doi:10.6041/j.issn.1000-1298.2022.07.024

基于 YOLO v5s 的自然场景油茶果识别方法

宋怀波[1,2]　王亚男[1,2]　王云飞[1,2]　吕帅朝[1,2]　江　梅[1,2]

(1.西北农林科技大学机械与电子工程学院,陕西杨凌 712100;
2.农业农村部农业物联网重点实验室,陕西杨凌 712100)

摘要:针对油茶果体积小、分布密集、颜色多变等特点,为实现自然复杂场景下油茶果的快速精准定位,并依据果实的疏密分布,确定恰当的自动振荡采收装置夹持位置,利用 YOLO v5s 卷积神经网络模型,开展了自然环境下油茶果图像检测方法研究,用 3 296 幅油茶果图像制作 PASCAL VOC 的数据集,对网络进行了 150 轮训练,得到的最优权值模型准确率为 90.73%,召回率为 98.38%,综合评价指标为 94.4%,平均检测精度为 98.71%,单幅图像检测时间为 12.7 ms,模型占内存空间为 14.08 MB。与目前主流的一阶检测算法 YOLO v4 - tiny 和 RetinaNet 相比,其精确率分别提高了 1.99 个百分点和 4.50 个百分点,召回率分别提高了 9.41 个百分点和 10.77 个百分点,时间分别降低了 96.39% 和 96.25%。同时结果表明,该模型对密集、遮挡、昏暗环境和模糊虚化情况下的果实均能实现高精度识别与定位,具有较强的鲁棒性。研究结果可为自然复杂环境下油茶果机械采收及小目标检测等研究提供借鉴。

关键词:油茶果;自然场景;目标检测;YOLO v5s;深度学习;数据增强

中图分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1000-1298(2022)07-0234-09　　**OSID**:

Camellia oleifera Fruit Detection in Natural Scene Based on YOLO v5s

SONG Huaibo[1,2]　WANG Ya'nan[1,2]　WANG Yunfei[1,2]　LÜ Shuaichao[1,2]　JIANG Mei[1,2]

(1. *College of Mechanical and Electronic Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*
2. *Key Laboratory of Agricultural Internet of Things*, *Ministry of Agriculture and Rural Affairs*, *Yangling*, *Shaanxi* 712100, *China*)

Abstract: In view of the characteristics of small size, dense distribution and changeable color of *Camellia oleifera* fruit, in order to realize the rapid and accurate identification of *Camellia oleifera* fruit in complex natural scene, and determine the appropriate clamping position for the automatic oscillating harvesting device according to the density distribution of the fruit, the YOLO v5s convolutional neural network model was used to carry out research on the image detection method of *Camellia oleifera* fruit in the natural scene. Through data enhancement, totally 3 296 *Camellia oleifera* fruit images were obtained to make the PASCAL VOC data set. After 150 rounds of training, the optimal weight model was got. The accurate rate was 90.73%, the recall rate was 98.38%, the comprehensive evaluation index was 94.4%, the average detection accuracy was 98.71%, the single image detection time was 12.7 ms, and the memory size of the model was 14.08 MB. Compared with the current mainstream first-stage detection algorithms YOLO v4 - tiny and RetinaNet, its accuracy rate was increased by 1.99 percentage points and 4.50 percentage points, the recall rate was increased by 9.41 percentage points and 10.77 percentage points, and the time was reduced by 96.39% and 96.25%, respectively. In addition, the weight file of the YOLO v5s model was small, indicating that its network was simpler and had the advantage of rapid deployment. It could be transplanted to edge devices in the future to provide algorithm reference for the vision system of the *Camellia oleifera* fruit automatic harvesting device. Through comparative experiment, the results also showed that the model can achieve high-precision recognition and positioning of fruits in dense, occluded, dim environments and fuzzy blur conditions, and it had strong robustness. The research results can provide a reference for the research of mechanical harvesting of *Camellia oleifera* fruit under the natural complex environment.

Key words: *Camellia oleifera* fruits; natural scene; object detection; YOLO v5s; deep learning; data enhancement

收稿日期:2021-07-27　修回日期:2021-08-21

基金项目:国家重点研发计划项目(2019YFD1002401)和国家自然科学基金项目(31701326)

作者简介:宋怀波(1980—),男,教授,博士生导师,主要从事数字图像处理研究,E-mail:songhuaibo@nwsuaf.edu.cn

0　引言

油茶树是典型的长效型经济油料作物[1]，主要分布在我国南方的高山丘陵山地，是带动区域经济增长的优良树种。摇枝振荡式机械采收是目前最适合油茶果的采收方式，确定油茶果的疏密分布以确定振荡采收的夹持区域，是研发自动振荡采收机械的关键技术之一[2]，故实现复杂场景下油茶果果实目标的高效识别意义重大。

