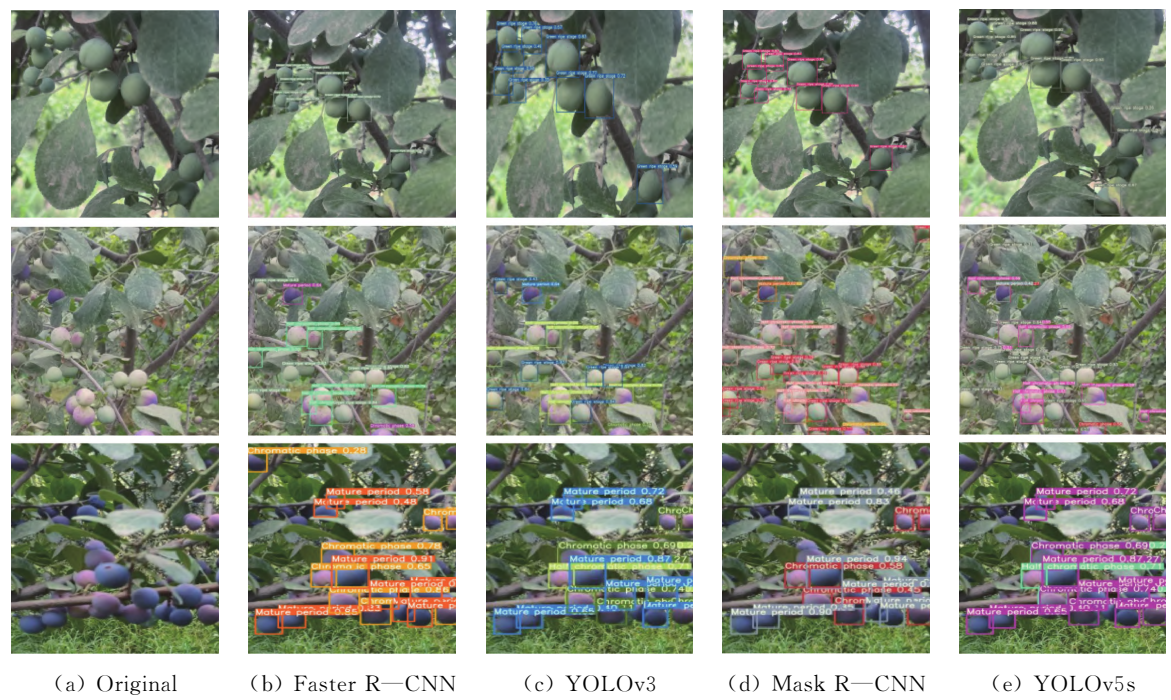


图 9 不同训练网络模型的乌梅不同成熟度检测效果(1)

Fig. 9 Different maturity detection effects black plum of different trained network models(1)

