

基于多目标采样和改进 Mask R-CNN 的 木瓜成熟度检测

Papaya maturity detection based on multi-target sampling
and improved Mask R-CNN

齐国红¹ 张云龙¹ 苏 曼²

QI Guohong¹ ZHANG Yunlong¹ SU Man²

(1. 郑州西亚斯学院, 河南 郑州 451100; 2. 河南大学, 河南 郑州 450046)

(1. Zhengzhou Sias University, Zhengzhou, Henan 451100, China;

2. Henan University, Zhengzhou, Henan 450046, China)

摘要:目的:提高木瓜成熟度检测准确率及鲁棒性。方法:提出一种基于多目标采样和改进 Mask R-CNN 的木瓜成熟度检测方法。通过均值平均精度、准确率、精确率—召回率曲线和计算时间等指标,验证所提方法的有效性和鲁棒性,并将其检测效果与 Faster R-CNN、RetinaNet 和 CenterMask 等方法进行对比。结果:试验方法对木瓜成熟度检测的平均精度均值、50% 平均精度均值、75% 平均精度均值分别为 98.43%, 98.67%, 98.68%, 对未成熟、半成熟和成熟木瓜成熟度的平均检测精度为 99.38%, 98.81%, 99.37%。结论:该方法可用于开发木瓜成熟度检测的电子系统,提升木瓜成熟度检测和木瓜分级的性能。

关键词:成熟度检测;多目标采样;Mask R-CNN;小数据集;木瓜

Abstract: Objective: Improve the accuracy and robustness of papaya ripeness detection. Methods: A method of papaya ripeness detection based on multi-target sampling and improved Mask R-CNN was proposed. In the process of data expansion, the method introduced multi-object sampling technology to generate enhanced images from small data sets taken under controlled conditions, which was conducive to extending the proposed method to data sets with complex features of actual papaya images. The effectiveness and robustness of the proposed method were verified

by means of average accuracy, accuracy, accuracy-recall curve and calculation time, and the results of papaya ripeness detection effect were compared with those of Faster R-CNN, RetinaNet and CenterMask. Results: The values of mean average precision, 50% mean average precision and 75% mean average precision for the papaya ripeness detection were 98.43%, 98.67% and 98.68%, respectively. The average accuracies for the ripeness detection of immature, semi-mature and mature papayas were 99.38%, 98.81% and 99.37%, respectively. Conclusion: This method can be used to develop an electronic system for papaya ripeness detection and improve the performance of papaya ripeness detection and grading.

Keywords: maturity detection; multi-target sampling; Mask R-CNN; small data set; papaya

随着农业科技的不断发展和人们对食品品质的日益关注,对农作物的精准监测与检测成为了现代农业中至关重要的研究领域^[1]。作为热带和亚热带地区的主要经济作物之一,木瓜因其丰富的营养价值和广泛的应用而备受关注^[2]。然而,木瓜的成熟度对于品质和产量的影响不容忽视^[3]。传统的木瓜成熟度评估方法通常依赖于主观经验和人工抽样,一定程度上限制了成熟度评估的准确性和效率^[4]。

近年来,计算机视觉技术的迅猛发展为农作物成熟度的非侵入式监测提供了新的可能,其中,深度学习技术在图像分析领域取得了显著成果。此外,目标检测和语义分割方法的结合在农作物成熟度检测方面展现出了巨大潜力。熊俊涛等^[4]基于轻量化 YOLO v5-Lite 模型,提出了一种自然环境下木瓜成熟度检测方法,该方法对木瓜成熟度检测的均值平均精度(mAP)为 92.4%。Suban

基金项目:河南省科技攻关项目(编号:232102110274);河南省高等学校重点科研项目(编号:24B210019);河南省教育厅第九批河南省重点学科(检测技术与自动化装置)建设项目(编号:教高[2018]119号)

作者简介:齐国红(1987—),女,郑州西亚斯学院讲师,硕士。

E-mail: qighong20@sina.com

收稿日期:2024-01-05 **改回日期:**2024-03-01

等^[5]提出了一种基于 k 最近邻算法的木瓜成熟度识别方法,该方法能够准确识别木瓜果实的成熟度。Ratha 等^[6]提出了一种基于 VGG16 和离散小波变换的番木瓜成熟度识别方法,该方法对木瓜成熟度检测的准确率达 98%。Behera 等^[7]提出了一种基于 VGG19 和迁移学习算法的木瓜成熟度分类方法,通过 VGG19 和迁移学习算法的结合,该方法对木瓜成熟度的分类准确率得到显著提升。

Mask R-CNN 作为一种强大的视觉识别模型,通过同时实现目标检测和像素级别的语义分割,为农作物成熟度的精准检测提供了可能^[8]。但现有的 Mask R-CNN 模型在应用于木瓜成熟度检测时仍存在一些挑战。首先,木瓜的生长状态和形态多样,生长环境复杂,导致成熟度检测变得更加复杂,而现有方法针对不同背景下木瓜成熟度检测的泛化能力不足。其次,由于木瓜的颜色和纹理在不同成熟度阶段呈现出明显变化,传统的目标检测方法可能无法准确捕捉这些细微的特征变化。