油茶果果实具有体积小、分布密集等特点，在自然环境下油茶果果实邻接、重叠、遮挡普遍存在，且易受不同光照条件的影响，识别难度较大。目前，国内外学者对果实目标识别与定位进行了大量研究[3-7]。刘天真等[8]提出了一种改进的 YOLO v3 - SE 模型进行冬枣果实识别，与传统的 YOLO v3 模型相比，该模型的综合评价指标提升了 2.38 个百分点，平均精度均值（Mean average precision，mAP）提升了 4.38 个百分点。刘小刚等[9]将卷积神经网络（Convolutional neural networks，CNN）用于草莓果实的识别之中，成熟草莓识别准确率为 97.14%，未成熟草莓识别准确率为 96.51%。陈斌等[10]利用 Faster - RCNN 模型开展了自然环境下油茶果检测方法研究，平均识别率为 92.39%，准确率为 98.92%，召回率为 93.32%，综合评价指标为 96.04%，平均每幅图像的识别时间为 0.21 s。李立君等[11]提出一种基于凸壳理论和颜色特征阈值分割法的遮挡油茶果定位检测方法，相比传统凸壳算法，该方法虽然耗时增加了 24.07%，但识别率提升了 7.47 个百分点。张习之等[12]提出了一种基于改进卷积自编码机神经网络的油茶果图像识别方法，利用非对称分解卷积核提升模型的训练速度，以输出端重构输入端的直连训练方式减少了信息损失，提升了模型收敛性能，识别准确率为 90.4%，单幅图像识别时间 2 s。陈斌等[13]提出了基于归一化傅里叶描述子、Hu 不变矩特征和 BP 神经网络的油茶果识别方法，识别率提升了 15%。

REDMON 等[14-15]提出的 YOLO 系列神经网络算法具有识别速度快、精度高等特点[16-22]。YOLO v5s 模型通过借鉴 CutMix 方法[23]，并采用 Mosaic 数据增强方式，可有效解决模型训练中的小目标问题。本文拟选择 YOLO v5s 作为油茶果目标识别模型，利用训练好的 YOLO v5s 网络实现油茶果目标的快速精准定位，以期为自然条件下油茶果目标识别提供一种有效方法。

1　材料与方法

1.1　材料

1.1.1　油茶果数据采集与预处理

试验数据采集于湖南省衡阳市衡南县大三湘油茶种植基地，油茶果成熟果实多为红色或红黄色，未成熟果实为青黄色和青色，受到品种和种植环境影响，部分油茶果表面颜色差异较大，增加了识别难度，部分示例图像如图 1 所示。数据采集时间为 2020 年 8 月 31 日—9 月 12 日，采集设备是华为 HMA - AL00 型相机，分辨率为 2 976 像素 × 3 968 像素，焦距 35 mm。

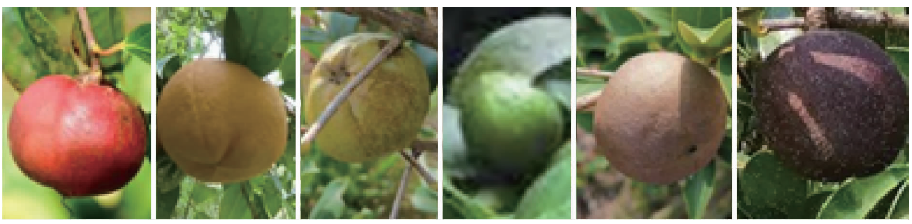

图 1　不同颜色的油茶果

Fig. 1　*Camellia oleifera* fruit with different colors

为保证所制作数据集的多样性，分别采集了白天、傍晚、遮挡、重叠、密集及逆光等环境下的油茶果图像 1 382 幅及视频 5 段，再从每段视频中随机选取 20 幅图像，共 1 482 幅图像作为原始数据，图像分辨率为 2 976 像素 × 3 968 像素。为了提高模型的泛化能力和鲁棒性，对原始图像进行了数据增强以增加训练的基础数据量。经典的数据增强[24]方法包括：图像翻转、图像旋转、图像缩放、图像裁剪、图像平移、添加噪声和修改对比度等。选择对部分图像添加椒盐噪声和高斯噪声进行了数据集增强，此操作可用来模拟特殊天气和镜头不清晰等情况，通过引入噪声点降低图像的清晰度，有利于验证算法的鲁棒性。图 2b 是对图 2a 原始图像添加了 2.5% 椒盐噪声之后的图像，图 2d 是对图 2c 原始图像添加了 12% 高斯噪声之后的图像，可以看出图像平滑度降低。同时，增加了 Mosaic 拼接法，即将多幅待检测图像各截取一部分合成一幅图像，随即遮挡图像部分进行整体检测，这种方式能够同时有效提升扰动条件下模型的检测准确性，拼接结果如图 2e 所示。

