

基于残差块与注意力机制的果蔬自动识别方法

余琼¹ 张瑞^{1,2} 李德豪¹ 员玉良^{1,3} 王至秋^{1,3}

(1. 青岛农业大学机电工程学院, 青岛 266109; 2. 哈尔滨工程大学智能科学与工程学院, 哈尔滨 150001;
3. 青岛农业大学山东省主要农作物机械化生产协同创新中心, 青岛 266109)

摘要: 针对果蔬识别中识别效率低、成本高等问题,本文提出了基于残差块和注意力机制的果蔬识别模型,并成功部署于果蔬智能识别设备。果蔬自动识别装置由 Raspberry Pi、STM32F103ZET6、摄像头、称量传感器、处理器、显示屏、微型打印机、扎口机以及电源等部分组成。中央控制器与显示屏进行交互实时显示各种参数,通过摄像头与称量传感器采集待测物体图像与待测物体质量,由部署于 Raspberry Pi 的果蔬自动识别模型对果蔬进行精准识别,同时协同单片机 STM32F103ZET6 将果蔬相关信息打印并控制扎口机进行封口打包。本文以 YOLO v5 网络为基础,通过增加残差块与注意力机制构建果蔬自动识别模型 RB + CBAM - YOLO v5。以自制的数据集训练网络,将 6 种网络进行对比试验,并选择最优网络进行设备端检测试验。试验结果表明,RB + CBAM - YOLO v5 的精确率、召回率与 mAP_{0.5} 分别为 83.55%、96.08%、96.20%,较 YOLO v5 提升 4.47、1.10、0.90 个百分点。将 RB + CBAM - YOLO v5 模型部署于嵌入式设备 Raspberry Pi 中,设备可实现精准识别、自动称量、打印凭条以及快速打包等功能,可满足果蔬识别以及无人售卖装置的需求。

关键词: 果蔬识别; 残差块; 注意力机制; YOLO v5

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2023)S2-0214-09

Method for Fruit and Vegetable Automatic Recognition Based on Residual Block and Attention Mechanism

YU Qiong¹ ZHANG Rui^{1,2} LI Dehao¹ YUN Yuliang^{1,3} WANG Zhiqiu^{1,3}

(1. College of Mechanical and Electrical Engineering, Qingdao Agricultural University, Qingdao 266109, China
2. College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China
3. Shandong Collaborative Innovation Center of Major Crop Mechanized Production Equipment, Qingdao Agricultural University, Qingdao 266109, China)

Abstract: To solve the problems of low efficiency and high cost in fruits and vegetables recognition, a fruit and vegetable recognition model based on residual block and attention mechanism was proposed, and successfully deployed in fruit and vegetable intelligent recognition equipment. The fruit and vegetable automatic recognition device was composed of Raspberry Pi, STM32F103ZET6, camera, weighing sensor, processor, display screen, micro printer, binding machine and power supply. The central controller interacted with the display screen to display various parameters in real time. The image and quality of the object to be measured were collected through the camera and weighing sensor. The fruit and vegetable automatic recognition model deployed in the Raspberry Pi could accurately identify the fruits and vegetables. At the same time, it cooperated with MCU STM32F103ZET6 to print fruit and vegetable related information and control the tying machine to seal and pack. Based on YOLO v5 network, an automatic recognition model RB + CBAM - YOLO v5 was constructed by adding residual blocks and attention mechanism. The network was trained with the self-made data set, and six kinds of networks were compared, and the optimal network was selected for the device side detection test. The experimental results showed that the accuracy rate, recall rate and mAP_{0.5} of RB + CBAM - YOLO v5 were 83.55%,

收稿日期: 2023-06-30 修回日期: 2023-08-30

基金项目: 国家自然科学基金项目(52275258)

作者简介: 余琼(1998—),男,硕士生,主要从事机器学习研究,E-mail: yuqiong1998@163.com

通信作者: 王至秋(1968—),男,高级实验师,主要从事智能仪器与监测装备研究,E-mail: njdh@163.com

96.08% and 96.20% , respectively , which were 4.47 percentage points , 1.10 percentage points and 0.90 percentage points higher than those of YOLO v5. The RB + CBAM – YOLO v5 model was deployed in the embedded device Raspberry Pi , and the device could realize accurate identification , automatic weighing , printing slip and fast packaging functions , which could meet the needs of fruits and vegetables identification and unsold devices.

Key words: fruits and vegetables recognition ; residual block ; attention mechanism ; YOLO v5

0 引言

快速精准识别果蔬种类是实现无人售卖的基本前提。相较于传统人工识别的方式,诸多研究中采用图像处理识别技术进行果蔬识别^[1],一定程度上推动了果蔬销售的自动化发展进程。BHARGAVA 等^[2]分析了计算机视觉和图像处理技术在农业食品行业的应用,提出了概述预处理、分割、特征提取和使用这些特征进行分类的不同方法。陶华伟等^[3]针对果蔬图像纹理特征提出了一种颜色完全局部二值模式纹理特征提取算法,利用 HSV 颜色直方图、外点/内点颜色直方图提取图像颜色特征,采用匹配得分融合算法将颜色和纹理特征相融合,采用最近邻分类器实现果蔬农产品分类,识别率提升 5%。DUBEY 等^[4]对果蔬分类和果病分类进行研究,根据颜色和纹理线索来识别水果和蔬菜的种类,最高平均准确率达到 90%。PENG 等^[5]利用形状不变矩等方法对水果的颜色和形状特征进行特征提取,根据提取的特征向量,应用支持向量机分类器对水果进行分类识别。传统图像处理方法利用纹理和颜色线索来进行识别,将图像分割为前后景,通过颜色纹理特征搭建针对于该特征的分类器进行识别。但此方法下,往往分类器难以搭建,或需要大量数据来支持分类器的搭建依据,并且对图像特征提取能力有待提高,不足以满足目前需求。

