

DOI: 10.13733/j.jcam.issn.2095-5553.2026.04.019

马保建, 陈棒棒, 何婉凌, 等. 基于改进 YOLOv8s 的轻量化冬枣识别方法[J]. 中国农机化学报, 2026, 47(4): 126-132, 146

Ma Baojian, Chen Bangbang, He Wanling, et al. Lightweight winter jujube recognition method based on improved YOLOv8s [J]. Journal of Chinese Agricultural Mechanization, 2026, 47(4): 126-132, 146

基于改进 YOLOv8s 的轻量化冬枣识别方法*

马保建¹, 陈棒棒¹, 何婉凌¹, 张香玉¹, 蒋焕煜²

(1. 新疆理工学院机电工程学院, 新疆阿克苏, 843100;

2. 浙江省农业智能装备与机器人重点实验室, 杭州市, 310058)

摘要:为解决自然环境下冬枣果实的快速识别以及边缘设备部署计算资源有限的问题,提出基于 YOLOv8s 模型的轻量化冬枣识别方法 YOLOv8s-SF1,以便后续将模型部署到边缘设备上。将 StarNet 轻量化网络结构替代 YOLOv8s 的骨干提取网络,以降低原模型的复杂度;将 FasterNet 引入 C2f 模块中构建一个新的结构 C2f-Faster,提取更丰富的特征的同时降低计算资源的消耗;提出一种新的网络检测头 Detect_LSCD,进一步降低存储量和模型复杂度。试验结果表明,改进 YOLOv8s-SF1 模型的帧率为 313.5 帧/s,相比于原模型提高 14.66%,而模型大小和参数量分别为 6.8 MB 和 3.471 M,相对原模型分别减少 68.37% 和 68.80%,并且精确率、召回率和平均精度均值相对原模型仅下降 1% 左右。改进的 YOLOv8s-SF1 模型大小相比于 YOLOv5s、YOLOv8s、YOLOv9s 和 YOLOv10s 分别减少 61.58%、68.37%、53.10% 和 56.96%,参数量分别减少 61.90%、68.80%、51.57% 和 51.91%,而且浮点计算量分别降低 50.42%、58.45%、55.80% 和 44.86%。

关键词:冬枣;自然环境;轻量化模型;YOLOv8s;智能化采收

中图分类号:TP391.4 文献标识码:A 文章编号:2095-5553(2026)04-0126-08

Lightweight winter jujube recognition method based on improved YOLOv8s

Ma Baojian¹, Chen Bangbang¹, He Wanling¹, Zhang Xiangyu¹, Jiang Huanyu²

(1. College of Mechanical and Electrical Engineering, Xinjiang Institute of Technology, Aksu, 843100, China;

2. Key Laboratory of Intelligent Equipment and Robotics for Agriculture of Zhejiang Province, Hangzhou, 310058, China)

Abstract: To solve the problem of rapid recognition of winter jujube in natural environment and limited computing resources for edge devices, a lightweight recognition method of winter jujube, YOLOv8s-SF1, based on YOLOv8s model was proposed, so as to deploy the model to edge devices later. Firstly, the backbone extraction network of YOLOv8s is replaced by StarNet lightweight network to reduce the complexity of the original model. Secondly, FasterNet is introduced into C2f module to build a new structure, C2f-Faster, to reduce the consumption of computing resources. Finally, a new network detection head, Detect_LSCD, is proposed to further reduce the storage capacity and model complexity. The experimental results show that the frame rate of YOLOv8s-SF1 model is 313.5 frames/s, which is 14.66% higher than that of the original model, while the weight size and parameter number are 6.8 MB and 3.471 M, respectively, which are reduced by 68.37% and 68.80% compared with the original model. The Precision rate, recall rate and average accuracy are only reduced by about 1% compared with the original model. The model size of the improved YOLOv8s-SF1 is reduced by 61.58%, 68.37%, 53.10% and 56.96% compared to the YOLOv5s, YOLOv8s, YOLOv9s and YOLOv10s, respectively. The number of parameters decreased by 61.90%, 68.80%, 51.57% and 51.91% respectively, and the floating-point calculation amount decreased by 50.42%, 58.45%, 55.80% and 44.86%.

Keywords: winter jujube; natural environment; lightweight models; YOLOv8s; intelligent harvesting

收稿日期:2024 年 10 月 16 日 修回日期:2025 年 10 月 16 日

* 基金项目:新疆维吾尔自治区自然科学基金(2022D01C357)

