

黄彤镔, 黄河清, 李震, 等. 基于YOLOv5改进模型的柑橘果实识别方法[J]. 华中农业大学学报, 2022, 41(4): 170-177.
DOI: 10.13300/j.cnki.hnlkxb.2022.04.021

基于YOLOv5改进模型的柑橘果实识别方法

黄彤镔^{1,2}, 黄河清^{1,2}, 李震^{1,2,3,4}, 吕石磊^{1,2,3,4},
薛秀云^{1,2}, 代秋芳^{1,2}, 温威^{2,5}

1. 华南农业大学电子工程学院(人工智能学院), 广州 510642; 2. 国家柑橘产业技术体系机械化研究室, 广州 510642;
3. 广东省农情信息监测工程技术研究中心, 广州 510642; 4. 人工智能与数字经济广东省实验室(广州), 广州 510330;
5. 华南农业大学工程基础教学与训练中心, 广州 510642

摘要 为实现在自然环境下对柑橘果实的识别, 提出一种基于YOLOv5改进模型的柑橘识别方法。通过引入CBAM(convolutional block attention module, 卷积注意力模块)注意力机制模块来提高网络的特征提取能力, 改善遮挡目标与小目标的漏检问题; 采用 α -IoU损失函数代替GIoU损失函数作为边界框回归损失函数, 提高边界框定位精度。结果显示: 本研究提出的模型平均精度AP值达到91.3%, 在GPU上对单张柑橘果实图像的检测时间为16.7 ms, 模型占用内存为14.5 Mb。结果表明, 本研究基于YOLOv5的改进算法可实现在自然环境下快速准确地识别柑橘果实, 满足实时目标检测的实际应用需求。

关键词 YOLOv5; 柑橘识别; 自动采摘; CBAM; 损失函数; 注意力机制; α -IoU

中图分类号 S661.1; TP391.41 **文献标识码** A **文章编号** 1000-2421(2022)04-0170-08

随着智慧农业的快速普及, 我国柑橘产业智能化正快速发展, 果园产量预测和机器人自动采摘^[1-2]等成为研究热点。快速精确识别柑橘果实目标对实现果园柑橘自动采摘、柑橘产量预测以及柑橘产业智能化管理具有重要意义。只有当柑橘果实识别效果达到一定精度, 机器人采摘才能优于手动采摘, 因此, 提高检测精度至关重要。

近年来, 针对柑橘果实识别问题, 国内外相关文献基于传统机器学习算法提出许多解决方法。熊俊涛等^[3]利用K-means聚类算法实现柑橘果实分割, 该方法复杂情况下识别效果较差。卢军等^[4]提出分层轮廓分析算法检测树上绿色柑橘, 其算法时间效率不高。毕松等^[5]利用多重分割法提出了柑橘视觉识别模型, 模型平均准确率为86.6%, 复杂情况下检测效果较差。邹继欣等^[6]、周洪刚等^[7]利用色差分析实现柑橘分割, 其算法受噪声影响较大。

传统机器视觉算法在复杂场景下鲁棒性差, 很难满足实际工作需要。近年来, 越来越多研究人员

将卷积神经网络应用于农业检测领域, 体现出巨大的算法优势, 标志性算法有RCNN^[8]、Faster-RCNN^[9]、SSD^[10]、YOLO^[11]等。顾伟等^[12]利用改进SSD模型对群体棉籽进行识别, 提高了对尺寸较小物体的检测精度。杨万里等^[13]运用SegNet和Faster-RCNN对盆栽水稻稻穗进行分割与计数。薛月菊等^[14]提出基于改进YOLOv2^[15]的未成熟芒果识别方法。唐熔钗等^[16]针对识别果园百香果提出改进YOLOv3网络模型^[17], 识别精度达到97.5%。吕石磊等^[18]提出基于改进YOLOv3-Lite轻量型柑橘识别方法, 提高模型识别精度, 增强算法的鲁棒性。陈文康等^[19]提出改进YOLOv4^[20]的柑橘检测方法, 但该方法模型较大, 检测速度较慢。

为解决上述模型存在的检测速度较慢、检测条件要求较高等问题, 本研究通过改进YOLOv5网络模型, 融入CBAM(convolutional block attention module, 卷积注意力模块)注意力机制模块^[21]增强网络提取图像特征的能力, 并引入 α -IoU^[22]损失函数, 提

收稿日期: 2022-01-13

基金项目: 国家重点研发计划(2020YFD1000107); 国家自然科学基金项目(31971797); 国家现代农业产业技术体系建设专项(CARS-26); 广东省科技厅项目(2021A1515010923); 广东省省级乡村振兴战略专项(粤财农[2021]37号); 广东省大学生科技创新培养专项(pdjh2020a0083)