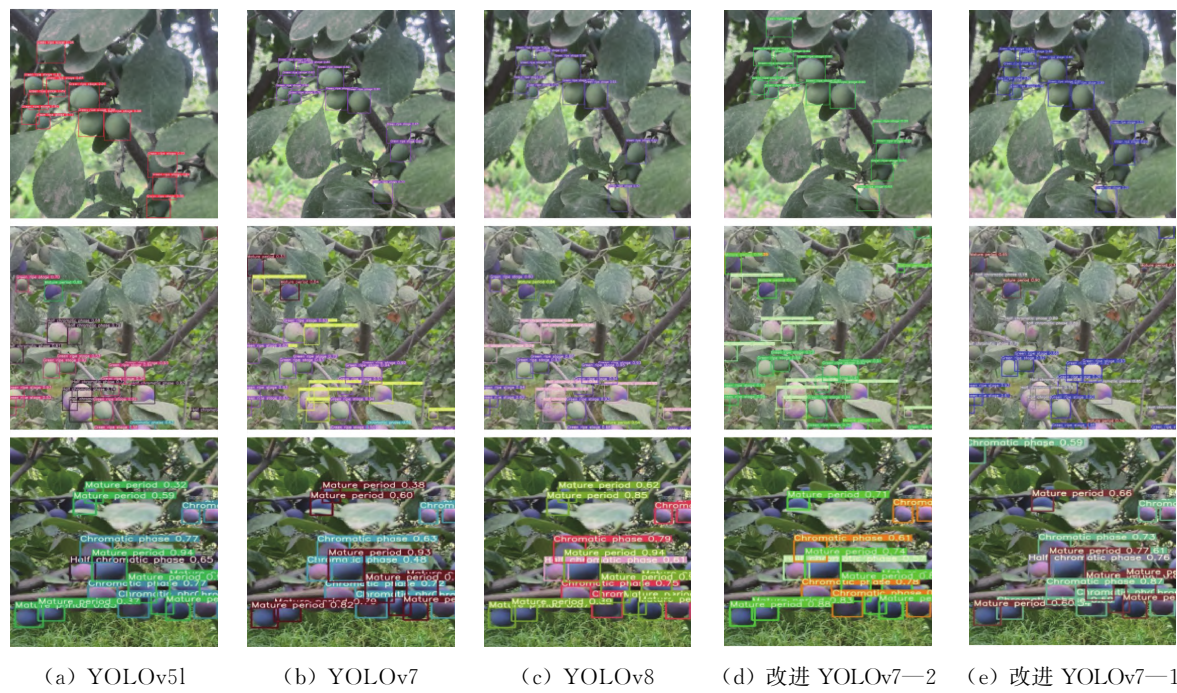


图 10 不同训练网络模型的乌梅不同成熟度检测效果(2)

Fig. 10 Different maturity detection effects black plum of different trained network models(2)

由图 9 和图 10 可知,只有 YOLOv7 模型对乌梅果实成熟度的检测识别效果较好,因此,选择 YOLOv7 模型进行改进,改进之后的模型效果更好。改进的 YOLOv7—1 模型与其他模型在验证集上对不同成熟度乌梅检测的试验结果如表 2 所示。改进的 YOLOv7—1 模型在检测中的总体平均精度均值比改进的 YOLOv7—2 模型、Faster R—CNN 模型、YOLOv3 模型、Mask R—CNN 模型、YOLOv5s 模型、YOLOv5l 模

型、YOLOv7 模型和 YOLOv8 模型分别高 4.8%、12.4%、0.9%、0.7%、12.6%、1.7%、5.8%和 12.3%,且检测速度也较快。其中,改进的 YOLOv7—2 模型的检测精度较高,并且出现漏检误检的情况较少;Faster R—CNN 模型检测精度较低,不仅参数量大,还耗时;YOLOv3 模型内存较大且检测时间较长,出现漏检较多;Mask R—CNN 模型检测精度较低,不仅参数量大还耗时;YOLOv5s 检测精度较差,且容易出现漏检、错检、

多检的现象;YOLOv5l 模型在检测精度效果上表现较好,但过拟合现象比较严重;YOLOv7 模型虽内存较小,但检测精度较差,且容易出现漏检、错检、多检的现象;YOLOv8 模型虽内存较大,但检测精度较差,且容易出现漏检、错检、多检的现象。因此,综合对比可以看出,改进的 YOLOv7—1 模型在乌梅各种成熟度的检测精度与检测速度上都有更大优势。

表 2 不同训练网络模型的乌梅成熟度检测结果

Tab. 2 Results of black plum maturity detection with different trained network models

模型	精确率 P/%	召回率 R/%	mAP /%	检测时间 /ms	参数量 /MB
改进的 YOLOv7—1	74.7	77.6	80.5	11	135
改进的 YOLOv7—2	72.4	75.0	75.7	20	74.9
Faster R—CNN	68.3	67.8	68.1	63	121.91
YOLOv3	71.3	75.0	79.6	89.7	198
Mask R—CNN	71.9	76.0	79.8	125	202.2
YOLOv5s	66.9	66.0	67.9	98	90.15
YOLOv5l	71.1	74.6	78.8	71.3	101
YOLOv7	70.8	68.9	74.7	10.45	71.4
YOLOv8	68.6	68.1	68.2	82	88.8