近年来机器学习研究愈发深入,特别是食品工程领域,机器学习在果蔬产品检测方面的应用优势也开始显现。DILL 等^[6]提出了 II 型模糊,利用深度学习模型来识别和分类水果图像。NARANJO-TORRES 等^[7]提出采用卷积神经网络进行水果识别的方法,采用 AlexNet、VGG16、MobileNet、InceptionV3、ResNet5 等多种模型进行水果识别,在 Testing 上 VGG16 的准确率达到 89.58%。ZHANG 等^[8]设计了一个 13 层的卷积神经网络,使用 3 种类型数据增强方法(图像旋转、Gamma 矫正以及噪声注入)进行数据扩充。比较了最大池化和平均池化,最终模型整体准确率为 94.94%。WANG 等^[9]提出一种改进性的深度卷积网络应用于水果类别分类,创建了一个 8 层的深度卷积神经网络,利用参数

整流线性单元替代普通整流线性单元,准确率达到 95.67%。KATARZYNA 等^[10]提出了一种用于零售业应用的水果品种分类的双轨方法。使用 2 个具有相同架构但权重矩阵不同的 9 层卷积神经网络。对测试集的整体图像分类准确率达到 99.78%。

在上述研究中,均采用了深度学习进行水果或者蔬菜的识别和检测,证明了运用深度学习在水果、蔬菜检测领域上的可行性。但其运行端仍停留在 PC 机上,具有一定的使用局限性,另外在训练数据方面,上述研究大都采用普通 CNN 网络,训练模型对于水果、蔬菜的识别准确率有待提高,以及对复杂情况下的适应能力有待提升。

嵌入式设备以轻巧、便携为特点。随着开源硬件的快速发展,为深度学习在嵌入式设备上运行提供了可能。HOLALAD 等^[11]以 SPARTAN 3 为平台,开发了一种简单有效的水果识别系统,利用 3D 特征向量对最小距离分类器进行 Matlab 分析,成功率达到 85%。通过将 JAVA ECLIPSE IDE 生成的测试图像 COE 文件烧录到 IP 核上,开发了 Verilog 编码的硬件平台。ZHANG 等^[12]提出了一种基于水果颜色和形态特征的单次多盒检测器(SSD)改进模型。使用 RealSense D435i 型双目深度相机收集田间待采摘水果的图像。FAKHROU 等^[13]提出了一种基于智能手机的视觉障碍儿童食物和水果识别系统,在由 29 种菜肴和水果组成的数据集上取得的准确率达 95.55%,最后将 CNN 部署于手机端。上述方法一方面未考虑果蔬套袋问题,多数包装袋都为半透明状态,从而弱化果蔬特征,对识别率产生影响。另一方面这些方法对数据集要求较高,从一些地方采集的图像往往不适用于其他地方的水果。

基于以上研究背景,本文利用自建水果、蔬菜数据集,引入自设计残差块,自设计残差块可以将经过残差组件的输入特征信息跨越多层卷积,将特征信息传送给后层,解决随着网络层数的加深,而效果却下降的问题,同时引入注意力机制提升网络特征提取能力,提出针对水果、蔬菜的检测网络,采用开源硬件,搭建外围电路,制作果蔬识别装备。面向实际销售环境的真实背景,完成对果蔬的识别、打印凭条、称量、人机交互、打包一体化功能。

1 材料与方法

1.1 果蔬数据集

1.1.1 数据采集

数据集来源于山东省青岛市城阳区春阳路大润发超市(36°19.28'N,120°24.25'E)、山东省济南市历下区经十路1111号华润济南万象城(36°39.92'N,117°6.40'E)、山东省青岛市城阳区正阳中路136号家佳源购物中心(36°18.32'N,120°25.60'E)、公开数据集 fruits-360_dataset、网络爬虫,将5部分图像合并、筛选构建果蔬数据集,其中,实地拍摄图像与网络下载图像比例为7:1。

数据集主要由移动设备 iPhone XR 型手机(摄

像头 1.2×10^7 像素、后置 $f/1.8$ 光圈)、HuaWei Nova4 型手机(后置三摄: 4.8×10^7 像素主摄, $f/1.8$ 光圈 + 1.6×10^7 像素超广角, $f/2.2$ 光圈 + 2×10^7 像素景深, $f/2.4$ 光圈)实地拍摄得到,经过裁剪得到 640 像素 $\times$ 640 像素图像。为了增强数据兼容性另采取部分网络爬虫所得图像与开源数据集 fruits-360_dataset,将两类图像构成数据库,对数据库进行人工筛选排除相似度高的图像,之后进行标注工作。果蔬位置随机摆放,果蔬之间包含堆叠、粘连、遮挡等复杂情况。数据集包含不同光照、背景的情况,可以使网络在真实环境中具有更强的感知学习能力^[14]。部分果蔬样本见图1。



