

基于改进 YOLOv8 的朝天椒轻量级检测方法

李瑞¹, 张富贵^{1,2†}, 吴雪梅¹, 袁奎¹, 郑乐¹, 李鑫¹, 罗歆妍¹

(1. 贵州大学机械工程学院, 贵州 贵阳 550025; 2. 现代制造技术教育部重点实验室, 贵州 贵阳 550025)

摘要:为提高朝天椒果实的检测精度, 实现将模型方便快捷部署到移动端, 提出了一种基于改进 YOLOv8 的轻量化朝天椒果实检测方法。在骨干网络中使用 GhostConv 和全新设计的 C2fGhost 替换传统卷积与 C2f 模块, 减少了网络参数和计算量, 提升了模型的检测性能; 为提高模型在复杂背景下朝天椒果实的检测效果, 添加全局注意力机制模块于颈部网络, 提高模型的特征融合能力; 使用 SIoU 边界损失函数替代原损失函数, 提升了网络边界框回归性能和对小目标果实的检测效果。实验结果表明, 对自然环境下建立的朝天椒果实数据集, 改进后模型的参数量、计算量和权重大小分别为 2.46×10^6 、10.5 GFLOPS 和 6.02 MB, 仅为原始网络模型的 81%、37% 和 96%, 且检测速度和平均精度分别达到了 96 帧/s 和 88.5%。不仅提高了检测精度, 还大幅减少了模型的整体体积, 为后续朝天椒果实采摘机器人移动端的部署和应用提供参考和依据。

关键词:图像处理; 朝天椒; YOLOv8; 轻量化网络; 小目标; 自然环境; 注意力机制; 损失函数

中图分类号: TP391

文献标识码: A

Lightweight Detection of Clustered Pepper Based on Improved YOLOv8

LI Rui¹, ZHANG Fugui^{1,2†}, WU Xuemei¹, YUAN Kui¹, ZHENG Le¹, LI Xin¹, LUO Xinyan¹

(1. School of Mechanical Engineering, Guizhou University, Guiyang Guizhou 550025, China;

2. Key Laboratory of Advanced Manufacturing Technology of the Ministry of Education, Guiyang Guizhou 550025, China)

Abstract: In order to improve the detection accuracy of clustered pepper fruits and realize the convenient and rapid deployment of the model to mobile terminal, a lightweight clustered pepper fruits detection method based on improved YOLOv8 is proposed. In the backbone network, GhostConv and the newly designed C2fGhost are used to replace the traditional convolution and C2f modules, which reduces the network parameters and computation volume, and improves the detection performance of the model; in order to improve the model's effect of detecting artichoke fruits in complex backgrounds, the global attention mechanism module is added in the neck network to improve the feature fusion capability of the model; and the SIoU boundary loss function is used to replace the original loss function, which improves the model's feature fusion capability. function is used to replace the original loss function, which improves the performance of network bounding box regression and the detection effect of small target fruits. The experimental results show that for the clustered pepper fruits dataset established in the natural environment, the number of parameters, the computational volume and the weight size of the improved model are 2.46×10^6 , 10.5 GFLOPS and 6.02 MB, respectively, which are only 81%, 37% and 96% of the original network model, and the detection speed and the average accuracy reach 96 frame/s and 88.5%, respectively, which not only improves the detection accuracy and average accuracy not only improve the detection accuracy, but also significantly reduce the overall volume of the model, which can provide reference and basis for the subsequent deployment and application of clustered pepper fruits picking robot mobile.

收稿日期: 2024-05-09

基金项目: 贵州省科技创新基地建设项目(黔合中引地 2023010)

作者简介: 李瑞(1998—), 男, 贵州遵义人, 硕士研究生, 研究方向: 智能农机装备与智能识别检测。

† 通信联系人, E-mail: zhufugui@vip.163.com

Key words: image processing; clustered pepper; YOLOv8; lightweight network; small target; natural environment; attention mechanism; loss function

辣椒作为我国最大的单品蔬菜,在食品加工、医疗美容、休闲农业、出口贸易等领域发挥重要作用^[1]。在现代农业生产中,辣椒检测与采摘越来越趋向于规模化和智能化。其中,贵州朝天椒因其独特的果实向天生长形态以及辛辣香浓的风味而闻名。由于贵州朝天椒具有花果同期特性,其果实颜色随成熟度变化,及时识别和检测朝天椒成为农业生产中的一项重要任务^[2]。