黄彤镔, E-mail: 386138502@qq.com

通信作者: 李震, E-mail: lizhen@scau.edu.cn

出一种改进YOLOv5神经网络的柑橘果实识别方法,旨在为快速精确识别柑橘果实目标提供新的思路。

1 材料与方法

1.1 数据采集与数据增强

柑橘图像的采集地为湖南省郴州市宜章县柑橘种植园,使用数码相机、高清手机等设备在自然光照下拍摄柑橘树冠图像,并采集树冠下方含阴影的果实。为获得不同距离下的柑橘图像,分别在距离果树0.5、1.0、2.0 m地方对每棵果树进行拍摄,共采集889张柑橘原始图片,图片中柑橘目标均为成熟果实。图片包括阴天、晴天,涵盖顺光、逆光等所有光照情况。

为防止训练数据量过少发生过拟合现象,本研究利用Matlab工具对原始数据集进行数据扩增,通过采用随机翻转、局部裁剪、长宽缩放、限制对比度自适应直方图均衡化、中值滤波、加入高斯噪声、加入伽马噪声等方法进行数据增强。每张原始图像都生成5张新图片,总共得到4 445张。经过数据增强会导致部分图片质量变化较为严重,需进行人工筛选,挑选出符合条件的数据作为最终数据集,去除质量较差图片560张,最终数据集为3 885张。

1.2 数据集准备

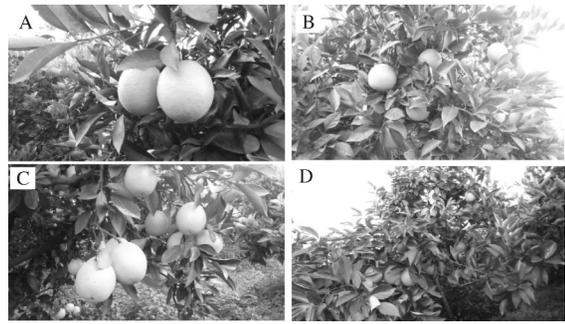
由于摄像机与柑橘目标距离远近的不同,拍摄到的柑橘图像中柑橘大小不同,本研究根据微软公开数据集MS COCO标准,将图片中分辨率小于32像素 \times 32像素的柑橘目标定义为小目标。

通过Labelimg数据标注工具对每张图片进行标注,并将数据集按照70%、20%、10%的比例划分为训练集、测试集和验证集,图片数量分别为2 721、776、388张。将最后数据按照PASCAL VOC数据集的格式存储,再将测试集分为两部分:测试集A包含目标平均遮挡程度小于30%的图片以及小目标物体,共492张图片;测试集B包含目标平均遮挡程度大于30%的图片以及大目标物体,共284张图片。其中,测试集A包含标注样本2 751个,测试集B包含标注样本3 635个。

数据集样本图片基本涵盖在自然场景下进行目标检测易出现问题的情况,数据集部分样本图片示例如图1所示。

1.3 柑橘果实检测模型设计

1)YOLOv5模型改进。YOLOv5主要由Backbone和Head两部分组成,主要用于提取图像特征的



A:大目标样本 Large target sample; B:强曝光样本 Strong exposure sample; C:密集轻度遮挡样本 Dense and lightly occluded samples; D:密集重度遮挡样本 Dense and heavily occluded samples.

图1 部分样本图片示例

Fig.1 Some sample image examples

模块为C3模块,在Backbone部分,C3模块包含的位置信息、细节信息较多,但语义信息较少。而Head部分,C3模块主要进行纹理特征提取,包含的位置信息、细节信息较少,而语义信息较多。小目标特征信息在经过多个C3模块处理后,位置信息粗糙,特征信息易丢失,从而引起网络模型对小目标的误检和漏检。为提高原始模型对柑橘小目标果实的检测精度,本研究提出改进YOLOv5网络模型,具体结构如图2所示。在Backbone中增加CBAM注意力机制模块,置于C3模块之后,CBAM模块通过通道和空间注意力模块能在网络训练过程中加强对小目标果实的点特征、线特征等特征信息的学习。

2)注意力机制。在柑橘目标检测任务中,因小目标柑橘占据像素数较少,其特征信息极易在深层网络中丢失,出现漏检、错检。而CBAM注意力机制模块通过通道注意力模块和空间注意力模块可有效增大小目标在整张特征图中的权重,促使小目标信息更易被网络学习。CBAM注意力机制模块结构如图3所示。

CBAM模块同时关注空间和通道信息,在通道注意力模块上分别经过最大值池化和平均池化将尺寸为 $C \times H \times W$ 的特征图变成 $C \times 1 \times 1$,经过MLP转换,压缩通道数,再通过加和操作进行合并结果,并进行sigmoid归一化,使得提取到的高层特征更全面更丰富。

将通道注意力模块的输出当作输入进入空间注意力模块,同样经过最大值池化和平均池化,再通过Concat操作将两者堆叠,只压缩通道维度而不压缩空间维度。关注重点在目标的位置信息,并专注于有用的目标对象。

