

郑文轩,杨 瑛. 基于改进轻量化 YOLO v7 的成熟期香梨目标检测方法[J]. 江苏农业科学,2024,52(20):121-128.

doi:10.15889/j.issn.1002-1302.2024.20.015

基于改进轻量化 YOLO v7 的成熟期香梨目标检测方法

郑文轩,杨 瑛

(江苏第二师范学院物理与信息工程学院,江苏南京 211200)

摘要:为在自然环境下准确地识别和检测香梨果实,以 YOLO v7 为基础网络模型,针对果园中香梨果实、果叶、枝干之间相互遮挡问题,提出一种改进的轻量化 YOLO v7 梨果目标检测方法。该方法将 MobileNet v3 引入 YOLO v7 模型中作为其骨干特征提取网络,从而减少网络的参数量,使其更容易部署在移动端和生产实际,在特征融合层引入协同注意力机制 CA(coordinate attention)模块,以提高网络的特征表达能力,将原 YOLO v7 中的损失函数 CIoU 替换为 SIoU,从而提高模型的检测速度和性能。最后利用 Grad-CAM 方法产生目标检测热力图,进行特征可视化。结果表明,改进的轻量化 YOLO v7 模型的平均精度均值(mAP)、精确率、召回率指标分别为 96.33%、94.36%、89.28%,检测速度为 87.71(帧/s),模型内存占用量与原 YOLO v7 相比减少 21.45 MB;其检测平均精度均值(mAP)与 Faster R-CNN、YOLO v3、MobileNet v3-YOLO v4、YOLO v5s、YOLO v7 模型相比分别提高 28.37%、9.66%、13.14%、4.58%、3.20 百分点。研究表明,改进的轻量化 YOLO v7 模型对成熟期香梨具有很好的目标检测效果和鲁棒性,可为香梨自动化采摘提供有效的技术支持。

关键词:目标检测;香梨;YOLO v7;轻量化;注意力机制

中图分类号:S126;TP391.41 **文献标志码:**A **文章编号:**1002-1302(2024)20-0121-08

香梨作为世界五大水果之一,由于其经济价值、营养保健价值高,可鲜食、加工增值等多种用途,

深受广大消费者的喜爱。2022 年我国梨产量超过 1 800 万 t,是果农重要的经济来源之一。目前香梨的采摘方式主要是依靠农民手工采摘,费时费力,效率低下,很难满足作业周期短、采摘任务重、采摘需求急的现实要求^[1]。在现代农业生产中,农业采摘越来越偏向机械化和智能化,越来越多的采摘机器人投入生产使用,并逐步代替人力劳动。但

收稿日期:2023-11-23

基金项目:新疆生产建设兵团科技创新项目(编号:2021CB021)。

作者简介:郑文轩(1980—),男,河南南阳人,博士,教授,从事图像分析、计算机视觉研究。E-mail:wenzhuangdx@163.com。

通信作者:杨 瑛,博士,教授,从事农业信息技术研究。E-mail:yangyingtmdx@163.com。

分割方法[J]. 农业工程学报,2018,34(12):193-201.

[20]和兴华. 基于卷积神经网络的玉米冠层图像分割与生育期鉴定方法[D]. 南昌:江西农业大学,2018.

[21]万园洁,卿琳波,何小海,等. 基于改进全卷积网络的小麦图像分割[J]. 计算机系统应用,2018,27(3):221-227.

[22]Xiong X,Duan L F,Liu L B,et al. Panicle-SEG:a robust image segmentation method for rice panicles in the field based on deep learning and superpixel optimization[J]. Plant Methods,2017,13:104.

[23]Ronneberger O,Fischer P,Brox T. U-net:convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham:Springer,2015:234-241.

[24]Chen L C,Zhu Y K,Papandreou G,et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//European Conference on Computer Vision. Cham:Springer,2018:833-851.

[25]Zhao H S,Shi J P,Qi X J,et al. Pyramid scene parsing network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu:IEEE,2017:6230-6239.

[26]Zhou Z W,Siddiquee M M R,Tajbakhsh N,et al. UNet:re designing skip connections to exploit multiscale features in image segmentation [J]. IEEE Transactions on Medical Imaging,2020,39(6):1856-1867.

[27]Huang H M,Lin L F,Tong R F,et al. UNet3:a full-scale connected UNet for medical image segmentation[C]//2020 IEEE International Conference on Acoustics,Speech and Signal Processing. Barcelona:IEEE,2020:1055-1059.

[28]Oktay O,Schlemper J,Le Folgoc L,et al. Attention U-net: learning where to look for the pancreas[EB/OL]. (2018-04-11)[2023-11-01]. <http://arxiv.org/abs/1804.03999v3>.

[29]Chen J N,Lu Y Y,Yu Q H,et al. TransUNet:transformers make strong encoders for medical image segmentation[EB/OL]. (2021-02-08)[2023-11-01]. <http://arxiv.org/abs/2102.04306v1>.