近年来,随着基于深度学习的目标检测算法在农业领域的广泛应用,根据检测阶段的数量不同,目标检测模型主要有两类,一类是以 R-CNN^[3]、Fast R-CNN^[4] 和 Faster R-CNN^[5] 为代表的两阶段检测算法,该类型算法首先需要生成候选框区域,来对候选区域进行识别。例如,Zhong 等^[6] 利用改进 Faster R-CNN 识别尺寸小且成簇生长的朝天椒,通过使用了 RoI Align 技术和特征金字塔网络(FPN),提升小目标的识别精度,改进的模型平均精度达到了 87.5%。但其网络模型庞大且检测时间长,不能满足实时检测需求。另一类是以 SSD^[7] 和 YOLO 系列^[8] 为代表的单阶段目标检测算法,不需要生成候选框,通过将边界问题转换为回归问题,利用从网络中提取特征来预测果实位置和类别,具有准确率高、速度快、训练时间短、计算量小等特点。例如商高高等^[9] 针对现有的模型过于冗余、检测速度与精度不平衡等问题,提出一种轻量型果园环境柑橘果实检测方法。改进的模型平均检测精度均值为 94.22%,相较于原始 YOLOv4 模型提高了 1.1 个百分点,模型参数减少了 95.22%,检测速度为原来的 4.03 倍,具有较高的识别精确度和较强的鲁棒性。Sun 等^[10] 提出一种基于改进 YOLOv5n 模型的轻量化苹果识别方法,实验结果显示改进后的轻量化模型推理速度为 6.25 MB,为原始 YOLOv5n 模型的 80%,并且精确度和平均精度达到了 97.81%和 99.1%,相比原始 YOLOv5n 模型分别提高了 11.89 和 23.12 个百分点。Wang 等^[11] 提出一种小米辣青椒轻量化检测模型 SM-YOLOv5,采用 MobileV3-Large 作为骨干网络,与基准模型 YOLOv5s 相比,其平均精度均值提升了 1.4 个百分点,参数量仅为原模型的 42.48%,计算量仅需 7.6GLOPS,且模型鲁

棒性更强。张震等^[12] 采用部分卷积(Partial Convolution, PConv)替换多分支堆叠模块中的部分常规卷积进行轻量化改进,以降低模型的参数量和计算量。实验结果显示,相较于 YOLOv7 原始模型,改进后的模型精确度和平均精度均值分别提升了 4.15 和 1.39 个百分点,其参数量和计算量分别降低了 22.93%和 27.41%,实现了检测模型精度、速度和轻量化方面的平衡。陈仁凡等^[13] 为解决温室环境下草莓果实快速准确识别问题,提出一种基于改进 YOLOv5s 的草莓成熟度检测方法。在主干中引入 Shuffle_Block 和颈部结构中使用全维度动态卷积模块(Omni-dimensional Dynamic Convolution, ODConv),从而实现模型轻量化。实验结果显示,改进后的 YOLO-ODM 模型的平均精度均值达到了 97.4%,模型体积仅为 7.79 MB。与原网络相比,轻量化的 YOLO-ODM 模型大小减少了 43%,浮点运算量降低了 52%。

针对辣椒果实检测方法的研究,尽管已有文献实现了较为准确的检测目标,但大多数仅限于实验室环境中,较少考虑在复杂农田自然环境和多目标环境下,实现计算量更小、时延更低且检测精度更高的果实检测。为解决此类问题,该研究基于 YOLO 系列中较新发布的、具有更高精度以及更小参数量和模型大小的 YOLOv8 模型,对朝天椒果实检测方法进行改进和优化,为算力有限的移动设备研制一种精确度高和检测速度快的朝天椒果实检测模型。

1 YOLOv8 算法

YOLOv8 是 YOLO 系列中的一个新型且高效的模型,它在检测精度和速度上相较流行的 YOLOv5 网络^[14] 有大幅的提升。各个版本根据网络深度和宽度的不同,分为五个不同的规模:YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x,以适应不同的运行环境和精度需求。本研究选择使用体积小、精度高的 YOLOv8n 网络。YOLOv8n 模型检测网络主要由四个部分组成(如图 1 所示):Input、Backbone、Neck、Head^[15]。

输入端(Input)采用了 Mosaic 数据增强技术,

但在训练的最后 10 个周期会将其关闭。同时,引入了 Anchor-Free 机制来直接预测目标的中心位置而非传统的锚框偏移量,这种方法有效改善了目标定位的精度并解决了正负样本不平衡问题。

骨干网络(Backbone)负责从输入图像中提取高级语义特征,这一部分包含 Conv、C2f 和 SPPF 模块。Conv 模块通过卷积操作、批量归一化(BN)和 SiLU 激活函数处理图像;YOLOv8 引入了一种新的 C2f 结构,这是对残差特征进行学习的核心模块,它借鉴了 YOLOv7 中的 ELAN 并行梯度流思想,并改进了 YOLOv5 的 C3 模块,以增强梯度流的同时维持模型的轻量化;SPPF(Spatial Pyramid Pooling Fusion),也称为空间金字塔池化模块,可以将不同大小的特征图统一转换为固定大小的特征向量,以便进一步处理。

颈部网络(Neck)的主要任务是融合多尺度特征并生成特征金字塔,使用的是 PANet 结构,包括特征金字塔网络(FPN)和路径聚合网络(PAN)两个部分。FPN 通过在卷积神经网络中提取特征图构建特征金字塔,并采用自顶向下的方法通过上采样与粗粒度特征图进行融合,实现多层特征的结合。然而,FPN 单独使用可能会丢失部分位置信息,因此引入 PAN 作为补充。PAN 使用自底向上的结构,在卷积层中融合不同层次的特征图,更精确地保存空间信息。FPN 和 PAN 的结合有效地实现了信息的上下流动,提升了网络的检测性能。