第一作者:马保建,男,1987 年生,山东曹县人,博士,副教授;研究方向为农业机器人。E-mail: 1009384790@qq.com

通讯作者:陈棒棒,男,1992 出生,甘肃陇西人,硕士,讲师;研究方向为智能农业装备。E-mail: 1359129289@qq.com

0 引言

冬枣的人工采收严重制约该产业规模化发展^[1]。虽然针对鲜食冬枣选择性采摘田间机器人还未见报道,但随着人工智能和机器人技术的发展,鲜食冬枣的自动化采收成为可能。枣园中鲜食冬枣生长比较密集、体积小并且遮挡比较严重,对于冬枣采摘机器人首要解决的问题是在复杂自然环境下实现冬枣果实快速识别。与枣类相关的小目标果实识别,比如樱桃、小番茄、草莓、油茶果、苹果幼果、青皮核桃等已经得到了广泛研究^[2]。苗荣慧等^[3]通过改变YOLOv7的骨干网络,在不降低准确率的前提下使模型轻量化,以提高樱桃的检测速度,单幅樱桃图像的检测时间为82 ms,模型大小为66.5 MB。毕松等^[4]利用YOLOv3模型对高架栽培种植的草莓进行识别并实现采摘点的定位,单个草莓识别耗时为92 ms,定位误差为2.3 mm,基本满足温室条件下草莓自动采摘要求。宋怀波等^[5]利用YOLOv5s模型对自然条件下的油茶果进行检测,相比于YOLOv4-Tiny模型,检测速度和精度都有提升,单幅油茶果图像检测速度为12.7 ms,模型大小为14.08 MB。郝建军等^[6]提出了基于MobilNetv3骨干网络的YOLOv3模型,提升了模型轻量化程度,改进模型对青皮核桃检测平均精度均值*mAP*为94.52%,检测速度为31帧/s。

目前,枣类果实识别也得到许多学者的关注,比如王铁伟等^[7]提出改进的Faster R-CNN对自然环境下的红枣进行识别,通过数据平衡提高红枣识别精度。王昱潭等^[8]结合改进的Faster R-CNN对灵武长枣进行识别,识别准确率进一步提高到98.26%。同样地,Wu等^[9]利用Faster R-CNN与传统图像处理算法分别对红枣进行识别,对比试验发现Faster R-CNN有较高的检测精度,但未涉及单幅图像的检测速度。虽然基于两阶段的检测算法对枣识别精度高,但整个网络比较复杂而且检测过程消耗的时间较长,而单阶段检测方法基于回归的原理,模型轻量化且检测速度相对较快,在枣果实识别方面也得到应用。王昱潭等^[10]提出了改进SSD轻量化的模型对灵武长枣进行识别,结果表明,平均准确率为96.6%,且检测速度为28.05帧/s。刘天真等^[11]提出YOLOv3-SE模型对自然环境下的冬枣进行检测,冬枣识别的准确率、召回率分别为88.71%和83.80%,但未提及模型大小和参数量。王菁等^[12]利用YOLOv4-Tiny模型对不同品种和不同成熟度的枣进行检测,结果表明,该模型的平均精度均值*mAP*为94%。同时,也有学者利用图像分割的方法对枣进行识别研究,比如薛君蕊等^[13]利用改进的FCN-8s对灵武长枣图像进行分割,其中分割

结果的评价指标相比于原FCN-8s模型都有提高,图像分割速度为16.20帧/s。

针对自然场景下冬枣小目标识别过程中存在模型计算量大、检测速度慢、不利于部署到边缘设备的问题,本文设计一种基于YOLOv8s改进的轻量化方法。改进的YOLOv8s利用StarNet轻量化网络替换原来的骨干网络,降低模型的复杂度,并将FasterNet引入C2f模块中,减少计算量。同时,提出一种全新的检测头Detect_LSCD,在保证模型检测精度的前提下进一步减小模型参数以及计算量。

1 改进模型构建

1.1 数据集采集与制作

利用佳能相机EOS R50采集冬枣图像,单幅图像分辨率为6 000像素×4 000像素。采集地点为浙江义乌某农场冬枣种植基地。在自然条件下获取冬枣图像,采用不同的获取角度和距离对具有不同遮挡程度、不同视场的冬枣进行获取,共采集1 783幅冬枣图像,图1为部分采集的冬枣数据。数据集采用PASCAL VOC的格式,采用Labelimg软件对图像中的冬枣进行标注,得到对应的xml格式文件,标注过程中,对遮挡非常严重的冬枣不进行标注,且标注框尽可能贴合冬枣轮廓边缘,将标注后的图像样本按照7:2:1的比例进行划分,其中训练集图像1 248幅,测试集图像357幅,验证集图像178幅。



图1 冬枣样本数据

Fig. 1 Winter jujube images

1.2 YOLOv8s模型改进

1.2.1 YOLOv8s轻量化网络结构设计

YOLOv8^[14]模型基于尺度缩放因子提供n、s、m、l、x不同大小的模型,以满足不同的实际应用需求。为降低网络结构的计算冗余和参数量,并权衡模型检测速度和精度,选用YOLOv8s作为基线模型,该模型主要包括特征提取Backbone模块、特征融合Neck模块和检测头Head模块。相比于YOLOv5模型^[15],主

要改变是将特征提取和特征融合模块中的 C3 模块替换为 C2f 模块,也是仿照 ResNet 模型^[15]的跳层连接的思路,且去除采样过程中的卷积操作,进一步轻量化。SPPF 模块借鉴空间金字塔池化模块(SPP)进行设计,实现多尺度信息融合以提高模型的性能,同时能够加快处理速度。Head 模块中采用解耦头的结构(Decoupled-Head),把分类和检测分开处理,同时采用无锚点方法(Anchor-Free),减少对锚框的依赖,以满足检测形状大小变化的目标。基于基线模型设计轻量化的 YOLOv8s-SF1 结构,如图 2 所示。