图1 果蔬样本示例
Fig.1 Examples of fruit and vegetables images

果蔬数据集由15种不同的水果和蔬菜组成,分别是:生菜、芸豆、枣、葡萄、茄子、玉米、芹菜、油菜、辣椒、草莓、火龙果、菠萝、橙子、蘑菇、香蕉,如表1所示。选取其中8138幅图像作为数据集,训练集

每幅图像包含多个检测目标,目标框个数达到26197,满足大数据量需求。

数据集标签分布如图2所示,图2中代表训练集的各类目标框数量,图2a中的数据点分布呈正方形,表示图像整个正方形区域内均有目标框,当果蔬位置偏离图像中心时亦可有效检测。

表1 实验数据
Tab.1 Experimental data

种类	图像数量	标签数量
火龙果	607	837
草莓	554	1 299
菠萝	522	519
橙子	572	2 017
蘑菇	531	1 675
生菜	508	608
芸豆	512	4 109
枣	506	3 549
辣椒	557	2 353
葡萄	538	2 992
茄子	557	1 020
玉米	531	994
芹菜	514	1 516
油菜	512	1 412
香蕉	617	1 297
总计	8 138	26 197

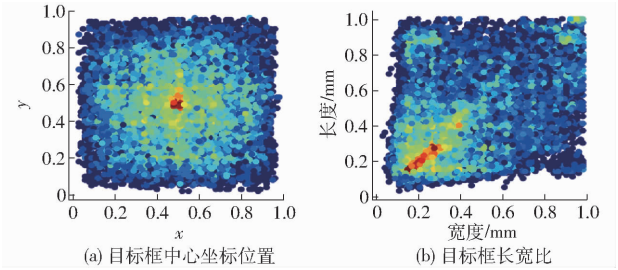


图2 数据集标签分布图
Fig.2 Label distribution of data set

1.1.2 数据增强

对输入图像采用 Mosaic 数据增强,将每一轮的迭代周期(Epochs)中的图像进行随机替换,替换方法为随机缩放、随机裁剪和随机排布,通过上述方法获得图像,更换训练中的图像,下一轮重新更换不同方式组合的图像,从而增强网络的泛化能力^[15]。经

过 Mosaic 数据增强之后被标注的图像如图 3 所示。

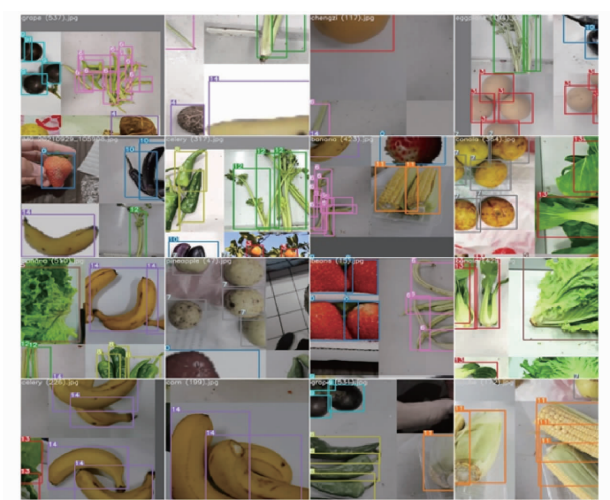


图 3 数据增强

Fig. 3 Data enhancement

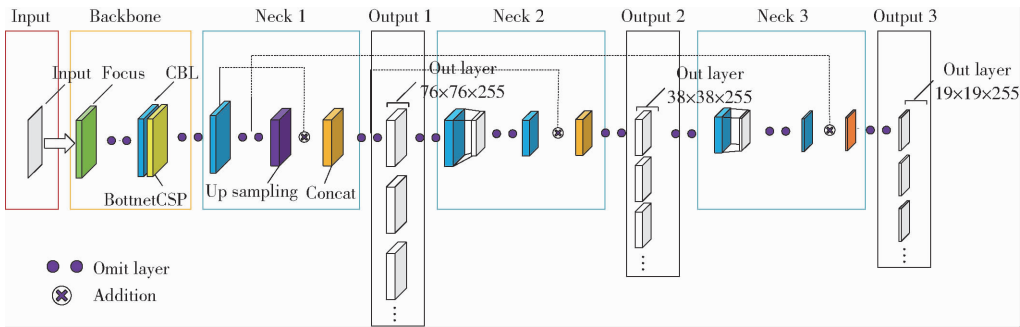


图 4 YOLO v5 简图

Fig. 4 Schematic of YOLO v5

1.3 构建果蔬识别模型

1.3.1 卷积改进方案

为了提高 Backbone 结构针对果蔬的深层次特征提取能力,本文加入残差块结构,提出第 1 种改进方式,在 BottnetCSP 中将第 2、3 层卷积替换为自设计残差块(简称 RB - YOLO v5),设计了新的 Backbone 结构,如图 5 所示。网络层数对特征提取能力有很大影响,随着网络层数的增加,对特征提取能力应当更强,但研究表明,并不是网络层数越多,特征提取能力越强。为了解决这个问题,引入残差块突破深层次的瓶颈^[18]。

该残差块包含 3 个卷积层、2 个归一化层、1 个激活层,分支经过 1 次卷积、1 次归一化,最后两者输出相加再经过 1 层激活。该结构可以将经过残差组件的输入特征信息跨越多层卷积,将特征信息传送给后层,解决了随着网络层数的加深,而效果却下降的问题,有效加深了网络层数,增强了其 Backbone 结构的特征提取能力^[19-20]。