头部网络(Head)作为最后的预测部分,YOLOv8 借鉴了 YOLOX 的解耦思想实现类别特征和位置特征的解耦合,根据不同尺寸的特征图来获取不同大小目标物体的类别和位置信息,提高检测效率。

2 改进的 YOLOv8 算法

针对当前朝天椒果实检测中存在的模型参数量大、权重文件大以及检测精度低等问题,提出了一种基于改进 YOLOv8 的轻量级朝天椒果实检测方法。由于朝天椒果实生长密集,果实间存在相互重叠或遮挡,且受不同光照条件影响,果实特征在图像中存在表征不清晰的情况,因此需要通过改进网络模型,增强对果实重叠、遮挡等不同光照等情况下的特征提取能力,主要在以下 3 个方面进行优化和改进,以实现更高的检测性能和更低的计算成本。首先,为解决模型权重文件过大的问题,引入轻量级 GhostNet^[16]网络中的 GhostConv 以及全

新设计模块 C2f 来减少模型的参数量;为确保模型算法集中于小目标的检测,在颈部网络中引入全局注意力机制(Global Attention Mechanism, GAM)^[17];为提升网络模型的边界框回归性能,引入 SIoU(Spatial Intersection over Union)损失函数解决了预测框漂移问题。改进后的 YOLOv8n 网络结构如图 2 所示。

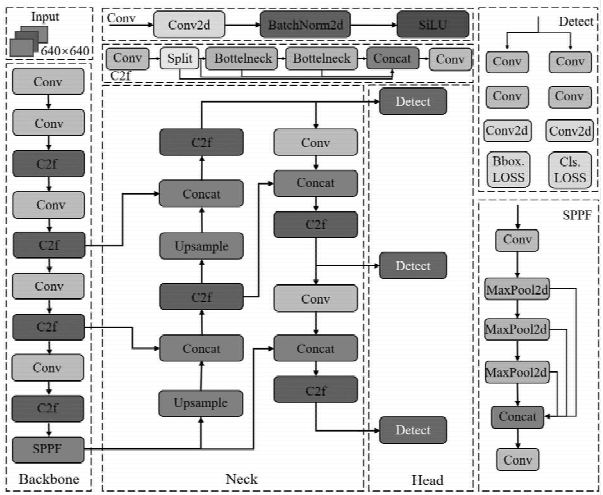


图 1 YOLOv8n 网络结构

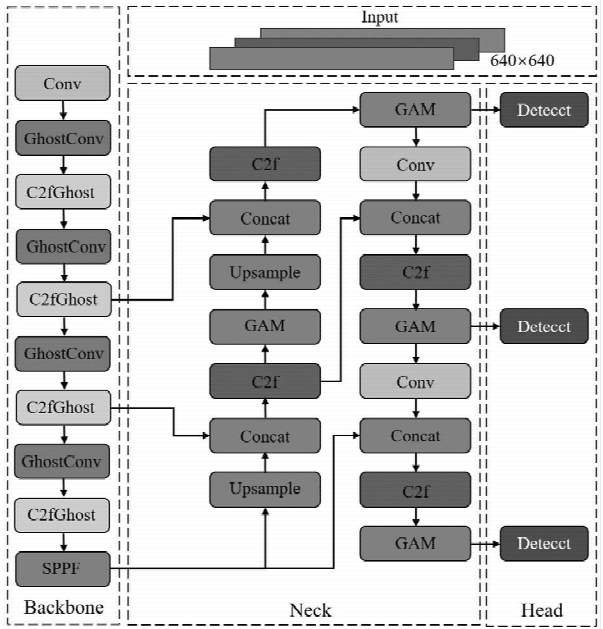


图 2 改进的 YOLOv8n 网络结构

2.1 GhostNet 骨干网络

GhostNet 是某实验室 2020 年设计的轻量级网络, GhostConv 是 GhostNet 网络中的卷积模块,可以替代普通的卷积。如图 3 所示, GhostNet 轻量级网络模型可以保持原有卷积输出特征图的尺寸以及通道大小的前提下,降低网络的计算量和

参数量。首先从少量的非线性卷积中提取特征,然后使用线性卷积操作生成 Ghost 特征图。接着,将两个卷积阶段的特征图进行组合,得到通道数更多的特征图。该过程消除了冗余特征,降低了模型的计算量。

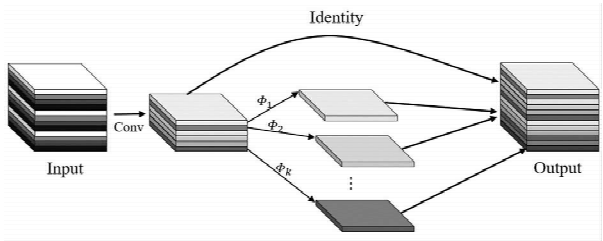


图 3 GhostNet 设计核心

如图 4 所示,“廉价操作”是一种低成本的线性操作。GhostConv 首先应用一个与原始 3×3 卷积大小相同的卷积来生成一半的特征图。随后,在经济有效的计算(Cheap Operation)中执行另一个步长为 1 的 3×3 卷积,以获得另一半的特征图。最后将这两组特征图拼接在一起,形成完整的特征图。