4 结论

1) 选择网络模型改进的 YOLOv7 模型用于乌梅成熟度检测。实验结果表明,改进的 YOLOv7—1 模型具有较强的鲁棒性,同时保证乌梅成熟度的检测精度和检测速度,满足在自然环境中对乌梅成熟度的检测。当乌梅处于绿熟期、半转色期、转色期、成熟期阶段时,改进 YOLOv7—1 模型的识别正确率分别为 98.66%、94.33%、94.66%、98.33%。

2) 为验证改进的 YOLOv7 模型的性能,设置 4 组网络进行实验验证,并对结果进行定量分析,试验结果表明改进的 YOLOv7—1 模型更好。其平均精确均值比改进 YOLOv7—2 模型、Faster R—CNN 模型、YOLOv3 模型、Mask R—CNN 模型、YOLOv5s 模型、YOLOv5l 模型、YOLOv7 模型和 YOLOv8 模型分别高 4.8%、12.4%、0.9%、0.7%、12.6%、1.7%、5.8% 和 12.3%。

3) 根据乌梅生长过程中各种激素与其颜色关联机制,通过机器视觉和图像处理方法,获得果实颜色特征参数。研究结果有助于完善乌梅成熟度评价指标,确定乌梅成熟度适时采摘,为后续保鲜、包装和运输方式选择提供依据。

参 考 文 献

[1] 黄海. 浅析我国乌梅产业发展现状、存在的问题及发展建

议[J]. 科学种养, 2018(8): 9—10.

[2] 张君成, 梁华, 王燕, 等. 乌梅药理作用研究进展[J]. 辽宁中医药大学学报, 2021, 23(8): 122—126.

[3] 刘宇飞, 汤晓华. 基于机器视觉的番茄颜色分级方法的研究[J]. 机电产品开发与创新, 2013, 26(2): 99—100.

[4] 周文静, 查志华, 吴杰. 改进圆形 Hough 变换的田间红提葡萄果穗成熟度判别[J]. 农业工程学报, 2020, 36(9): 205—213.

Zhou Wenjing, Zha Zhihua, Wu Jie. Ear maturity of red grape in field based on improved circular Hough transform [J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(9): 205—213.

[5] 王立舒, 秦铭霞, 雷洁雅, 等. 基于改进 YOLOv4—Tiny 的蓝莓成熟度识别方法[J]. 农业工程学报, 2021, 37(18): 170—178.

Wang Lishu, Qin Mingxia, Lei Jieya, et al. Blueberry maturity recognition method based on improved YOLOv4—Tiny [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(18): 170—178.

[6] 李竹, 牟昌红, 嵇康轩, 等. 基于深度学习的蓝莓成熟度预测[J]. 安徽农业科学, 2023, 51(5): 232—236.

[7] 王勇, 陶兆胜, 石鑫宇, 等. 基于改进 YOLOv5s 的不同成熟度苹果目标检测方法[J]. 南京农业大学学报, 2024, 47(3): 602—611.

[8] 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 Ronghui, Li Zhiwei, Wu Jinlong, et al. Lightweight cherry tomato maturity detection method based on improved YOLOv7 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 225—233.

[9] Wang C, Bochkovskiy A, Liao H. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464—7475.

[10] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16 × 16 words: Transformers for image recognition at scale [J]. arXiv preprint arXiv: 2010.11929, 2020.

[11] Zhu L, Wang X, Ke Z, et al. BiFormer: Vision transformer with Bi-Level Routing Attention [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 10323—10333.

[12] 李天华, 孙萌, 丁小明, 等. 基于 YOLOv4+HSV 的成熟期番茄识别方法[J]. 农业工程学报, 2021, 37(21): 183—190.

Li Tianhua, Sun Meng, Ding Xiaoming, et al. Recognition method of mature tomato based on YOLOv4+HSV [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(21): 183—190.