

基于注意力机制与改进 YOLO 的温室番茄快速识别

张俊宁¹ 毕泽洋¹ 闫英¹ 王鹏程¹ 侯冲² 吕树盛²

(1. 北京信息科技大学机电工程学院, 北京 100092; 2. 中国农业机械化科学研究院集团有限公司, 北京 100083)

摘要: 为了实现复杂环境下农业机器人对番茄果实的快速准确识别, 提出了一种基于注意力机制与改进 YOLO v5s 的温室番茄目标快速检测方法。根据 YOLO v5s 模型小、速度快等特点, 在骨干网络中加入卷积注意力模块 (CBAM), 通过串联空间注意力模块和通道注意力模块, 对绿色番茄目标特征给予更多的关注, 提高识别精度, 解决绿色番茄在相似颜色背景中难识别问题; 通过将 CIoU Loss 替换 GIoU Loss 作为算法的损失函数, 在提高边界框回归速率的同时提高果实目标定位精度。试验结果表明, CB-YOLO 网络模型对温室环境下红色番茄检测精度、绿色番茄检测精度、平均精度均值分别为 99.88%、99.18% 和 99.53%, 果实检测精度和平均精度均值高于 Faster R-CNN 模型、YOLO v4-tiny 模型和 YOLO v5 模型。将 CB-YOLO 模型部署到安卓手机端, 通过不同型号手机测试, 验证了模型在移动终端设备上运行的稳定性, 可为设施环境下基于移动边缘计算的机器人目标识别及采收作业提供技术支持。

关键词: 温室番茄; 目标检测; YOLO; 注意力机制; 损失函数

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2023)05-0236-08

OSID:



Fast Recognition of Greenhouse Tomato Targets Based on Attention Mechanism and Improved YOLO

ZHANG Junning¹ BI Zeyang¹ YAN Ying¹ WANG Pengcheng¹ HOU Chong² LÜ Shusheng²

(1. Mechanical and Electrical Engineering College, Beijing Information Science and Technology University, Beijing 100092, China

2. Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing 100083, China)

Abstract: In order to realize the rapid and accurate recognition of greenhouse tomato fruit by agricultural picking robot in the complicated environment of greenhouse, a fast target detection method for greenhouse tomato fruit based on attention mechanism and improved YOLO v5s was proposed. According to the characteristics of small size and fast speed of YOLO v5s (You only look once v5s) model, the convolutional block attention module (CBAM) was added into the backbone network. By concatenating spatial attention module and channel attention module, the problem of color similarity between green tomato fruit and its background was solved. More attention was paid to the target features of green tomato fruit to improve the recognition accuracy. Replacing GIoU Loss with CIoU Loss as the new loss function of the algorithm contributed to improve the positioning accuracy while improving the bounding box regression rate. The test results showed that the recognition accuracy of the CB-YOLO network model for red tomato fruit detecting precision and green tomato fruit detecting precision and mean average precision in greenhouse environment was 99.88%, 98.18% and 99.53%, respectively. Compared with Faster R-CNN network model, YOLO v4-tiny network model and YOLO v5 network model, the detection accuracy and the mean average precision were improved. The CB-YOLO model was deployed to Android system of mobile phones after being tested by different mobile phones, which verified the stability of the performance detection of the deployment model under actual working condition. It will provide technical support for target detection and harvesting based on robotic mobile edge computing in facility environments.

Key words: greenhouse tomato; objection detection; YOLO; attention mechanism; loss function

收稿日期: 2022-09-23 修回日期: 2022-12-15

基金项目: 北京高校重点研究培育项目(2021YJYP201)

作者简介: 张俊宁(1980—),男,教授,博士生导师,主要从事农业机器人测控技术研究, E-mail: zhangjunning@bistu.cn

0 引言

番茄是最受欢迎的蔬果之一,而我国更是全球最大的生产国和出口国^[1]。番茄机械化采摘可降低人工成本和强度,提高采摘效率^[2]。在农业采摘机器人技术领域,温室番茄果实识别达到一定的工作速度和精度时,机器人采摘效率才能优于手动采摘,因此提高采摘目标的检测精度和速度至关重要。

以往基于传统图像处理方法的番茄果实识别研究,普遍采用颜色、纹理和边缘等特征进行识别。王丽丽^[3]提出基于 HIS 颜色特征的目标识别方法,对于相互分离的田间成熟番茄进行识别。梁喜凤等^[4]利用双目立体视觉灰度化番茄彩色图像,通过腐蚀、膨胀的方法对直方图进行处理,然后运用曲线拟合法分割图像,提取出成熟番茄目标。冯青春等^[5]在 2R-G-B 色差模型下提取红色成熟番茄果实颜色特征,利用动态阈值分割的方法识别红色成熟番茄果实,但识别耗时较长,未考虑叶片遮挡等复杂环境下番茄果实的识别。马翠花等^[6]提出了基于显著性检测与圆形随机 Hough 变换的目标识别方法,对于未成熟番茄果实的正确识别率为 77.6%。以上研究一定程度上实现了番茄的识别,但识别速度较慢,识别一幅图像至少需 5 s,达不到实时需求,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。