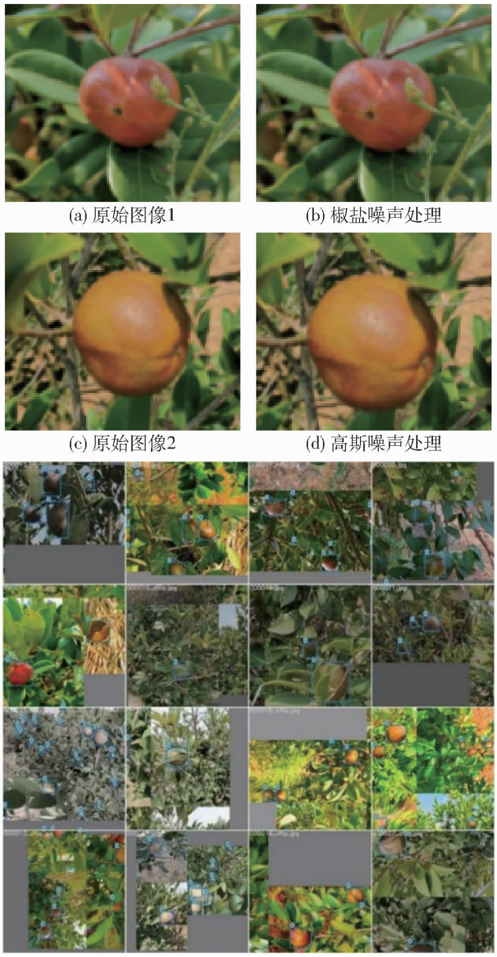

图2 数据增强

Fig. 2 Data enhancement

1.1.2 数据集制作

通过数据增强处理,结合所标注的原始样本,共获得3 296幅图像,按照比例4:4:1将图像分为简单、中等和困难3个等级:图像平滑且果实稀疏无阴影遮挡为简单样本,图像果实密集或轻微遮挡或有阴影为中等难度样本,图像果实超密集或严重遮挡或严重噪声、虚化为困难样本。采用LabelImg工具进行标注,生成.xml标签文件。由于本研究仅进行油茶果果实目标识别,因此在标签制作时只将图像分为果实目标与背景两类,仅需标注果实目标,图像其他部分LabelImg自动标注为背景,数据保存为PASCAL VOC格式[25]。

为测试算法的有效性和环境适应性,采集了695幅图像作为测试样本。在自然条件下进行果实识别,影响因素主要为果实重叠或枝叶遮挡、果实密集程度、光照环境和图像平滑度4方面,具体分类如表1所示。果实轮廓或果实表面被遮挡1/3以下属于轻微遮挡,2/3以上属于严重遮挡;光照因素按照图像光线明暗进行分类;图像中目标果实为1~3个属于稀疏图像,4~8个属于密集程度中等,9个果实及以上属于密集分布。果实数量增多会伴随着拍摄距离增加和果实目标变小的问题;噪声处理和虚化图像导致图像不清晰,进一步增加了识别难度。

表1 测试集分类

Tab. 1 Classification of test set

影响因素	图像类别	数量/幅
遮挡程度	未遮挡	60
	轻微遮挡	65
	严重遮挡	65
光照	亮光	60
	阴影	65
密集程度	稀疏	60
	中等	65
	密集	65
图像平滑度	清晰	60
	虚化或噪声	130

1.1.3 试验平台

本研究模型训练和测试均在同一环境下运行,硬件配置为英特尔Xeon E5-1620 v4 @ 3.50 GHz四核,GPU为11 GB的NVIDIA GeForce RTX 2080Ti,32 GB运行内存,软件环境为64位Windows 10系统,PyTorch深度学习框架。本文对比算法均在该硬件环境下进行。

1.2 方法

1.2.1 基于YOLO v5s的油茶果目标识别算法

YOLO系列深度学习框架利用回归的思想,以一阶网络直接完成目标位置检测和目标物体分类[26]。本试验基于YOLO v5s深度学习网络实现油茶果的检测,经过150轮训练得到最优模型。为了检测模型的优良性,选择当下认可度较高的YOLO v4-tiny和RetinaNet进行对比试验,用相同的数据集依次对网络进行训练,分别得到其最优模型。最后选取不同特征的低质量图像组成验证集对得到的模型分别进行验证,通过相关指标比较网络性能,验证模型的有效性,具体操作过程如图3所示。

1.2.2 YOLO v5s网络模型

YOLO v5s沿用了YOLO系列的整体布局,整个网络结构分成Input、Backbone、Neck和Output 4部分,网络结构如图4所示。在网络输入端自适应缩放图像,采用Mosaic数据增强方式,可自动计算数据集的最佳锚框值。Backbone部分在YOLO v4-tiny基础上增加了Focus结构和CSPNet(Cross stage partial networks)结构,实现了跨阶段局部融合网络。Focus结构共包含4次切片操作和1次32个卷积核的卷积操作,切片操作将1幅图像中每隔1个像素取1个值,类似于近邻下采样,得到4幅图像,但是4幅图像的信息并没有丢失,而是将原图像宽度W、