当输入与输出尺寸不匹配时,该残差块的残差连接(Skip Connection)路径中的 Conv4 层还可将输入图像调整为不同尺寸。

1.2 YOLO v5

目标检测网络普遍复杂、网络体积较大,其中 YOLO v5 速度快、网络参数量少、准确度高、方便在嵌入式设备上部署。如图 4 所示,YOLO v5 的整个网络结构分为 4 部分,分别为 Input、Backbone、Neck、Output。Backbone 阶段采用 Focus 切片操作以及 BottnetCSP 结构^[16],BottnetCSP 为提取特征的主要部分。Neck 阶段采用 FPN + PAN 结构。FPN 层自顶向下传达强语义特征,而特征金字塔则自底向上传达强定位特征,两者结合从不同的主干层对不同的检测层进行特征融合。YOLO v5 网络的输出层采用 GIoU 作为损失函数^[17]。尽管 YOLO v5 有着数据增强、金字塔架构等优点,但其特征提取能力还有待提高。因此,在此基础上提出改进型 YOLO v5 网络,增加对果蔬的检测效果。

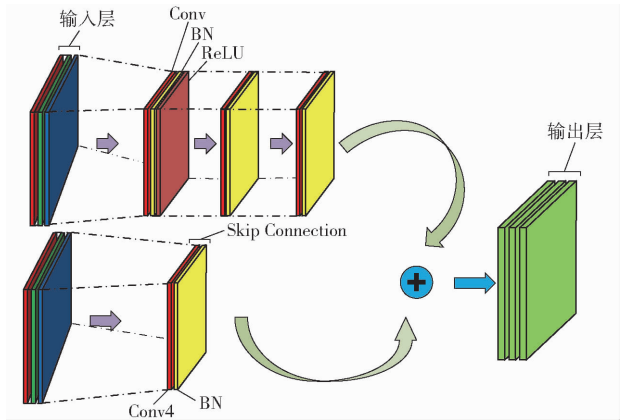


图 5 自设计残差块结构图

Fig. 5 Structure diagram of self-designed residual block

对于残差块可以进行跨层传递信息给出如下解释,普通 2 层网络之间的传递过程为

$$Z^{[l+1]} = W^{[l+1]} a^{[l]} + b^{[l+1]} \quad (1)$$

$$a^{[l+1]} = Z^{[l+1]} \quad (2)$$

$$Z^{[l+2]} = W^{[l+2]} a^{[l+1]} + b^{[l+2]} \quad (3)$$

$$a^{[l+2]} = Z^{[l+2]} \quad (4)$$

式中 $a^{[l]}$ ——输入的原图像或者特征图
 $W^{[l+1]}、W^{[l+2]}$ ——第 1、2 层的权重
 $b^{[l+1]}、b^{[l+2]}$ ——第 1、2 层的偏置

$Z^{[l+1]}$ 、 $Z^{[l+2]}$ ——经过第 1、2 层输出的特征图

$a^{[l+1]}$ 、 $a^{[l+2]}$ ——输入第 2、3 层特征图

当加入 Skip Connection 通道时, $a^{[l]}$ 经过 $Y(a^{[l]})$ 非线性函数得到 $E^{[l]}$, 本文中的 $Y(x)$ 为一个 Conv4 加 BN 层

$$E^{[l]} = Y(a^{[l]})$$

(5)

进行相加得到 $a^{[l+2]}$ 最后的输出, 与式 (4) 比较, 加入了上一层的特征, 实现了跨层传递信息^[21]。

$$a^{[l+2]} = G(Z^{[l+1]} + E^{[l]})$$

(6)

在网络提取特征的过程中, 网络对每个特征都进行相同的操作, 无论是否对后续推理有用, 而注意力机制可以将无用的特征与有用的特征在空间或者通道上乘以不同的权重, 有用的权重系数高于无用的权重系数, 即可在网络的计算过程中有效减少对无用特征的提取, 增强对有用特征的关注^[22]。

1.3.2 添加 SE 注意力机制改进方案

SE 为通道注意力机制模块中的一种。输入 $W \times H \times C$ 的特征图, 经过平均池化裁剪将特征图变成 $1 \times 1 \times C$, 只保留通道特征, 将该特征图送入 FC 网络, 由 FC 网络给特征图乘以权重系数, 输出该特征图, 之后与原网络特征相乘, 得到加以修改的特征图^[23]。

基于此, 提出第 2 种改进方式: 在主干提取网络 Backbone 结构最后 1 层添加 SE (简称 SE - YOLO v5), 以及第 3 种改进方式: 在 BottnetCSP 中将第 2、3 层卷积替换为自设计残差块的同时, 并在主干提取网络 Backbone 结构最后 1 层添加 SE (简称 RB + SE - YOLO v5)。

1.3.3 添加 CBAM 注意力机制改进方案

CBAM 由通道注意力模块与空间注意力模块共同组成。特征的每一个通道都代表一个专门的检测器, 通道注意力机制会给每一个特征赋予不同的权重。CBAM 通过学习的方式自动获取每个特征通道、空间的重要程度, 利用得到的重要程度来提升特征所占比重并抑制对当前任务不重要的特征。首先对输入信息采用了全局平均池化和最大池化两种方法来分别利用不同的信息^[24], 公式为

$$M_c(F) = \text{Sigmoid}(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F)))$$

(7)

$$F' = M_c(F) F$$

(8)

在通道注意力模块之后, 再将特征图送入空间注意力模块, 公式为

$$M_s(F') = \text{Sigmoid}(f([\text{AvgPool}(F'), \text{MaxPool}(F')]))$$

(9)

$$F'' = M_s(F') F'$$

(10)