3)损失函数。YOLOv5原始模型中采用Loss_{GIoU}

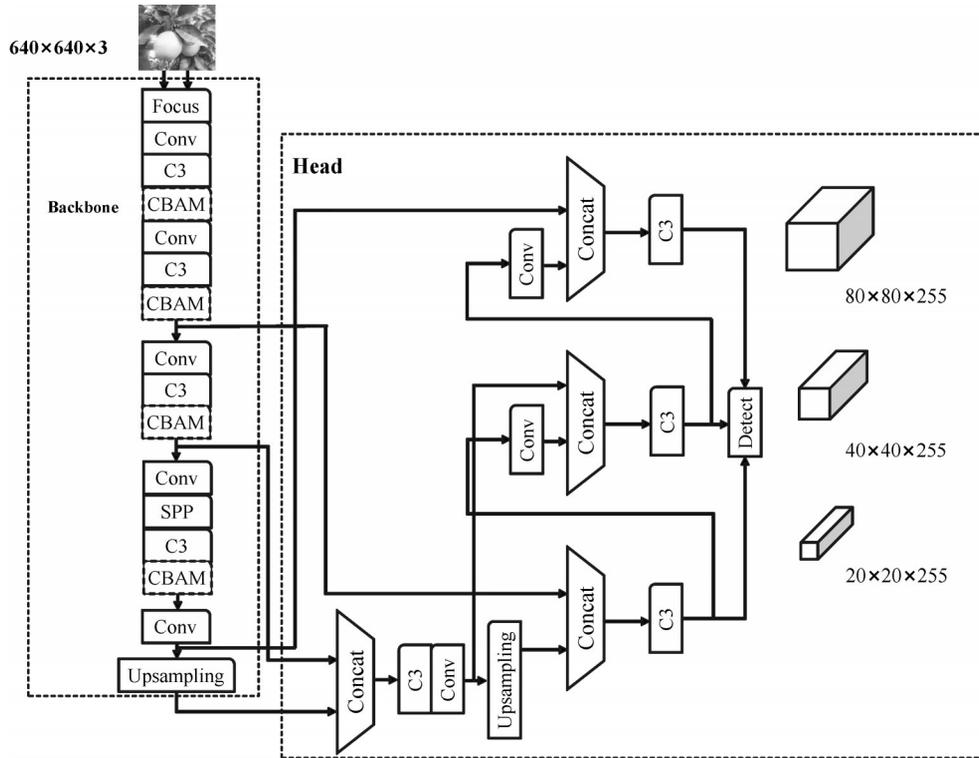


图2 改进YOLOv5网络总体框图

Fig.2 Improve the overall block diagram of the YOLOv5 network

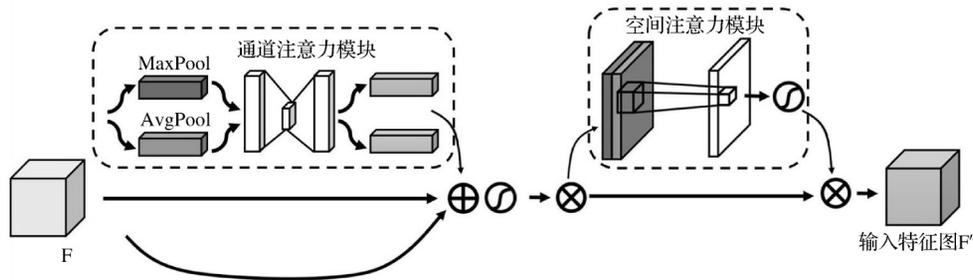


图3 CBAM总体结构图

Fig.3 CBAM overall structure diagram

做bbox的损失函数, $Loss_{GIoU}$ 存在一种问题: 如图4所示实线矩形框为目标框, 虚线框为预测框, 图4中所示3种情况为预测框在目标框内部, 该状态下GIoU值是相同的, 但其预测框位置是不同的, 无法区分相对位置关系。因此, 无法精确定位到效果最佳的预测框。

为解决 $Loss_{GIoU}$ 存在的问题, 本研究引入 $Loss_{\alpha-IoU}$ 用以改进, $Loss_{\alpha-IoU}$ 保留原损失函数全部性质, 同时更加关注高IoU目标, 并为优化所有层次的目标创造了更多空间, 实现不同水平的检测框回归精度。

$Loss_{\alpha-IoU}$ 在原有的 $Loss_{IoU}$ 上增加一个包含 α 的幂化指标, 其定义如公式(1)所示。

$$Loss_{\alpha-IoU} = \frac{1 - IoU^\alpha}{\alpha}, \alpha > 0 \quad (1)$$

通过对公式中 α 取值进行调制, 可以得到现有的大多数IoU损失函数, 更容易用于改进检测器的效果, 同时不会引入额外的参数, 也不增加训练时间。

本研究中采用目前常用效果最好的 $Loss_{CIoU}$, 改进为 $Loss_{\alpha-Clou}$, 其定义如公式(2)所示。

$$Loss_{\alpha-Clou} = 1 - IoU^\alpha + \frac{\rho^{2\alpha}(b, b^{st})}{c^{2\alpha}} + (\beta v)^\alpha \quad (2)$$