在自然环境下,天气、光照以及水果果实大小,成簇状生长,果实颜色与背景相近,果实和枝干之间相互遮挡等问题,大大增加了水果果实目标检测的难度。

随着计算机算力的不断提升,深度学习技术在目标检测领域有了巨大的优势^[2]。与传统的目标检测算法相比,深度卷积神经网络可以自动从训练数据中学习多层次的特征,模型具有较强的泛化能力和更强的特征提取能力^[3-4]。傅隆生等提出一种对多簇猕猴桃具有 94.78% 的识别准确率的果实识别算法^[5]。崔潇等在 2019 年提出的桃子检测模型,是以 MTCNN(多任务卷积神经网络算法)为主干的一种网络,对成熟桃子的检测准确率达到 89.3%^[6]。He 等提出 SPP-Net 网络,从之前的单一尺度输入变为了效率更高的多尺度输入^[7]。Ren 等提出 Faster-RCNN 网络去弥补目标检测速度慢的问题^[8]。为了进一步改善目标检测速度上的问题,Redmon 等提出了速度更快的网络模型 YOLO,但方法存在检测精度低的缺陷^[9]。后续研究者陆续提出 YOLO v2、YOLO v3、YOLO v4 和 YOLO v5 等来改进检测精度低的问题,但这些网络模型在小目标的检测上表现效果欠佳。随着目标检测技术的不断发展,Tian 等提出了一种改进的 YOLO v3 模型,用于在不同光照条件、复杂背景以及遮挡等情况下检测不同成熟度的苹果^[10]。王立舒等提出一种改进 YOLO v4-Tiny 的蓝莓成熟度识别方法,该方法在遮挡与光照不均等复杂场景中平均精度很高^[11]。周桂红等在 YOLO v4 主干特征提取网络的 ResBlock 模块中加入 SC-SE 注意力机制,并将 PA-Net 模块中的部分卷积替换为深度可分离卷积,改进后的 YOLO v4 网络模型检测成熟苹果的平均精度达到 97.27%^[12]。Lyu 等提出了一种轻量级的目标检测 YOLO v5-CS 模型,在 YOLO v5 网络中加入了注意力模块,自然环境中对绿色柑橘识别准确率 98.23%^[13]。YOLO v7 模型是一个具有检测精度高、速度快、支持高分辨率图像的目标检测模型,已被广泛应用到智慧农业的各个领域。虽然,目前对水果检测方面的研究已经取得较好的成效,但是针对水果目标检测的检测精度和速度还有待提高,在检测距离较远及存在遮挡的目标检测时会出现漏检和误检的问题。

针对以上问题,本研究提出一种改进的轻量化 YOLO v7 香梨目标检测模型。首先,在试验数据采集时针对现实生产中存在的果实采摘距离远近、光

照、有无枝干或树叶遮挡等问题进行有针对性的图像采集。然后,将 YOLO v7 的骨干特征提取网络替换为轻量级的 MobileNet v3 骨干网络,大幅降低模型的参数量和计算量。在特征融合层引入 CA 注意力模块来加强网络的特征表达能力,提高相邻果实和被枝叶遮挡果实的检测精度,通过试验对比将原 YOLO v7 中的损失函数 CIoU 替换为 SIoU,从而提高模型的检测速度和精度等。

1 材料与方法

1.1 图像数据采集

本试验开展于 2023 年 8 月,试验所用的香梨数据集拍摄于新疆塔里木大学园艺试验站新梨 7 号试验区,研究对象为成熟期新梨 7 号香梨。采集图像信息见表 1,不同场景香梨图像见图 1。考虑到智能机器人在对梨果采摘时会遇到采摘距离远近、光照强弱不均匀等问题,试验数据采集设计了远距(距离为 50~100 cm)和近距(距离为 10~30 cm)2 种图片拍摄模式,拍摄时间选择在 07:00—08:00、13:00—14:00、18:00—19:00 等 3 个时间段,拍摄时期为 8 月 15—30 日,此时段为香梨成熟期但还未采摘。试验共采集香梨图片 2 586 张,其中远距图片为 1 167 张,近距图片为 1 419 张。

表 1 采集图像信息

采集距离	采集条件	图片数量 (张)
近距离采集	顺光	457
	逆光	382
	枝叶遮挡	361
	光照不均	219
远距离采集	顺光	367
	逆光	312
	枝叶遮挡	339
	光照不均	149

1.2 香梨图像数据集制作

由于新梨 7 号香梨数据集的数据样本量不足以让模型在训练时达到很好的收敛状态,为改善网络训练效果和提高模型的泛化能力,采用数据增强方法增加样本数量,防止网络因训练样本不足导致过拟合问题。本研究采用水平翻转、垂直翻转、非等比例缩放、随机平移、随机裁剪和 mosaic 等数据增强方式随机组合,对数据集进行增广。利用 Labelimg 标注软件对香梨数据集进行人工标注,标