式中 F ——输入的特征图

MLP——两层神经网络处理

F' ——经过通道注意力机制输出后的结果

$M_s(F')$ ——权重系数

$f()$ ——卷积操作

Sigmoid——激活函数

AvgPool(F)——平均池化

MaxPool(F)——最大池化

F'' 为经过空间注意力机制输出之后的结果。与通道注意力相似, 给定一个 $H \times W \times C$ 的特征 F' , 先分别进行通道维度的平均池化和最大池化得到 2 个 $H \times W \times 1$ 的通道, 并将这 2 个通道拼接在一起。然后, 经过一个 7×7 的 Conv 层, 激活函数为 Sigmoid, 得到权重系数 $M_s(F')$ 。最后, 将权重系数和特征 F' 相乘即可得到缩放后的新特征图 F'' , 如图 6 所示。

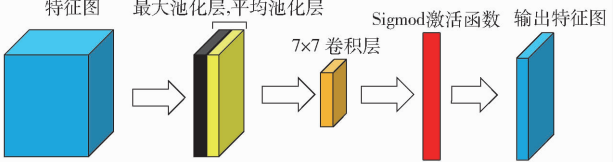


图 6 空间注意力
Fig. 6 Spatial attention

基于此, 提出第 4 种改进方式: 在主干提取网络 Backbone 结构每层 Conv 之后加 1 层 CBAM, 最后 1 层添加 CBAM (简称 CBAM - YOLO v5); 以及第 5 种改进方式: BottnetCSP 中将第 2、3 层卷积替换为自设计残差块的同时, 并在主干提取网络 Backbone 结构每层 Conv 之后加 1 层 CBAM, 以及最后 1 层添加 CBAM (简称 RB + CBAM - YOLO v5)。

1.4 参数设置与网络训练

采用开源 PyTorch 深度学习框架, 以 Python 为编程语言, 选用谷歌 Colab Pro 平台进行训练, GPU 型号为 Tesla P100 - PCIE - 16GB, 软件环境: CUDA Toolkit 11.0, Python 3.7, PyTorch 1.5。试验部分超参数如表 2 所示。

表 2 超参数设置
Tab.2 Hyperparameter settings

超参数	数值/类型
Epochs	80
Batch_size	16
Depth_multiple	0.33
Width_multiple	0.50
Optimizer	Adam
Lr0	0.01
GIoU	0.05

学习率是训练过程中重要的超参数,过大则会波动大,容易发散,以至不收敛;过小则会导致训练过慢,收敛效果不好等问题。较原有固定学习率的方式,本试验采用余弦退火实时更改学习率的方法,适应 Loss 梯度寻找最佳学习率,经验证该方式比原有方式具有更好的收敛效果和更少的训练时间^[25]。

根据系统需求,本文提出的果蔬自动识别装置硬件方案如图 7 所示,果蔬识别设备硬件包括 Raspberry Pi、STM32F103ZET6、摄像头、称量传感器、处理器、显示屏、微型打印机、扎口机、电源。实物如图 8 所示。

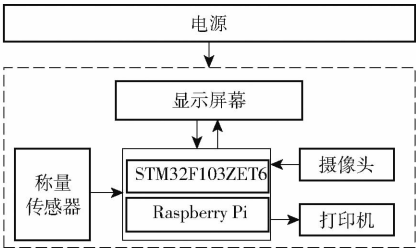


图 7 系统硬件结构图

Fig. 7 System hardware structure diagram

中央控制器由 Raspberry Pi 与 STM32F103ZET6 组成^[26-27]。中央控制器与显示屏进行交互实时显示各种参数,通过摄像头与称量传感器采集待测物体图像和质量,将结果通过打印机打印凭条并利用扎口机进行封口打包。为满足系统精简稳定的需求,采用称量传感器与支架一体化的设计,同时将固定板固定在称量传感器上,然后将包装盒通过扎口机进行打包操作,摄像头利用支架固定在设备上方,对水果进行俯拍。为增加装置的可操作性和提升操

作的可视化程度,采用串口屏绘制人工交互界面。当检测到有待测物体放置时,开始对待测物体进行检测,将该种类的水果、蔬菜的价格显示出来,点击“打印凭条”,该类果蔬的价格、总价、种类、时间被打

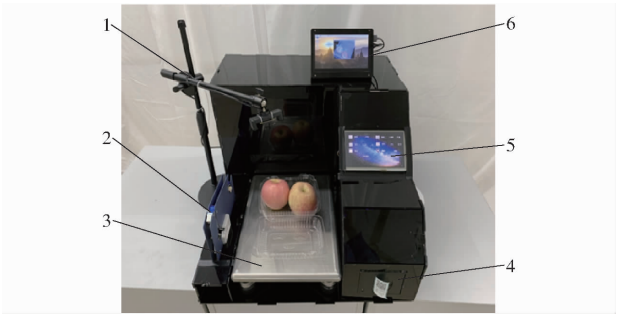


图 8 果蔬自动识别装置实物图

Fig. 8 Fruit and vegetable automatic identification device

1. 摄像头与支架 2. 手动捆扎装置 3. 称量平台 4. 凭条打印机 5. 人机交互界面 6. 云平台实时显示

首先,在自建果蔬数据集的基础上,通过 Mosaic 数据增强对果蔬数据集进行处理,之后在 BottleneckCSP 结构中将原有卷积替换成残差块,同时在主干提取网络不同位置添加多种注意力机制,达到增强特征提取能力的目的。对于多种改进方式,提出 5 种针对果蔬检测的识别网络,将所得实验数据依据客观评价标准进行衡量,分析网络的综合性能,将 5 种改进型网络与原网络部署在嵌入式设备上,对 6 种网络进行对比实验,并确定最优网络,进行实际预测。其次,Raspberry Pi 通过 USART 协议将预测结果发送给 STM32F103ZET6,通过单片机控制外围电路完成后续打包、打印凭条、获取质量、人机交互等功能。工作流程如图 9 所示。