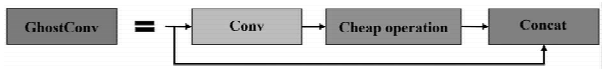


图 4 幻影卷积模块

如图 5 所示,设计了一个全新的模块 C2fGhost,通过在原网络中 C2f 模块中用 GhostBottleneck 替换所有的 Bottleneck 而得到,该结构采用了跨阶段特征融合策略和截断梯度流技术,以增强网络层之间特征的可变性,减少冗余梯度的影响,从而提升了网络的学习效率。此外,通过引入 GhostConv 和 C2fGhost 模块,减少了原网络结构中大量的 3×3 的普通卷积,大幅压缩了网络的模型大小,降低了参数量和计算量,使得模型能够部署于在移动设备,便于实现移动端的朝天椒果实检测。



图 5 幻影瓶颈模块

2.2 GAM 全局注意力机制

在复杂背景下进行朝天椒果实检测时,常遇到背景占据图片大部分的情况,同时需要识别图片中的小目标果实。为增强模型对这些小型果实特征的识别能力,引入了 GAM(Global Attention Mechanism)注意力机制,GAM 是继 CBAM 后全

新的注意力模块,其结构如图 6 所示,不仅是一种结合通道注意力机制和空间注意力机制的模块,还可以在减少弥散信息的同时放大全局维度特征交互的注意力机制。通过给定一个特征图 F_1 ,经过两个独立的注意力子模块,先通过通道注意力机制模块对原特征图进行校正,得到中间状态 F_2 ,然后经过空间注意力机制模块再次校正,得到最终特征图 F_3 。

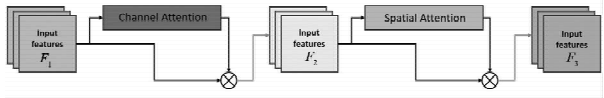


图 6 GAM 注意力模块

通道注意力子模块如图 7 所示,使用三维布局来捕获三维数据。它利用双层多层感知机(MLP)以增进跨通道的空间相关性。

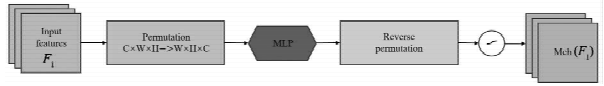


图 7 通道注意力子模块

中间阶段的定义如式(1)所示(其中 Mch 表示通道注意力图, $\otimes$ 表示逐元素乘法):

$$F_2 = Mch(F_1) \otimes F_1 \tag{1}$$

空间注意力子模块如图 8 所示,对尺寸为 $C\times H\times W$ 的输入特征图 F_2 进行操作。它使用两个 3×3 卷积来对 3×3 块中像素之间的非线性关系进行建模,以便参数可以更好地捕获像素空间之间的关系。

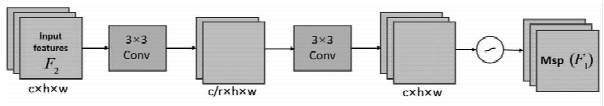


图 8 空间注意力子模块

中间状态的定义如式(2)所示(其中 Msp 表示空间注意力图):

$$F_3 = Msp(F_2) \otimes F_2 \tag{2}$$

2.3 改进损失函数

原始 YOLOv8 模型使用的是 DFL Loss+CIoU Loss 作为回归损失,但是 CIoU 在纵横交并比描述的相对值方面存在一定的模糊,并且没有考虑难易样本的平衡问题。因此本文引入 SIoU Loss 代替 CIoU Loss,SIoU 作为 IoU 的改进形式,不仅考虑边界框之间的交并,还考虑果实真实框和预测框之间的向量角度、距离和形状信息。这使得 SIoU 损失更适合检测和定位果实,因为它可以更

准确地捕获目标的详细特征。

SIoU 损失函数是一种综合性的回归损失函数，它包含角度成本、距离成本、形状成本和 IoU 成本四个组成部分。这种多元素结构使得 SIoU 损失函数能够更细致地调整边界框的位置和形状，从而提高目标检测的精确度。通过采用 SIoU 损失，改进的检测算法能够更有效地处理小目标和在复杂环境中的检测任务。SIoU 的计算如式(3)所示。

$$L_{\text{SIoU}} = 1 - \text{IoU} + \frac{4}{\pi^2} \arctan\left(\frac{w_{gt}}{h_{gt}} - \frac{w_p}{h_p}\right) \quad (3)$$