公式(2)中, c 代表的是能够同时包含预测框和真实框的最小闭包区域的对角线距离, 而 v 衡量长宽比的一致性, 其定义如公式(3); β 是一个正的权衡参数, 其值如公式(4), $\rho^{2\alpha}(b, b^{st})$ 分别代表了预测框和真实框的中心点的欧式距离。

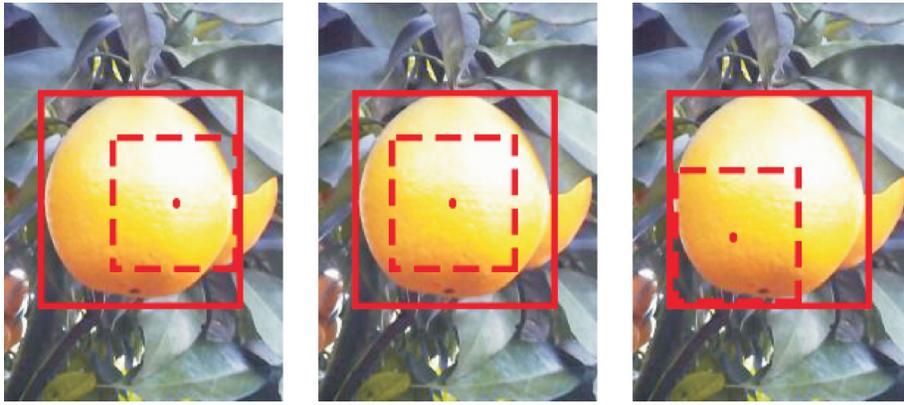


图4 预测框位于目标框内部

Fig.4 The prediction box is inside the target box

$$v = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h} \right)^2 \quad (3)$$

$$\beta = \frac{v}{(1 - IoU) + v} \quad (4)$$

公式(3)中, $\frac{\omega^{gt}}{h^{gt}}$ 和 $\frac{\omega}{h}$ 分别代表目标框和预测框各自的宽高比。

2 结果与分析

2.1 模型训练

本研究训练使用的操作系统为Ubuntu 18.04。CPU 型号为 Intel (R) Core I7-10700F, GPU 型号为 Nvidia GeForce RTX2060。测试的框架为 Pytorch1.9.0, 并使用 CUDA10.1 版本并行计算框架配合 CUDNN7.6.5.32 版本的神经网络加速库。

本研究所有结果利用推理速度作为模型推理速度的评价指标, 推理速度指对单张柑橘果实图像的检测时间。采用平均精度 P_A (average precision)、准确率 P (precision)、召回率 R (recall) 和权重大小等角度衡量检测效果, 具体计算如式(5)~(7)所示。