图1 不同场景香梨图像

注框选择果实最小外接矩形,标注后产生的标注信息文件为 xml 类型文件,储存了果实图像文件名、标注区域矩形框 4 个角的位置信息以及标注种类等信息。图像增广后的图片共 4 316 张,以 7:2:1 的比例划分训练集(3 021 张)、测试集(863 张)、验证集(432 张)进行模型的训练和测试。

1.3 试验环境配置

本研究使用 Win 10 操作系统,显卡型号为 Nvidia GeForce RTX - 3090,处理器型号为 Intel® Core™ i7 - 12700KF 3.60 GHz,深度学习框架为 PyTorch 2.0,编程平台为 PyCharm,编程语言为 Python 3.8,所有对比算法均在相同环境下运行。为提高网络训练的效率,本试验采用冻结训练和解冻训练 2 种方法进行。冻结骨干网络训练时初始学习率设为 0.01,Batchsize 设为 16,动量设为 0.93;解冻之后的网络训练学习率设为 0.01,Batchsize 设为 8,动量设为 0.93,迭代次数为 300 次。

1.4 模型评价指标

本研究采用 5 个指标评价目标检测模型的性能。在模型检测精度方面,选用精确率(Precision,简称 P)、召回率(Recall,简称 R)和平均精度均值(mean Average Precision,简称 mAP)作为评价指标;在模型检测性能方面,选用检测速度(detection speed)和模型内存占用量作为评价指标,这些评价指标的具体计算方法可见参考文献[14-16]。

2 研究方法

2.1 YOLO v7 网络结构介绍

YOLO v7 网络结构主要由主干网络(Backbone)、特征融合网络(FPN)和检测头(YOLO Head)等 3 个部分组成。在 YOLO v7 特征提取网络中使用了多分支堆叠模块(Multi_Concat_Block),如图 2 所示,它是将多个卷积标准化激活函数进行堆叠后进行特征融合,从而提高准确率。同时使用创新的过渡模块 Transition_Block 来进行采样,YOLO v7 提出了重参数化卷积(RepConv)和动态标签分配等策略,使其在检测精度和效率上取得了较好的效果。在原 YOLO v7 中使用的 CIoU 损失函数,只考虑了检测框尺度的损失,没有考虑预测框与真实框之间的不匹配,因此本研究使用 SIoU 损失函数代替 CIoU 损失函数,将方向尺度的损失考虑到模型训练中。

2.2 YOLO v7 模型的主干网络改进

为了使改进的 YOLO v7 模型更符合生产实际,试验使用 MobileNet v3 骨干网络作为 YOLO v7 模型的骨干网络实现特征提取。MobileNet v3 网络与传统卷积神经网络相比在准确率小幅降低的前提下大大减少了模型参数量与运算量^[17],实现模型的轻量化,更有利于部署在移动端,其网络参数设计如表 2 所示。MobileNet v3 综合了 MobileNet v1 和