式中： w_{gt} 和 h_{gt} 是真实边界框的宽度和高度， w_p 和 h_p 是预测边界框的宽度和高度。

3 实验结果与分析

3.1 实验环境

该研究使用的算法模型的训练环境如表 1 所示。

表 1 训练和测试平台环境		
配置	配置名称	详细信息
硬件配置	CPU	Intel Core i5-13600KF
	内存大小	32 GB
	GPU	英伟达 RTX 4070
	显存大小	12 GB
软件配置	操作系统	Windows11 23H2
	Python	3.9.0
	框架	Pytorch 1.11
	CUDA	Cuda 11.3

3.2 实验数据集与实验参数设置

本研究主要研究对象是“遵辣九号”品种的朝天椒果实，实验数据采集地为某农业科技有限公司以及辣都(虾子)辣椒智慧产业园。使用 iPhone 14 手机作为主要的拍摄工具，采集时间为 2023 年 6 月 29 日至 8 月 19 日。为保证实验数据的多样性，分别在晴天、阴天和雨后的不同天气条件下模拟机械采摘的作业方式，通过三脚架固定

设备位置，并以俯视、平视、仰视三种视角采集图像。如图 9 所示，朝天椒数据集共 3 种成熟度果实，包括绿熟期 (Green_Pepper)、转色期 (Mix_Pepper) 以及完熟期 (Red_Pepper) 的高分辨率的照片。由于采集的图像是在各农田复杂的环境下随机拍摄的，对光照不充足、严重逆光、果实遮挡严重及模糊的照片予以删除，整理保留共 2400 张图像，采用 JPEG 格式进行存储，如表 2 所示为采集

图像详细信息。使用 LabelImg 软件对 2400 张朝天椒果实图像进行目标标注，对目标以最小外接矩形框进行标注处理。随后将 2400 张图像按照 8 : 2 的比例将数据集划分为训练集(1920 张)和测试集(480 张)进行模型训练和实验。

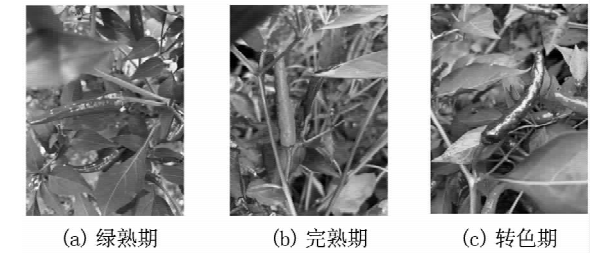


图 9 不同成熟度朝天椒果实

表 2 采集图像详细信息

影响因素	采集条件	图像数量
天气情况	晴天, 阴天, 雨天	939787674
不同时段	07:00—09:00, 13:00—14:00, 16:00—17:00	995963442
遮挡情况	果实遮挡, 枝叶遮挡, 无遮挡	6215791200
不同视角	俯视, 平视, 仰视	1143809448

实验中输入图像的像素被缩放至大小为640×640,并配合使用 Mosaic 数据增强,但在训练最后 10 个周期将其关闭。在模型训练的过程中,初始学习率(Learning Rate, LR)是 0.0001,配合优化器 Adam 调整学习率, batch_size 与 epoch 分别设置为 16 与 300。在进行性能测试时,设置 IoU 阈值为 0.5,预测的结果采用非极大值抑制(Non-Maximum Suppression, NMS),并且 NMS 的 IoU 阈值同为 0.5。

3.3 实验评价指标

将召回率(Recall)、精确率(Precision)、检测速度、模型体积与计算量等相关参数对本研究网络模型的性能进行评价。其中模型检测速度以每秒处理的帧数 Frames Per Second(FPS)作为评价指标,模型检测各类朝天椒果实的准确度采用精度 AP(Average Precision)作为评价指标, mAP 则表示数据集中所有类别 AP 的均值。

3.4 实验结果分析

3.4.1 主干网络对比实验

以 YOLOv8n 模型为基础,通过使用不同主干网络的实验比较,验证引入 GhostConv 与 C2fGhost 模块的优点,实验结果如表 3 所示。从表 3 中可知,首先,采用轻量级网络 ShuffleNetv2

代替原 YOLOv8n 的 DarkNet-53 网络,检测速度提升了 $56 \text{ 帧} \cdot \text{s}^{-1}$,但是检测精确率没有太大变化;其次,用 MobileNetv3 代替原 YOLOv8n 的 DarkNet-53 网络,检测精确率提升了 1.02% ,但是检测速度没有太大提升;最后用 GhostNet 代替 DarkNet-53 网络,精确率提升 5.31% ,帧率提升接近一倍。因此,GhostNet 网络提升了 YOLOv8n 模型的检测性能。

3.4.2 消融实验

为了验证本文提出的 YOLOv8n 改进方法对朝天椒果实检测的有效性,进行了消融实验。在相同的训练过程、超参数设置和训练部署环境下得到了表 4 中的实验结果。在表 4 中,YOLOv8n 在没有改进的情况下,达到了 $36 \text{ 帧} \cdot \text{s}^{-1}$ 的检测速度和 80.7% 的 mAP。