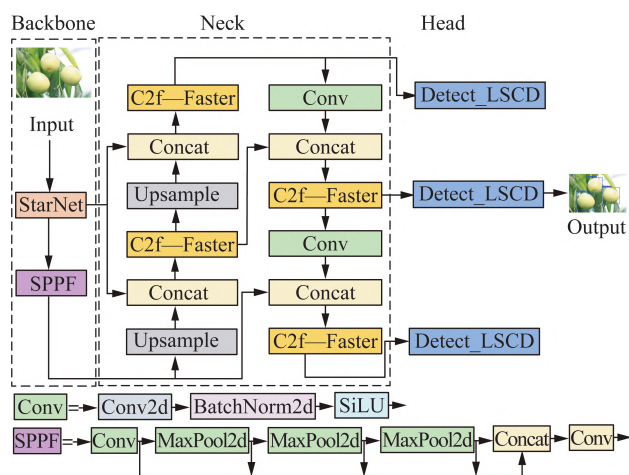


图 2 YOLOv8s-SF1 结构

Fig. 2 YOLOv8s-SF1 structure

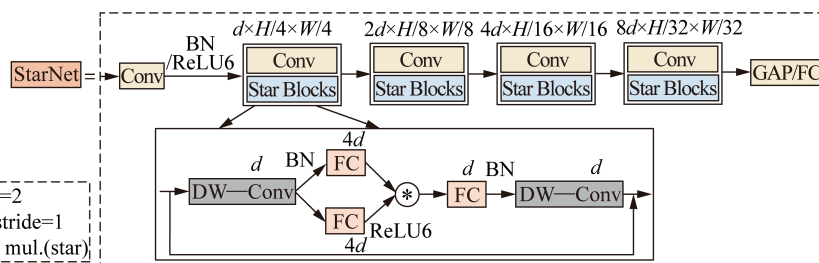


图 3 StarNet 结构

Fig. 3 StarNet structure

1.2.3 C2f-Faster 模型的改进

YOLOv8s 模型采用的 C2f 网络结构中包含大量的串联 Bottleneck,导致网络的过度冗余。利用 FasterNet 模型中的部分卷积模块(PCConv)设计 FasterNet Block,并以类继承的形式替换原来 C2f 网络中的所有 Bottleneck,构建 C2f-Faster 模型,在特征融合 Neck 部分替换全部的 C2f 模块,使整体网络结构轻量化,进一步减少特征融合过程中的计算量和浮点数,C2f-Faster 网络结构如图 4 所示。其中 PCConv 仅在部分网络通道上进行部分卷积并不影响其余网络通道,在保持其他通道不变的条件充分进行特征提取,这样既利用了设备的计算能力,又减小计算复杂度。

将原特征提取模块替换为 StarNet,同时引入 Faster 模块到 C2f 模块中,并提出一种轻量化检测头(Detect_LSCD)替换原有的检测头,进一步降低模型参数量和计算复杂度,有效提升检测速度。

1.2.2 StarNet 轻量化网络

YOLOv8s 的骨干网络利用 DarkNet5 框架自上而下的特征提取方式,为后续在边缘设备上轻量化部署并保证有效地识别冬枣精度,引入更高效的轻量化 StarNet 网络进行特征提取,StarNet 网络利用分层网络思想,包括 4 个阶段,每个阶段首先利用 Conv 卷积进行下采样操作(卷积核大小为 3,步长为 2),然后经过 Star Blocks 层进行特征提取,并在最后阶段加入深度卷积(卷积核大小为 7,步长为 1),使得每个阶段的通道数量增加一倍(从通道数 d 扩展到 $8d$),其中通道扩展因子设置为 4。而特征图的尺寸从 $(H/4) \times (W/4)$ 减小到 $(H/32) \times (W/32)$,其中 H 、 W 分别表示原特征图的高和宽。每个阶段网络结构宽度加倍,同时为提升特征提取的效率,将层归一化(Layer Normalization)替换成批归一化(Batch Normalization)操作。

为避免特征提取的复杂运算,通过*操作将输入映射到高维空间,在网络单层中通过矩阵元素逐点相乘进行融合特征,StarNet 网络结构如图 3 所示。

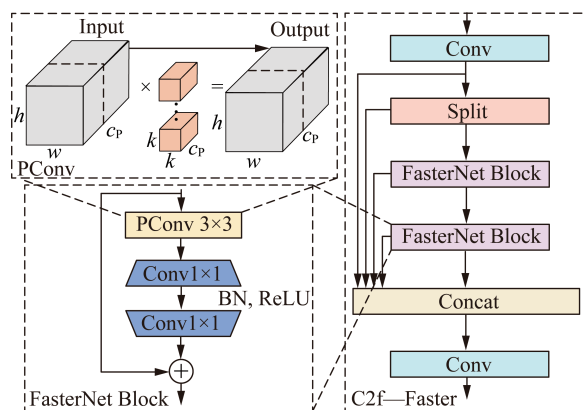


图 4 C2f-Faster 结构

Fig. 4 C2f-Faster structure

图 4 中,假设 PCConv 中 h 、 w 为输入特征图的高与宽, c 代表输入的通道数量, c_p 代表参与计算的部分

通道,由于只有部分通道参与提取特征,在保证通道信息时,可减少计算量和访问内存量。 k 表示卷积核的尺寸, r 为参与的卷积率,计算如式(1)所示,则 PConv 的计算量 F_{PConv} 计算如式(2)所示,而内存的访问量 M_{AC} 计算如式(3)所示。

$$r = \frac{c_p}{c} \quad (1)$$

$$F_{PConv} = h \cdot w \cdot k^2 \cdot c_p^2 \quad (2)$$

$$M_{AC} = h \cdot w \cdot 2c_p + k^2 \cdot c_p^2 \approx h \cdot w \cdot 2c_p \quad (3)$$

