

嵌入式设备的轻量化百香果检测模型

罗志聪^{1,2} 李鹏博¹ 宋飞宇¹ 孙奇燕³ 丁昊凡¹

(1. 福建农林大学机电工程学院, 福州 350002; 2. 福建省农业信息感知技术重点实验室, 福州 350002;
3. 福建农林大学计算机与信息学院, 福州 350002)

摘要: 为在有限的嵌入式设备资源下达到实时检测要求,提出一种基于改进 YOLO v5 的百香果轻量化检测模型 (MbECA-v5)。首先,使用 MobileNetV3 替换主干特征提取网络,利用深度可分离卷积代替传统卷积减少模型的参数量。其次,嵌入有效通道注意力网络 (ECANet) 关注百香果整体,引入逐点卷积连接特征提取网络和特征融合网络,提高网络对百香果图像的特征提取能力和拟合能力。最后,运用跨域与域内多轮训练相结合的迁移学习策略提高网络检测精度。试验结果表明,改进后模型的精确率和召回率为 95.3% 和 88.1%;平均精度均值为 88.3%,较改进前提高 0.2 个百分点。模型计算量为 6.6 GFLOPs,体积仅为 6.41 MB,约为改进前模型的 1/2,在嵌入式设备实时检测速度为 10.92 f/s,约为 SSD、Faster RCNN、YOLO v5s 模型的 14 倍、39 倍、1.7 倍。因此,基于改进 YOLO v5 的轻量化模型提高了检测精度和大幅降低了计算量和模型体积,在嵌入式设备上能够高效实时地对复杂果园环境中的百香果进行检测。

关键词: 百香果; 嵌入式设备; YOLO v5; 轻量化; 迁移学习

中图分类号: TP391.4; S24 **文献标识码:** A **文章编号:** 1000-1298(2022)11-0262-08

OSID:



Lightweight Passion Fruit Detection Model Based on Embedded Device

LUO Zhicong^{1,2} LI Pengbo¹ SONG Feiyu¹ SUN Qiyao³ DING Haofan¹

(1. College of Mechanical and Electronic Engineering, Fujian Agriculture and Forestry University, Fuzhou 350002, China
2. Fujian Key Laboratory of Agricultural Information Sensing Technology, Fuzhou 350002, China
3. College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

Abstract: In order to meet the real-time detection requirements under the limited resources of embedded devices, a passion fruit detection model based on improved YOLO v5 lightweight network (MbECA-v5) was proposed. Firstly, MobileNetV3 was used to replace the feature extraction network, the depth separable convolution was used to replace the traditional convolution to reduce the number of model parameters. Secondly, the effective channel attention network (ECANet) was embedded to focus on the whole passion fruit. Point-by-point convolution connection feature extraction network and feature fusion network were introduced to improve the feature extraction ability and fitting ability of the network for passion fruit images. Finally, the transfer learning strategy combined with cross-domain and within-domain multi-training was used to improve the network detection accuracy. Experimental results showed that the accuracy and recall of the improved model were 95.3% and 88.1%, respectively. The mAP value of 88.3%, compared with the model before the improvement, it was increased by 0.2 percentage points. And the number of calculations was 6.6 GFLOPs. The model volume was only 6.41 MB, which was about half of the improved model. The real-time detection speed in embedded device was 10.92 f/s, the detection speed in embedded device was about 14 times, 39 times and 1.7 times of SSD, Faster RCNN and YOLO v5s. Therefore, the lightweight model based on improved YOLO v5 greatly reduced the amount of calculation and model volume, and it can detect passion fruit in complex orchard environment efficiently on embedded devices, which was of great significance to improve the intelligent level of orchard production.

Key words: passion fruit; embedded device; YOLO v5; lightweight; migration study

0 引言

随着农业智能化和农业数字革命的快速发展,利用计算机视觉进行果实识别和产量预估等精细化管理成为发展智慧农业的研究热点^[1-2]。百香果是福建富民产业的“福果”,百香果产业不仅有高效的经济效益,还有助于推动福建农业的供给侧结构改革^[3]。目前,百香果的收获仍局限于人工采摘,具有成本高、作业效率低、劳动强度大和风险高等缺点^[4]。为了机械化采摘,实现百香果目标精准检测与定位尤为重要。

针对复杂环境下的百香果果实识别与定位,传统的方法效果不佳,而深度学习网络^[5-7]可以自动提取特征,将简单的特征融合为更复杂的特征,通过复杂特征解决识别定位问题,在目标检测方面精度和效率都得到了大幅提升。TU 等^[8]通过改进 RCNN 提出多尺度快速的区域卷积神经网络 (MS-FRCNN)对叶片严重遮挡的百香果进行检测和计数。该方法识别能力较强,但网络结构复杂、训练速度过慢。唐熔钗等^[9]通过在 YOLO v3 模型中加入密集网络 DenseNet 来提高模型对自然环境下百香果的检测精度,但加深网络结构使得计算损耗增加。WU 等^[10]在 YOLO v3 网络基础上增强特征融合网络和减少多尺度网络检测自然环境中的百香果,但对百香果的识别和定位具有局限性。TU 等^[11]提出的基于 RGB-D 检测器与深度学习网络结合的模型,在检测百香果成熟度方面具有较高的检测精度,但该网络只存在理论研究,并没有部

署到嵌入式设备。虽然深度学习网络的优化和改进在百香果的检测和定位方面取得优异效果,但同时巨大参数量会导致硬件存储成本和计算成本过高^[12],限制了在精准农业化管理中应用和推广。寻找一种精度高、体积小的深度学习算法应用于嵌入式设备尤为重要。