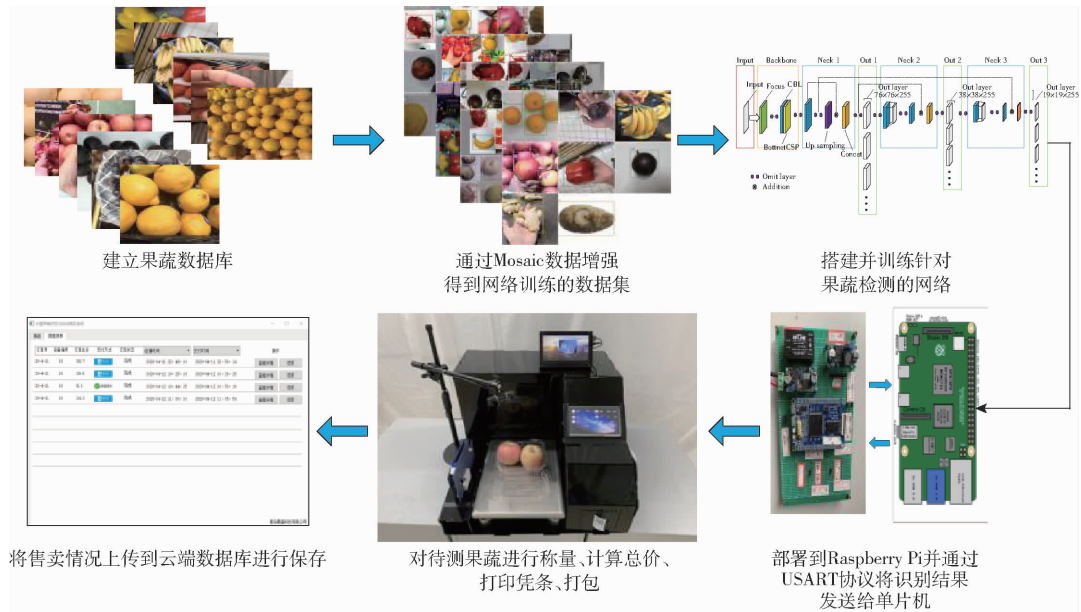


图 9 整体工作流程图

Fig. 9 Overall working flowchart

2 结果与分析

2.1 评价指标

使用自建果蔬数据集,对6种网络进行训练。引入多项质量指标进行客观评价,使用该评价指标对每种网络以及推理结果进行分析与验证,完成对果蔬检测网络的评估。

为验证网络效果,采用精确率(Precision)、召

回数(Recall)、平均精度均值(Mean average precision, mAP)、GIoU损失值、置信度损失值等评价指标对网络性能进行客观评价。mAP_{0.5}表示交并比阈值为0.5时的mAP,mAP_{0.5:0.95}表示交并比阈值在[0.5,0.95]之间每隔0.05取一次mAP的平均值^[28-30]。

2.2 不同模型对比

对6种检测网络性能对比,如图10所示。

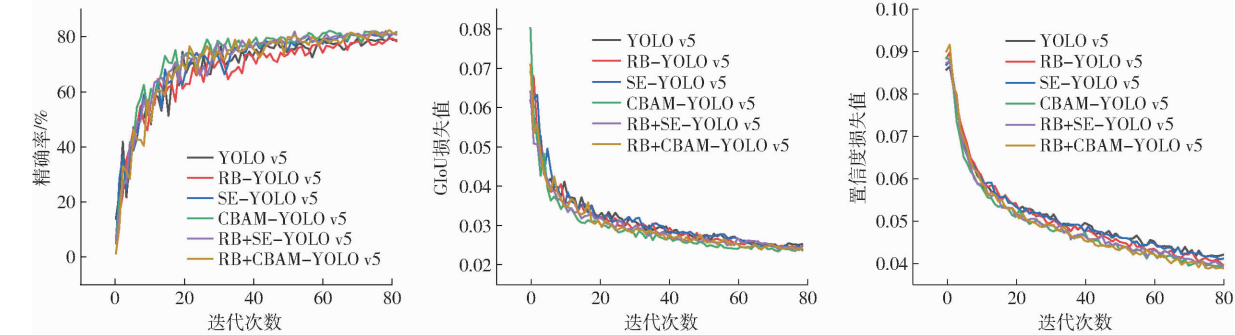


图 10 训练曲线

Fig. 10 Training curves

按照网络评估指标对网络进行训练结果对比与分析,结果见表3。

表 3 训练结果
Tab.3 Training results %

网络	精确率	召回率	mAP _{0.5}	mAP _{0.5:0.95}
YOLO v5	79.08	95.97	95.30	72.07
RB-YOLO v5	79.03	95.98	96.14	71.74
SE-YOLO v5	77.69	93.58	95.64	71.56
CBAM-YOLO v5	80.46	93.58	96.04	71.51
RB+SE-YOLO v5	82.39	97.52	97.11	72.85
RB+CBAM-YOLO v5	83.55	96.08	96.20	73.20