1.2.4 轻量化检测头 Detect_LSCD

原 YOLOv8s 模型的每个单一尺度的检测头需要 2 个 3×3 的卷积和 1 个 1×1 卷积提取信息,导致大量模型参数,基于此,利用组归一化(GN)^[19]对卷积层进行重新设计(Conv_GN),并提出利用共享卷积层 Conv_GN(3×3)减少参数量,使得检测头更轻。同时,为解决检测头所检测的目标物尺度不一的问题,利用 Scale 层对提取特征进行不同程度的缩放处理,以提升检测头的性能。新的轻量化检测头 Detect_LSCD 结构如图 5 所示。

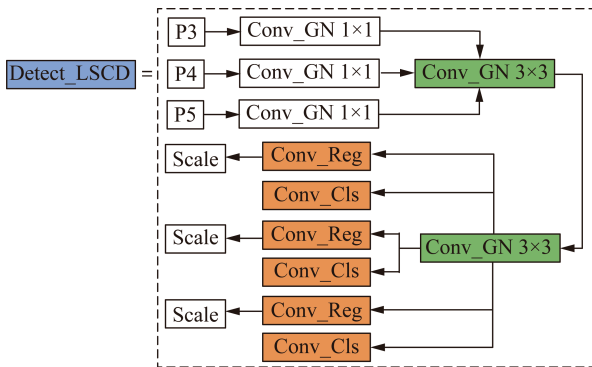


图 5 检测头 Detect_LSCD 结构
Fig. 5 Detect_LSCD structure

Detect_LSCD 同样采用解耦头结构,通过 3 个 1×1 Conv_GN 卷积层输出到共享的 3×3 Conv_GN 层,然后再通过一次共享卷积并使用 Scale 层对特征进行缩放,最后以不同缩放尺度分支输出分类和检测结果。

1.3 模型训练条件

整个模型训练在 Windows 10 系统上进行,电脑硬件配置内存 32 GB, GTX3080 和 i7 处理器。模型配置需要搭配的库有 Pytorch 2.2、Anaconda 3.8、Python 3.9、CUDA 12.1 和 cuDNN 11.3。在模型训练过程中没有选择 ImageNet 预训练权重,训练过程中数据批量设置为 4,初始学习率为 0.01,采用的优化器为 SGD,通过预试验确定训练轮次为 150 个 Epochs,训练结束后保存精度最优以及最后的训练权重,利用最优的权重对红枣测试数据集进行验证,得到检测结果。

1.4 模型评价指标

为验证改进模型的有效性,试验采用不同的模型对相同的测试数据集进行测试试验,利用精确率 P 、召回率 R 、平均精度均值 $mAP@0.5$ 来评估模型的精度;采用模型大小、参数量和浮点计算量 $GFLOPs$ 对模型进行复杂程度评价;以每秒传输帧数 FPS 评估模型的检测速度。

2 结果与分析

2.1 训练结果评估

使用 YOLOv8-SF1 模型在数据集上进行训练与测试,训练过程如图 6(a)所示,在前 30 个 Epochs 中,模型损失值 box_loss 快速下降,当迭代到 100 个轮次时,损失曲线基本趋近平缓并稳定在最小值,完成训练过程,表明没有出现拟合现象。如图 6(b)所示,在前 20 个 Epochs 迭代过程中,模型的精确率 P 上升最快,随着训练轮次的增加,在约 40 个 Epochs 时,模型精确率 P 趋于稳定,最终稳定在 95% 左右,改进的模型训练结束并达到最优训练结果,表明改进模型的有效性。

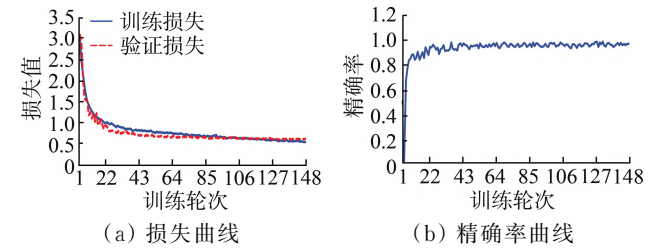


图 6 训练过程
Fig. 6 Training process

2.2 轻量化模型消融实验

轻量化模型消融实验结果如表 1 所示。从实验 2 可知,利用 StarNet 替换原有的骨干网络后,模型大小和参数量分别为基线模型的 59.07% 和 58.81%,而且模型的检测速度也提升 9.9 个百分点,原因在于 StarNet 利用高效的 * 操作运算直接将特征映射到高维空间,避免了复杂的计算过程,降低了网络的复杂度。实验 3 在 YOLOv8s 模型中把原来的 C2f 替换为 C2f-Faster 模块,模型大小为基线模型的 74.88%,且精确率、平均精度均值均有所提升,原因在于通过部分卷积 PConv 能够减少不必要的网络通道并有效地提取特征,进而减少模型大小和计算冗余。实验 4 替换为 Detect_LSCD 检测头后,模型精确率有略微下降,但其他性能指标都得到了提升,说明提出的检测头结构能够在参数量与计算量更少的前提下确保检测精度。实验 5~实验 7 分别是 StarNet 和 C2f-Faster 模块组合、StarNet 和 Detect_LSCD 检测头组合以及 C2f-Faster 和 Detect_LSCD 检测头组合,从