近年来基于深度学习的卷积神经网络成为研究热点,也应用到温室番茄的识别中^[7-9]。目前,目标检测大致分为两大主流算法。第 1 类为两阶段算法(Two-stage):先计算候选区域(Region proposal),然后进行 CNN 目标分类,如 R-CNN^[10]、Fast R-CNN^[11]、Faster R-CNN^[12]和 Mask R-CNN^[13]等,此类算法准确率较高,然而神经网络模型较大,计算速度慢,无法满足实时检测的场景需求^[14-15],无法满足实际场景的实时性;第 2 类为一阶段算法(One-stage):在检测过程中可以直接同时输出定位和分类结果,如 YOLO^[16](You only look once)和 SSD^[17]等,此类算法模型小,计算速度快,对于小目标识别有较好的效果,更容易部署到智能手机和嵌入式平台等边缘计算设备,但是此类模型识别准确率低于第 1 类。受温室复杂的环境和温室番茄个体表征因素的影响,温室番茄目标检测任务需考虑检测精度和速度,而一步走算法可以在检测的过程中同时输出类别的目标框和检测精度,能更好地区分目标对象和背景,因此更加适应该任务场景。YOLO v3^[18]能以 3 倍的推理速度取得与 SSD 相近的检测结果,且不需要人工设置各种参数阈值,鲁棒性高。而 YOLO v5 作

为 YOLO 系列新的网络模型,相对于 YOLO v3 和 YOLO v4^[19]在检测速度和检测精度上都有相对的提升,但为实现在温室环境中番茄目标独特性检测,需对 YOLO v5 的网络结构进行相应的改进。

本文提出一种改进的 YOLO v5 目标检测算法,通过引入卷积注意力模块(Convolutional block attention module, CBAM)^[20],提取番茄目标检测的轮廓特征。通过改进损失函数,使得预测框位置更加精准,提高目标检测精度。根据温室番茄场景的复杂性,拍摄不同天气、不同光照角度和不同拍摄角度的番茄图像,确保数据集的丰富性和多样性;通过对数据样本进行预处理和图像增强,提高数据集质量的同时更加适用于温室番茄的果实目标检测。部署改进的 YOLO v5 网络模型到手机端,以期为基于移动边缘计算的设施环境机器人作业提供技术支撑。

1 数据样本采集和预处理

1.1 数据样本采集

温室番茄果实图像采集于北京市通州区农业机械研究所京鹏智慧温室,番茄种植方式为吊蔓式栽培。为增加样本多样性,在农艺专家指导下,分别于 2021 年 12 月 23 日中午(阴天)和 2022 年 1 月 9 日下午(晴天)采集不同成熟度(绿熟期和坚熟期)、不同拍摄角度(俯拍、仰拍和平拍)和不同光照角度(侧光、逆光和顺光)的番茄图像,如图 1 所示。采集可见光图像分辨率为 4 928 像素×3 264 像素,颜色表示为 sRGB,图像保存格式为 jpg。采用 Nikon D7000 型相机和手机进行拍摄,共采集图像 1 000 幅,通过图像筛选选择其中 752 幅图像作为训练样本。通过将番茄的最小外接矩形框作为真实框(Ground truth),提高真实框内数据参数准确率。

1.2 数据样本预处理和图像增强

在复杂情况下的温室番茄图像中,不同天气和不同角度下光照情况也不同,进而导致采集到的番茄果实图像颜色差异很大。另外,温室番茄果实生长的姿态各异、重叠遮挡严重等情况导致果实的形状特征难以提取。为了获取准确的数据参数,将筛选分类后的番茄图像通过标注软件 LabelImg 对其数据集进行人工标注。直方图均衡化是一种增强图像对比度的方法,本文采用自适应直方图均衡化和随机调整图像饱和度(-5%~5%)的方法来减小光照强度对图像质量的影响,同时也会增加番茄与番茄、番茄与背景的对比度,提高数据集质量;通过缩放、水平/垂直翻转或旋转±90°等方法,解决番茄果实生长姿态不同的问题,改善样本集数据不均衡。