YOLO v5 训练检测精确率为 79.08%,RB + CBAM-YOLO v5 的网络的精确率为 83.55%,提升 4.47 个百分点,较其他 4 种改进方式有明显的提升,有效提高模型的检测精确率;改进前的网络召回率为 95.97%,5 种改进方式与原网络持平,保证了高召回率,避免了漏检的问题;YOLO v5 的 mAP_{0.5} 为 95.30%左右,5 种改进方式下 mAP_{0.5} 均高于原网络。

通过回归损失以及置信度损失曲线可以看出,在保证高召回率、mAP_{0.5}、mAP_{0.5:0.95}的前提下,RB + CBAM-YOLO v5 的训练损失低于 YOLO v5、CBAM-YOLO v5、RB-YOLO v5、RB+SE-YOLO v5、SE-YOLO v5。

由训练损失与训练精度可以看出,RB + CBAM-YOLO v5 的精确率最高、训练损失最低,精确率、召回率、mAP_{0.5} 分别为 83.55%、96.08%、96.20%,较

YOLO v5 分别提高 4.47、1.1、0.9 个百分点。因此选取 RB + CBAM-YOLO v5 作为果蔬检测网络并部署到 Raspberry Pi。

结合前文分析,最终选取 RB + CBAM-YOLO v5 作为最终检测网络并进行部署。选取山东省青岛市城阳区正阳路 157 号万象汇(36°18.73'N, 120°25.25'E)和山东省青岛市城阳区正阳中路 136 号家佳源购物中心(36°18.32'N,120°25.60'E)的 15 种果蔬样本,在实际设备端进行检测测试,测试效果满足正常需求。



图 11 装置现场测试图

Fig. 11 Device field test diagram

3 结论

(1)自建果蔬数据集,提出了针对果蔬检测的改进网络,经验证与原网络相比其性能更高,有效增强了网络的特征提取能力。在保证原网络性能的前提下,通过对主干网络卷积之后的特征图进行加深目标权重以及在 BottleneckCSP 中嵌入残差块替代原有的卷积,使网络能更好地对特征进行提取。同时,该改进方法对在目标检测网络中引入注意力机制以及残差块提供了一定的参考意义。

(2)通过实验对比,改进后模型精确率、召回率 1.1、0.9 个百分点。将模型部署于设备端进行现场和平均精度均值较原 YOLO v5 网络分别提高 4.47、测试,测试效果满足日常果蔬自动识别需求。

参 考 文 献

[1] 潘梅. 图像识别在果蔬分类与识别中的应用[J]. 现代农业科技, 2021(16): 257–259.
PAN Mei. Application of image recognition in classification and recognition of fruits and vegetables[J]. Modern Agricultural Science and Technology, 2021(16): 257–259. (in Chinese)

[2] BHARGAVA A, BANSAL A. Fruits and vegetables quality evaluation using computer vision: a review[J]. Journal of King Saud University-Computer and Information Sciences, 2018, 33(3): 243–257.

[3] 陶华伟,赵力,奚吉,等. 基于颜色及纹理特征的果蔬种类识别方法[J]. 农业工程学报, 2014, 30(16): 305–311.
TAO Huawei, ZHAO Li, XI Ji, et al. Fruits and vegetables recognition based on color and texture features[J]. Transactions of the CSAE, 2014, 30(16): 305–311. (in Chinese)

[4] DUBEY R S, JALAL S A. Application of image processing in fruit and vegetable analysis: a review[J]. Journal of Intelligent Systems, 2015, 24(4): 405–424.

[5] PENG H, SHAO Y, CHEN K, et al. Research on multi-class fruits recognition based on machine vision and SVM[J]. IFAC PapersOnLine, 2018, 51(17): 817–821.

[6] DILL H S, MURUGESAN G, KHEHRA B S, et al. Fruit recognition from images using deep learning applications[J]. Multimedia Tools and Applications, 2022, 81(23): 33269–33290.

[7] NARANJO-TORRES J, MORA M, HERNANDEZ-GARCIA R, et al. A review of convolutional neural network applied to fruit image processing[J]. Applied Sciences, 2020, 10(10): 3443.

[8] ZHANG Y, DONG Z, CHEN X, et al. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation[J]. Multimedia Tools and Applications, 2019, 78(3): 3613–3632.

[9] WANG S, CHEN Y. Fruit category classification via an eight-layer convolutional neural network with parametric rectified linear unit and dropout technique[J]. Multimedia Tools and Applications, 2020, 79: 15117–15133.

[10] KATARZYNA R, PAWEL M. A vision-based method utilizing deep convolutional neural networks for fruit variety classification in uncertainty conditions of retail sales[J]. Applied Sciences, 2019, 9(19): 3971.

[11] HOLALAD S H, WARRIER P, SABARAD D A. An FPGA based efficient fruit recognition system using minimum distance classifier[J]. Journal of Information Engineering and Applications, 2012, 2(6): 1–10.

[12] ZHANG X, GAO Q, PAN D, et al. Research on spatial positioning system of fruits to be picked in field based on binocular vision and SSD model[J]. Journal of Physics: Conference Series, 2021, 1748(4): 042011.

[13] FAKHROU A, KUNHOTH J, MAADEED S A. Smartphone-based food recognition system using multiple deep CNN models [J]. Multimedia Tools and Applications, 2021, 80: 33011–33032.