研究拟提出一种非破坏性的木瓜成熟度检测框架,开发一种基于多目标采样和改进 Mask R-CNN 的木瓜成熟度检测方法。该方法在数据增强过程中采用单图像多目标采样技术,从而将单目标训练模型推广到能够适应多目标和杂乱背景场景;由于小数据集导致模型对未见数据高度敏感,且不同成熟度阶段的木瓜具有相似的纹理和颜色模式,因此,在所提出的方法中评估和比较了分割模型的鲁棒性;提出了最佳置信度阈值试验性选择方法,以期小数据集、具有相似模式但具有强鲁棒性的水果分级模型开发提供依据。

1 Mask R-CNN

试验目的是利用目标检测和实例分割模型^[9]对图像中的木瓜果实进行检测,并对其成熟度进行识别。所提出的木瓜成熟度检测方法将不同卷积神经网络层次的特征结合到自下而上的区域方案中,在目标定位和分割方面有较大改进。检测准确度依赖于区域候选技术,该技术已被引入基于 R-CNN 的多种 CNN 方法,如快速 R-CNN 和 Mask R-CNN^[10]。

所改进的 Mask R-CNN 结构实现了对 Faster R-CNN 扩展,如图 1 所示。

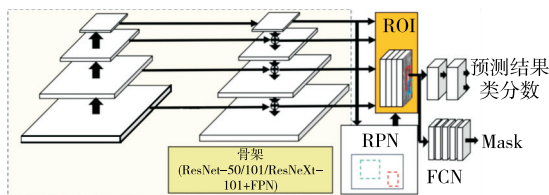


图 1 所提方法使用的 Mask R-CNN 架构

Figure 1 Mask R-CNN architecture used by the proposed method

改进的 Mask R-CNN 包括使用卷积神经网络(主干)的区域候选网络(RPN)和用于对象回归、分类的网络头部两个主要组件。与 Faster R-CNN 相比,Mask R-CNN 还输出每个感兴趣区域(ROI)的掩模检测。

改进的 Mask R-CNN 使用的卷积主干为 ResNeXt 网络^[11]和具有 50,101 层的残差网络 ResNet^[12]。

ResNeXt 重复了一个构建模块,该模块聚合了一组具有相同拓扑结构的转换。与 ResNet 相比,其引入了一个新的维度:基数(转换集的大小) C ,并将其作为除深度和宽度之外的一个关键因素。一组聚合转换可以表示为:

$$F(x) = \sum_{i=1}^C T_i(x), \quad (1)$$

式中:

$T_i(x)$ ——任意函数。

类似于简单神经元,函数 T_i 应将 x 投影到一个(可选的低维)嵌入中,并对其进行转换。

(1) 基本构建块:ResNeXt 的基本构建块称为“基数组”。“基数”指组内并行路径或分支的数量。在 ResNeXt 中,每个组包含多条并行路径,这些路径用于从输入数据中捕获不同的特征。使用多条路径与传统的 ResNet 架构不同,后者只使用两条路径(一条快捷路径和一条主路径)。基数参数(通常表示为“ C ”)确定组内有多少个并行路径。

(2) 基数组:在基数组内,并行路径被实现为“瓶颈块”。这些块包括 3 个主要组件: 1×1 卷积, 3×3 卷积和另一个 1×1 卷积。 1×1 卷积用于降低输入特征的维度,而 3×3 卷积负责捕获更复杂的模式。然后,这些并行路径的输出通常通过求和聚合在一起,以创建组的最终输出。

(3) 架构:ResNeXt 架构可表示为:“ResNeXt- $C(C \times 4)$ - Y ”。“ C ”表示基数,确定组内并行路径数。“ $C \times 4$ ”表示组内每条并行路径中的滤波器或通道数。“4”是原始 ResNeXt 中常用的乘数。“ Y ”表示整体网络架构,例如 ResNeXt-50 的“50”,ResNeXt-101 的“101”等。

(4) 更深和更宽的网络:ResNeXt 在增加深度(加深)和增加宽度(添加更多并行路径/组)之间提供了一个权衡。这种灵活性使研究人员能够根据任务的特定要求来定制架构。更深的网络可以捕获更复杂的特征,而更宽的网络(具有更多并行路径)可以捕获更多样的特征,使其对各种任务非常有效。

残差网络具有跳跃连接,解决了梯度消失问题。ROI 特征通过从特征金字塔网络(FPN)的不同层次中提取出来,FPN 具有自上而下的结构和横向连接^[13]。使用 ResNet-FPN 主干进行特征提取,其在准确率和计算时间上都具有很高的性能。区域候选网络用于提出候选对

象,使用注意机制在边界框中输出一组候选对象。

所改进 Mask R-CNN 的网络头部使用 ROI Align, ROI Align 使用双线性插值来计算采样位置的输入特征的精确值,然后对组合结果使用最大值或平均值^[14]。