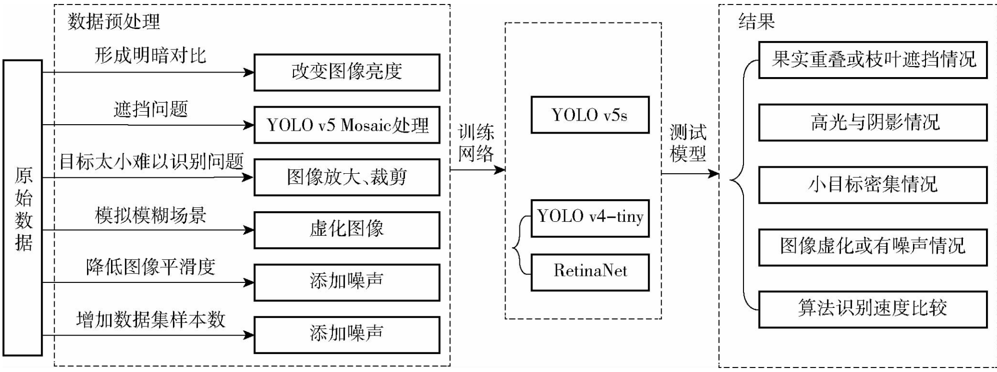

图 3　试验流程图

Fig. 3　Flow chart of experiment

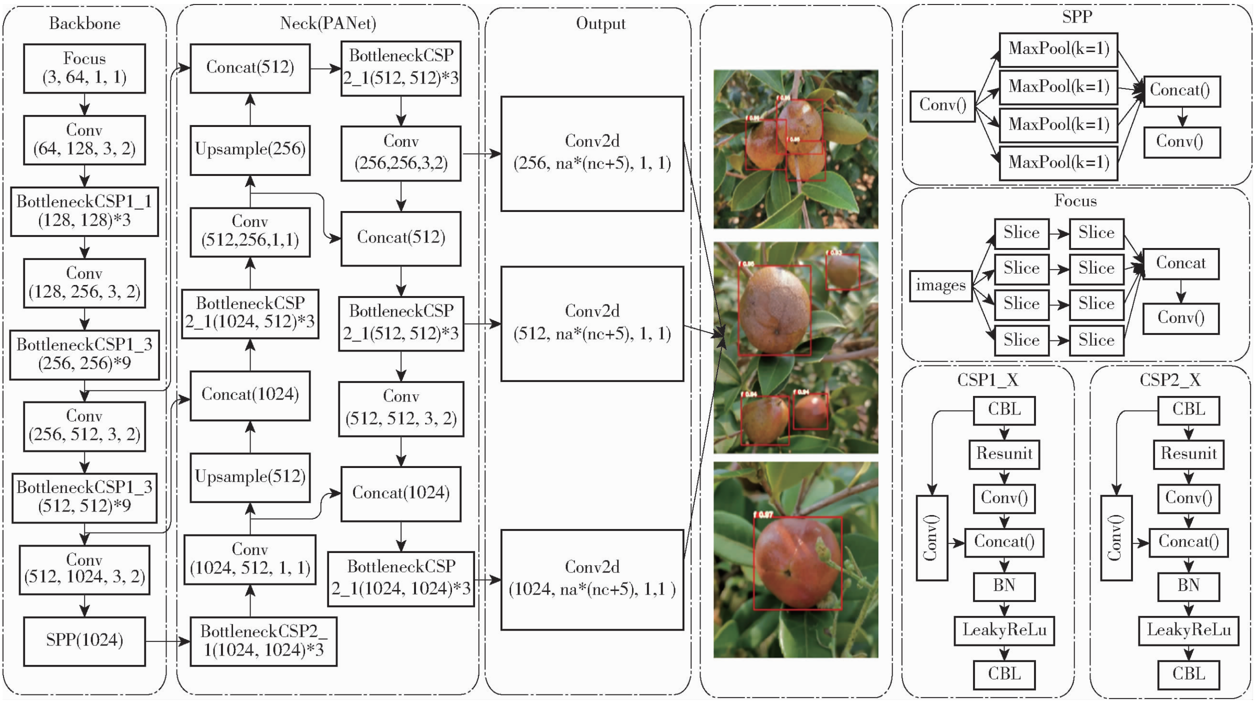

图 4　YOLO v5s 网络模型

Fig. 4　Network model of YOLO v5s

高度 H 信息集中到了通道空间，输入通道扩充了 4 倍，即拼接图像相对于原 RGB 3 通道模式变成了 12 个通道。然后将新得到的图像进行卷积操作，原始 $608 \times 608 \times 3$ 的图像变成信息完整的 2 倍 $304 \times 304 \times 32$ 的特征图。CSPNet 结构仿照 DenseNet 的密集跨层跳层连接思想，进行局部跨层融合，利用不同层的特征信息获得更为丰富的特征图。Neck 部分包含了路径聚合网络（Path aggregation network，PANet）[27] 和空间金字塔池化（Space pyramid pooling，SPP）[28] 模块。自顶向下将高层特征信息与不同层 CSP 模块的输出特征进行聚合，再通过自底向上路径聚合结构聚合浅层特征，从而充分融合不同层的图像特征。SPP 模块采用 1×1、5×5、9×9、13×13 共 4 种不同的池化核进行处理，处理后再进行张量拼接。SPP 处理可有效增加感受野，分离出显著的上下文特征，同时不损失原有检测速度。输出层使用 GIOU_Loss 代替 YOLO v3 的 IOU_Loss 作为损失函数，增加了相交尺度的衡量，缓解了 IOU_Loss 无法优化 2 个框不相交的情况。