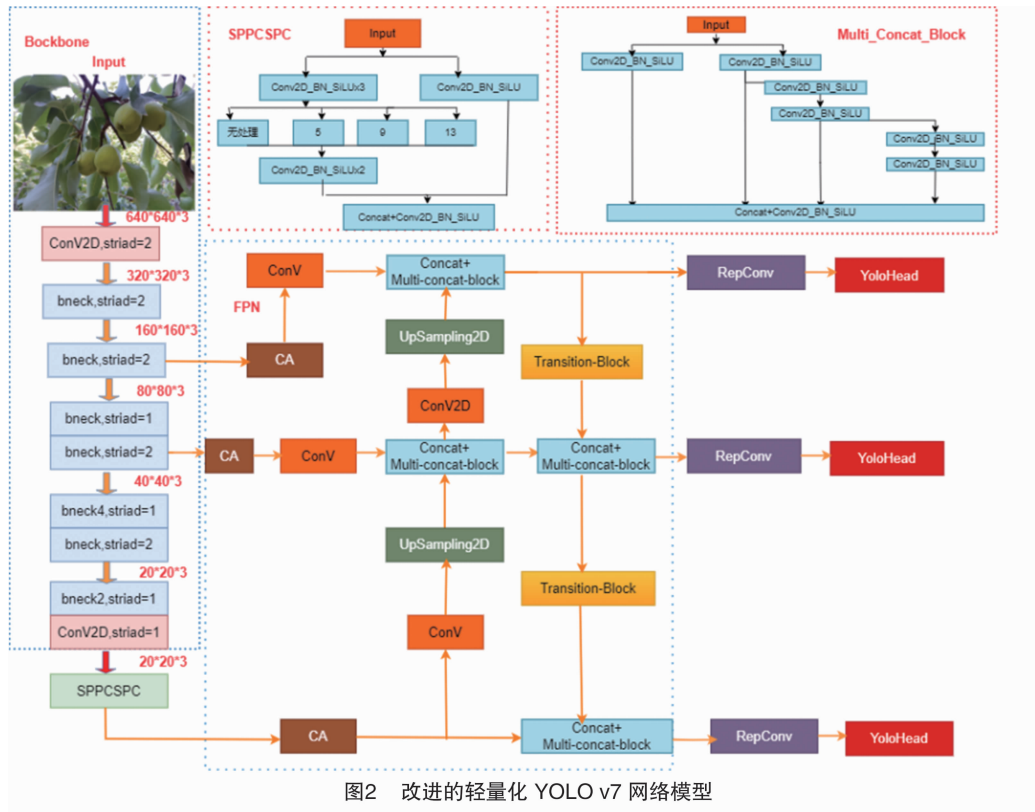


图2 改进的轻量化 YOLO v7 网络模型

表 2 MobileNet v3 (small) 网络参数

输入	操作	升维通道	输出通道	SE	NL	步长
$224^2 \times 3$	Conv2d, 3×3	—	16	—	HS	2
$112^2 \times 16$	bneck, 3×3	16	16	✓	RE	2
$56^2 \times 16$	bneck, 3×3	72	24	—	RE	2
$28^2 \times 24$	bneck, 5×5	88	24	—	RE	1
$28^2 \times 24$	bneck, 5×5	96	40	✓	HS	2
$14^2 \times 40$	bneck, 5×5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5×5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5×5	120	48	✓	HS	1
$14^2 \times 48$	bneck, 5×5	144	48	✓	HS	1
$14^2 \times 48$	bneck, 5×5	244	96	✓	HS	2
$7^2 \times 96$	bneck, 5×5	576	96	✓	HS	1
$7^2 \times 96$	bneck, 5×5	576	96	✓	HS	1
$7^2 \times 96$	Conv2d, 1×1	—	576	✓	HS	1
$7^2 \times 576$	pool, 7×7	—	—	—	—	1
$7^2 \times 576$	Conv2d, 1×1	—	1 024	—	HS	1
$1^2 \times 1 024$	Conv2d, 1×1	—	k	—	—	1

注：“✓”表示本层引入了 SE 模块，“—”表示本层未引入 SE 模块。NL 代表激活函数的类型，包括 HS (H-Swish)、RE (ReLU)。

MobileNet v2 的特点，其在主干网络中加入了注意力机制 SE (squeezeand excitation networks) 模块^[18]，使用了深度可分离卷积 (depthwise separable convolutions) 和逆残差结构 (the inverted residual

with linear bottleneck)。同时，使用 H-Swish 激活函数代替 Swish 激活函数，减少运算量，提高了模型性能。

2.3 引入 SE 通道注意力机制和 CA 协同注意力机制

注意力机制是一种即插即用的模块，引入注意力机制可以更有效地提取特征，进而提升模型的检测性能。在 MobileNet v3 主干网络中加入 SE 注意力机制模块，SE 注意力机制模块主要考虑的是通道注意力 (channel attention)，如图 3 所示。在 YOLO v7 特征融合层引入协同注意力机制 CA (coordinate attention,) 模块^[19]，该注意力模块不仅可以把通道注意力和空间注意力进行有效的结合，更值得注意的是它同时在通道注意力中融入了位置信息，避免全局 pooling-2D 操作造成的位置信息丢失，将注意力分别放在宽度和高度 2 个维度上，有效利用输入特征图的空间坐标信息，具体如图 4 所示。

2.4 SIOU 损失函数分析

YOLO v7 模型算法中，预测框回归损失采用 CIoU，但 CIoU 并没有考虑到预测框与真实框方向间的不匹配。这种方向的不匹配，可能导致收敛速度减慢且收敛效率低的问题。针对 CIoU 存在的问题，本研究经过试验对比，采用 SIOU 损失函数替换

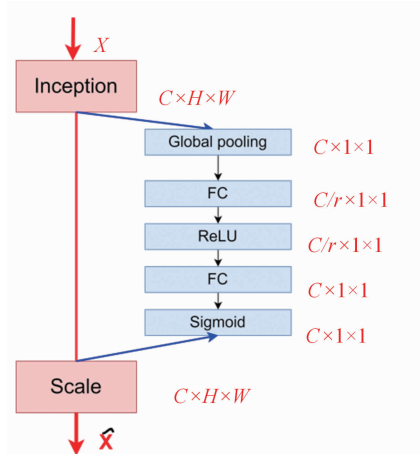


图3 SE 注意力模块

YOLO v7 使用的 CIoU 损失函数^[20]。SIoU 损失函数主要优点在于进一步考虑了真实框和预测框之间的向量角度,重新定义相关损失函数,SIoU 主要包含 4 个部分:角度损失 (angle cost)、距离损失 (distance cost)、形状损失 (shape cost)、IoU 损失 (IoU cost),图 5 为 SIoU 参数示意图。