双线性插值,又称为双线性内插。在数学上,双线性插值是有两个变量的插值函数的线性插值扩展,其核心思想是在两个方向分别进行一次线性插值。

如图 2 所示,若想得到未知函数 f 在点 $P=(x, y)$ 的值,假设已知函数在 $Q_{11}=(x_1, y_1)$ 、 $Q_{12}=(x_1, y_2)$ 、 $Q_{21}=(x_2, y_1)$ 和 $Q_{22}=(x_2, y_2)$ 4 个点的值。首先在 x 方向进行线性插值,得到

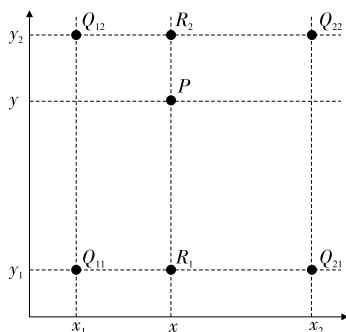


图 2 双线性插值示例图

Figure 2 Bilinear interpolation example diagram

$$f(R_1) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}), R_1 = (x, y_1), \quad (2)$$

$$f(R_2) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}), R_2 = (x, y_2). \quad (3)$$

然后在 y 方向进行线性插值,得到

$$f(P) \approx \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2). \quad (4)$$

得到所要的结果:

$$f(x, y) \approx \frac{f(Q_{11})}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y_2 - y) + \frac{f(Q_{21})}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y_2 - y) + \frac{f(Q_{12})}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y - y_1) + \frac{f(Q_{22})}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y - y_1). \quad (5)$$

如果选择一个坐标系使得 f 的 4 个已知点坐标分别为 $(0, 0)$ 、 $(0, 1)$ 、 $(1, 0)$ 和 $(1, 1)$, 那么插值公式就可以化简为

$$f(x, y) \approx f(0, 0)(1-x)(1-y) + f(1, 0)x(1-y) + f(0, 1)(1-x)y + f(1, 1)xy. \quad (6)$$

ROI Align 的输出用于根据框检测和类分数对每个

候选对象执行对象回归和分类。全卷积网络(FCN)的头部网络包括卷积掩码检测分支,该分支具有用于使用 sigmoid 和二进制损失检测掩码的一些滤波器。总的训练损失包括分类损失、边界框损失和平均二进制交叉熵损失^[15]。

所改进 Mask R-CNN 参数使用均值平均精度(mAP)与准确率[真实木瓜标签数量与所有检测到的数量(假阳性标签和真实标签之和)之比]之间的关系进行选择。

2 数据采集和数据增强

所构建的木瓜图像数据集包括原始数据集 #1 和增强数据集 #2 两部分。数据集 #1 中的木瓜图像在白色背景下拍摄,然后在像素级别对每个图像进行标注。由于数据集 #1 中木瓜图像数量有限,使用旋转、缩放和平移技术进行图像增强,并从像素级别的标签生成附加的图像数据集,得到增强数据集 #2。数据集 #1 和数据集 #2 用于训练和测试木瓜图像检测和分割模型、最佳成熟度检测阈值的选择和验证基于多目标采样和改进 Mask R-CNN 的木瓜成熟度检测模型。

实际上,木瓜的成熟度分为未成熟、半成熟和完全成熟 3 个级别(表 1),该指标可用于确定目标在送达客户之前仍保持新鲜的时间长度。

表 1 木瓜成熟度等级分组

Table 1 Classification group of papaya maturity

等级	成熟度指数	描述
未成熟	0	全绿色
	1	轻微变色
半成熟	2	金色区域比例 $< 1/4$
	3	金色区域比例为 $1/4 \sim 1/2$
成熟	4	金色区域比例为 $1/2 \sim 3/4$
	5	金色区域比例为 $3/4 \sim 1$
	6	全金色

未成熟的木瓜可以在成熟前存放 6~7 d,而部分成熟的木瓜可以再保鲜 3~4 d。完全成熟的木瓜应尽快食用或加工。因此,未成熟和部分成熟的木瓜可以通过包装以便远距离输送或出口,而完全成熟的木瓜将在当地销售。

2.1 数据采集

为了获得图像,木瓜被放置在一个由漫射荧光灯源均匀照明的相机中。加入白色背景以提高对比度,并在此背景上放置尺子以允许在开发的代码中进行校准(见图 3)。使用高分辨率数码相机(尼康 AF-S DX Nikkor 18~55 mm)获取木瓜图像,无闪光灯和变焦($55\times$)。