目前学者在深度学习网络轻量化方面做了多种尝试,如设计轻量化网络^[13]、网络剪枝^[14]、知识蒸馏^[15]、模型压缩^[16]等。轻量化卷积神经网络通过减少计算量和参数量设计出更加高效的特征提取网络结构,在减少参数的同时几乎不影响网络模型性能。如 MobileNet^[17]、ShuffleNet^[18]、GhostNet^[19]。

由于 YOLO v5 模型较大,在嵌入式平台部署后实时检测百香果的性能低。本文基于改进 YOLO v5 模型,以复杂环境下绿色百香果为研究对象,在确保检测精度较高的情况下,使得模型轻量化且可部署于嵌入式平台,节约边缘计算设备资源从而达到实时检测要求。

1 数据集建立

本文图像采集地点为福建省漳州市诏安县福果农业综合开发有限公司的黄金百香果基地,采集时间为 2021 年 6 月底至 8 月初。选用尼康数码相机拍摄,拍摄时相机距离百香果为 80 ~ 100 cm,图像尺寸为 1 920 像素×1 080 像素。分别采集了百香果顺光和逆光时独立果实、多个果实、叶片遮挡和果实重叠下共 1 000 幅图像,复杂环境下的百香果图像如图 1 所示。

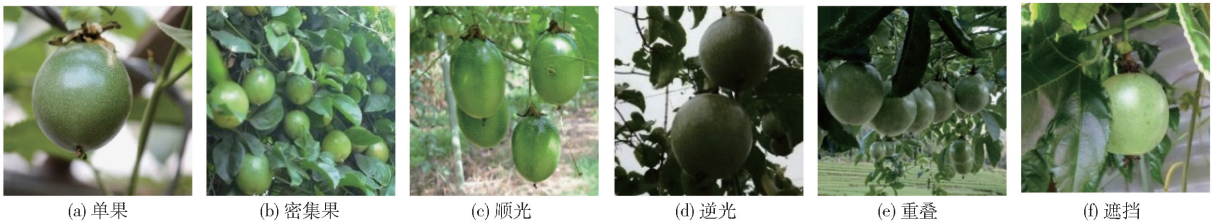


图 1 复杂环境下百香果图像

Fig. 1 Passion fruit images in complex environment

为了降低因训练样本多样性不足而产生的模型过拟合现象,采用对原始图像进行旋转、镜像、加入高斯噪声和改变对比度等处理扩充数据集,增广的百香果图像如图 2 所示。为了与原图有所区别,高斯噪声平均值和标准方差分别设置为 3 和 40。对比度调节是将原图像 R、G、B 3 通道的强度值[0.4, 1]映射到新图像[0,1]中。最终获得 2 000 幅复杂环境下百香果图像作为数据集并保存为 .jpg 格式。使用人工标注工具 LabelImg 对百香果目标进行标注,获取准确的百香果位置数据。为了减少复杂背

景像素对检测结果的影响,标注时将百香果的最小外接矩形框作为真实框 (Ground truth),最终得到 VOC 格式的 xml 文件。文件中包含百香果名称、尺寸和位置。对以上标注好的数据集划分训练集和测试集,并按照比例 8 : 2 进行随机分配。训练集和测试集分别为 1 600 幅和 400 幅。

2 百香果检测模型改进与部署

2.1 YOLO v5 模型构建

YOLO 系列算法^[20-21]是目前应用最为广泛的



图 2 百香果增广数据集

Fig.2 Passion fruit augmented dataset

一阶段目标算法之一,相比 SSD 算法,YOLO 算法提取物体特征更充分,检测精度高、速度快。相比 Fast-RCNN^[22]网络,YOLO v5 把检测问题作为回归问题处理,可以直接从图像中得到边界框坐标和所属类别的概率,实现端到端的检测性能优化。YOLO v5 网络结构可以分为 4 部分:输入端(Input)、主干网络(Backbone)、颈网络(Neck)、检测网络(Detect)。主要结构如图 3 所示。

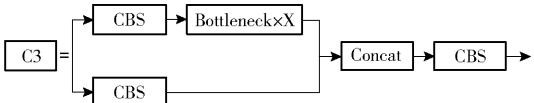


图 4 C3 网络结构

Fig.4 C3 network structure

效果。PAN 结构从下往上提取特征层信息,得到丰富的位置信息。最后将提取的特征层进行融合,使得 PANet 网络携带更强的语义特征,增强了定位能力。

检测网络以目标对象的概率类别、置信度得分、包围框位置作为结果输出。YOLO v5 主要计算 3 类损失函数:边框损失(Bounding box loss)、类间损失(Class loss)和目标损失(Object loss)。其中类间损失采用 Focal loss 损失函数^[23]。

2.2 改进的百香果检测模型

2.2.1 模型轻量化

本研究在特征提取网络层选取 MobileNetV3^[24]轻量化模型替换 YOLO v5 主干网络,MobileNetV3 结合了深度可分离卷积、线性瓶颈和倒置残差结构,为了学习使用全局信息选择性地强调具有信息性的特征,抑制不明显特征而引入了 SE 结构。MobileNetV3 特征提取网络采用一个 3×3 的标准卷积和多个倒置残差结构,不仅能够减少网络参数还能提取百香果的丰富特征信息。文中主干网络选取的 MobileNetV3 结构如表 1 所示。