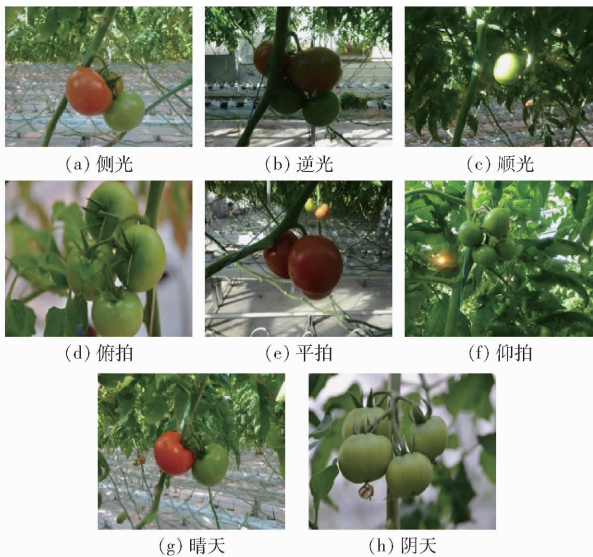


图1 复杂情况下温室番茄图像

Fig.1 Tomato images in complex environments in greenhouse

此外,本文还通过给采集到的图像添加噪声,降低拍摄过程中的噪声影响,避免深度学习模型在训练过程中出现过拟合现象,通过将图像尺寸按一定比例缩放,降低检测过程中距离的影响。通过以上方法,最终扩增样本集图像至4 101幅,其中70%用于训练集、10%用于验证集、20%用于测试集。

2 理论基础

2.1 YOLO v5 算法

YOLO v5 的网络结构主要由 Input、Backbone、Neck 和 Prediction 组成。Input 为温室番茄图像的输入环节,通过 Mosaic 数据增强、自适应锚框计算与自适应图像缩放等方法,对输入的温室番茄图像进行预处理,将原始温室番茄图像统一缩放到一个标准尺寸。Backbone 结构由 Focus 模块、CBL 模块、CSP 模块和 SPP 模块组成,主要实现温室番茄图像的特征提取。CSP 模块可以使网络深度加深,特征

提取能力大幅增强,Focus 模块主要为图像切片操作,能够进一步提取温室番茄果实目标特征。Neck 结构主要由 FPN + PAN 模块构成,其功能为增强网络特征学习能力,提升网络对语义信息和定位信息的学习性能,实现温室番茄的特征融合。在 Prediction 结构中,通过3个检测头分别对原图像进行8倍、16倍和32倍下采样,生成3个不同大小的特征向量,实现温室番茄果实目标预测。

2.2 注意力机制模块

由于番茄图像的枝叶背景和绿色果实的相似程度极高,从而使得 YOLO v5 的网络模型中骨干网络 (Backbone) 的隐含层易被网络忽视小目标的特征信息,而 CBAM 注意力机制模块对这部分特征信息进行注意力重构,给予绿色番茄果实目标信息更多的注意力。CBAM 是一种结合空间和通道的卷积注意力机制模块,将给定的空间坐标图作为输入,CBAM 模块会沿着两个独立的维度(空间和通道)依次推断注意力图,然后将注意力图和输入特征图相乘进行自适应特征优化,并且可很好地和 YOLO v5 网络算法进行融合,使提取到的高层特征更加全面更加丰富,更加适合于温室番茄特定场景。CBAM 结构图如图2所示。