Model-1 在 YOLOv8n 的基础上引入 GhostNet 替换了 Backbone。GhostNet 由于引入了廉价操作,大幅降低了模型 Backbone 部分的参数和计算量,使得 Model-1 的检测速度达到 $98 \text{ 帧} \cdot \text{s}^{-1}$,比 YOLOv8n 提高了 $32 \text{ 帧} \cdot \text{s}^{-1}$ 。Model-2 将 GAM 注意力机制模块添加到颈部结构中,增加了网络对小目标的兴趣,提高了网络的特征提取和融

合能力,最终提高了小物体检测的准确性。相较于原始 YOLOv8n 网络,平均精度均值提高了 2.9% ,达到了 83.6% 。但是模型的复杂度依旧较高,导致检测速度降低了 $14 \text{ 帧} \cdot \text{s}^{-1}$ 。Model-3 在原有 YOLOv8n 的基础上,用 SIoU 损失代替原模型的 CIoU 损失函数。由于 SIoU 解决了 CIoU Loss 函数中预测框漂移的问题,Model-3 的平均精度提高了 1.7% ,达到了 82.4% ,同时保持了原有的检测速度。Model-4 在引入 GhostNet 的同时加入 GAM 注意力机制,增强了特征融合能力并降低模型成本,使得推理准确率显著提高了 5.1% ,达到了 85.2% ,检测速度也提高到了 $103 \text{ 帧} \cdot \text{s}^{-1}$ 。Model-5 是本文改进的朝天椒果实检测模型,命名为 YOLO-GGS。通过在 Model-4 的基础上使用 SIoU,进一步增强了 YOLOv8n 模型的损失函数,使得检测精度相较于 Model-4 有 3.3% 的提升。此外,SIoU 的计算复杂度降低了模型的检测速度,降低了 $7 \text{ 帧} \cdot \text{s}^{-1}$ 。消融实验表明,本文对 YOLOv8n 的改进使得检测速度提高了 60 帧/s ,达到了 $96 \text{ 帧} \cdot \text{s}^{-1}$,平均精度提高了 7.8% ,达到了 88.5% 。这验证了 YOLO-GGS 算法的有效性。

表 3 YOLOv8n 不同主干网络对比实验

模型	主干网络	精确率/%	召回率/%	模型体积/MB	检测速度/帧 · s ⁻¹
YOLOv8n	DarkNet-53	80.39	80.21	6.27	36
YOLOv8n	ShuffleNetv2	80.86	80.79	2.3	92
YOLOv8n	MobileNetv3	81.41	81.33	4.9	59
YOLOv8n	GhostNet	85.7	85.4	5.1	68

表 4 消融实验结果

模型	GhostNet	GAM	SIoU	mAP@0.5/%	计算量/G	检测速度/(帧 · s ⁻¹)
YOLOv8n				80.7	28.2	36
Model-1	✓			80.1	8.9	98
Model-2		✓		83.6	29.4	22
Model-3			✓	82.4	28.9	36
Model-4	✓	✓		85.2	9.5	103
YOLO-GGS	✓	✓	✓	88.5	7.95	96

从图 10 中可看出,本研究提出的模型收敛速度较快,在 20 次 epoch 左右就趋于收敛,并且能够在短时间内训练出较优的模型。可以看出,引入 GhostNet、GAM 和 SIoU 后的 YOLO-GGS 与原始 YOLOv8n 模型相比,收敛速度明显加快。YOLO-

GGS 和 YOLOv8n 在测试集上的识别结果如表 4 所示,YOLO-GGS 算法的检测速度明显高于 YOLOv8n,达到了 $96 \text{ 帧} \cdot \text{s}^{-1}$ 。表 4 和图 10 共同表明 YOLO-GGS 比 YOLOv8n 具有更高的检测精度。此外,由于网络 Neck 组件中增强的特征融合能力,

该算法对小目标类果实的识别准确率显著提高。

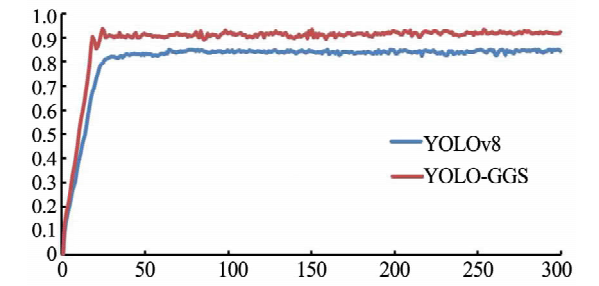


图 10 训练结果

3.4.3 不同环境条件模型性能对比实验

由于使用机械手采摘朝天椒果实时需要在各种环境条件下进行高精度的目标识别,因此模型必须表现出高鲁棒性。为了评估模型的鲁棒性,本文在朝天椒测试集中的图像子集上进行了实验,通过在不同真实场景下进行模型训练,如不同天气条件、各种拍摄视角与距离以及多目标等特殊情况。随后,分别采用 YOLOv8n 模型和 YOLO-GGS 模型进行测试,实验结果如表 5 所示,效果如图 11 所示。