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (5)$$

$$R = \frac{T_p}{T_p + F_n} \times 100\% \quad (6)$$

$$P_A = \int_0^1 P(R) dR \times 100\% \quad (7)$$

公式(5)~(6)中, T_p 为真实的正样本数量, F_p 为虚假的正样本数量, F_n 为虚假的负样本数量。

网络模型训练阶段, 迭代批量大小为16, 总迭代次数设置为300次。初始学习率设置为0.001。从图5可以看出, 当迭代次数接近100次时, 模型开始逐渐

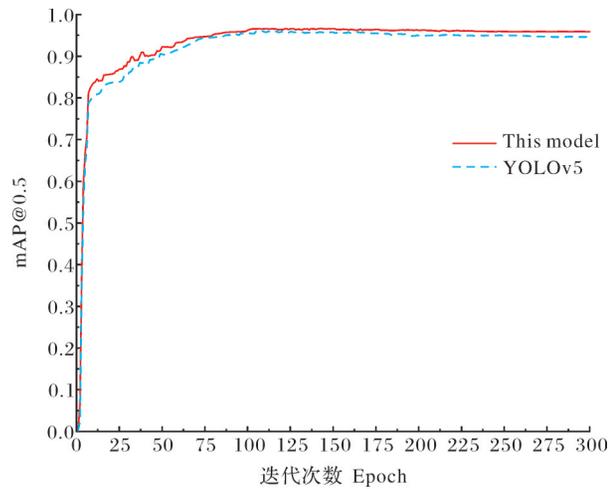


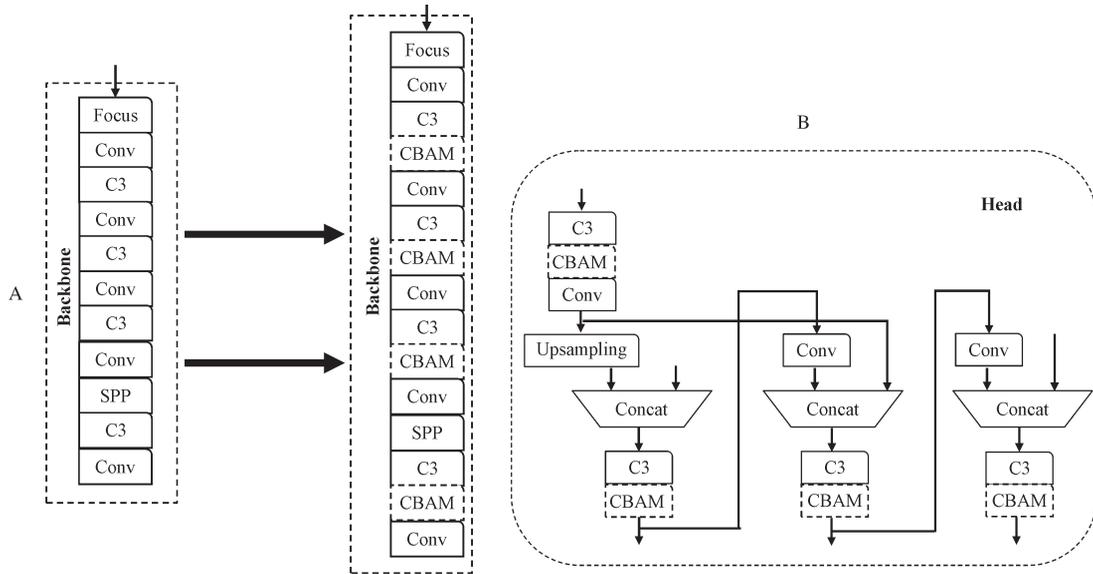
图5 YOLOv5与本研究模型mAP@0.5值比较

Fig.5 Comparison of mAP@0.5 value between YOLOv5 and this model

收敛; 相较YOLOv5模型, 改进的YOLOv5模型收敛更加快速, 最终精度较高。

2.2 CBAM 模块融合位置对比

将CBAM模块融合到网络模块的不同位置, 并对检测结果展开研究。在YOLOv5的Backbone与Head两个区域分别融入CBAM注意力模块, 生成2种基于YOLOv5算法的新网络模型: CBAM-YOLOv5-BackBone、CBAM-YOLOv5-Head。图6展示了2种CBAM注意力模块融合网络的具体位置。检测试验结果如表1所示, 将CBAM模块融合于Backbone的检测效果较好, 我们认为YOLOv5中提取特征的关键网络在Backbone, 在Backbone中隐含着易被网络忽视的小目标特征信息, CBAM注意力机制模块对这部分特征信息进行注意力重构, 突出重要信息。而在网络更深的Head部分, 小目标的特征信息被淹没, 语义信息较为粗糙, 注意力模块难以区分空间特征和通道特征。



A: CBAM-YOLOv5-Backbone; B: CBAM-YOLOv5-Head.

图6 2种融合CBAM模块的YOLOv5模型

Fig.6 Two YOLOv5 models incorporating CBAM modules

表1 CBAM模块融合结果对比

Table 1 Comparison of CBAM module fusion results

网络模型 Network models	准确率 Precision	召回率 Recall	平均精度 Average precision	%
YOLOv5	92.6	84.1	89.7	
CBAM-YOLOv5-Backbone	92.7	85.9	91.0	
CBAM-YOLOv5-Head	91.3	83.7	87.0	

2.3 YOLOv5消融试验及对比试验

为验证本研究提出的关于YOLOv5的2种改进策略,在数据集上展开消融试验,以判断每个改进点的有效性,依次在原有模型基础上加上CBAM与 α -IoU损失函数,使用改进用“√”表示,未使用改进用“-”表示。训练过程使用相同的参数配置,结果如表2所示。

表2 YOLOv5消融试验

Table 2 YOLOv5 ablation experiment

CBAM	α -IoU	准确率 Precision	召回率 Recall	平均精度 Average precision	%
-	-	92.6	84.1	89.7	
√	-	92.7	85.9	91.0	
-	√	91.2	85.4	90.4	
√	√	93.1	84.3	91.3	

引入CBAM之后可以看出,召回率提高了1.8个百分点。同时,平均精度提高了1.3个百分点,而在

引入 α -IoU损失函数的提升表现较弱。分析认为引入注意力机制增强了骨干网络对特征提取的能力,关注到许多易被淹没的语义信息,表现在检测精度上的提升较为明显,而 α -IoU损失函数降低了预测框的回归损失,提高了回归精度,因此仅有小幅提高。

为了验证本研究提出的方法具有更好的效果,选择Faster-RCNN、YOLOv4、YOLOv4-tiny 3种网络模型在相同的配置环境下利用相同的数据集,进行对比试验。试验结果如表3所示。本研究的模型减少了许多额外算力与内存开销,保持了较高的识别精度和较快的检测速率,同时占用内存资源也较少,适合部署于可移动嵌入式设备平台,相比其他3种网络有较为明显的优势。