试验数据可以看出,精确率、召回率以及平均精度均值都有略微下降,但模型大小和参数量均大幅减小,其中减少最多的是 StarNet 和 Detect_LSCD 检测头组合,模型大小为基线模型的 44.18%,且参数量为基线模型的 43.56%。实验 8 为结合 StarNet、C2f—Faster 和 Detect_LSCD 检测头的 YOLOv8s—SF1 模型的试

验结果,可以看出,相对于基线模型,精确率、召回率以及平均精度均值分别下降 1.7 个百分点、0.9 个百分点和 0.8 个百分点,改进后 YOLOv8s—SF1 模型的这 3 个指标略微降低,但模型大小、参数量和 GFLOPs 分别下降 68.37%、68.80%、58.45%,且 FPS 提高 14.66%,表明所提轻量化方法的有效性。

表 1 消融实验

Tab. 1 Ablation test

实验	StarNet	C2f—Faster	Detect_LSCD	$P/\%$	$R/\%$	$mAP@0.5/\%$	模型大小/MB	参数量/M	GFLOPs	FPS/(帧·s ⁻¹)
1	×	×	×	95.8	77.1	89.7	21.5	11.126	28.4	273.4
2	√	×	×	96.4	75.3	87.7	12.7	6.543	17.3	300.5
3	×	√	×	97.3	76.2	89.8	16.1	8.305	21.4	280
4	×	×	√	94.9	77.7	89.7	18.2	9.430	25.8	289.4
5	√	√	×	94.3	76.6	88.4	10.1	5.167	14.5	271
6	√	×	√	95.4	75.9	87.8	9.5	4.847	14.7	298.8
7	×	√	√	93.2	79.2	89.8	12.8	6.609	18.7	286.7
8	√	√	√	94.1	76.2	88.9	6.8	3.471	11.8	313.5

2.3 不同轻量化模型性能对比

试验以 YOLOv8s 为基线模型,利用目前主流轻量化主干网络 MobileNetV4^[16]、FasterNet、EfficientViT^[17]、GhostNet^[18]和 StarNet 替换基线模型的主干,在训练时,除了主干网络之外的其他参数都保持一致,试验结果如表 2 所示。使用 StarNet 进行骨干改进的方法相较于其

他网络,精确率 P 、召回率 R 和平均精度均值 $mAP@0.5$ 没有明显降低,在部分性能上还优于其他网络,而且模型大小、参数量和 GFLOPs 都达到最小,分别为 12.7 MB、6.543 M 和 17.3, FPS 达到最大值 300.5 帧/s。结果表明,StarNet 网络在不牺牲网络性能的基础上,降低模型的计算量和计算复杂度。

表 2 不同轻量化模型性能对比

Tab. 2 Performance comparison of different lightweight models

轻量化骨干模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	模型大小/MB	参数量/M	GFLOPs	FPS/(帧·s ⁻¹)
MobileNetV4	95.1	77.5	90.2	20.6	10.626	33.9	225.4
FasterNet	96.6	75.6	87.5	16.7	8.616	21.7	237.9
EfficientViT	96.0	75.1	89.4	16.7	8.382	20.4	93.1
GhostNet	94.7	77.4	90.2	16.1	8.270	20.6	264.7
StarNet	96.4	75.3	87.7	12.7	6.543	17.3	300.5

2.4 不同检测模型性能对比试验

将改进的 YOLOv8s—SF1 模型与 YOLOv5s、YOLOv8s、YOLOv9s 和 YOLOv10s 进行对比,结果如表 3 所示。YOLOv8s—SF1 模型的精确率、召回率、平均精度均值相比于其他模型有略微的下降,但 YOLOv8s—SF1 模型的模型大小、参数量、GFLOPs 和 FPS 都有大幅度提升。其中,模型大小相比于 YOLOv5s、

YOLOv8s、YOLOv9s 和 YOLOv10s 分别减少 61.58%、68.37%、53.10% 和 56.96%,同时参数量分别减少 61.90%、68.80%、51.57% 和 51.91%,浮点计算量分别降低 50.42%、58.45%、55.80% 和 44.86%,FPS 分别提升 9.3 个百分点、14.7 个百分点、33.6 个百分点和 20.7 个百分点。说明改进 YOLOv8s—SF1 模型轻量化效果较好,便于后续模型的部署。

表 3 不同模型的性能对比

Tab. 3 Performance comparison of different models

模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	模型大小/MB	参数量/M	GFLOPs	FPS/(帧·s ⁻¹)
YOLOv5s	96.6	75.8	89.8	17.7	9.112	23.8	286.8
YOLOv8s	95.8	77.1	89.7	21.5	11.126	28.4	273.4
YOLOv9s	95.6	77.9	89.1	14.5	7.167	26.7	234.6
YOLOv10s	94.4	76.8	89.1	15.8	7.218	21.4	259.8
YOLOv8s—SF1	94.1	76.2	88.9	6.8	3.471	11.8	313.5