角度损失计算公式:

$$\begin{aligned} \Lambda &= 1 - 2 \cdot \sin^2 \left[\arcsin \left(\frac{c_h}{\sigma} \right) - \frac{\pi}{4} \right] \\ &= \cos \left\{ 2 \cdot \left[\arcsin \left(\frac{c_h}{\sigma} \right) - \frac{\pi}{4} \right] \right\} \end{aligned} \quad (1)$$

其中:

$$\frac{c_h}{\sigma} = \sin(\alpha) \quad (2)$$

根据角度损失 Λ 公式,定义距离损失 Δ 计算公式如下:

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho^t}) = 2 - e^{-\gamma \rho^x} - e^{-\gamma \rho^y}; \quad (3)$$

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w} \right)^2, \rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h} \right)^2, \gamma = 2 - \Lambda \quad (4)$$

形状损失 Ω 定义如下:

$$\Omega = \sum_{t=w,h} (1 - e^{-w_t})^\theta = (1 - e^{-w_w})^\theta + (1 - e^{-w_h})^\theta \quad (5)$$

其中:

$$w_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})}, w_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (6)$$

式中: w 、 h 、 w^{gt} 、 h^{gt} 分别表示预测框和真实框的宽和高。 θ 控制对形状损失的关注程度,为避免过于关注形状损失而降低对预测框的移动, θ 参数范围为 $[2, 6]$ 。

SIoU 损失的定义为:

$$Loss_{SIoU} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (7)$$

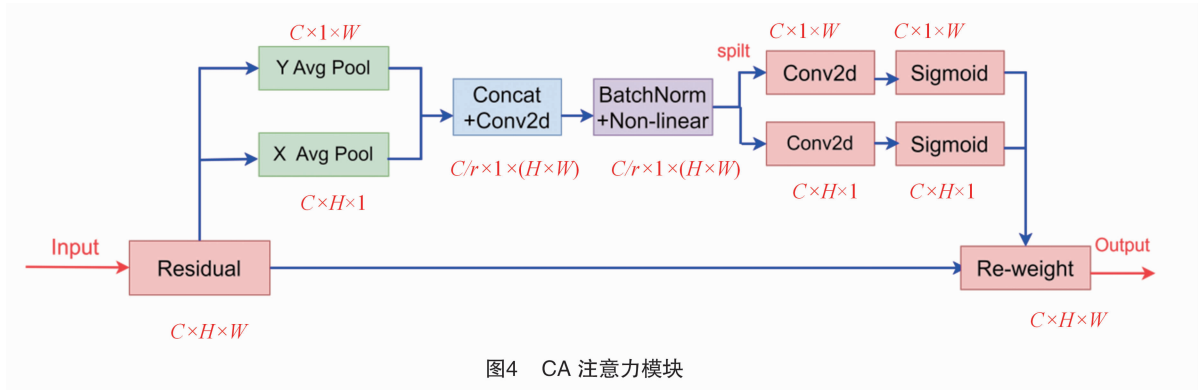


图4 CA 注意力模块

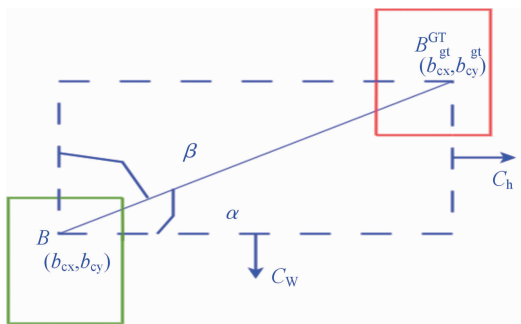


图5 SIoU 参数示意

3 试验结果与分析

3.1 改进的 YOLO v7 模型消融试验结果

为验证改进的 YOLO v7 模型的性能,本研究选用 4 组消融试验验证网络模型的性能,分别为 YOLO v7、在 YOLO v7 特征融合层加入 CA 注意力模块、替换骨干网络为 MobileNet v3 的 YOLO v7 模型和改进的轻量化 YOLO v7 模型。选用平均精度均值 (mAP) 和检测速度等 2 个评价指标,对验证集

中的香梨图像进行评价。由表 3 可知,改进后的 YOLO v7 模型的平均精度均值 (mAP) 和检测速度均高于其他 3 组模型。

表 3 改进的 YOLO v7 模型消融试验结果

网络模型	平均精度均值 (%)	检测速度 (帧/s)
YOLO v7	93.13	75.33
YOLO v7 + CA	93.69	74.75
MobileNet v3 - YOLO v7	90.77	86.93
改进 YOLO v7	95.32	87.71