表 5 不同环境下两种目标检测模型性能对比		
场景	YOLOv8nmAP@0.5	YOLO-GGSmAP@0.5
	80.7%	88.5%
晴天	75.1%	83.3%
阴天	63.9%	78.2%
俯视	68.7%	76.8%
仰视	69.1%	73.9%
远距离	70.4%	82.7%

从结果中可以看出,仰视和远距离对算法识别性能的影响最为明显。在这两种环境中, YOLOv8n 和 YOLO-GGS 都经历了 mAP 的降低。然而,与 YOLOv8n 相比, YOLO-GGS 表现出更小的精度下降。其中,在俯视条件下, YOLOv8n 和 YOLO-GGS 的 mAP 分别降低了 12% 和 11.7%, 而在仰视条件下, YOLOv8n 和 YOLO-GGS 的 mAP 分别降低了 11.6% 和 14.6%。这表明 YOLO-GGS 具有更好的泛化能力和鲁棒性,在具有挑战性的环境条件下表现更可靠。

3.4.4 不同网络模型性能对比实验

为了验证本文提出的 YOLO-GGS 算法在目标检测领域的检测效果,将 YOLO-GGS 与其他常用工程算法 Faster R-CNN、SSD、YOLOv5s、YOLOv7-tiny 在相同的设备上训练。算法对

比实验结果如表 6 所示,与 SSD 和 YOLOv5s 相比, Faster R-CNN 算法取得了更高的准确率。然而,由于其两阶段的性质,它实现的检测速度只有 $4 \text{ 帧} \cdot \text{s}^{-1}$ 。与 Faster R-CNN 相比, SSD 虽然在速度上略有提升,但在移动设备仍无法满足实时检测要求,其检测精度仅为 71.2%。YOLOv5s 算法在速度和检测精度之间取得了平衡,在保持 72.9% 的检测精度的同时,实现了 $37 \text{ 帧} \cdot \text{s}^{-1}$ 的检测速度。然而,其检测速度和精度依旧达不到预期的要求。YOLOv7-tiny 和 YOLOv8n 都是在 YOLOv5 的基础上进行的改进,提高了检测精度。然而,引入的改进增加了模型复杂度和计算量,导致检测速度相比 YOLOv5 明显下降。

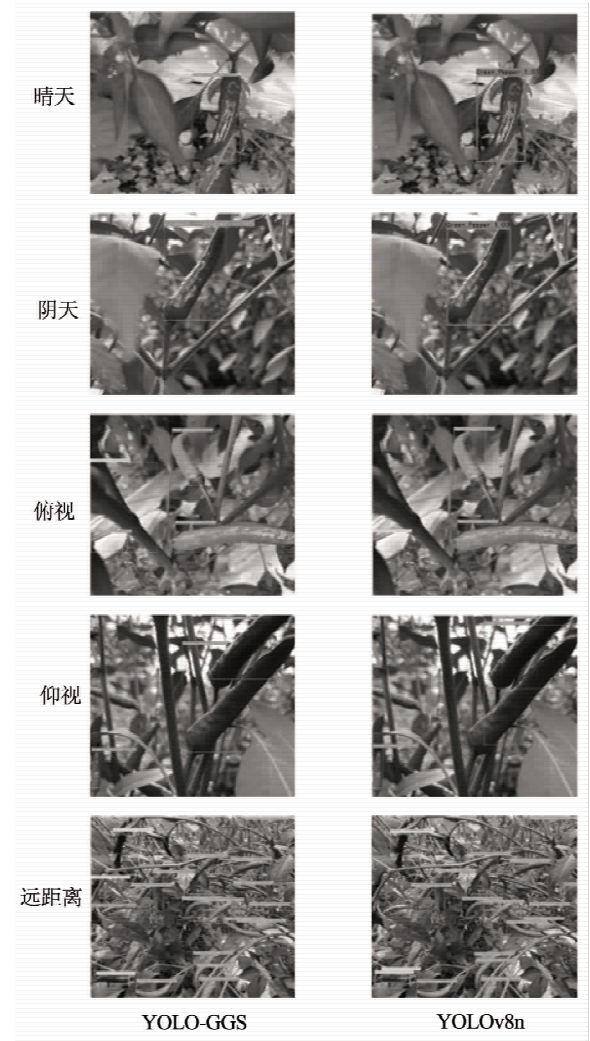


图 11 各种场景下改进前后模型检测效果对比

相比之下,本文提出的 YOLO-GGS 模型在检测精度和速度上都优于其他对比算法,能够满足朝天椒果实检测的实时检测需求。

表 6 不同检测模型对比

模型	参数量/M	计算量/G	模型体积/MB	mAP@0.5	检测速度/(帧·s ⁻¹)
Faster-RCNN	28.39	47.1	108.69	74.1%	4
SSD	24.01	69.2	91.6	71.2%	18
YOLOv5s	7.01	15.8	13.7	72.9%	37
YOLOv7-tiny	6.2	12.7	11.3	79.2%	23
YOLOv8n	3.019	28.2	6.27	80.7%	36
YOLO-GGS	2.46	10.5	6.02	88.5%	96