2.5 不同场景下冬枣检测结果对比试验

一般情况下,对模型进行轻量化改进会导致模型精度下降,为进一步验证轻量化后的 YOLOv8s—SF1 模型的性能,分别构建无遮挡、有遮挡、顺光、逆光和大视场 5 种自然条件下的数据集各 50 张。不同场景下冬枣检测性能试验结果如表 4 所示。图 7 为不同模型在 5 种场景下的检测可视化效果。

在无遮挡情况下, YOLOv8s—SF1 模型与原模型在 P 、 R 和 $mAP@0.5$ 指标上基本接近,相对其他模型, P 、 R 和 $mAP@0.5$ 下降最大分别为 1.2 个百分点、1.1 个百分点和 0.6 个百分点。

在有遮挡情况下, YOLOv8s—SF1 模型相对原模型, R 和 $mAP@0.5$ 分别下降 0.2 个百分点和 1.6 个百分点,而相对其他模型, P 、 R 和 $mAP@0.5$ 下降最大分别为 1.4 个百分点、1.2 个百分点和 1.6 个百分点。

在顺光和逆光条件下, YOLOv8s—SF1 模型与其他模型在 P 、 R 和 $mAP@0.5$ 指标上相差不大,表明提出的改进模型对于不同光照条件下的冬枣识别具有一定的鲁棒性。但 YOLOv8s—SF1 模型在顺光条件下的各个性能指标均略高于逆光条件下的结果,原因在于逆光条件下冬枣图像的亮度较低,识别难度较大,而且 YOLOv5s 和 YOLOv8s 模型还出现错检的情况。

在大视场情况下, YOLOv8s—SF1 模型与原模型在 P 、 R 和 $mAP@0.5$ 指标上基本接近,而相对其他模型, P 、 R 和 $mAP@0.5$ 下降最大分别为 2.5 个百分点、1.0 个百分点和 1.3 个百分点。

表 4 不同场景下冬枣检测性能

Tab. 4 Detection performance under different scenarios

场景	模型	$P/\%$	$R/\%$	$mAP@0.5/\%$
无遮挡情况	YOLOv5s	99.2	72.7	86.5
	YOLOv8s	98.7	72.7	87.1
	YOLOv9s	99.1	72.1	87.1
	YOLOv10s	99.2	72.3	87.0
	YOLOv8s—SF1	98.0	71.6	86.5
有遮挡情况	YOLOv5s	92.8	87.8	94.0
	YOLOv8s	89.0	88.4	94.3
	YOLOv9s	87.8	87.8	92.4
	YOLOv10s	88.5	89.4	94.3
	YOLOv8s—SF1	91.4	88.2	92.7
顺光	YOLOv5s	99.1	71.5	85.3
	YOLOv8s	98.6	73.1	86.4
	YOLOv9s	98.9	71.9	86.9
	YOLOv10s	98.5	72.1	86.9
	YOLOv8s—SF1	98.1	72.1	86.6
逆光	YOLOv5s	98.8	72.2	85.1
	YOLOv8s	97.8	74.1	85.4
	YOLOv9s	98.4	72.9	86.2
	YOLOv10s	98.3	73.4	85.9
	YOLOv8s—SF1	97.9	73.2	86.1
大视场情况	YOLOv5s	87.6	83.7	93.1
	YOLOv8s	86.8	84.9	91.0
	YOLOv9s	89.1	84.0	93.0
	YOLOv10s	88.8	85.2	93.4
	YOLOv8s—SF1	86.6	84.2	92.1