通过引入通道注意力机制模块,可以有效检测目标轮廓特征,获得更多目标检测的内容,其计算方法为

$$M_c(F) = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad (1)$$

式中 $M_c(F)$ ——通道注意力输出权值

σ ——激活函数 sigmoid

F_{avg}^c ——平均池化后在空间上的特征映射

F_{max}^c ——最大池化后在空间上的特征映射

W_0 ——第1个全连接层的权重矩阵

W_1 ——第2个全连接层的权重矩阵

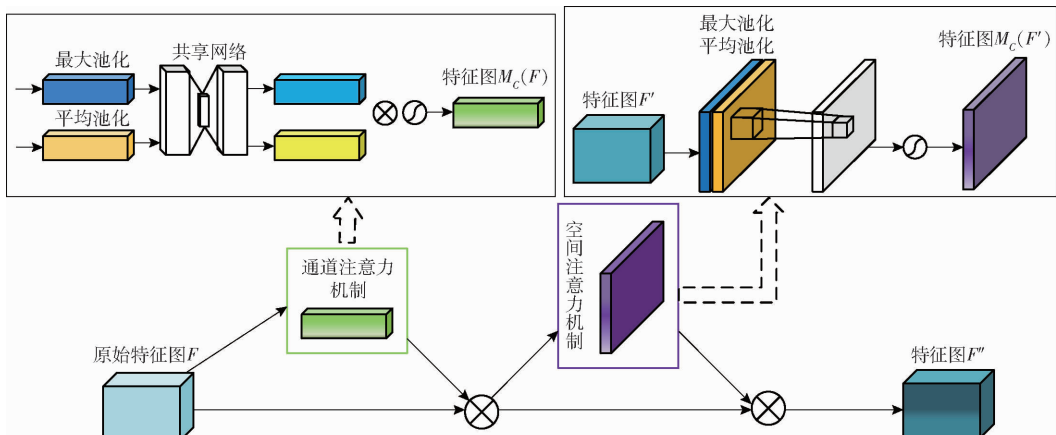


图2 CBAM 模块结构图

Fig.2 Structure diagram of CBAM module

通过引入空间注意力机制模块,可以有效定位检测目标的位置,提高目标检测的检测率,其计算方法为

$$M_s(F) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (2)$$

式中 $M_s(F)$ ——空间注意力输出权值

$f^{7 \times 7}$ ——尺寸为 7×7 卷积运算滤波器

F_{avg}^s ——在通道上平均池化后特征映射

F_{max}^s ——在通道上最大池化后特征映射

总之,输入特征图 F 首先与通道注意力模块逐元素相乘,得到的特征结果再和空间注意力机制模块点乘,经过 CBAM 处理之后得到最终的特征图 F'' 。其计算方法为

$$F' = M_c(F) \times F \quad (3)$$

$$F'' = M_s(F) \times F' \quad (4)$$

式中 F ——输入特征图

F' ——经过通道注意力加权得到的特征图

F'' ——经过空间注意力加权得到的特征图

2.3 损失函数改进

YOLO v5 的原始模型中采用 GIoU Loss^[21],其使用 Bounding box(目标位置)的损失函数,使用二进制交叉熵和 Logits 损失函数计算类概率和目标得分的损失。但是,GIoU Loss 存在的问题是,当预测框在目标框的内部时,如图 3 所示,GIoU 的值相同,其预测框位置不同于目标框,损失函数收敛慢,结合温室番茄复杂的环境,无法区分相对位置关系。因此这种方法无法精确地确定效果最佳的预测框。

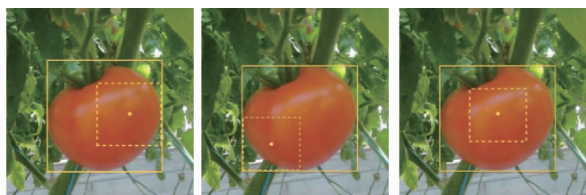


图3 预测框位于目标框内部

Fig. 3 Prediction box inside target box

为了解决上述问题,本文引入 CIoU Loss^[22] 作为预测框偏差的偏差指标,其计算公式为

$$CIoU = IoU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \quad (5)$$

$$L_{CIoU} = 1 - CIoU \quad (6)$$

$$\text{其中 } v = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h} \right)^2 \quad (7)$$

$$\alpha = \frac{v}{1 - IoU + v} \quad (8)$$

式中 b, b^{gt} ——预测框和目标框中心点

ρ —— b 和 b^{gt} 欧氏距离

c ——同时覆盖预测框和目标框的最小矩形对角线距离

α ——权重

v ——预测框和目标框对角线倾斜角的方差

L_{CIoU} ——CIoU 损失值

IoU ——预测框与目标框的交并比

ω^{gt}, h^{gt} ——预测框宽度和高度

ω, h ——目标框宽度和高度

2.4 模型构建

YOLO v5 网络包括 YOLO v5s、YOLO v5m、YOLO v5l、YOLO v5x 和 YOLO v5p 共 5 个版本,不同版本之间的区别在于模型深度,随着模型深度的增加,模型进行特征提取的操作次数也逐步增加,模型参数也随之增多,从而使得计算量变多,精度会相应地提升,但是提取图像信息的时间增加,会导致检测时间增加。本文研究对象为绿色和红色的番茄,识别目标仅为两种,检测目标为番茄的坐标信息而不是个别类的信息,并且其个体在图像中占用的像素比适中,因此本研究采用目标检测网络模型深度较浅的 YOLO v5s 网络模型,该模型小、检测速度快,且容易部署在嵌入式设备中。

YOLO v5 模型的主干网络(Backbone)包括 Darknet-53 主干网络、路径聚合网络^[23](Path aggregation network, PANet),可以实现对温室番茄图像特征的提取,在 CSP1_X 结构和 CSP2_1 结构后添加 CBAM 注意力机制模块(图中浅绿色虚线框),使提取到的目标特征更加全面更加丰富,改进后的 YOLO v5 模型(CB-YOLO)如图 4 所示。

3 检测网络结果与分析

3.1 试验平台

本试验软件环境为基于 Windows 10、Python 3.7.2、PyTorch 1.10.1 和 Cuda 11.3 搭建的深度学习框架,试验硬件环境(CPU)为 Intel I9-12900KF,显卡为 NVIDIA GeForce RTX 3090 24G。最大学习率选取 0.01,采用余弦退火动态调整学习率可以较快达到局部收敛效果,批量大小为 16 时,训练速度较快。相关模型参数如表 1 所示。