1.2.3　对比算法及评价指标

为了验证 YOLO v5s 模型的有效性和实用性，本试验采用相同的油茶果数据集与 YOLO v4 - tiny 和 RetinaNet 网络进行了对比测试。选取精确率 P（Precision）、召回率 R（Recall）、调和平均值 F1（F-measure）和平均精度均值（mAP）作为评价指标进行性能对比。

1.2.4　权重模型训练

本试验共训练了 150 轮，权重模型序号记为 1 ~ 150。每 25 轮取一个权重模型，加上 P 值最高的模型，共 7 个权重模型，模型评价指标如表 2 所示。从表 2 可知，模型 24 综合性能最差；模型 49、74、99 评价指标逐渐变好，性能逐渐上升；模型 124、149 性

能逐渐稳定。模型134的F1值最高,说明该模型综合表现最好,油茶果目标识别的准确率为90.73%,召回率为98.39%,平均精度为98.71%,故选择利用模型134进行油茶果果实的识别。

表2 权值模型的性能比较

Tab.2 Performance comparison of weight models

%

模型序号	P	R	mAP	F1
24	56.25	94.45	93.44	70.5
49	76.27	96.17	96.61	85.1
74	80.81	98.45	98.07	88.8
99	86.83	98.51	98.44	92.3
124	88.46	98.45	98.64	93.2
149	88.63	98.78	98.85	93.0
134	90.73	98.39	98.71	94.4

2 试验结果与讨论

2.1 试验结果

通过用相同的数据集进行训练和测试,YOLO v5s、YOLO v4 - tiny和RetinaNet 3种网络的最优模型评估指标如表3所示。从表3可以看出,YOLO v5s所占内存空间最小,说明其网络更加简洁,模型具有更好的移植潜力。其他各评估指标均优于YOLO v4 - tiny和RetinaNet 2种网络。相比二者,YOLO v5s的P分别提高了1.99个百分点和4.50个百分点,R分别提高了9.41个百分点和10.77个百分点,mAP分别提高了9.48个百分点和10.79个百分点。

表3 评估指标对比

Tab.3 Comparison of evaluation indicators

模型	P/%	R/%	mAP/%	F1/%	模型占内存空间/MB
YOLO v5s	90.73	98.39	98.71	94.4	14.08
YOLO v4 - tiny	88.74	88.98	89.23	88.9	138.92
RetinaNet	86.23	87.62	87.92	86.9	138.08

测试时分别记录了3个模型的检测时间,并将每幅图像的平均耗时作为其检测速度。试验结果表明,针对单幅图像,YOLO v5s检测耗时12.7 ms,YOLO v4 - tiny检测耗时351.8 ms,RetinaNet检测耗时338.5 ms。相比2种模型,YOLO v5s检测速度明显提升,时间分别降低了96.39%和96.25%,所占内存空间分别减少124.84 MB和124 MB,表明YOLO v5s在速度方面更具优越性,更有利于将其移植至边缘设备中。

2.2 讨论

2.2.1 遮挡目标识别对比

为了进一步探究YOLO v5s对遮挡问题的处理能力,选取轻微遮挡和严重遮挡2种图像各65幅,分别输入2.1节得到的3个模型进行检测,部分图像的试验结果如表4所示。由表4可知,遮挡程度越大,检出目标的置信度越低。由于YOLO v5s采用了多层神经特征网络结构,对于更小的栅格预测更占优势,轻微遮挡和严重遮挡情况下均无漏检情况;YOLO v4 - tiny在2种情况下也没有漏检;RetinaNet虽无漏检,但严重遮挡情况下错检1个。

表4 遮挡情况检测结果对比

Tab.4 Comparison of detection results under occlusion

模型		实际果实个数	正确检出个数	平均置信度
YOLO v5s	轻微遮挡	4	4	0.96
	严重遮挡	5	5	0.92
YOLO v4 - tiny	轻微遮挡	4	4	0.97
	严重遮挡	5	5	0.67
RetinaNet	轻微遮挡	4	4	0.91
	严重遮挡	5	5	0.70