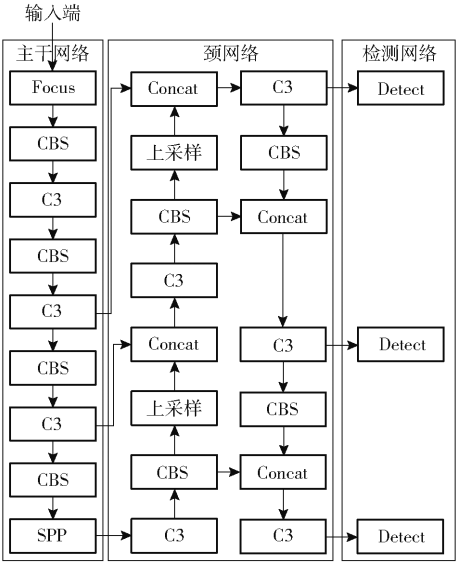


图 3 YOLO v5 网络结构图

Fig.3 YOLO v5 network structure diagram

输入端由 Mosaic 数增强方式将百香果图像进行随机缩放、裁剪、排布后输入网络。主干网络主要实现百香果的特征提取,由 Focus 模块、C3 模块、标准卷积(CBS)模块、SPP 模块组成。Focus 模块对输入的百香果图像进行切片操作,通过扩张通道数来提取百香果目标特征。C3 模块在 CSP 结构上改进得到,由于网络优化中梯度信息重复,使计算量过高,去掉了残差输出后的批量归一化层和激活函数层,减少了运算量,如图 4 所示。CBS 模块包括二维卷积、批量标准化和激活函数(SiLU)。SPP 模块通过不同核尺寸提取百香果特征之后叠加进行特征融合。颈网络主要实现百香果的特征融合,以路径聚合网络(PANet)为主,分为 FPN 结构和 PAN 结构。FPN 结构从上往下提取特征层信息,提升了目标检测

表 1 主干网络中 MobileNetV3 结构

Tab.1 MobileNetV3 structure in the trunk network

层数	操作	注意力模块	非线性激活函数	步长
1	Conv2d,3×3		HS	2
2	Bneck,3×3	有	RE	2
3	Bneck,3×3		RE	2
4	Bneck,3×3		RE	1
5	Bneck,5×5	有	HS	2
6	Bneck,5×5	有	HS	1
7	Bneck,5×5	有	HS	1
8	Bneck,5×5	有	HS	1
9	Bneck,5×5	有	HS	1
10	Bneck,5×5	有	HS	2
11	Bneck,5×5	有	HS	1
12	Bneck,5×5	有	HS	1

在主干网络中由于使用 MobileNetV3 深度可分离卷积替代了 YOLO v5 传统卷积,在有效提取百香果特征的同时减少了参数量与计算量。假设 D_k 表示卷积核尺寸, M 表示输入特征图通道数, N 表示卷积核个数, D_w 和 D_h 表示输出特征图的长和宽。深度可分离卷积参数量 P_D 计算式为

$$P_D = D_k^2 M + MN \quad (1)$$

深度可分离卷积计算量 Q_D 计算式为

$$Q_D = D_k^2 M D_w D_h + M N D_w D_h \quad (2)$$

传统卷积参数量 P_N 计算式为

$$P_N = D_k^2 M N \quad (3)$$

传统卷积计算量 Q_N 计算式为

$$Q_N = D_k^2 M N D_w D_h \quad (4)$$

深度可分离卷积和传统卷积参数量比 R_p 、计算量比 R_q 计算式为

$$R_p = \frac{P_D}{P_N} = \frac{1}{N} + \frac{1}{D_k^2} \quad (5)$$

$$R_q = \frac{Q_D}{Q_N} = \frac{1}{N} + \frac{1}{D_k^2} \quad (6)$$

由式(5)、(6)可知,深度可分离卷积的参数量和计算量是传统卷积的 $\frac{1}{N} + \frac{1}{D_k^2}$ 倍。

2.2.2 有效通道注意力机制

深度学习中注意力机制用来强调百香果目标处理对象的重要信息,以选择出比当前任务目标更关键的信息。MobileNetV3 在颈部采用 SENet 轻量化结构,SENet 对每个通道独立使用全局平均池化操作,使其生成的特征层具有全局感受野;使用 2 个非线性的全连接层,使得网络提取丰富的特征信息。虽然在全连接层捕捉非线性通道交互采用降维方式控制模型复杂性,但是降维会使得通道注意力产生效率低和不必要的依赖关系,也会使所有通道产生消极影响。

本文采用 ECANet (Efficient channel attention net) [25] 替换 SENet 构建 MvECA 模块,如图 5 所示。

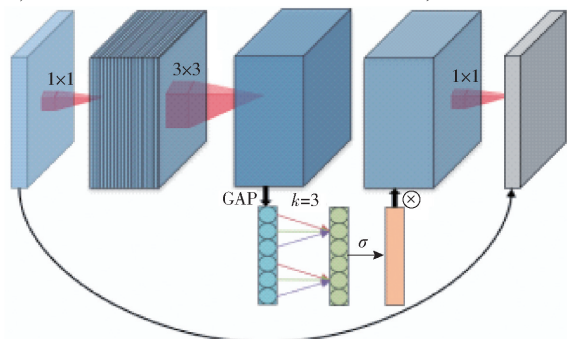


图 5 MvECA 结构

Fig. 5 MvECA structure

该模块避免了降维引起的检测效果不理想,同时采用交互覆盖率 K 的快速一维卷积有效地捕获了跨通道交互信息,更好地捕捉百香果图像重要特征,既提升模型效果又降低模型复杂度,该模块引入很少的附加参数,并且在局部有适当的跨通道交互,更有效的学习通道注意力,带来了显著的性能增益。其中交互覆盖率 K 与通道维数 C 之间的关系为

$$K = \psi(C) = \left\lfloor \frac{\text{lb}C}{2} + \frac{1}{2} \right\rfloor_{\text{odd}} \quad (7)$$