对 60 个番木瓜果实采集了 240 张图像(每个木瓜 4 张图像),不同成熟阶段的木瓜图像如图 4 所示。

由于木瓜图像中含有背景像素,对于未成熟、半成熟和完全成熟的不同成熟度的原始图像,用白色像素标记木瓜,如图 5 所示。

2.2 数据增强

通过在原始带标签的真实图像基础上生成更多的数据,以提高训练数据的效率,使其更加真实和实用^[16]。使

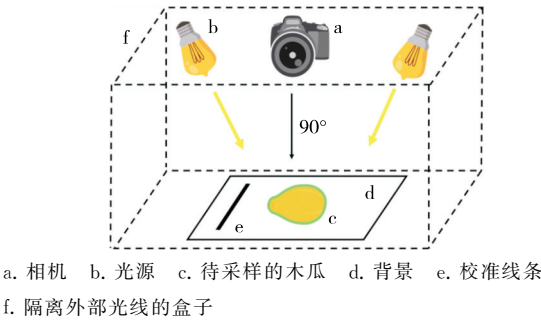


图 3 木瓜图像捕捉试验装置

Figure 3 An experimental device for capturing papaya images

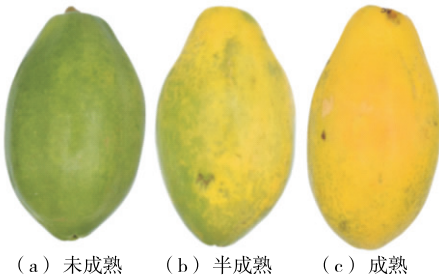


图 4 不同成熟阶段的木瓜图像

Figure 4 Papaya in different stages of maturity

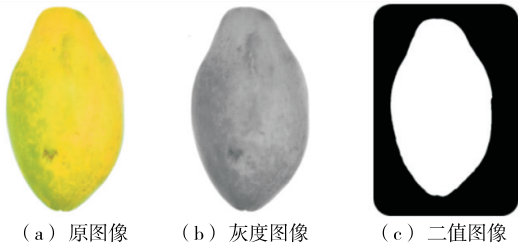


图 5 木瓜图像样本及其标记的木瓜所在范围实况

Figure 5 Papaya image samples and their real marked papaya range

用几何变换来区分 50% 原始图像中不同的木瓜外皮、朝向和位置。此外,图像中包括了背景元素。通过像素级别的标注,将所有的木瓜对象存储在数据库中。将另外的 25% 原始图像添加到不同的背景图像中,包括天然木瓜田、超市、工厂等,另外 25% 的原始图像添加到带有多个不同尺寸、朝向和位置的多木瓜背景图像中,以模拟更真实的包含多个对象的图像^[17]。每个木瓜都在像素级别和实例级别上进行了标注。用于训练的木瓜图像包括原始图像和增强图像,完全与测试数据集分开。

3 试验与结果分析

为验证所提出的木瓜成熟度检测方法 Mask R-CNN (ResNeXt-101) 的有效性和鲁棒性,将其对木瓜成熟度的检测结果与 R-CNN 的其他模型进行对比,所选择对比的方法对水果进行检测时具有较高的平均检测精度^[18]。ResNeXt-101 简记为 X101。

该试验使用 detectron2 框架^[19]和配备有 Intel(R) 2 核 Xeon(R) CPU @ 2.20 GHz、NVIDIA Tesla V100-SXM2 显卡和 16 GB HBM2 内存的 Google Colab 机器,使用 PyTorch 1.8 运行程序。参数配置是根据每个模型的总损失经验进行选择;所有基于 R-CNN 模型均使用相同的参数配置,例如最大迭代次数为 1 000,基本学习率为 0.000 25,每批处理 2 张图像。原始图像尺寸为 4 128 像素×3 096 像素,增强后的图像尺寸从 855 像素×1 280 像素到 6 016 像素×4 000 像素不等。每个模型均使用在 COCO 数据集^[20]上预训练的原始模型进行初始化,然后在两个不同数据集上进行微调:原始的木瓜数据集(数据集#1)和带有多目标采样的木瓜数据集(数据集#2),如表 2 所示。

3.1 R-CNN 模型和残差网络

所对比的 R-CNN 模型包括 RetinaNet、CenterMask、Mask R-CNN、Faster R-CNN,所采用的残差网络包括具有 99,57 层的 VoVNetV2 和具有 50,101 层的 ResNet。根据 R-CNN 模型和残差网络的不同组合,所对比的方法主要包括 CenterMask (VoVNet-99)、CenterMask (VoVNet-57)、Faster R-CNN (ResNet-50)、Faster R-CNN (ResNet-101)、RetinaNet (ResNet-101)、RetinaNet (ResNet-50)、Mask R-CNN (ResNet-101)、Mask R-CNN

表 2 原始数据集和增强数据集中不同成熟阶段木瓜图像数

Table 2 The number of papaya images at different stages of maturity in the original dataset and the enhanced dataset

数据集	训练集			测试集			总图像数
	未成熟	半成熟	成熟	未成熟	半成熟	成熟	
原始数据集	56	56	56	24	24	24	240
增强数据集	105	118	117	41	48	42	360

(ResNet-50)。后续分析中,VoVNet-99 简记为 v99,VoVNet-57 简记为 v57,ResNet-101 简记为 r101,ResNet-50 简记为 r50。