3.2 模型评估指标

为了验证模型有效性,从定性和定量两方面进行评估。对于定性评价,通过对比 CB-YOLO 模型和其他模型的检测图像差异来评估模型性能,即比较目标框的定位精确度,以及是否存在漏检、误检情况。定量评价方面,选取指标:准确率 P (Precision)、召回率 R (Recall)、红色和绿色番茄果实的检测精度 A_{PR} 和 A_{PG} 以及平均精度均值 M_{AP} (Mean average precision)。

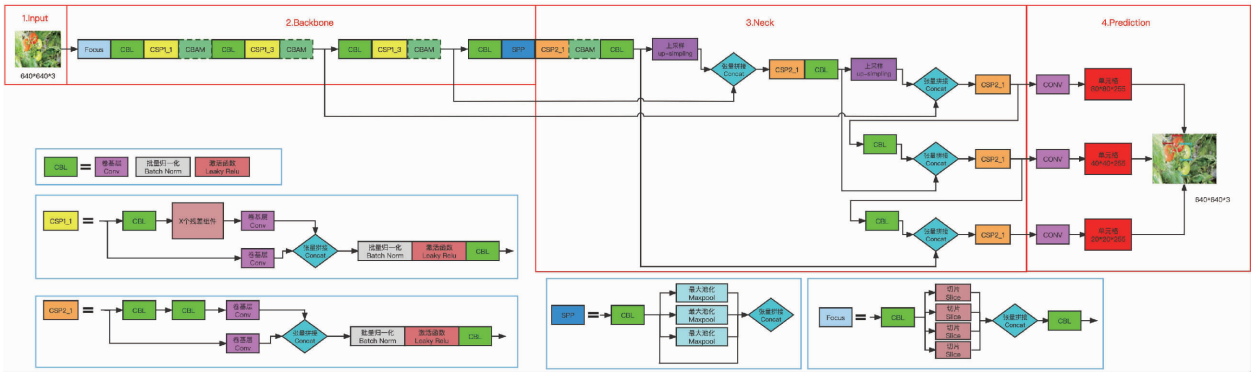


图4 改进后 YOLO v5 主要结构

Fig.4 Main structure of improved YOLO v5

表 1 试验相关模型参数

Tab.1 Test-related model parameters

参数	数值/类型
图像尺寸/(像素×像素)	224×224, 416×416, 640×640、800×800、960×960
优化器	Adam
批量大小	16
最大学习率	0.01
迭代次数	200
学习率调整方式	余弦退火

表 3 输入不同尺寸图像的绿色果实训练结果

Tab.3 Training results of green fruit inputting images of different sizes

输入图像尺寸/ (像素×像素)	GPU 参数量	准确率 P/%	召回率 R/%	平均精度均 值 $M_{AP}/%$
224×224	5.80×10^8	96.51	93.89	98.33
416×416	1.83×10^9	97.61	94.42	99.03
640×640	4.52×10^9	98.88	98.21	99.52
800×800	6.46×10^9	98.81	97.06	99.34
960×960	9.47×10^9	98.20	99.50	99.50

3.3 不同尺寸图像输入对 YOLO v5 模型的影响

不同尺寸图像输入对红色和绿色果实的训练结果如表 2、3 所示。由表 2、3 可得,输入图像尺寸越大,所需的 GPU 参数量就越大,使得推理训练所需的时间越长。在输入图像尺寸为 800 像素×800 像素时,红色果实准确率最高,为 97.85%,平均精度均值最高,为 99.34%。在输入图像尺寸为 640 像素×640 像素时,绿色果实准确率最高,为 98.88%,平均精度均值最高,为 99.52%。但在召回率方面,在图像输入尺寸为 224 像素×224 像素和 416 像素×416 像素情况下,红色果实和绿色果实的召回率相对其他两种相差很大,原因为该图像尺寸输入到网络模型的深度太浅,不利于目标特征的提取,导致目标特征的召回率低,因此需要更深的网络进行训练识别。红色果实和绿色果实的平均精度均值几乎一致。综合上述分析,在输入图像尺寸为 640 像素×640 像素

表 2 输入不同尺寸图像的绿色果实训练结果

Tab.2 Training results of red fruit inputting images of different sizes

输入图像尺寸/ (像素×像素)	GPU 参数量	准确率 P/%	召回率 R/%	平均精度均 值 $M_{AP}/%$
224×224	5.80×10^8	96.90	96.59	98.71
416×416	1.83×10^9	97.32	98.78	98.88
640×640	4.52×10^9	97.84	99.36	99.13
800×800	6.27×10^9	97.85	99.48	99.34
960×960	9.47×10^9	94.12	99.71	99.06

和 800 像素×800 像素时,识别效果较好,但是图像输入尺寸为 800 像素×800 像素情况下占用 GPU 参数量大,因此,在输入图像尺寸为 640 像素×640 像素时效率更高,更加适合。