3.2 香梨果实图像检测结果

图 6 为不同距离、是否有遮挡、不同光照条件下 YOLO v7 和改进的 YOLO v7 模型对香梨目标检测结果。

由图 6 可知,无论是在近距离还是远距离,无遮挡顺光和逆光的情况下(图 6 - a、图 6 - b、图

6 - d),YOLO v7 模型和改进的 YOLO v7 模型均可准确检测梨果,没有漏检和误检的情况,但是在有枝叶遮挡的情况下,YOLO v7 的检测效果(图 6 - a、图 6 - f)均出现了漏检或误检的情况。而改进的 YOLO v7 模型仅在光照不均匀且有相邻梨果遮挡(图 6 - c)和远距离有遮挡(图 6 - e)的情况下出现个别香梨漏检情况,整体检测更加准确。

由表 4 可知,改进的 YOLO v7 模型的平均精度均值(mAP)、精确率和召回率在近景、远景 2 种场景下均高于 YOLO v7 模型,在测试近景数据集时,改进 YOLO v7 模型的平均精度均值(mAP)、精确率和召回率分别比 YOLO v7 高 1.35、1.98、1.63 百分点;在测试远景数据集时,改进的 YOLO v7 模型的平均精度均值(mAP)、精确率和召回率分别比 YOLO v7 模型高 1.49、0.12、1.68 百分点。

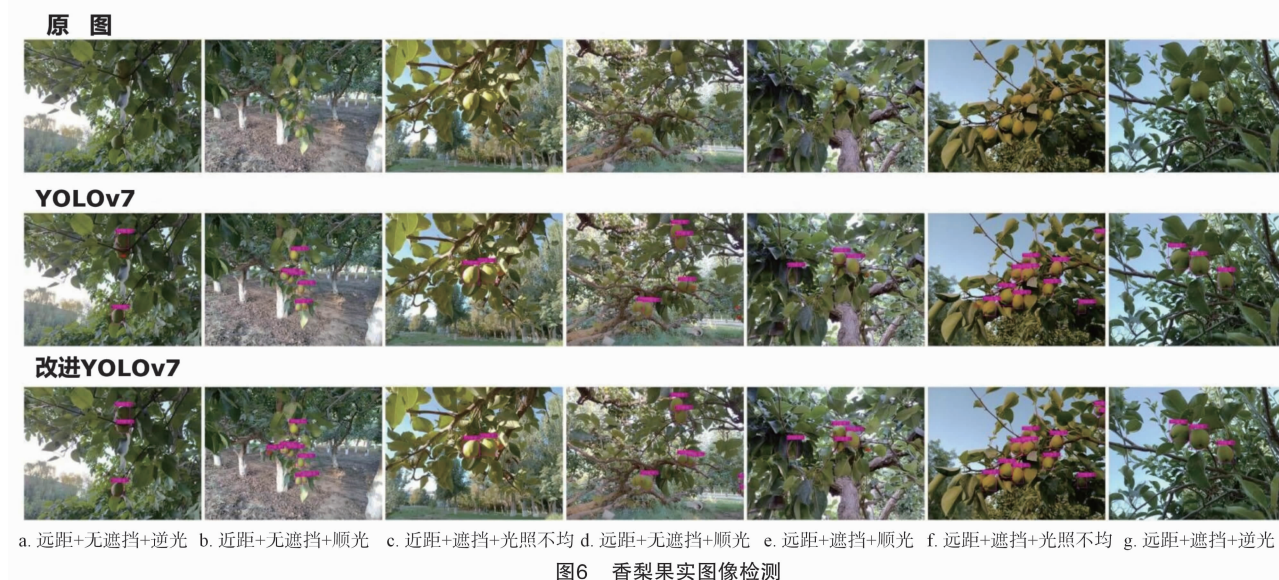


图6 香梨果实图像检测

表 4 YOLO v7 和改进 YOLO v7 模型在不同场景下的性能比较

网络模型	场景	平均精度均值 (%)	精确率 (%)	召回率 (%)
YOLO v7	远景	91.23	91.35	89.46
	近景	95.34	93.31	90.73
改进 YOLO v7	远景	92.72	91.47	91.14
	近景	96.69	95.29	92.36

3.3 不同目标检测网络模型对比结果分析

为定量比较不同模型的性能,将改进的 YOLO v7 模型与 Faster R - CNN、YOLO v3、MobileNet v3 - YOLO v4、YOLO v5s 以及原 YOLO v7 模型在验证集上进行性能比较。由表 5 可知,改进的 YOLO v7 模

型平均精度均值 (mAP)、精确率、召回率分别为 96.33%、94.36%、89.28%。其均值平均精度 (mAP) 与 Faster R - CNN、YOLO v3、MobileNet v3 - YOLO v4、YOLO v5s 以及原 YOLO v7 模型相比分别提升 28.37、9.66、13.14、4.58、3.20 百分点。检测速度为 87.71 帧/s,模型内存占用量与原 YOLO v7 相比减少了 21.45 MB。