表3 不同网络模型检测能力对比

Table 3 Comparison of different networks

网络模型 Network models	大小 Size (pixels)	平均精度/% Average precision	检测速率/ (帧/ms) FPS	模型大小/ Mb Memory size
Faster-RCNN	640	90.3	3.70	195.3
YOLOv4	640	88.3	0.07	256.0
YOLOv4-tiny	640	85.1	3.19	23.5
改进的YOLOv5 Improved YOLOv5	640	91.3	23.6	14.5

2.4 改进前后模型精度对比试验

将改进前后2个模型进行精度测试,结果如表4

所示。YOLOv5模型识别测试集B与数据集A+B的平均精度分别为75.9%和89.7%。改进的YOLOv5模型检测平均精度相比原模型分别提高了4.5个百分点和1.6个百分点。根据试验结果可知,本研究提出的改进模型在环境较为复杂的情况下,模型精度提高效果较为明显。

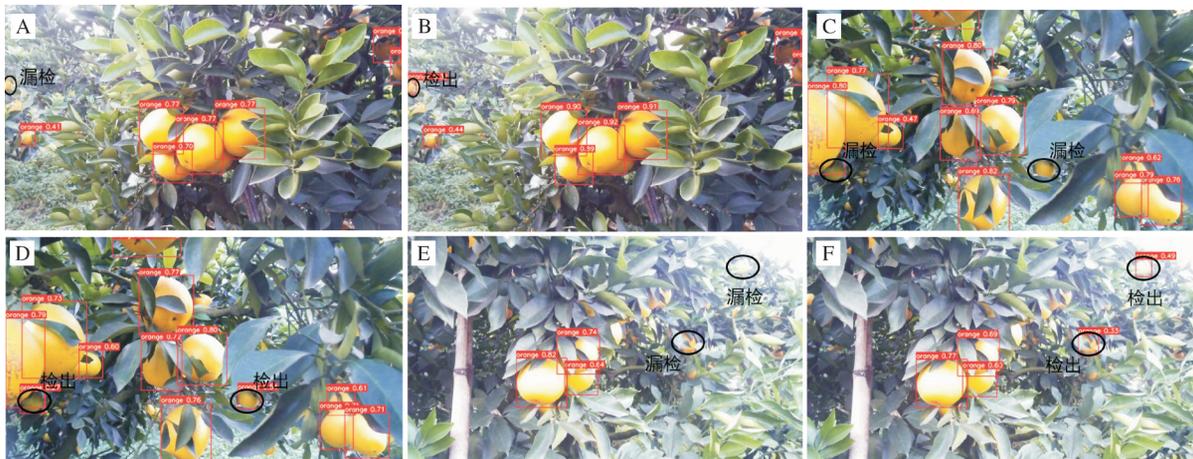
为了更加直观比较改进前后模型检测效果的差异,将测试集分别通过YOLOv5以及改进后YOLOv5算法检测输出,部分检测结果如图7所示。其中分为小目标检测、遮挡目标检测以及强曝光目标检测3种检测效果对比。图7A、C、E为YOLOv5模型的检测结果,图7B、D、F为改进YOLOv5模型的

表4 不同遮挡程度数据集检测结果对比

Table 4 Comparison of detection results of different occlusion degree datasets

模型 Network models	平均精度/% Average precision			检测速率/ (ms/帧)FPS
	A	B	A+B	
YOLOv5	97.3	75.9	89.7	14.8
改进的YOLOv5 Improved YOLOv5	97.4	80.4	91.3	16.7

检测结果。针对3种特殊目标的检测,YOLOv5模型均存在漏检的情况,而改进的YOLOv5模型则有明显改善。



A, B: 小目标的检测 Small target detection; C, D: 遮挡目标的检测 Detection of occluded targets; E, F: 强曝光目标的检测 Detection of strongly exposed targets; A, C, E: YOLOv5模型检测结果 YOLOv5 model detection results; B, D, F: 改进YOLOv5模型的检测结果 Improve the detection results of the YOLOv5 model.