其中 $\left\lfloor \frac{\text{lb}C}{2} + \frac{1}{2} \right\rfloor_{\text{odd}}$ 表示最接近 $\frac{\text{lb}C}{2} + \frac{1}{2}$ 的奇数。为了诠释局部跨通道交互方法能够保证模型有效性,使用矩阵 $\mathbf{W}_k$ 表示学习通道注意力,包含 $k \times C$ 个参数,通道注意力可以表示为

$$\omega_i = \sigma \left(\sum_{j=1}^k w_i^j y_i^j \right) \quad (y_i^j \in \Omega_i^k) \quad (8)$$

式中 Ω_i^k —— y_i 的 k 个相邻通道的集合

y_i ——当前目标的聚合特征

w_i^j —— y_i^j 权重,与其相邻的 k 分通道共同作用产生

而所有通道共享相同的学习参数可以让模型性能进一步提升。共享后通道注意力计算式为

$$\omega_i = \sigma \left(\sum_{j=1}^k w^j y_i^j \right) \quad (y_i^j \in \Omega_i^k) \quad (9)$$

最后,通过 k 个参数的一维卷积快速实现有效获取局部跨通道的交互信息。

2.2.3 逐点卷积网络

通过轻量化网络 MvECA 将主干网络替换后,在与特征融合网络堆叠时通道出现不匹配的情况。针对这一情况,再次对主干网络进行修改。在主干网络的第 4 层和第 10 层增加逐点卷积网络 (Pointwise conv net),方便特征提取和特征融合网络在维度上更好连接,同时也会提高网络对百香果检测的平均精度,与标准卷积不同的是逐点卷积采用 1×1 的卷积核对上一层的特征层进行单点上的特征提取。该设计首先采用二维卷积提取百香果特征,接着归一化处理特征使其保留原始的表达能力,最后输入到非线性激活函数中。非线性激活函数使用 h_swish 函数

$$\text{h_swish}(x) = x \frac{\text{ReLU6}(x+3)}{6} \quad (10)$$

其中 ReLU6 函数在 16 位浮点数或者 8 位整型低精度时有较优的数值分辨率,易于量化部署,有利于在嵌入式移动平台中使用。

综上,最终改进后 MbECA-v5 结构如图 6 所示。

2.2.4 跨域与域内多轮训练迁移学习策略

对于网络模型参数的大量减少引起检测精度下

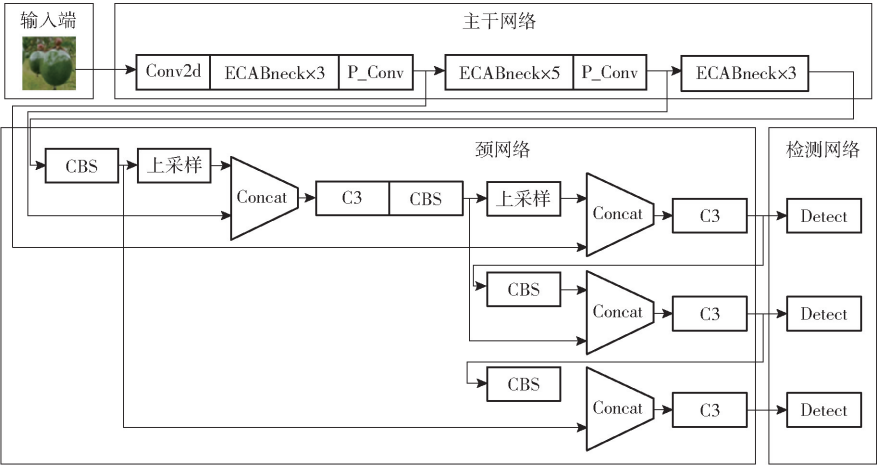


图 6 MbECA-v5 结构图

Fig. 6 Structure diagram of MbECA-v5

降问题,本文采用了跨域与域内多轮训练相结合的学习策略进行知识迁移。即利用跨域与域内相结合的迁移方式在原来训练 70 轮的基础上再训练 70 轮。虽然扩大复杂情况下百香果训练数据集会更容易学习到百香果特征,提高检测精度,但是由于受百香果成熟的季节性限制,大规模的数据集实地获取较为困难,时间成本高。通常情况下,域内和跨域多轮训练迁移方式也可以提高检测精度,而对于识别对象为复杂环境下的百香果,迁移前后的源域和目标域相似度有差异,精度提升有限。因此在源域与目标域之间添加仅含百香果特征的过渡域,以减少源域和目标域之间的差异,大幅提高模型的检测精度。

本文使用 VOC2007 数据集作为源域,原扩充的数据集为目标域,记作数据集 2。由于通用百香果数据集少,因此将扩充的百香果数据集中 500 幅图像随机挑选为过渡域,记为数据集 1。根据 OTOVIC 等^[26]提出预训练模型在不相关的学习任务上比随机初始化模型效果要好的观点,将 VOC2007 数据集训练模型的结果作为数据集 2 的预训练模型,使用 VOC2007 数据集网络模型特征参数学习复杂环境下百香果的特征,该策略称为跨域迁移学习;将数据集 1 训练模型结果作为数据集 2 的预训练模型,用部分百香果特征参数调整复杂环境的百香果特征参数,该策略称为域内迁移学习。将 VOC2007 数据集和数据集 1 共同训练模型结果作为数据集 2 的预训练模型,进一步学习复杂环境下的百香果特征,提升迁移学习的效果。该策略称为跨域与域内结合学习。

2.3 训练与嵌入式平台

模型训练平台硬件配置为 AMD Ryzen 7 5800H CPU @ 3.70 GHz, GeForce GTX 3060 Laptop GPU, 16 GB 运行内存,软件环境为 64 位 Windows 10 系统,采用 Pytorch 1.7 深度学习框架和 Python 3.7 语