3.4 不同损失函数应用于 YOLO v7 的性能比较

为分析不同损失函数的性能表现,本研究将原 YOLO v7 中的损失函数 CIOU 分别替换为 AlphaIoU、DIOU、SIOU^[21-24],4 种损失函数应用于 YOLO v7 中的性能表现(表 6)。使用 4 种损失函数训练模型的损失曲线,由损失曲线图(图 7)可以看出,DIOU 损

表 5 不同检测模型在验证集中的性能比较

网络模型	平均精度均值 (%)	精确率 (%)	召回率 (%)	模型大小 (MB)	检测速度 (帧/s)
Faster R-CNN	67.96	66.75	62.65	96.26	10.96
YOLO v3	86.67	88.24	77.17	61.81	57.17
MobileNet v3-YOLO v4	83.19	85.51	71.46	11.47	68.93
YOLO v5s	91.75	93.11	83.72	46.79	72.62
YOLO v7	93.13	94.83	88.63	37.27	75.32
改进 YOLO v7	96.33	94.36	89.28	15.82	87.71

表 6 不同损失函数的性能比较

损失函数	精确率 (%)	召回率 (%)	平均精度均值 (%)	训练时间 (h)
AlphaIoU	92.51	88.91	92.53	6.25
CIoU	92.72	91.56	93.81	8.31
DIoU	92.18	87.43	92.22	6.54
SIoU	93.61	90.87	94.13	5.53

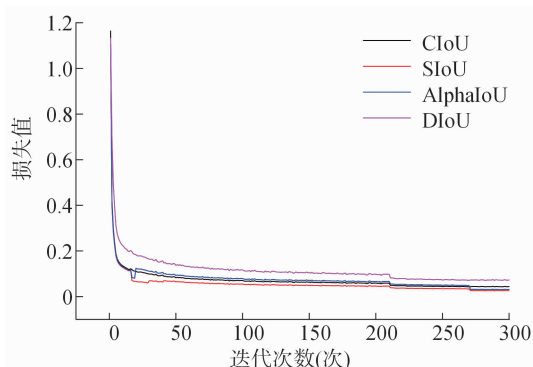


图7 不同损失函数损失值随迭代轮次变化的曲线

失函数在第 17~21 个轮次时略有“波动”,但使用 AlphaIoU、DIoU、CIoU 和 SIoU 这 4 种损失函数训练模

型都可以快速收敛并维持在较低的损失值。

对比 4 种损失函数性能,SIoU 与 AlphaIoU 相比,其平均精度均值(mAP)、精确率和召回率分别提高 1.60、1.10、1.96 百分点,训练时长缩短了 0.72 h;使用 SIoU 与 DIoU 相比,其平均精度均值(mAP)、精确率和召回率分别提高 1.91、1.43、3.44 百分点,训练时长缩短了 1.01 h;使用 SIoU 与 CIoU 相比,其平均精度均值(mAP)、精确率分别提高了 0.32、0.89 百分点,召回率降低了 0.69 百分点,训练时长缩短 2.78 h。根据上述分析可知,使用 SIoU 损失函数进行 YOLO v7 模型训练,检测精度更高。

3.5 检测热力图分析

深度学习热力图可视化是一种直观展示模型学习过程和结果的方法。通过使用热力图,可以更好地理解模型的内部工作原理和模型对输入数据的敏感程度,理解模型的关注区域。本研究利用 Grad-CAM 产生目标检测热力图^[25]。由图 8 可知,通过香梨图像检测热力图,可以看出改进的 YOLO v7 模型的目标检测热力图更贴近真实梨果区域。通过图 8-b、图 8-d、图 8-e、图 8-f 圆形圈出的区域可以看出,原 YOLO v7 目标检测热力值相对改进的 YOLO v7 模型的目标检测热力值更低,表现出漏检现象。通过目标检测热力图可以看出,改进的 YOLO v7 模型可以从弱语义的目标中提取更强的目标特征,有着较好的鲁棒性和泛化能力^[26]。

4 结论

基于 YOLO v7 构建了改进的轻量化 YOLO v7 香梨目标检测模型,实现了对自然环境下香梨的快

原图



YOLO v7



改进YOLO v7



a. 远距+无遮挡+逆光 b. 近距+无遮挡+顺光 c. 近距+遮挡+光照不均 d. 远距+无遮挡+顺光 e. 远距+遮挡+顺光 f. 远距+遮挡+光照不均 g. 远距+遮挡+逆光

图8 香梨果实图像检测热力

速准确检测。该模型采用 MobileNet v3 骨干网络替换 YOLO v7 原模型中的骨干网络,并在特征融合层引入 CA 注意力机制,使用 SIoU 损失函数替换 YOLO v7 中使用的 CIoU 损失函数,从而提高梨果的目标检测精度。改进的轻量化 YOLO v7 模型的平均精度均值(mAP)、精确率、召回率指标分别为 96.33%、94.36%、89.28%。检测速度为 87.71 帧/s,模型内存占用量与原 YOLO v7 相比减少了 21.45 MB。