对应的检测图像如图5所示,其中图5a~5c为轻微遮挡图像检测结果,图5d~5f为严重遮挡图像结果。本研究检测结果统一使用红色方框标记,未识别结果采用蓝色方框标记,误识别结果采用蓝色箭头标记。从图5e和图5f可以看出,YOLO v4 - tiny和RetinaNet均存在定位不准确、严重遮挡情况下无法识别出果实目标等问题。可见,在果实目标重叠或者严重被遮挡情况下YOLO v5s依旧能准确分割果实边缘,对于果实目标预测取得了更好的效果。

2.2.2 阴影图像检测对比

亮度对果实目标识别的效果具有较大影响,昏暗环境下油茶果表面变得模糊,边缘特征清晰度降低,果实目标特征下降,增加了识别难度。为探究YOLO v5s对此类问题处理能力,选取逆光、阴天及光线不足环境下油茶果图像共65幅输入3个模型进行测试。部分试验结果如图6所示,其中不同部分图像从上到下、从左到右依次记为1、2、3、4,对应检测结果如表5所示,4幅图像中油茶果目标分别是4、2、2、1个。由表5可以看出,YOLO v5s将目标全部正确识别出,YOLO v4 - tiny漏检1个,RetinaNet漏检1个。从图6b来看,虽然YOLO v4 - tiny的单个置信度高,但其存在目标定位不准确的问题。由此可以看出,在同样的阴天或者傍晚环境下,YOLO v5s对油茶果目标识别在准确率和置信度上均更有优势。

2.2.3 小目标密集图像识别对比

为了验证YOLO v5s对密集小目标的识别效

(a) YOLO v5s, 轻微遮挡　(b) YOLO v4-tiny, 轻微遮挡　(c) RetinaNet, 轻微遮挡

(d) YOLO v5s, 严重遮挡　(e) YOLO v4-tiny, 严重遮挡　(f) RetinaNet, 严重遮挡

图 5　遮挡情况图像检测结果对比

Fig. 5　Comparison charts of detection results under occlusion

(a) YOLO v5s　(b) YOLO v4-tiny　(c) RetinaNet

图 6　逆光或昏暗环境图像检测结果对比

Fig. 6　Comparison charts of detection results in backlit or dim environment

表 5　阴影、逆光情况检测结果对比

Tab. 5　Detection results under shadow and backlight conditions

模型	实际果实个数	正确检出个数	平均置信度
YOLO v5s	9	9	0.93
YOLO v4 - tiny	9	8	0.88
RetinaNet	9	8	0.92

果，选取不同密集程度的油茶果图像共 190 幅分别输入 3 个模型进行检测。部分检测结果如图 7 所示，其中前 3 幅为密集分布图像，第 4 幅为超密集分布图像，对应的试验结果如表 6 所示。由表 6 可以看出，YOLO v5s 的识别效果最好。对于果实密集分布的图像，YOLO v5s 漏检 5 个，错误识别 1 个；YOLO v4 - tiny 漏检 23 个，错误识别 1 个；RetinaNet 漏检 22 个。对于果实超密集分布的图像，YOLO v5s 漏检 6 个，YOLO v4 - tiny 漏检 26 个，RetinaNet 漏检 26 个。综上可知，YOLO v5s 在小目标识别方面更具优势，可用于密集小目标识别之中。但也可以发现，YOLO v5s 在油茶果果实识别方面仍存在一定的不足：在识别小目标且伴随有果实超密集分布

图7 果实密集、小目标情况图像检测结果对比

Fig. 7 Comparison of detection results of dense fruits and small targets

的图像时，存在个别漏检情况。这是由于YOLO v5s神经网络通过多层卷积和池化等方式提取图像特征，图像的细节信息会随着网络深度的增加而有一定丢失。

表6 果实密集、小目标情况检测结果对比

Tab. 6 Comparison of detection results of dense fruits and small targets

模型		实际果实个数	正确检出个数	平均置信度
YOLO v5s	密集分布	45	40	0.89
	超密集分布	35	29	0.75
YOLO v4 - tiny	密集分布	45	22	0.84
	超密集分布	35	9	0.70
RetinaNet	密集分布	45	23	0.86
	超密集分布	35	9	0.33

2.2.4 虚化及噪声图像检测对比

为了探究YOLO v5s对模糊图像的检测能力，选取虚化、模糊和经过噪声处理后的图像共130幅分别输入3个模型进行测试。部分检测效果如图8所示，对应检测结果如表7所示。由表7可以看出，YOLO v5s的识别效果最好，无漏检；YOLO v4 - tiny漏检4个；RetinaNet漏检1个。结果表明，在识别不平滑图像时YOLO v5s更有优势，可以完整地检测出目标果实，具有良好的鲁棒性。

3 结论

（1）本研究将YOLO v5s网络用于油茶果的识别，经过训练得到的最优模型检测的 P 为90.73%，R 为98.38%，mAP为98.71%。与一阶检测算法

图8 虚化、噪声情况图像检测结果对比

Fig. 8 Comparison of detection results under blur and noise

表 7　虚化、噪声情况检测结果对比

Tab. 7　Comparison of detection results under blur and noise

模型	实际果实个数	正确检出个数	平均置信度
YOLO v5s	10	10	0.93
YOLO v4 − tiny	10	6	0.97
RetinaNet	10	9	0.88