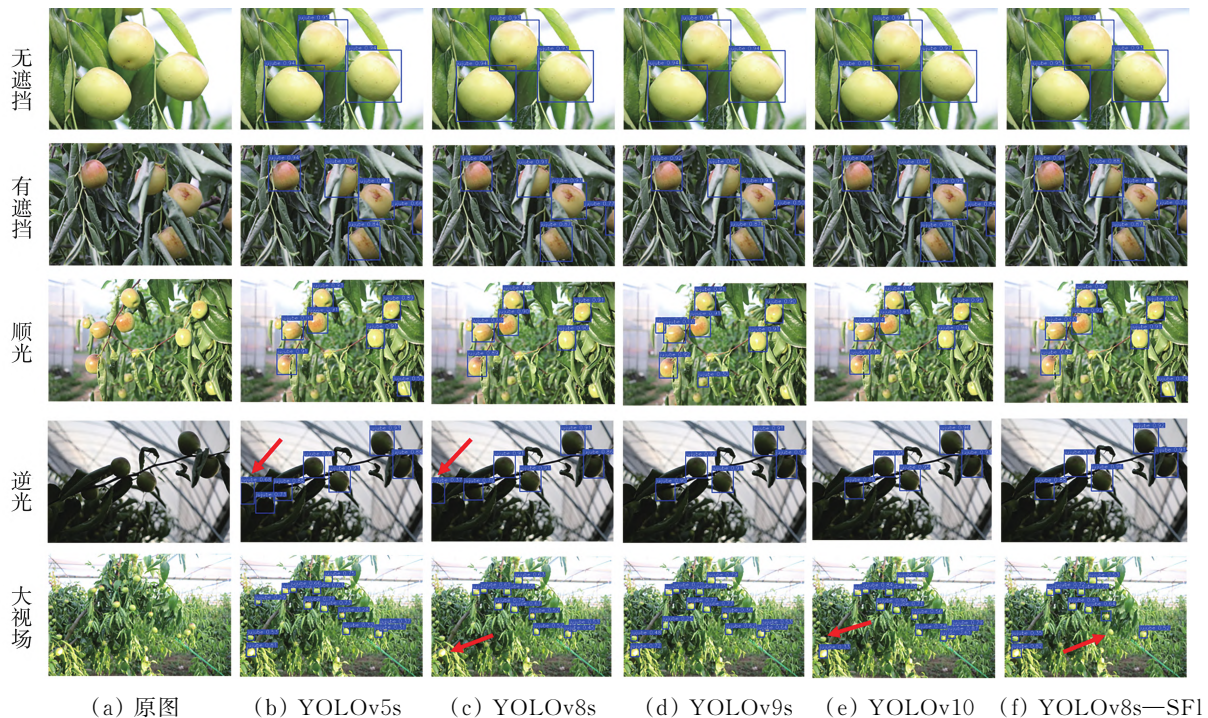


图 7 不同模型检测结果

Fig. 7 Detection results for different models

虽然在大视场情况下, YOLOv8s、YOLOv10s、改进的 YOLOv8s—SF1 存在一定的漏检情况(红色箭头),但整体上改进的 YOLOv8s—SF1 能够在复杂的自然环境下以较高的精度识别出冬枣。综上,说明提出的 YOLOv8s—SF1 模型能够在保证模型精度的前提下进行有效轻量化。

3 结论

1) 提出基于 YOLOv8s 的改进模型 YOLOv8s—SF1,改进模型在验证集上的精确率、召回率和平均精度均值分别为 94.1%、76.2% 和 88.9%,相比于原模型略微下降,但模型大小、参数量和 GFLOPs 分别减少 68.37%、68.80% 和 58.45%,且 FPS 提升 14.66 个百分点。

2) 改进的 YOLOv8s—SF1 模型大小相比于 YOLOv5s、YOLOv8s、YOLOv9s 和 YOLOv10s 分别减少 61.58%、68.37%、53.10% 和 56.96%,同时参数量减少 61.90%、68.80%、51.57% 和 51.91%,而且浮点计算量降低 50.42%、58.45%、55.80% 和 44.86%,说明改进的模型轻量化效果较好,更利于后续模型的边缘设备部署。

在后续的研究中需对冬枣进行多类别标注,进一步探究多类别标注对冬枣检测结果的影响,同时考虑实际应用情况,后期需结合深度相机在特定距离范围内获取冬枣图像,以确保与冬枣采摘机器人的数据一致。

参 考 文 献

- [1] 徐慧群,李阳,张建军. 冬枣果实特性及机械化采摘现状研究[J]. 中国农机化学报, 2023, 44(2): 53—59.
Xu Huiqun, Li Yang, Zhang Jianjun. Research status on fruit characteristics of winter jujube and mechanized harvesting [J]. Journal of Chinese Agricultural Mechanization, 2023, 44(2): 53—59.
- [2] 孔德刚,江达,江东博,等. 小型水果多功能采收装置设计与试验[J]. 中国农机化学报, 2023, 44(5): 19—25.
Kong Degang, Jiang Da, Jiang Dongbo, et al. Design and test of multifunctional harvesting device for small fruits [J]. Journal of Chinese Agricultural Mechanization, 2023, 44(5): 19—25.
- [3] 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. Lightweight maturity detection of cherry tomato based on improved YOLOv7 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 225—233.
- [4] 毕松,隗朋峻,刘仁学. 温室高架栽培草莓空间姿态识别与采摘点定位方法[J]. 农业机械学报, 2023, 54(9): 53—64.
Bi Song, Wei Pengjun, Liu Renxue. Spatial posture recognition and picking point location method for greenhouse raised-frame strawberry cultivation [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 53—64.
- [5] 宋怀波,王亚男,王云飞,等. 基于 YOLOv5s 的自然场景油茶果识别方法[J]. 农业机械学报, 2022, 53(7): 234—242.
Song Huaibo, Wang Ya'nan, Wang Yunfei, et al. Camellia oleifera fruit detection in natural scene based on YOLOv5s [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 234—242.
- [6] 郝建军,邴振凯,杨淑华,等. 采用改进 YOLOv3 算法检测青皮核桃[J]. 农业工程学报, 2022, 38(14): 183—190.
Hao Jianjun, Bing Zhenkai, Yang Shuhua, et al. Detection of green walnut by improved YOLOv3 [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(14): 183—190.
- [7] 王铁伟,赵瑶,孙宇馨,等. 基于数据平衡深度学习的不同成熟度冬枣识别[J]. 农业机械学报, 2020, 51(S1): 457—463.
Wang Tiewei, Zhao Yao, Sun Yuxin, et al. Recognition approach based on data-balanced Faster R—CNN for winter jujube with different levels of maturity [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(S1): 457—463.
- [8] 王昱潭,朱超伟,赵琛,等. 基于 Faster R—CNN 的灵武长枣图像检测方法[J]. 计算机工程与应用, 2021, 57(4): 216—224.
Wang Yutan, Zhu Chaowei, Zhao Chen, et al. Image detection method of Lingwu long jujube based on Faster R—CNN [J]. Computer Engineering and Applications, 2021, 57(4): 216—224.
- [9] Wu J, Wu C, Guo H, et al. Research on red jujubes recognition based on a convolutional neural network [J]. Applied Sciences, 2023, 13(11): 6381.
- [10] 王昱潭,薛君蕊. 改进 SSD 的灵武长枣图像轻量化目标检测方法[J]. 农业工程学报, 2021, 37(19): 173—182.
Wang Yutan, Xue Junrui. Lightweight object detection method for Lingwu long jujube images based on improved SSD [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(19): 173—182.
- [11] 刘天真,滕桂法,苑迎春,等. 基于改进 YOLOv3 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.