(1) RetinaNet 是一种应用于密集采样候选对象的单级检测器。RetinaNet 由一个主干网络和两个子网络组成,分别用于框分类和框回归。由于背景和前景之间的类不平衡是影响检测准确性的主要问题,RetinaNet 通过对负面示例进行重点训练来计算焦点损失。焦点损失由所有候选对象的焦点损失之和计算得出。训练损失包括分类损失和边界框损失。

(2) 基于最新分割模型 CenterMask 的检测方法,CenterMask 是一个单阶段和无锚点实例分割模型,CenterMask 通过增加一个空间注意引导的掩码网络(SAG-mask)用于检测物体掩码。SAG-mask 网络包含空间注意图(SAM),在聚焦于信息像素并减少噪声方面起着至关重要的作用。

3.2 评估指标

为了评估不同方法对木瓜果实图像成熟度检测的性能,考虑检测到的目标与真实边界框之间的相似性,包括重叠面积和检测到的木瓜成熟度。交并比(IoU) I_{IoU} 可用于衡量检测目标的边界框(A_p)和真实边界框(A_{gt})之间重叠区域的比率^[21]:

$$I_{IoU} = \frac{A_p \cap A_{gt}}{A_p \cup A_{gt}} \quad (7)$$

当 $I_{IoU} = 1$ 时,IoU 分数为最佳,而 IoU 阈值通常设

置为识别到的目标是否正确的限制程度。此外,精确率可以用来识别所有检测到的物体中的正确阳性样本数,而召回率用于识别图像中所有真实边界框对象^[22]。

$$P_{Precision} = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (8)$$

$$R_{Recall} = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (9)$$

式中:

N_{TP} ——正确检测到的样本数量;

N_{FP} ——错误检测到的样本数量或不存在的样本数量;

N_{FN} ——未检测到的真实样本数量。

此外,目标检测器的置信度分数也可以考虑在检测指标内。较大的置信度分数可以被视为正的检测结果。将精确度和召回率重新表示为置信度阈值(τ)的函数^[23]:

$$P_{Precision}(\tau) = \frac{N_{TP}(\tau)}{N_{TP}(\tau) + N_{FP}(\tau)}, \quad (10)$$

$$R_{Recall}(\tau) = \frac{N_{TP}(\tau)}{N_{TP}(\tau) + N_{FN}(\tau)} \quad (11)$$

理想的目标检测器应具有精确率和召回率均为 1 的性能。精确率—召回率曲线(PR 曲线)呈现了从检测开始的不同阈值的精确率—召回率关系,试验方法的 PR 曲线呈锯齿状(图 6)。此外,平均精度(AP)是一种常用的指标,为 PR 曲线下的面积;mAP 为 AP 在类别和阈值上的平均值。

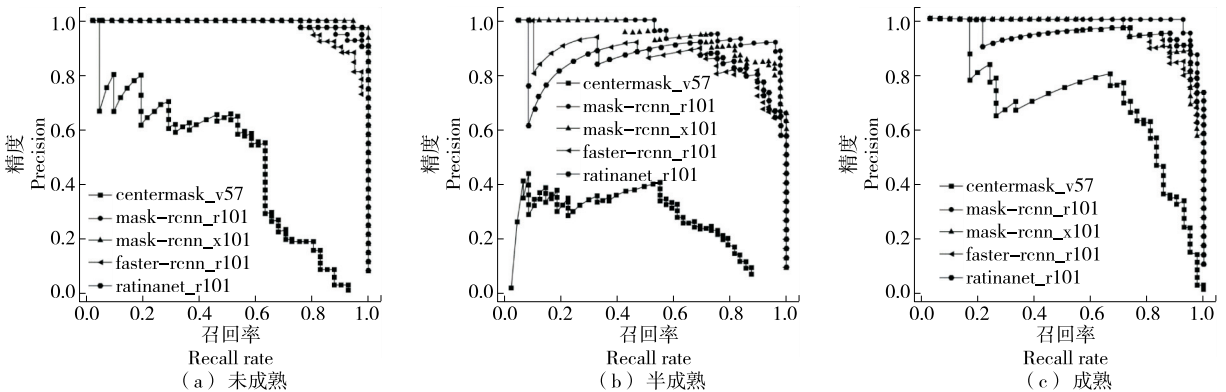


图 6 不同方法对未熟、半熟和全熟木瓜检测的 PR 曲线

Figure 6 PR curves of unripe, half-ripe and fully ripe papaya by different methods

3.3 成熟度阶段检测及最优阈值选择

3.3.1 成熟度阶段检测 mAP 从 50% 变化到 95%,步长为 5%,AP50 和 AP75 分别对应 50% 和 75% 的平均检测精确度。由表 3 可知,试验方法在所有指标得分上的表现最好,而 Mask R-CNN 和 Faster R-CNN 在考虑特定重叠阈值时在 AP50 和 AP75 上表现出较高的精确度。CenterMask 在 AP50 阈值下展现出良好的检测性能。