言,配合 CUDA 11.0 进行训练。

模型验证和迁移部署设备采用英伟达公司 Jetson Nano, GPU 为 128core Maxwell, CPU 为四核 ARM, A57@1.43 GHz, 内存为 4 GB、64 位 LPDDR4。相应配置系统为 ARM 版 Ubuntu 18.04 操作系统,模型运行环境的配置为 Jetpack 4.5、Python 3.6、Pytorch 1.8、Torchvision 0.9、Cuda 10.2、Cudnn 8.0。Jetson Nano 测试图如图 7 所示。



图 7 Jetson Nano 测试图

Fig. 7 Jetson Nano test diagram

1. 运行界面 2. 图像检测界面 3. Jenson Nano

2.4 模型评价指标

模型评价指标为精确率 P 、召回率 R 、平均精度均值(mAP)和浮点数计算量。精确率 P 表示预测的正确样本占有所有样本的比例,召回率 R 表示预测的正确样本占实际正样本的比例,平均精度均值便于从宏观角度评价不同卷积神经网络模型对图像识别性能。浮点数计算量用于衡量模型计算复杂度。

嵌入式设备实时检测性能对在百香果园的应用非常关键,检测速度也就成为了目标检测算法的一个重要评价指标,用于评价模型部署在嵌入式平台的性能。对于速度这一指标本文使用实时帧率 FPS 表征。

3 结果与分析

3.1 模型训练结果与分析

试验采用 Adam 优化算法进行优化,图像尺寸

调整为 640 像素 × 640 像素,学习率 (Learning rate) 为 0.001,初始动量 (Momentum) 设置为 0.937,初始权重衰减系数 0.1,批尺寸 (Batch size) 设为 8,训练轮数 (Epoch) 为 70。测试时设置交并比 (IOU) 为 0.5,置信度为 0.6。使用算法在数据集上进行训练,训练结束后保存精度最高的和最后一次的模型参数。

训练过程的参数指标如图 8 所示。随着训练轮数增加,精确率和召回率均不断上升,训练前 15 轮精确率和召回率曲线上升速度最快,当训练达到 70 轮时,精确率和召回率曲线趋近于平缓,最终分别达到 94.2% 和 83.6%。mAP 曲线也随着训练次数的增加逐渐上升,最终复杂环境下百香果检测平均精度均值达到 92.5%,上述训练曲线趋势说明改进后的模型具有有效性和高识别性。

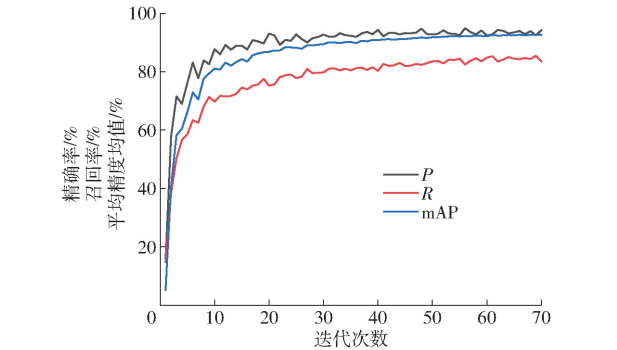


图 8 训练模型参数指标变化趋势

Fig. 8 Parameters of training model

3.2 消融试验

消融试验结果如表 2 所示。从表 2 可以看出,利用轻量化网络替换主干网络减少了参数量,使用注意力模块关注百香果整体,同时减小了网络模型。添加设计的逐点卷积拟合通道和提升精度,使用迁移学习策略调整初始化网络模型,大幅提升了模型的精确率、召回率和平均精度均值。

表 2 消融试验结果
Tab. 2 Results of ablation test

基线网络	轻量化模型	注意力模块	逐点卷积	迁移学习	精确率/%	召回率/%	平均精度均值/%	模型体积/MB
					96.3	87.6	88.1	14.40
	✓				97.2	77.9	78.2	7.00
YOLO v5s	✓	✓			95.7	81.7	81.7	6.12
	✓	✓	✓	✓	96.0	82.0	82.1	6.41
	✓	✓	✓	✓	95.3	88.1	88.3	6.41

注:“✓”表示改进 YOLO v5s 使用的方法。

由表 2 可知,在 YOLO v5s 网络中将原来主干网络替换为轻量化网络后精确率、召回率和平均精

度均值都有所下降,造成上述原因是改进的模型参数量和卷积层数减少,不利于复杂环境下百香果的特征提取,导致网络提取特征能力下降。为了弥补替换主干网络所带来的精度大幅下降,在网络中添加注意力机制和逐点卷积,可以看出,对召回率和平均精度均值都有明显提升。最后利用跨域与域内多训练相结合的迁移学习策略对模型初始化参数调整,减少前后源域与目标域的差异性,提升迁移训练效果。与基线网络相比,经过上述改进后的模型平均精度均值提高 0.2 个百分点,模型占用存储空间 (模型体积) 仅为 6.41 MB,减少为基线网络的 55.5%。

3.3 不同注意力机制对比试验

在网络轻量化模型的基础上施加不同的注意力机制,展现不同注意力机制对模型检测能力的影响,如表 3 所示。

表 3 不同注意力机制试验结果
Tab. 3 Results of different attention tests

注意力机制	交互覆盖率 K	平均精度均值/%	模型体积/MB
SE		78.2	7.00
CBAM		80.4	6.35
CA		80.9	6.63
ECA	1	81.3	6.41
	3	81.7	6.41
	5	81.0	6.41