图7 不同场景下模型部分检测结果对比

Fig.7 Comparison of partial detection results of models in different scenarios

3 讨论

本研究针对柑橘果实检测任务,提出一种基于YOLOv5改进模型的柑橘检测方法。将注意力机制CBAM与原始YOLOv5融合,提高模型网络对特征信息的提取能力;使用 α -IoU损失函数替换原始Loss_{GIoU},降低了损失函数的回归损失,提高了回归精度,改善了多遮挡等复杂情况下漏检以及小目标物体识别效果差的问题。本研究通过拍摄与数据增强制作了自然场景下的柑橘图像数据集,用于模型的训练和测试。试验结果表明,在自然环境下,改进模型检测的平均精度达到91.6%,对单张柑橘果实图像的检测时间为16.7 ms,模型大小为14.5 Mb。为了验证本研究模型的优越性,以检测精度、检测速度等

为判断依据,与原始模型进行对比试验,其检测平均精度比原始模型高出1.6个百分点,而在重叠遮挡较多以及小目标较多的情况下,检测的平均精度提高了4.5个百分点。同时在远景小目标、重叠遮挡以及强曝光等情况下的柑橘目标预测,本研究提出的改进型YOLOv5算法模型对原始模型出现的漏检情况有明显改进。

柑橘果实检测对自然环境下的柑橘机械化、智能化操作具有一定指导性,由于柑橘主要种植于复杂的山地果园环境,生长过程较难控制,产量预测、机械化自动采摘都离不开快速准确的柑橘目标检测。本研究方法改进的模型较为轻量,适合移植到各种便携式的可移动平台,满足复杂环境下的户外作业。在今后的研究中可以综合考虑不同大小、不

同颜色的柑橘果实特征建立算法模型,以实现更精确的检测效果。

参考文献 References

- [1] QIN X Y, HUANG R W, HUA B. Research and implementation of yield recognition of *Citrus reticulata* based on target detection [J/OL]. *Journal of physics: conference series*, 2021, 1820(1): 012128 [2022-01-13]. <https://iopscience.iop.org/article/10.1088/1742-6596/1820/1/012128>. DOI:10.1088/1742-6596/1820/1/012128.
- [2] HARRELL R C, ADSIT P D, MUNILLA R D, et al. Robotic picking of citrus [J]. *Robotica*, 1990, 8(4): 269-278.
- [3] 熊俊涛, 邹湘军, 彭红星, 等. 扰动柑橘采摘的实时识别与采摘点确定技术 [J]. *农业机械学报*, 2014, 45(8): 38-43. XIONG J T, ZOU X J, PENG H X, et al. Real-time identification and picking point localization of disturbance citrus picking [J]. *Transactions of the CSAM*, 2014, 45(8): 38-43 (in Chinese with English abstract).
- [4] 卢军, 胡秀文. 弱光复杂背景下基于MSER和HCA的树上绿色柑橘检测 [J]. *农业工程学报*, 2017, 33(19): 196-201. LU J, HU X W. Detecting green citrus fruit on trees in low light and complex background based on MSER and HCA [J]. *Transactions of the CSAE*, 2017, 33(19): 196-201 (in Chinese with English abstract).
- [5] 毕松, 高峰, 陈俊文, 等. 基于深度卷积神经网络的柑橘目标识别方法 [J]. *农业机械学报*, 2019, 50(5): 181-186. BI S, GAO F, CHEN J W, et al. Detection method of citrus based on deep convolution neural network [J]. *Transactions of the CSAM*, 2019, 50(5): 181-186 (in Chinese with English abstract).
- [6] 邹继欣, 崔天时, 邢伟. 基于DSP的柑橘果实自动识别系统 [J]. *农机化研究*, 2015, 37(2): 166-170. ZOU J X, CUI T S, XING W. Automatic identification system of citrus fruit based on DSP [J]. *Journal of agricultural mechanization research*, 2015, 37(2): 166-170 (in Chinese with English abstract).
- [7] 周洪刚, 康敏. 基于机器视觉的成熟柑橘自动识别研究 [J]. *江苏农业科学*, 2013, 41(6): 380-381. ZHOU H G, KANG M. Research on automatic recognition of mature citrus based on machine vision [J]. *Jiangsu agricultural sciences*, 2013, 41(6): 380-381 (in Chinese with English abstract).
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. [S.l.]: IEEE, 2014: 580-587.
- [9] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE transactions on pattern analysis and machine intelligence*, 2017, 39(6): 1137-1149.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector [C/OL]//LEIBE B, MATAS J, SEBEN, et al. *Computer Vision -ECCV 2016: Lecture Notes in Computer Science Cham*. [S.l.]: Springer International Publishing, 2016: 21-37. [2022-01-13] https://doi.org/10.1007/978-3-319-46448-0_2.
- [11] 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, June 27-30, 2016, Las Vegas, NV, USA. [S.l.]: IEEE, 2016: 779-788.
- [12] 顾伟, 王巧华, 李庆旭, 等. 基于改进SSD的棉种破损检测 [J]. *华中农业大学学报*, 2021, 40(3): 278-285. GU W, WANG Q H, LI Q X, et al. Cotton seed damage detection based on improved SSD [J]. *Journal of Huazhong Agricultural University*, 2021, 40(3): 278-285 (in Chinese with English abstract).
- [13] 杨万里, 段凌凤, 杨万能. 基于深度学习的水稻表型特征提取和穗质量预测研究 [J]. *华中农业大学学报*, 2021, 40(1): 227-235. YANG W L, DUAN L F, YANG W N, et al. Research on rice phenotypic feature extraction and panicle quality prediction based on deep learning [J]. *Journal of Huazhong Agricultural University*, 2021, 40(1): 227-235 (in Chinese with English abstract).
- [14] 薛月菊, 黄宁, 涂淑琴, 等. 未成熟芒果的改进YOLOv2识别方法 [J]. *农业工程学报*, 2018, 34(7): 173-179. XUE Y J, HUANG N, TU S Q, et al. Immature mango detection based on improved YOLOv2 [J]. *Transactions of the CSAE*, 2018, 34(7): 173-179 (in Chinese with English abstract).
- [15] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. [S.l.]: IEEE, 2017: 6517-6525.
- [16] 唐熔钊, 伍锡如. 基于改进YOLO-V3网络的百香果实时检测 [J]. *广西师范大学学报(自然科学版)*, 2020, 38(6): 32-39. TANG R C, WU X R. Real-time detection of passion fruit based on improved YOLO-V3 network [J]. *Journal of Guangxi Normal University (natural science edition)*, 2020, 38(6): 32-39 (in Chinese with English abstract).
- [17] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. 2018: arXiv: 1804.02767 [cs.CV] [2022-01-13]. <https://arxiv.org/abs/1804.02767>.
- [18] 吕石磊, 卢思华, 李震, 等. 基于改进YOLOv3-LITE轻量级神经网络的柑橘识别方法 [J]. *农业工程学报*, 2019, 35(17): 205-214. LÜ S L, LU S H, LI Z, et al. Orange recognition method using improved YOLOv3-LITE lightweight neural network [J]. *Transactions of the CSAE*, 2019, 35(17): 205-214 (in Chinese with English abstract).
- [19] 陈文康, 陆声链, 刘冰浩, 等. 基于改进YOLOv4的果园柑橘检测方法研究 [J]. *广西师范大学学报(自然科学版)*, 2021, 39(5): 134-146. CHEN W K, LU S L, LIU B H, et al. Real-time citrus recognition under orchard environment by improved YOLOv4 [J]. *Journal of Guangxi Normal University (natural science edition)*, 2021, 39(5): 134-146 (in Chinese with English abstract).
- [20] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. 2020: arX-