YOLO v4 − tiny 和 RetinaNet 相比，模型 P 分别提高了 1.96 个百分点和 4.47 个百分点，R 分别提高了 9.40 个百分点和 10.76 个百分点。表明将 YOLO v5s 网络用于油茶果目标的识别是可行的。

（2）对比试验表明，YOLO v5s 模型具有检测精度高、抗干扰能力强的优点。在果实重叠或大面积被遮挡，逆光或昏暗环境，密集小目标等自然情况下都能达到良好的识别效果，在处理模糊、虚化、不平滑图像时均可准确识别定位果实目标。

（3）YOLO v5s 模型检测速度快，检测单幅图像的平均耗时仅为 12.70 ms，相比 YOLO v4 − tiny 和 RetinaNet 2 种算法，时间分别降低了 96.39% 和 96.25%。且模型体积小，具备更好的移植潜力。

参　考　文　献

[1] 伍德林，赵恩龙，姜山，等. 基于双摆模型的油茶果冠层振动参数优化与试验[J]. 农业机械学报，2021，52(12)：96 − 104.
WU Delin, ZHAO Enlong, JIANG Shan, et al. Optimization analysis and test of canopy vibration parameters of camellia fruit based on double pendulum model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(12): 96 − 104. (in Chinese)

[2] 伍德林，李超，曹成茂，等. 摇枝式油茶果采摘装置作业过程分析与试验[J]. 农业工程学报，2020，36(10)：56 − 62.
WU Delin, LI Chao, CAO Chengmao, et al. Analysis and experiment of the operation process of branch-shaking type camellia oleifera fruit picking device[J]. Transactions of the CSAE, 2020, 36(10): 56 − 62. (in Chinese)

[3] 李寒，陶涵虓，崔立昊，等. 基于 SOM − K − means 算法的番茄果实识别与定位方法[J]. 农业机械学报，2021，52(1)：23 − 29.
LI Han, TAO Hanhu, CUI Lihao, et al. Recognition and localization method of tomato based on SOM − K − means algorithm [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 23 − 29. (in Chinese)

[4] 苑进. 选择性收获机器人技术研究进展与分析[J]. 农业机械学报，2020，51(9)：1 − 17.
YUAN Jin. Research progress analysis of robotics selective harvesting technologies[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(9): 1 − 17. (in Chinese)

[5] 王文杰，贡亮，杨智宇，等. 基于多源图像融合的自然环境下番茄果实识别[J]. 农业机械学报，2021，52(9)：156 − 164.
WANG Wenjie, GONG Liang, YANG Zhiyu, et al. Tomato fruit recognition based on multi-source fusion image segmentation algorithm in open environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 156 − 164. (in Chinese)

[6] 杨长辉，刘艳平，王毅，等. 自然环境下柑橘采摘机器人识别定位系统研究[J]. 农业机械学报，2019，50(12)：14 − 22，72.
YANG Changhui, LIU Yanping, WANG Yi, et al. Research and experiment on recognition and location system for citrus picking robot in natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(12): 14 − 22, 72. (in Chinese)

[7] 熊俊涛，何志良，汤林越，等. 非结构环境中扰动葡萄采摘点的视觉定位技术[J]. 农业机械学报，2017，48(4)：29 − 33，81.
XIONG Juntao, HE Zhiliang, TANG Linyue, et al. Visual localization of disturbed grape picking point in non-structural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(4): 29 − 33, 81. (in Chinese)

[8] 刘天真，滕桂法，苑迎春，等. 基于改进 YOLO v3 的自然场景下冬枣果实识别方法[J]. 农业机械学报，2021，52(5)：17 − 25.
LIU Tianzhen, TENG Guifa, YUAN Yingchun, et al. Winter jujube fruit recognition method based on improved YOLO v3 under natural scene[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 17 − 25. (in Chinese)

[9] 刘小刚，范诚，李加念，等. 基于卷积神经网络的草莓识别方法[J]. 农业机械学报，2020，51(2)：237 − 244.
LIU Xiaogang, FAN Cheng, LI Jianian, et al. Identification method of strawberry based on convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 237 − 244. (in Chinese)

[10] 陈斌，饶洪辉，王玉龙，等. 基于 Faster − RCNN 的自然环境下油茶果检测研究[J]. 江西农业学报，2021，33(1)：67 − 70.
CHEN Bin, RAO Honghui, WANG Yulong, et al. Study on detection of camellia fruit in natural environment based on Faster − RCNN[J]. Acta Agriculturae Jiangxi, 2021, 33(1): 67 − 70. (in Chinese)