4 结 论

提出了一种名为 YOLO-GGS 的轻量级网络架构,以解决现有朝天椒果实检测模型中识别速度慢、计算量高、识别准确率低等问题。以 YOLOv8n 为基础模型,添加了 GAM 注意力模块,增加了网络对小物体的兴趣,提高了网络的特征提取和融合能力。此外,引入 GhostConv 模块以及设计的全新模块 C2fGhost,减少了模型的参数量及模型大小;在提高特征提取能力的同时减少了网络参数数量。最后,使用 SIoU 作为损失函数,提升了网络模型的边界框回归性能。

改进后的 YOLO-GGS 网络应用于不同成熟度朝天椒果实的实际案例时,表现出参数和模型尺寸的减少,并在精度、mAP@0.5、参数量、FPS、模型大小 5 个方面与原始 YOLOv8n 网络模型相比表现更好。本研究提出的方法在不同成熟度朝天椒数据集上实现了 96 帧·s⁻¹ 的 FPS,参数量和计算量分别减少了 19.6% 和 71.9%,而 mAP 提高了 7.8%。与其他目标检测算法相比,YOLO-GGS 在保证高效检测速度和控制模型大小的同时,表现出更高的检测精度。它有效解决了不同成熟度朝天椒果实检测任务中的漏检等问题,进一步证实了本文提出的算法的强大泛化能力。为未来朝天椒采摘机器人的设计和开发提供了宝贵的参考依据。

关于未来的研究目标和方向,需要进一步研究模型在小目标检测中的性能优化。由于朝天椒果实自身的生长环境,特征图中细节明显丢失,这表明了潜在的改进途径。本研究将研究改进卷积策略和特征提取方法,以更好地捕捉微妙的细节,这对于在更广泛的评估指标中实现更高的准确性至关重要。

参考文献

[1] 乔立娟,赵帮宏,宗义湘,等. 我国辣椒产业发展现状、趋势及

对策[J]. 中国蔬菜, 2023(11): 9—15.
[2] 赵靛瑜,王洪亮,吴康云. 贵州辣椒育种研究进展[J]. 耕作与栽培, 2022, 42(3): 56—58.
[3] DONG C,ZHANG K,XIE Z Y,et al. An improved cascade rcnn detection method for key components and defects of transmission lines[J]. IET Generation, Transmission & Distribution, 2023, 17(19): 4277—4292.
[4] WANG X L,SHRIVASTAVA A,GUPTA A. A-fast-rcnn: hard positive generation via adversary for object detection[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (cvpr), [S.l.]:IEEE, 2017.
[5] LIU R,YU Z H,MO D L,et al. An improved faster-rcnn algorithm for object detection in remote sensing images[C]//2020 39th Chinese Control Conference (CCC), [S.l.]: IEEE, 2020.
[6] ZHONG S H,XU W P,ZHANG T H,et al. Identification and depth localization of clustered pod pepper based on improved Faster R-CNN[J]. IEEE Access, 2022, 10: 93615—93625.
[7] ZHAI S P,SHANG D R,WANG S H,et al. DF-SSD: an improved SSD object detection algorithm based on densenet and feature fusion[J]. IEEE Access, 2020, 8:24344—24357.
[8] JIANG P J,ERGU D,LIU F Y,et al. A review of YOLO algorithm developments[J]. Procedia Computer Science, 2022, 199: 1066—1073.
[9] 商高高,姜锬,韩江义,等. 一种轻量型果园环境果实检测方法[J]. 江苏大学学报(自然科学版), 2024, 45(1): 46—52+59.
[10] SUN Y,ZHANG D W,GUO X D,et al. Lightweight algorithm for apple detection based on an improved YOLOv5 model[J]. Plants, 2023, 12(17): 3032.
[11] WANG X F,WU Z W,JIA M,et al. Lightweight Sm-YOLOV5 tomato fruit detection algorithm for plant factory[J]. Sensors, 2023, 23(6): 3336.
[12] 张震,周俊,江自真,等. 基于改进 YOLO v7 轻量化模型的自然果园环境下苹果识别方法[J]. 农业机械学报, 2024, 55(3): 231—242+262.
[13] 陈仁凡,谢知,林晨. 基于 YOLO-ODM 的温室草莓成熟度的快速检测[J]. 华中农业大学学报, 2023, 42(4): 262—269.
[14] 孙丰刚,王云露,兰鹏,等. 基于改进 YOLOv5s 和迁移学习的苹果果实病害识别方法[J]. 农业工程学报, 2022, 38(11): 171—179.
[15] XIE S,SUN H W. Tea-yolov8s: a tea bud detection model based on deep learning and computer vision. [J]. Sensors, 2023, 23(14): 6576.
[16] HAN K,WANG Y H,TIAN Q,et al. Ghostnet: more features from cheap operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 1577—1586.
[17] LIU Y C,SHAO Z R,HOFFMANN N. Global attention mechanism: retain information to enhance channel-spatial interactions [EB/OL]. [2021—12—13]. <https://arxiv.org/abs/2112.05561>.