- iv: 2004.10934 [cs. CV] [2022-01-13]. <https://arxiv.org/abs/2004.10934>.
- [21] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C/OL].//FERRARI V, HEBERT M, SMIN-CHISESCU C, et al. ECCV 2018: proceedings of the 2018 European conference on computer vision: lecture notes in computer science. Cham. [S. l.] : Springer, 2018: 3-19. [2022-01-13]. https://doi.org/10.1007/978-3-030-01234-2_1.
- [22] HE J B, ERFANI S, MA X J, et al. Alpha-IoU: a family of power intersection over union losses for bounding box regression [EB/OL]. 2021: arXiv: 2110.13675 [cs. CV] [2022-01-13]. <https://arxiv.org/abs/2110.13675>.

Citrus fruit recognition method based on the improved model of YOLOv5

HUANG Tongbin^{1,2}, HUANG Heqing^{1,2}, LI Zhen^{1,2,3,4}, LÜ Shilei^{1,2,3,4},
XUE Xiuyun^{1,2}, DAI Qiufang^{1,2}, WEN Wei^{2,5}

1. College of Electronic Engineering (College of Artificial Intelligence),
South China Agricultural University, Guangzhou 510642, China;

2. Division of Citrus Machinery, China Agriculture Research System, Guangzhou 510642, China;

3. Guangdong Engineering Research Center for Monitoring Agricultural Information,
Guangzhou 510642, China;

4. Guangdong Artificial Intelligence and Digital Economy Laboratory (Guangzhou),
Guangzhou 510330, China;

5. Engineering Fundamental Teaching and Training Center, South China Agricultural University,
Guangzhou 510642, China

Abstract The rapid and accurate identification of citrus fruit is of great significance for the realization of automatically picking citrus in orchards, the prediction of citrus yield and the intelligent management of citrus industry. A citrus recognition method based on the improved model of YOLOv5 was proposed to realize the recognition of citrus fruits in natural environment. The feature extraction ability of the network and the problem of missed detection of occluded targets and small targets was improved by introducing the CBAM attention mechanism module. The α -IoU loss function instead of the GIoU loss function was used as the bounding box regression loss function to improve the positioning accuracy of the bounding box. The results showed that the average accuracy AP value of the proposed model reached 91.3%, with the detection time of a single citrus fruit image on the GPU of 16.7 ms and the model occupying 14.5 Mb of memory. It is indicated that the algorithm improved can quickly and accurately identify citrus fruits in the natural environment, meeting the practical application requirements of real-time target detection. It will provide new ideas for the intelligent citrus industry.

Keywords YOLOv5; citrus recognition; automatically picking; convolutional block attention module; loss function; mechanism of attention; α -IoU

(责任编辑:陆文昌)