从表 3 可以看出,相比 SE、CBAM^[27] 和 CA^[28] 注意力模块,施加 ECA 模块的 MobileNetV3 模型检测精度最为理想,原 MobileNetV3 模型使用 SE 注意力模块,检测平均精度均值为 78.2%。较 SE 注意力模块,施加 CBAM 模块检测精度上升 2.2 个百分点,模型体积减少 0.65 MB;施加 CA 模块检测精度提高 2.7 个百分点,体积减少 0.37 MB。在不同的交互覆盖率下施加 ECA 模块检测精度均有大幅提高,体积减少 0.59 MB。说明在轻量化网络中施加不同注意力机制会减少模型的体积,适合的注意力模块会使得特征提取能力增强和检测精度提高,需要根据不同的检测目标加以选择。与此同时,对比不同交互覆盖率对 ECA 注意力模块的影响,在交互覆盖率为 3 时平均精度均值最高,为 81.7%;说明上述组合方式能更为准确地检测到复杂环境下百香果果实目标。

本文所引入的轻量化 ECA 模块,使用捕获跨通道交互的学习方式学习有效的百香果通道特征,提高了模型对于复杂环境下百香果目标的敏感程度,进而改善百香果检测任务中果实重叠和密集时对检测精度带来的负面影响。

3.4 不同迁移学习方式训练试验

不同的迁移学习方式对模型的检测精度有明显的影响,结果如表 4 所示。

表 4 不同迁移学习方式检测能力对比
Tab.4 Comparison of detection ability of different transfer learning modes

学习方式	数据集	平均精度均值/%
跨域	VOC 数据集 + 数据集 2	82.9
域内	数据集 1 + 数据集 2	83.3
跨域与域内相结合	VOC 数据集 + 数据集 1 + 数据集 2	87.4
跨域与域内多轮训练	VOC 数据集 + 数据集 1 + 数据集 2	88.3

从表 4 可以看出,跨域与域内多轮训练的迁移学习方式使模型精度达到最佳,为 88.3%,结合图 8 可以看出训练轮次到达 70 轮时,平均精度均值曲线趋于稳定,但还具有上升的趋势,因此采取多轮训练方式可使平均精度均值达到相对较高水平。相比于

跨域迁移、域内迁移和跨域与域内相结合迁移方式的平均精度均值分别提高 5.4、5.0、0.9 个百分点。由于跨域与域内相结合的方式在掌握通用特征后学习部分百香果特征,进而学习复杂环境下的百香果特征。相比于跨域迁移,利用部分百香果数据集作为过渡域可以减少源域和目标域相似度不同所引起的负面影响;而相比于域内迁移,仅含少量的百香果特征模型作为网络初始化无法完成复杂环境下百香果检测任务,通过通用数据集和少量百香果数据集融合的特征对网络进行初始化,弥补不完全的百香果特征所造成信息的损失,而跨域与域内多轮训练学习方式不仅具有跨域与域内相结合方式的优点,同时还可以将训练达到饱和,获得最佳的模型检测精度和最强的泛化能力。

3.5 不同检测模型综合对比试验

将改进模型与 SSD、Faster RCNN、YOLO v5s、ShuffleNet-v5 和 GhostNet-v5 模型在 Jetson Nano 上进行测试对比,试验结果如表 5 所示。

表 5 不同模型综合对比结果

Tab.5 Comprehensive comparison results of different models

模型	精确率/%	召回率/%	平均精度均值/%	计算量/ GFLOPs	计算机实时检 测帧率/(f·s ⁻¹)	嵌入式实时检测 帧率/(f·s ⁻¹)	模型体积/ MB
SSD	90.5	83.3	83.5	61.2	25.4	0.79	90.60
Faster RCNN	90.6	86.2	86.6	273.4	12.2	0.28	521.00
YOLO v5s	96.3	87.6	88.1	16.3	102.6	6.39	14.40
ShuffleNet-v5	97.0	80.9	81.1	7.3	116.4	9.85	7.04
GhostNet-v5	96.4	81.4	81.5	6.5	114.8	9.75	6.87
MbECA-v5	95.3	88.1	88.3	6.6	124.7	10.92	6.41

ShuffleNet-v5 和 GhostNet-v5 模型是将 YOLO v5s 模型结构的主干网络分别替换为 ShuffleNet 和 GhostNet 网络。从表 5 可知,本文模型的平均精度均值为 88.3%,相比于 YOLO v5s 提高 0.2 个百分点。相比于 SSD、Faster RCNN、ShuffleNet-v5、GhostNet-v5 目标检测网络,提高 4.8、1.7、7.2、6.8 个百分点。在 6 种模型中 MbECA-v5 的平均精度均值最高,说明预测为正确百香果能力强,减少百香果采摘过程中误检率。本文模型体积仅为 6.41 MB,约为 SSD、Faster RCNN、YOLO v5s 的 1/14、1/81、1/2;从计算机和嵌入式平台实时检测速度考虑,改进后的模型 MbECA-v5 比其他 6 类模型速度均快。在计算机实时检测速度为 124.7 f/s,嵌入式平台实时检测速度为 10.92 f/s,约为 SSD、Faster RCNN、YOLO v5s 的 14、39、1.7 倍;计算量相较于 SSD 及 Faster RCNN 减少 89.2% 和 97.6%,相较于 YOLO v5s 及 ShuffleNet-v5 减少 59.5% 和 4.8%,相较于 GhostNet-v5 增加 1.5%。综上所

述,改进后的轻量化网络 MbECA-v5 具有良好的测试性能,在边缘设备上部署具有更好的实时性。

为了验证不同模型在复杂环境下百香果的检测效果,选逆光、重叠和密集百香果图像进行测试。选取置信度阈值为 0.6,IOU 阈值为 0.5。图 9 为 YOLO v5s 和 MbECA-v5 2 种模型在 Jetson Nano 上