为了验证改进的轻量化 YOLO v7 模型的性能,本研究选用 YOLO v7、YOLO v7 + CA、MobileNet v3 - YOLO v7 和改进的 YOLO v7 等 4 组网络模型进行消融试验。试验结果表明,改进后的 YOLO v7 模型在检测精度和检测速度方面均优于其他 3 组模型。

通过与 Faster R - CNN、YOLO v3、MobileNet v3 - YOLO v4、YOLO v5s 以及原 YOLO v7 模型进行对比发现,改进后的 YOLO v7 模型在香梨验证集上取得了最优的目标检测效果,其平均精度均值(mAP)分别提高 28.37、9.66、13.14、4.58、3.20 百分点。

参考文献:

- [1] 邵园园,王永贤,玄冠涛,等. 基于高光谱成像的肥城桃品质可视化分析与成熟度检测[J]. 农业机械学报,2020,51(8):344 - 350.
- [2] 宋怀波,尚钰莹,何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报,2023,54(1):1 - 19.
- [3] 林景栋,吴欣怡,柴毅,等. 卷积神经网络结构优化综述[J]. 自动化学报,2020,46(1):24 - 37.
- [4] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770 - 778.
- [5] 傅隆生,冯亚利, Tola E, et al. 基于卷积神经网络的田间多簇猕猴桃图像识别方法[J]. 农业工程学报,2018,34(2):205 - 211.
- [6] 崔潇,马原东,倪照凤,等. 基于深度学习的桃子检测研究[J]. 贵州科学,2019,37(2):82 - 86.
- [7] He K M, Zhang X Y, Ren S Q, 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.
- [8] Ren S Q, He K M, Girshick R, et al. Faster R - CNN: towards real - time object detection with region proposal networks[EB/OL]. (2015 - 06 - 04)[2023 - 11 - 20]. <https://arxiv.org/abs/1506.01497v3>.
- [9] Redmon J, Farhadi A. YOLO v3: an incremental improvement[EB/OL]. (2018 - 04 - 08)[2023 - 11 - 20]. <https://arxiv.org/abs/1804.02767v1>.
- [10] Tian Y N, Yang G D, Wang Z, 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.
- [11] 王立舒,秦铭霞,雷洁雅,等. 基于改进 YOLO v4 - Tiny 的蓝莓成熟度识别方法[J]. 农业工程学报,2021,37(18):170 - 178.
- [12] 周桂红,马帅,梁芳芳. 基于改进 YOLO v4 模型的全景图像苹果识别[J]. 农业工程学报,2022,38(21):159 - 168.
- [13] Lyu S L, Li R Y, Zhao Y W, et al. Green Citrus detection and counting in orchards based on YOLO v5 - CS and AI edge system[J]. Sensors, 2022, 22(2): 576.
- [14] Singh B, Najibi M, Davis S. Efficient multi - scale training[EB/OL]. [2023 - 11 - 20]. <https://arxiv.org/abs/1805.09300>.
- [15] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real - time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 779 - 788.
- [16] Redmon J, Farhadi A. YOLO 9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI: IEEE, 2017: 7263 - 7271.
- [17] Jiang B R, Luo R X, Mao J Y, et al. Acquisition of localization confidence for accurate object detection[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018: 816 - 832.
- [18] 彭心睿,潘晴,田妮莉. WCF - MobileNet v3: 轻量型新冠肺炎 CXR 图像识别网络[J]. 计算机工程与应用, 2023, 59(14): 224 - 231.
- [19] 刘雪,沈长盈,吕学泽,等. 基于改进 MobileNet v3 - Large 的鸡蛋新鲜度识别模型[J]. 农业工程学报,2022,38(17):196 - 204.
- [20] 林森,刘美怡,陶志勇. 采用注意力机制与改进 YOLO v5 的水下珍品检测[J]. 农业工程学报,2021,37(18):307 - 314.
- [21] He J, Sarah E, Ma X, et al. Alpha - IoU: a family of power intersection over union losses for boundingbox regression[C]//Proceedings of the Conference and Workshop on Neural Information Processing Systems. Long Beach, 2021: 1 - 10.
- [22] Zheng Z H, Wang P, Liu W, et al. Distance - IoU loss: faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993 - 13000.
- [23] Gevorgyan Z. SIoU loss: more powerful learning for bounding box regression[EB/OL]. (2022 - 05 - 25)[2023 - 11 - 20]. <https://arxiv.org/abs/2205.12740>.
- [24] 周绍发,肖小玲,刘忠意,等. 改进的基于 YOLO v5s 苹果树叶病害检测[J]. 江苏农业科学,2023,51(13):212 - 220.
- [25] Danell J, Khan F. Adaptive color attributes for real - time visual tracking[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014: 1090 - 1097.
- [26] 龙燕,杨智优,何梦菲. 基于改进 YOLO v7 的疏果期苹果目标检测方法[J]. 农业工程学报,2023,39(14):191 - 199.