3.4 改进后损失函数

CB-YOLO 算法检测结果如图 5 所示,前 20 轮迭代周期中网络快速拟合,损失值明显减小,后 180 迭代周期损失值基本平稳,表明模型训练效果不再随着迭代次数的增加而上升,表明该 CB-YOLO 网络模型损失函数收敛。

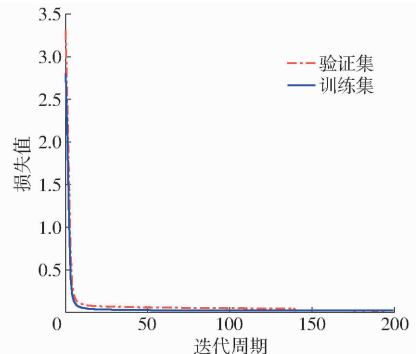


图5 损失值曲线

Fig.5 Loss curves

3.5 消融试验

为了验证改进后 YOLO v5 模型的合理性,开展消融试验,通过在 YOLO v5 原模型的基础上增加 CBAM 注意力机制和 CIOU 损失函数,来验证每

个改进点是否有效,消融试验结果如表 4 所示,其中“√”表示采用此项,“-”表示未采用此项。

表 4 YOLO v5 消融试验结果

Tab.4 YOLO v5 ablation experiment result

CBAM	CIOU	准确率	召回率	平均精度均值
		$P/\%$	$R/\%$	$M_{AP}/\%$
-	-	98.36	96.41	99.33
√	-	99.31	99.50	99.51
-	√	98.83	97.03	99.33
√	√	99.80	99.50	99.53

从表 4 可以看出,单独引入 CBAM 注意力机制,准确率提升 0.95 个百分点,召回率提升 3.09 个百分点,平均精度均值提升 0.18 个百分点,但是单独引入 CIOU 损失函数时,准确率、召回率和平均精度均值提升较小,分析认为 CBAM 注意力机制提高了骨干网络中的隐含层中被网络忽视特征信息的提取能力,使得效果提升明显,而 CIOU 损失函数使得预测框更加快速地收敛,并且得到更高的回归定位精度,因此效果提升幅度低。

为了更直观体现 CB-YOLO 模型的性能,本试验随机抽取 5 幅图像,将本文算法与 YOLO v4-tiny、YOLO v5 和 Faster R-CNN 3 种算法进行对比,结果如图 6 所示,在图像 1 中,在检测果实遮挡果实情况下,Faster R-CNN 存在漏检情况,YOLO v5、YOLO v4-tiny 和 Faster R-CNN 均可以有效识别小的目标,在图像 2 中,Faster R-CNN 在果实遮挡果实情况下存在漏检情况,在图像 3 中,YOLO v4-tiny 在枝干遮挡果实情况下存在漏检的情况,Faster R-CNN 可以有效识别小的目标,可以明显看出 YOLO v5 的检测效果优于 YOLO v4-tiny 和 Faster R-CNN;在图像 2 中,CB-YOLO 成功检测出图像 1 中小目标绿色番茄,并且在图像 3 和图像 5 等强曝光情况下,不但检测精度较高,也可以在图像 4 枝干遮挡果实情况下快速识别番茄,提升了小目标和被遮挡情况的检测率,鲁棒性高。

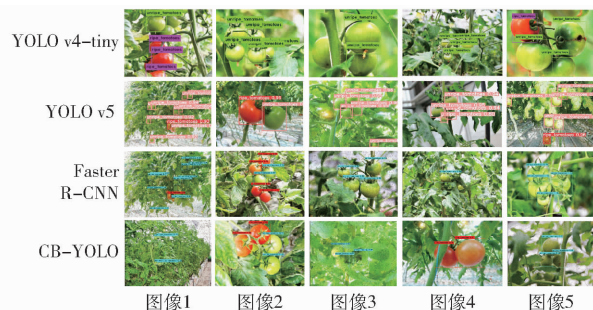


图 6 不同算法结果对比

Fig.6 Comparison of different algorithm results

从表 5 可以看出,CB-YOLO 模型对于红色番茄检测精度、绿色番茄检测精度和平均精度均值最高,分别为 99.88%、99.18% 和 99.53%,YOLO v5 的检测速率最高,达到 131.58 f/s,模型占用内存量最小,仅为 14.8 MB。相对于 YOLO v5 模型,CB-YOLO 模型虽然检测速度相对下降,但是仍然高于 Faster R-CNN 和 YOLO v4-tiny,并且相对于 YOLO v5 原模型,红色番茄检测精度提高 0.55 个百分点,绿色番茄检测精度提高 0.3 个百分点,平均精度均值提升 0.2 个百分点。