[11] 李立君，阳涵疆. 基于改进凸壳理论的遮挡油茶果定位检测算法[J]. 农业机械学报，2016，47(12)：285 − 292.
LI Lijun, YANG Hanjiang. Revised detection and localization algorithm for *Camellia oleifera* fruits based on convex hull theory [J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(12): 285 − 292. (in Chinese)

[12] 张习之，李立君. 基于改进卷积自编码机的油茶果图像识别研究[J]. 林业工程学报，2019，4(3)：118 − 124.
ZHANG Xizhi, LI Lijun. Research of image recognition of *Camellia oleifera* fruit based on improved convolutional auto-encoder [J]. Journal of Forestry Engineering, 2019, 4(3): 118 − 124. (in Chinese)

[13] 陈斌，饶洪辉，刘木华. 基于傅里叶描述子和 Hu 不变矩的油茶果识别[J]. 浙江农业科学，2020，61(9)：1876 − 1880.
CHEN Bin, RAO Honghui, LIU Muhua. Recognition of *Camellia oleifera* fruit based on Fourier descriptor and Hu invariant

moments[J]. Journal of Zhejiang Agricultural Sciences, 2020, 61(9): 1876 – 1880. (in Chinese)

[14] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017: 6517 – 6525.

[15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas: 2016: 779 – 788.

[16] 李菊霞,李艳文,牛帆,等. 基于 YOLO v4 的猪只饮食行为检测方法[J]. 农业机械学报,2021,52(3):251 – 256.
LI Juxia, LI Yanwen, NIU Fan, et al. Pig diet behavior detection method based on YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 251 – 256. (in Chinese)

[17] 赵德安,吴任迪,刘晓洋,等. 基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位[J]. 农业工程学报,2019,35(3):164 – 173.
ZHAO Dean, WU Rendi, LIU Xiaoyang, et al. Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background[J]. Transactions of the CSAE, 2019, 35(3): 164 – 173. (in Chinese)

[18] 薛月菊,黄宁,涂淑琴,等. 未成熟芒果的改进 YOLO v2 识别方法[J]. 农业工程学报,2018,34(7):173 – 179.
XUE Yueju, HUANG Ning, TU Shuqin, et al. Immature mango detection based on improved YOLO v2[J]. Transactions of the CSAE, 2018, 34(7): 173 – 179. (in Chinese)

[19] TIAN Yunong, YANG Guodong, WANG Zhe, et al. Apple detection during different growth stages in orchards using the improved YOLO – V3 model[J]. Computers and Electronics in Agriculture, 2019, 157: 417 – 426.

[20] 何东健,刘建敏,熊虹婷,等. 基于改进 YOLO v3 模型的挤奶奶牛个体识别方法[J]. 农业机械学报,2020,51(4):250 – 260.
HE Dongjian, LIU Jiamin, XIONG Hongting, et al. Individual identification of dairy cow based on improved YOLO v3 network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(4): 250 – 260. (in Chinese)

[21] 何斌,张亦博,龚健林,等. 基于改进 YOLO v5 的夜间温室番茄果实快速识别[J]. 农业机械学报,2022,53(5):201 – 208.
HE Bin, ZHANG Yibo, GONG Jianlin, et al. Fast recognition of tomato fruit in greenhouse at night based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 201 – 208. (in Chinese)

[22] 宋怀波,王云飞,段援朝,等. 基于 YOLO v5 – MDC 的重度粘连小麦籽粒检测方法[J]. 农业机械学报,2022,53(4):245 – 253.
SONG Huaibo, WANG Yunfei, DUAN Yuanchao, et al. Detection method of severe adhesive wheat grain based on YOLO v5 – MDC mothod[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 245 – 253. (in Chinese)

[23] YUN S, HAN D, CHUN S, et al. CutMix: regularization strategy to train strong classifiers with localizable features[EB/OL]. (2019 – 08 – 07)[2021 – 01 – 15]. https://arxiv.org/abs/1905.04899v2.

[24] 李永盛,何佳洲,刘义海,等. 基于图像检测识别的数据增强技术[J]. 舰船电子对抗,2021,44(1):66 – 70.
LI Yongsheng, HE Jiazhou, LIU Yihai, et al. Data enhancement technology based on image detection and recognition[J]. Shipboard Electronic Countermeasure, 2021, 44(1): 66 – 70. (in Chinese)

[25] MARK E, LUC VAN G, CHRISTOPHE K L, et al. The pascal visual object classes (VOC) challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303 – 338.

[26] WANG Xu, TANG Julie, MARK Whitty. Deep phenology: estimation of apple flower phenology distributions based on deep learning[J]. Computers and Electronics in Agriculture, 2021, 185: 106123.

[27] TRINH H C, LE D H, KWON Y K. PANET: a GPU-based tool for fast parallel analysis of robustness dynamics and feed-forward/feedback loop structures in large-scale biological networks[J]. PLoS ONE, 2014, 9(7): 1 – 9.

[28] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904 – 1916.