由表 4 可知,由于未知的背景干扰、重叠的目标、不

同尺寸的目标和其他目标变化,所有方法对木瓜成熟度检测的 mAP 均有不同程度的下降,但试验方法在所有指标得分上的表现仍是最好的。由表 5 可知,通过多目标采样技术,试验方法在不受限制的数据集(数据集 #2)上泛化能力更强,如 mAP 从 92.76% 提升至 98.43%。

3.3.2 计算时间 由图 7 可知,相比于 ResNet50 和 VoVNet59 等较小尺寸的层,r101(ResNet101)和 v99(VoVNet99)等较大尺寸的层通常需要更多的计算时间。

Mask R-CNN 通过全连接网络进行掩码检测,因此需要更多的计算时间。RetinaNet 和 Faster R-CNN 具有最高的计算速度,但 RatinaNet 出现了较多误检。

3.4 木瓜成熟度检测性能比较
为进一步验证试验方法的有效性,将试验方法的木瓜成熟度检测性能与基于轻量化YOLOv5-Lite模型的

表 3 使用原始数据集(数据集 #1)训练和测试不同方法时的木瓜成熟度检测结果
Table 3 Papaya ripeness detection results when trained and tested with different methods using the original dataset (dataset #1) %

方法	mAP	AP50	AP75	AP(未成熟)	AP(半成熟)	AP(成熟)
CenterMask (VoVNet-99)	53.63	83.36	63.58	90.33	66.20	93.82
CenterMask (VoVNet-57)	37.68	80.61	9.46	96.79	70.42	74.54
Faster R-CNN (ResNet-101)	91.01	98.94	98.94	99.18	98.80	99.07
Faster R-CNN (ResNet-50)	91.96	94.42	94.42	97.54	87.78	98.50
RetinaNet (ResNet-101)	91.96	94.42	94.42	97.54	87.78	98.50
RetinaNet (ResNet-50)	94.19	95.98	95.98	99.54	89.73	98.88
Mask R-CNN (ResNet-101)	87.47	97.47	97.47	98.92	94.65	99.17
Mask R-CNN (ResNet-50)	70.89	91.04	86.80	99.04	74.48	99.67
试验方法	98.43	98.71	98.71	99.54	99.00	99.86

表 4 使用原始数据集(数据集 #1)训练、增强数据集(数据集 #2)测试不同方法时的木瓜成熟度检测结果
Table 4 Papaya ripeness detection results of different methods trained using the original data set (dataset #1) and tested based on the enhanced data set (dataset #2) %

方法	mAP	AP50	AP75	AP(未成熟)	AP(半成熟)	AP(成熟)
CenterMask (VoVNet-99)	32.26	54.69	34.97	54.28	36.97	73.16
CenterMask (VoVNet-57)	24.02	54.29	5.56	64.75	38.47	59.53
Faster R-CNN (ResNet-101)	69.27	78.11	77.76	87.53	56.26	90.97
Faster R-CNN (ResNet-50)	71.12	73.58	73.58	79.12	57.78	84.23
RetinaNet (ResNet-101)	69.16	78.24	78.24	83.02	59.28	92.29
RetinaNet (ResNet-50)	75.22	77.07	77.07	81.65	63.97	86.40
Mask R-CNN (ResNet-101)	62.93	72.37	72.11	73.02	59.93	84.75
Mask R-CNN (ResNet-50)	56.29	73.62	72.26	76.06	53.46	92.00
试验方法	92.76	93.02	92.90	93.47	92.25	92.71

表 5 使用原始数据集(数据集 #1)+增强数据集(数据集 #2)训练和测试不同方法时的木瓜成熟度检测结果
Table 5 Papaya ripeness detection results of different methods trained and tested using the original data set (dataset #1) and the enhanced data set (dataset #2)simultaneously %

方法	mAP	AP50	AP75	AP(未成熟)	AP(半成熟)	AP(成熟)
CenterMask (VoVNet-99)	52.11	82.62	63.10	89.59	62.36	90.26
CenterMask (VoVNet-57)	30.19	78.86	7.19	94.16	70.27	73.00
Faster R-CNN (ResNet-101)	87.81	93.62	93.62	97.46	87.83	95.89
Faster R-CNN (ResNet-50)	87.72	92.67	93.79	97.52	84.31	97.65
RetinaNet (ResNet-101)	91.13	94.21	94.21	96.19	86.02	95.65
RetinaNet (ResNet-50)	90.11	93.09	93.09	98.73	88.96	89.90
Mask R-CNN (ResNet-101)	86.70	94.98	95.98	98.20	93.58	98.63
Mask R-CNN (ResNet-50)	62.81	90.76	84.76	97.18	69.36	98.00
试验方法	98.43	98.67	98.68	99.38	98.81	99.37

木瓜成熟度检测方法(YOLO v5-Lite)、基于 k 最近邻算法的木瓜成熟度识别方法(KNN)、基于 VGG16 和离散小波变换的番木瓜成熟度识别方法(VGG16-DWT)、基于 VGG19 和迁移学习算法的木瓜成熟度分类方法