(下转第 146 页)

- Theory, 1967, 13(1): 21—27.
- [23] Bangare S L, Dubal A, Bangare P S, et al. Reviewing Otsu's method for image thresholding [J]. International Journal of Applied Engineering Research, 2015, 10(9): 21777—21783.
- [24] Khan S, Naseer M, Hayat M, et al. Transformers in vision: A survey [J]. ACM Computing Surveys (CSUR), 2022, 54(10s): 41.
- [25] Hestness J, Narang S, Ardalani N, et al. Deep learning scaling is predictable, empirically [J]. arXiv preprint arXiv: 1712.00409, 2017.
- [26] Feng X, Shen Y, Wang D. A survey on the development of image data augmentation [J]. Comput. Sci. Appl, 2021, 11: 370—382.
- [27] Tan M, Le Q. EfficientNet: Rethinking model scaling for convolutional neural networks [C]. International Conference on Machine Learning, 2019: 6105—6114.
- [28] Gao S, Cheng M, Zhao K, et al. Res2Net: A new multi-scale backbone architecture [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(2): 652—662.
- [29] Zhang X, Zhou X, Lin M, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6848—6856.
- [30] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770—778.
- [31] Lin T, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context [C]. Computer Vision-ECCV 2014: 13th European Conference, 2014: 740—755.
- [32] Wang X, Zhang R, Kong T, et al. SOLOv2: Dynamic and fast instance segmentation [J]. Advances in Neural Information Processing Systems, 2020, 33: 17721—17732.
- [33] Cai Z, Vasconcelos N. Cascade R—CNN: Delving into high quality object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6154—6162.
- [34] Zhang J, Yang X, Jiang R, et al. RSAM—Seg: A SAM-based approach with prior knowledge integration for remote sensing image semantic segmentation [J]. arXiv preprint arXiv: 2402.19004, 2024.
- [35] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30: 6000—6010.
- [36] Peng B, Bai J, Li W, et al. Survey on visual transformer for image classification [J]. Journal of Frontiers of Computer Science & Technology, 2024, 18(2): 320—344.
- (上接第 132 页)
- Liu Tianzhen, Teng Guifa, Yuan Yingchun, et al. Winter jujube fruit recognition method based on improved YOLOv3 under natural scene [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 17—25.
- [12] 王菁, 范晓飞, 赵智慧, 等. 基于 YOLO 算法的不同品种枣自然环境下成熟度识别 [J]. 中国农机化学报, 2022, 43(11): 165—171.
- Wang Jing, Fan Xiaofei, Zhao Zhihui, et al. Maturity identification of different jujube varieties under natural environment based on YOLO algorithm [J]. Journal of Chinese Agricultural Mechanization, 2022, 43(11): 165—171.
- [13] 薛君蕊, 王昱潭, 曲爱丽, 等. 基于改进 FCN—8s 的灵武长枣图像分割方法 [J]. 农业工程学报, 2021, 37(5): 191—197.
- Xue Junrui, Wang Yutan, Qu Aili, et al. Image segmentation method for Lingwu long jujubes based on improved FCN—8s [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(5): 191—197.
- [14] 李茂, 肖洋轶, 宗望远, 等. 基于改进 YOLOv8 模型的轻量化板栗果实识别方法 [J]. 农业工程学报, 2024, 40(1): 201—209.
- Li Mao, Xiao Yangyi, Zong Wangyuan, et al. Detecting chestnuts using improved lightweight YOLOv8 [J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(1): 201—209.
- [15] 王磊磊, 王斌, 李东晓, 等. 基于改进 YOLOv5 的菇房平菇目标检测与分类研究 [J]. 农业工程学报, 2023, 39(17): 163—171.
- Wang Leilei, Wang Bin, Li Dongxiao, et al. Object detection and classification of pleurotus ostreatus using improved YOLOv5 [J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(17): 163—171.
- [16] Qin D, Lechner C, Delakis M, et al. MobileNetV4: Universal models for the mobile ecosystem [J]. arXiv preprint arXiv: 2404.10518, 2024.
- [17] Liu X, Peng H, Zheng N, et al. EfficientViT: Memory efficient vision transformer with cascaded group attention [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 14420—14430.
- [18] Han K, Wang Y, Tian Q, et al. GhostNet: More features from cheap operations [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580—1589.