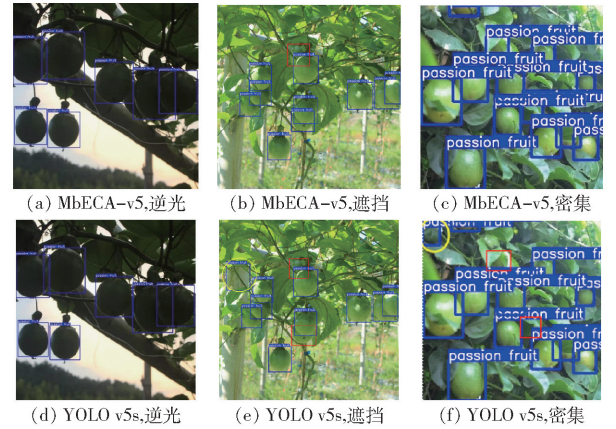


图 9 两种模型检测效果对比

Fig.9 Comparison of two models checking effects

检测结果。其中蓝色矩形框为预测结果,红色矩形框为漏检目标,黄色圆形框为误检目标。从图 9 可知,在逆光时,2 类模型预测结果均正确,无漏检和误检现象产生。在重叠时,2 种模型均存在漏检现象,MbECA-v5 模型预测仅漏检 1 个百香果目标,YOLO v5s 模型漏检 2 个百香果目标且产生误检现象。在密集时,MbECA-v5 检测效果优异,而 YOLO v5s 均有漏检和误检情况产生。综上,从识别的效果得出改进模型提高了在干扰环境下的检测能力,具有较好的鲁棒性和泛化能力。

4 结论

(1)提出一种基于嵌入式设备的轻量化模型 MbECA-v5,以 YOLO v5s 为基础模型,将原来的主干网络替换为 MobileNetV3 网络,在 MobileNetV3 网络中引入 ECA 注意力机制提取百香果全局信息,引入了逐点卷积进行通道调节。利用跨域与域内多训练相结合的迁移学习策略提高网络检测的精度。通过消融试验分析验证改进算法的有效性。

(2)复杂环境下百香果果实检测结果表明,

MbECA-v5 模型对百香果果实识别精确率为 95.3%,召回率为 88.1%,平均精度均值为 88.3%,与 YOLO v5s 相比,平均精度均值提高 0.2 个百分点。计算机和嵌入式实时检测的平均帧率分别为 124.7 f/s 和 10.92 f/s,检测速度相比原模型分别提高 22.1 f/s 和 4.53 f/s。在检测效果上,改进的模型比 YOLO v5 效果更好。说明改进的算法综合性能最佳,具有较好的鲁棒性和泛化能力。

(3)改进的算法与常用的目标检测网络和轻量化网络进行对比,本文模型的平均精度均值比 SSD 与 Faster RCNN 模型分别提高 4.8、1.7 个百分点。相比于 ShuffleNet-v5 与 GhostNet-v5 模型分别提高 7.2、6.8 个百分点。模型体积约为 SSD 与 Faster RCNN 模型的 1/14 和 1/81。在嵌入式设备的检测速度约为 SSD 与 Faster RCNN 模型的 14 倍和 39 倍,约为 ShuffleNet-v5 与 GhostNet-v5 模型的 1.12 倍和 1.11 倍。综上,MbECA-v5 模型减少嵌入式设备上的运行内存,可以兼顾精度和实时性,满足快速、准确地对田间百香果图像检测的要求,可为百香果采摘机器人视觉识别系统提供技术和理论支持。

参 考 文 献

- [1] TIAN H, WANAG T, LIU Y, et al. Computer vision technology in agricultural automation—a review[J]. Information Processing in Agriculture, 2020, 7(1): 1–19.
- [2] 陈媛媛,游炯,泽峰,等.世界主要国家精准农业发展概况及对中国的发展建议[J].农业工程学报,2021,37(11):315–324.
CHEN Yuanyuan, YOU Jiong, ZE Feng, et al. Review of precision agriculture development situations in main countries in the world and suggestions for China[J]. Transactions of the CSAE, 2021, 37(11): 315–324. (in Chinese)
- [3] WANG H, YANG C, LI X, et al. Digital technology boosting agricultural supply-side constitutive revolution in poor areas based on the intelligent environment of the internet of things[J]. Wireless Communications and Mobile Computing, 2022(5): 1–12.
- [4] WANG D, CAO W, ZHANG F, et al. A review of deep learning in multiscale agricultural sensing[J]. Remote Sensing, 2022, 14(3): 559–567.
- [5] 李振波,李萌,吴宇峰,等.基于优化 SSD 算法的冰鲜鲳鱼新鲜度评估方法研究[J].农业机械学报,2021,52(增刊):472–481.
LI Zhenbo, LI Meng, WU Yufeng, et al. Evaluation method of iced pomfret freshness based on improved SSD[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Suppl.): 472–481. (in Chinese)
- [6] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961–2969.
- [7] BENJUMEA A, TEETI I, CUZZOLIN F, et al. YOLO-Z: improving small object detection in YOLOv5 for autonomous vehicles[J]. arXiv e-prints; 2112.11798, 2021.
- [8] TU S, PANG J, LIU H, et al. Passion fruit detection and counting based on multiple scale Faster R-CNN using RGB-D images[J]. Precision Agriculture, 2020, 21(5): 1072–1091.
- [9] 唐熔钗,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 Rongchai, WU Xiru. Temporal 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)
- [10] WU X, TANG R. Fast detection of passion fruit with multi-class based on YOLOv3[J]. Lecture Notes in Electrical Engineering, 2021, 706: 818–825.
- [11] TU S, XUE Y, ZHENG C, et al. Detection of passion fruits and maturity classification using red-green-blue depth images[J]. Biosystems Engineering, 2018, 175: 156–167.
- [12] 李江昀,赵义凯,薛卓尔,等.深度神经网络模型压缩综述[J].工程科学学报,2019,41(10): 1229–1239.
LI Jiangyun, ZHAO Yikai, XUE Zhuoer, et al. Review of deep neural network model compression[J]. Journal of Engineering Science, 2019, 41(10): 1229–1239. (in Chinese)
- [13] 武星,齐泽宇,王龙军,等.基于轻量化 YOLOv3 卷积神经网络的苹果检测方法[J].农业机械学报,2020,51(8): 17–25.
WU Xing, QI Zeyu, WANG Longjun, et al. Apple detection method based on lightweight YOLOv3 convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 17–25. (in Chinese)