(VGG19-TLA)进行对比,结果见表 6。由表 6 可知,试验方法对木瓜成熟度检测的 mAP、AP50、AP75、未成熟的 AP、半成熟的 AP 和成熟的 AP 比其他方法至少提高了 0.07%,0.65%,0.56%,1.27%,0.78%,0.88%。

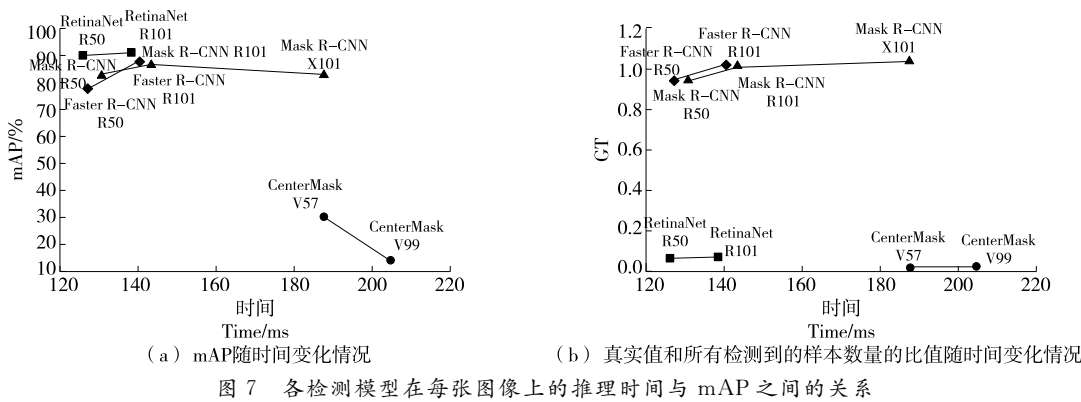


Figure 7 The relationship between the reason time of each test model on each image and the map

表 6 木瓜成熟度检测方法的检测性能对比

Table 6 Comparison of detection performance of various papaya ripeness detection methods %						
方法	mAP	AP50	AP75	AP(未成熟)	AP(半成熟)	AP(成熟)
试验方法	98.08	98.67	98.68	99.38	98.81	99.37
YOLO v5-Lite	92.40	92.43	94.59	95.32	92.69	93.98
KNN	90.91	90.98	92.35	92.16	91.80	92.24
VGG16-DWT	98.00	97.97	98.12	97.67	97.84	98.49
VGG19-TLA	98.15	98.02	97.98	98.11	98.03	98.32

综上,试验方法的木瓜成熟度检测 mAP、准确率和召回率均优于所对比的其他方法,验证了试验方法的有效性和实用性。

4 结论

研究提出了一种基于多目标采样和改进 Mask R-CNN 的木瓜成熟度检测方法。结果表明,通过多目标采样技术,试验方法对木瓜成熟度检测的泛化能力更强,可将均值平均精度、50%平均精度、75%平均精度、未成熟的平均精度、半成熟的平均精度和成熟的平均精度分别提高 5.67%,5.65%,5.78%,5.91%,6.56%,6.66%。与其他 4 种木瓜成熟度检测方法相比,试验方法对木瓜成熟度检测的均值平均精度、50%平均精度、75%平均精度、未成熟的平均精度、半成熟的平均精度和成熟的平均精度比其他方法至少提高了 0.07%,0.65%,0.56%,1.27%,0.78%,0.88%。试验方法的计算时间偏长,后续将进一步研究具有更高计算效率且可保持当前木瓜成熟度检测准确率、召回率等性能优势的方法。

参考文献

[1] 杨妮, 邓树林, 樊艳红, 等. 叶绿素荧光遥感反演及其农业监测

应用研究进展[J]. 江苏农业科学, 2023, 51(14): 1-13.
YANG N, DENG S L, FAN Y H, et al. Research progress of chlorophyll fluorescence retrieval by remote sensing and its application in agricultural monitoring [J]. Jiangsu Agricultural Sciences, 2023, 51(14): 1-13.
[2] 巨浩羽, 赵士豪, 赵海燕, 等. 光皮木瓜真空脉动干燥特性及神经网络模型[J]. 食品与机械, 2022, 38(3): 147-153.
JU H Y, ZHAO S H, ZHAO H Y, et al. Drying characteristics of Chaenomeles sinensis with vacuum pulsed drying technology based on BP neural network mode[J]. Food & Machinery, 2022, 38(3): 147-153.
[3] 周陈平, 杨敏, 郭金菊, 等. 番木瓜成熟过程中全基因组 DNA 甲基化和转录组变化分析[J]. 园艺学报, 2022, 49(3): 519-532.
ZHOU C P, YANG M, GUO J J, et al. Dynamic changes in DNA methylome and transcriptome patterns during papaya fruit ripening [J]. Acta Horticulturae Sinica, 2022, 49(3): 519-532.
[4] 熊俊涛, 韩咏林, 王潇, 等. 基于 YOLO v5-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 YOLO v5-lite[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 243-252.