表 5 4 种模型各项检测指标

Tab.5 Results of all test indexes of four models

检测模型	红色番茄检测精度	绿色番茄检测精度	平均精度均值	检测速率	模型占用内存
	$A_{PR}/\%$	$A_{PU}/\%$	$M_{AP}/\%$	($f \cdot s^{-1}$)	量/MB
CB-YOLO	99.88	99.18	99.53	113.90	15.40
YOLO v5	99.33	98.88	99.33	131.58	14.80
Faster R-CNN	97.67	97.59	97.63	25.99	195.80
YOLO v4-tiny	98.89	96.07	97.48	80.83	26.22

3.6 CB-YOLO 模型部署

为了探索研究智能手机作为边缘计算终端设备在农业机器人目标识别作业中的可行性和实用性,本文将 CB-YOLO 模型部署到不同型号的安卓手机端,验证其部署模型在自然情况下性能检测的稳定性。CB-YOLO 模型部署过程如下:首先,通过 ONNX(Open Neural Network Exchange)模型从训练好的 CB-YOLO 模型中 best.pt 权重文件生成 best.onnx、best.mlmodel 和 best.torchscript.pt 文件,然后对 ONNX 进行精简处理,其次通过腾讯 NCNN 框架将 ONNX 格式文件转换为 param 和 bin 格式文件,由于 NCNN 不支持切片(Slice)操作,因此将通过 Netron 可视化编译工具将 param 文件中的切片操作去除,最后,通过 Android Studio 工具将重新编译的文件写入 Android 系统,生成 YOLO v5s.apk 文件,通过手机下载安装 YOLO v5s.apk 安卓安装包,完成 CB-YOLO 安卓端的部署。

完成 CB-YOLO 模型安卓端部署后,对采集到的自然情况下的温室番茄图像用不同型号、不同运行内存和不同系统进行验证,效果如图 7 所示。从图 7 可以看出,CB-YOLO 模型部署到手机检测精度高,鲁棒性强。

为了更充分验证 CB-YOLO 模型部署到安卓手机端的番茄果实检测精度,通过随机选取 20 幅自然环境下的温室番茄图像,对不同型号、不同运行内存和不同系统进行统计对比,结果如表 6 所示。从

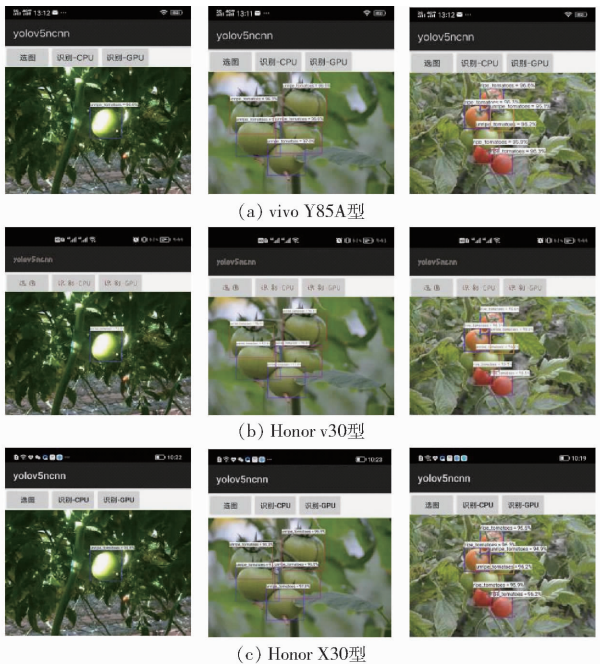


图7 不同型号手机对温室番茄果实的检测效果
Fig.7 Detection effects of different types of mobile phones on greenhouse tomato fruit

表6可以看出,不同类型手机对红色番茄精度、绿色番茄的精度和总精度相差在1%以内,证明CB-YOLO模型部署稳定性。

4 结论

(1)提出了一种改进的YOLO v5s目标检测算

表6 不同型号手机测试指标

Tab.6 Test indicators of different mobile phone models

手机型号	运行内存/GB	运行环境	红色番茄精度/%	绿色番茄精度/%	总精度/%
vivo Y85A	4	Android 8.0	94.8	95.3	95.1
Honor v30	8	Android 11.0	95.5	95.7	95.6
Honor X30	8	HarmonyOS 2.0.0	94.7	94.9	94.8

法(CB-YOLO)。在骨干网络中加入卷积注意力模块(CBAM),对目标特征给予更多的关注;通过将CIoU Loss替换GIoU Loss作为算法的损失函数,在提高边界框回归速率的同时提高定位精度。CB-YOLO网络模型对温室环境下番茄红色果实检测精度、绿色果实检测精度和平均识别精度分别为99.88%、99.18%和99.53%,对比原模型检测精度和平均精度均值均有提升。

(2)通过对比试验探究了在输入不同图像尺寸下对模型的影响,对图像输入尺寸为224像素×224像素、416像素×416像素、640像素×640像素、800像素×800像素、960像素×960像素进行试验,得出图像输入尺寸为640像素×640像素效果更优,鲁棒性更强。