(下转第 322 页)

- [27] CUI Yiming, CHE Wanxiang, LIU Ting, et al. Pre-training with whole word masking for Chinese BERT[J]. arXiv preprint arXiv:1906.08101, 2019.
- [28] 唐楠楠, 陈吉, 侯磊, 等. 融合注意力机制和指针标注的实体关系联合抽取方法[J]. 小型微型计算机系统, 2021:1-8. <https://kns.cnki.net/kcms/detail/21.1106.TP.20211217.1520.013.html>.
TANG Nannan, CHEN Ji, HOU Lei, et al. Joint extraction method of entity and relation based on attention mechanism and pointer annotation[J]. Journal of Chinese Computer Systems, 2021:1-8. (in Chinese)
- [29] 奚雪峰, 周国栋. 面向自然语言处理的深度学习研究[J]. 自动化学报, 2016, 42(10):1445-1465.
XI Xuefeng, ZHOU Guodong. A survey on deep learning for natural language processing[J]. Acta Automatica Sinica, 2016, 42(10):1445-1465. (in Chinese)
- [30] 李晓静, 林海伦, 贾岩涛, 等. 融合页面结构与内容的在线百科实体标注方法[J]. 计算机科学与探索, 2015, 9(10):1238-1246.
LI Xiaojing, LIN Hailun, JIA Yantao, et al. Online encyclopedia entities tagging method based on page structure and content [J]. Journal of Frontiers of Computer Science and Technology, 2015, 9(10):1238-1246. (in Chinese)
- [31] 鄂海红, 张文静, 肖思琪, 等. 深度学习实体关系抽取研究综述[J]. 软件学报, 2019, 30(6):1793-1818.
E Haihong, ZHANG Wenjing, XIAO Siqi, et al. Survey of entity relationship extraction based on deep learning[J]. Journal of Software, 2019, 30(6):1793-1818. (in Chinese)
- [32] MU Yang, DING Wei, ZHOU Tianyi, et al. Constrained stochastic gradient descent for large-scale least squares problem[C]//ACM, 2013.
- [33] 陈小倩, 曹伟召, 左军, 等. 不同施氮量对水稻南粳 46 产量及品质的影响[J]. 现代农业科技, 2020(2):17-18.
CHEN Xiaoqian, CAO Weizhao, ZUO Jun, et al. Effect of different nitrogen application on yield and quality of Nanjing 46[J]. Modern Agricultural Technology, 2020(2):17-18. (in Chinese)

(上接第 269 页)

- [14] MENG F, CHENG H, LI K, et al. Pruning filter in filter[J]. Advances in Neural Information Processing Systems, 2020, 33:17629-17640.
- [15] 方鹏, 郝宏运, 王红英, 等. 基于知识蒸馏的叠层笼养蛋鸡行为识别模型研究[J]. 农业机械学报, 2021, 52(10):300-306.
FANG Peng, HAO Hongyun, WANG Hongying, et al. Research on behavior recognition model of stacked cage laying hens based on knowledge distillation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10):300-306. (in Chinese)
- [16] TAN K, WANG D. Compressing deep neural networks for efficient speech enhancement[C]//ICASSP 2021—2021 IEEE International Conference on Acoustics. Speech and Signal Processing (ICASSP). IEEE, 2021: 8358-8362.
- [17] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:4510-4520.
- [18] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018:6848-6856.
- [19] 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.
- [20] 张勤, 王家辉, 李彬, 等. 基于 YOLOv3 目标检测的秧苗列中心线提取方法[J]. 农业机械学报, 2020, 51(8):34-43.
ZHANG Qin, WANG Jiahui, LI Bin, et al. Extraction of seedling line based on YOLOv3 target detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8):34-43. (in Chinese)
- [21] BOCHKOVSKIY A, WANG C Y, LIAO H, et al. Yolov4: optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.
- [22] 朱红春, 李旭, 孟扬, 等. 基于 Faster R-CNN 网络的茶叶嫩芽检测[J]. 农业机械学报, 2022, 53(5):217-224.
ZHU Hongchun, LI Xu, MENG Yang, et al. Detection of tea buds based on Faster R-CNN network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5):217-224. (in Chinese)
- [23] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [24] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019:1314-1324.
- [25] WANG Q, WU B, ZHU P, et al. Supplementary material for ECA-Net: efficient channel attention for deep convolutional neural networks[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2020:13-19.
- [26] OTOVIĆ E, NJIRJAK M, JOZINOVIĆ D, et al. Intra-domain and cross-domain transfer learning for time series data—how transferable are the features? [J]. Knowledge-Based Systems, 2022, 239: 107976.
- [27] 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.
- [28] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:13713-13722.