- [5] SUBAN I B, PARAMARTHA A, FORWONATUS M, et al. Identification the maturity level of carica papaya using the k-nearest neighbor[J]. Journal of Physics: Conference Series, 2020, 1 577 (1): 012028.
- [6] RATHA A K, BARPANDA N K, SETHY P K, et al. Papaya fruit maturity estimation using wavelet and ConvNET[J]. Ingenierie des Systemes d'Information, 2023, 28(1): 175-181.
- [7] BEHERA S K, RATH A K, SETHY P K. Maturity status classification of papaya fruits based on machine learning and transfer learning approach [J]. Information Processing in Agriculture, 2021, 8(2): 244-250.
- [8] SHIU Y S, LEE R Y, CHANG Y C. Pineapples' detection and segmentation based on faster and mask R-CNN in UAV imagery[J]. Remote Sensing, 2023, 15(3): 814.
- [9] 翁海勇, 李效彬, 肖康松, 等. 基于 Mask R-CNN 的柑橘主叶脉显微图像实例分割模型[J]. 农业机械学报, 2023, 54(7): 252-258, 271.
- WENG H Y, LI X B, XIAO K S, et al. Instance segmentation model for microscopic image of citrus main leaf vein based on mask R-CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(7): 252-258, 271.
- [10] 周燕, 韦勤彬, 廖俊玮, 等. 自然场景文本检测与端到端识别: 深度学习方法[J]. 计算机科学与探索, 2023, 17(3): 577-594.
- ZHOU Y, WEI Q B, LIAO J W, et al. Natural scene text detection and end-to-end recognition: Deep learning methods[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(3): 577-594.
- [11] 李聪, 喻国威, 张原嘉, 等. 基于 ResNeXt 与迁移学习的干制哈密大枣果梗/花萼及缺陷识别[J]. 食品与机械, 2022, 38(1): 135-140.
- LI C, YU G W, ZHANG Y J, et al. Research on recognition of stem/calyx and defects of dried Hami jujube based on ResNeXt and transfer learning[J]. Food & Machinery, 2022, 38(1): 135-140.
- [12] 魏冉, 裴悦琨, 姜艳超, 等. 基于改进 Faster R-CNN 模型的樱桃缺陷检测[J]. 食品与机械, 2021, 37(10): 98-105, 201.
- WEI R, FEI Y K, JIANG Y C, et al. Detection of cherry defects based on improved Faster R-CNN model[J]. Food & Machinery, 2021, 37(10): 98-105, 201.
- [13] 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.
- [14] 罗鑫, 王艳艳, 刘学渊, 等. 基于深度学习的林地澳洲坚果检测识别[J]. 森林工程, 2023, 39(2): 113-120.
- LUO X, WANG Y Y, LIU X Y, et al. Macadamia integrifolia detection and recognition in orchards based on deep learning[J]. Forest Engineering, 2023, 39(2): 113-120.
- [15] TANG Q, JIANG Y, XIN J, et al. A novel method for the recovery of continuous missing data using multivariate variational mode decomposition and fully convolutional networks[J]. Measurement, 2023, 220: 113366.
- [16] 任锐, 张淑娟, 赵华民, 等. 基于机器视觉的辣椒外部品质检测[J]. 食品与机械, 2021, 37(1): 165-168.
- REN R, ZHANG S J, ZHAO H M, et al. Study on external quality detection of pepper based on machine vision [J]. Food & Machinery, 2021, 37(1): 165-168.
- [17] 倪云峰, 叶健, 樊娇娇. 基于图像识别的水果分拣系统[J]. 江苏农业科学, 2021, 49(10): 170-176.
- NI Y F, YE J, FAN J J. Fruit sorting system based on image recognition[J]. Jiangsu Agricultural Sciences, 2021, 49(10): 170-176.
- [18] MINAEE S, BOYKOV Y, PORIKLI F, et al. Image segmentation using deep learning: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7): 3 523-3 542.
- [19] AKALIN F, YUMUŞAK N. Detection and classification of white blood cells with an improved deep learning-based approach[J]. Turkish Journal of Electrical Engineering and Computer Sciences, 2022, 30(7): 2 725-2 739.
- [20] HAMEED K, CHAI D, RASSAU A. Texture-based latent space disentanglement for enhancement of a training dataset for ANN-based classification of fruit and vegetables [J]. Information Processing in Agriculture, 2021, 10(1): 85-105.
- [21] PADILLA R, PASSOS W L, DIAS T L B, et al. A comparative analysis of object detection metrics with a companion open-source toolkit[J]. Electronics, 2021, 10(3): 279.
- [22] 杜国真, 卢明星, 季泽旭, 等. 基于改进 CNN 的苹果缺陷检测方法研究[J]. 食品与机械, 2023, 39(6): 155-160.
- DU G Z, LU M X, JI Z X, et al. Research on apple detection method based on improved CNN[J]. Food & Machinery, 2023, 39(6): 155-160.
- [23] RANGWANI H, TAKENORI S, TAKASHI K, et al. Cost-sensitive self-training for optimizing non-decomposable metrics [J]. Advances in Neural Information Processing Systems, 2022, 35: 26 994-27 007.