(3)将CB-YOLO模型部署到安卓手机端,通过不同手机型号测试,验证了部署模型在边缘计算终端设备上目标识别的稳定性,为基于移动边缘计算的设施环境机器人作业提供技术支持。

参考文献

- 霍建勇. 中国番茄产业现状及安全防范[J]. 蔬菜,2016(6):1-4.
HUO Jianyong. Status and safety prevention of tomato industry in China[J]. Vegetable,2016(6):1-4. (in Chinese)
- 李天华,孙萌,丁小明,等. 基于YOLO v4+HSV的成熟期番茄识别方法[J]. 农业工程学报,2021,37(21):183-190.
LI Tianhua,SUN Meng,DING Xiaoming,et al. Tomato recognition method at the ripening stage based on YOLO v4 and HSV[J]. Transactions of the CSAE, 2021,37(21):183-190. (in Chinese)
- 王丽丽. 番茄采摘机器人关键技术研究[D]. 北京:北京工业大学,2017.
- 梁喜凤,章艳. 串番茄采摘点的识别方法[J]. 中国农机化学报,2016,37(11):131-134,149.
LIANG Xifeng,ZHANG Yan. Recognition method of picking point for tomato cluster[J]. Journal of Chinese Agricultural Mechanization,2016,37(11):131-134,149. (in Chinese)
- 冯青春,程伟,杨庆华,等. 基于线结构光视觉的番茄重叠果实识别定位方法研究[J]. 中国农业大学学报,2015,20(4):100-106.
FENG Qingchun,CHENG Wei,YANG Qinghua,et al. Identification and localization of overlapping tomatoes based on linear structured light vision system[J]. Journal of China Agricultural University,2015,20(4):100-106. (in Chinese)
- 马翠花,张学平,李育涛,等. 基于显著性检测与改进Hough变换方法识别未成熟番茄[J]. 农业工程学报,2016,32(14):219-226.
MA Cuihua,ZHANG Xueping,LI Yutao,et al. Identification of immature tomatoes base on salient region detection and improved Hough transform method[J]. Transactions of the CSAE,2016,32(14):219-226. (in Chinese)
- 周云成,许童羽,郑伟,等. 基于深度卷积神经网络的番茄主要器官分类识别方法[J]. 农业工程学报,2017,33(15):219-226.
ZHOU Yuncheng,XU Tongyu,ZHENG Wei,et al. Classification and recognition approaches of tomato main organs based on DCNN[J]. Transactions of the CSAE,2017,33(15):219-226. (in Chinese)
- LIU Z, WU J, FU L, et al. Improved kiwifruit detection using pre-trained VGG16 with RGB and NIR information fusion[J]. IEEE Access, 2020,8:2327-2336.

- [9] 王文杰,贡亮,汪韬,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, WANG Tao, 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)
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014:580-587.
- [11] LI J, LIANG X, SHEN S M, et al. Scale-aware Fast R-CNN for pedestrian detection[C]//IEEE Transactions on Multimedia, 2015.
- [12] 孙龙清,孙希蓓,吴雨寒,等.基于DRN-Faster R-CNN的复杂背景多目标鱼体检测模型[J].农业机械学报,2021,52(增刊):245-251,315.
SUN Longqing, SUN Xibei, WU Yuhan, et al. Multi-target fish detection model based on DRN-Faster R-CNN in complex background[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.):245-251,315. (in Chinese)
- [13] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2):386-397.
- [14] 刘芳,刘玉坤,林森,等.基于改进型YOLO的复杂环境下蕃茄果实快速识别方法[J].农业机械学报,2020,51(6):229-237.
LIU Fang, LIU Yukun, LIN Sen, et al. Fast recognition method for tomatoes under complex environments based on improved YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6):229-237. (in Chinese)
- [15] 穆龙涛,高宗斌,崔永杰,等.基于改进AlexNet的广域复杂环境下遮挡猕猴桃目标识别[J].农业机械学报,2019,50(10):24-34.
MU Longtao, GAO Zongbin, CUI Yongjie, et al. Kiwifruit detection of far-view and occluded fruit based on improved AlexNet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(10):24-34. (in Chinese)
- [16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:779-788.
- [17] LIU Wei, DRAGOMIR A, DUMITRU E, et al. SSD: single shot multibox detector[C]//ECCV, 2016:21-37.
- [18] REDMON J, FARHADI A. YOLO v3: an incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.
- [19] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLO v4: optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2004.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018:3-19.
- [21] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019:658-666.
- [22] XIAO Liyun, ZHOU Peng, XU Ke, et al. Multi-directional scene text detection based on improved YOLO v3[J]. Sensors, 2021, 21(14):4870.
- [23] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018:8759-8768.