[14] 张志远,罗铭毅,郭树欣,等. 基于改进 YOLO v5 的自然环境下樱桃果实识别方法[J]. 农业机械学报, 2022, 53(增刊 1): 232–240.
ZHANG Zhiyuan, LUO Mingyi, GUO Shuxin, et al. Cherry fruit detection method in natural scene based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp.1): 232–240. (in Chinese)

[15] 陈翠琴,范亚臣,王林. 基于改进 Mosaic 数据增强和特征融合的 Logo 检测[J]. 计算机测量与控制, 2022, 30(10): 188–194.
CHEN Cuiqin, FAN Yachen, WANG Lin. Logo detection based on improved Mosaic data enhancement and feature fusion[J]. Computer Measurement & Control, 2022, 30(10): 188–194. (in Chinese)

[16] 宋怀波,王亚男,王云飞,等. 基于 YOLO v5s 的自然场景油茶果识别方法[J]. 农业机械学报, 2022, 53(7): 234–242.
SONG Huaibo, WANG Yanan, WANG Yunfei, et al. *Camellia oleifera* fruit detection in natural scene based on YOLO v5s [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 234–242. (in Chinese)

[17] 于娟,罗舜. 基于 YOLO v5 的违章建筑检测方法[J]. 计算机工程与应用, 2021, 57(20): 236–244.
YU Juan, LUO Shun. Detection method of illegal building based on YOLO v5[J]. Computer Engineering and Applications, 2021, 57(20): 236–244. (in Chinese)

[18] 李书琴,陈聪,朱彤,等. 基于轻量级残差网络的植物叶片病害识别[J]. 农业机械学报, 2022, 53(3): 243–250.
LI Shuqin, CHEN Cong, ZHU Tong, et al. Plant leaf disease identification based on lightweight residual network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 243–250. (in Chinese)

[19] 李滔,董秀成,林宏伟. 基于深监督跨尺度注意力网络的深度图像超分辨率重建[J]. 电子学报, 2023, 51(1): 128–138.
LI Tao, DONG Xiucheng, LIN Hongwei. Depth map super-resolution reconstruction based on deeply supervised cross-scale attention network[J]. Acta Electronica Sinica, 2023, 51(1): 128–138. (in Chinese)

[20] 何东健,王鹏,牛童,等. 基于改进残差网络的田间葡萄霜霉病病害程度分级模型[J]. 农业机械学报, 2022, 53(1): 235–243.

HE Dongjian, WANG Peng, NIU Tong, et al. Classification model of grape downy mildew disease degree in field based on improved residual network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(1): 235–243. (in Chinese)

[21] 韩冉,曾广森,王荣杰. 基于残差块网络的图像去雨算法[J]. 仪器仪表学报, 2021, 41(8): 175–182.
HAN Ran,ZENG Guangmiao,WANG Rongjie. An image derain algorithm based on the residual block network[J]. Chinese Journal of Scientific Instrument, 2021, 41(8): 175–182. (in Chinese)

[22] 王泽钧,马凤英,张瑜,等. 基于注意力机制和多尺度轻量化网络的农作物病害识别[J]. 农业工程学报, 2022, 38(增刊1): 176–183.
WANG Zejun,MA Fengying,ZHANG Yu,et al. Crop disease recognition using attention mechanism and multi-scale lightweight network[J]. Transactions of the CSAE, 2022, 38(Supp.1): 176–183. (in Chinese)

[23] 石伟,王德雨,张元良. 基于图像增强的鲳鱼识别[J]. 江苏海洋大学学报(自然科学版), 2022, 31(2): 64–71.
SHI Wei, WANG Deyu, ZHANG Yuanliang. Image enhancement- based gadus recognition[J]. Journal of Jiangsu Ocean University (Natural Science Edition),2022, 31(2): 64–71. (in Chinese)

[24] 王美华,吴振鑫,周祖光. 基于注意力改进 CBAM 的农作物病虫害细粒度识别研究[J]. 农业机械学报, 2021, 52(4): 239–247.
WANG Meihua, WU Zhenxin, ZHOU Zuguang. Fine-grained identification research of crop pests and diseases based on improved CBAM via attention[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 239–247. (in Chinese)

[25] 吴海滨,魏喜盈,刘美红,等. 结合空洞卷积和迁移学习改进 YOLO v4 的 X 光安检危险品检测[J]. 中国光学, 2021, 14(6): 1417–1425.
WU Haibin, WEI Xiying, LIU Meihong, et al. Improved YOLO v4 for dangerous goods detection in X-ray inspection combined with atrous convolution and transfer learning[J]. Chinese Optics, 2021, 14(6): 1417–1425. (in Chinese)

[26] 陈品岚,张小花,朱立学,等. 基于树莓派及深度学习的柑橘识别系统设计[J]. 中国农机化学报, 2022, 43(9): 131–136.
CHEN Pinlan, ZHANG Xiaohua, ZHU Lixue, et al. Citrus recognition system design based on Raspberry Pi and deep learning [J]. Journal of Chinese Agricultural Mechanization, 2022, 43(9): 131–136. (in Chinese)

[27] 杨断利,李今,陈辉,等. 基于 IPSO 优化 BP 神经网络的蛋鸡舍有害气体监测系统[J]. 农业机械学报, 2021, 52(4): 327–335.
YANG Duanli, LI Jin, CHEN Hui,et al. Monitoring system of harmful gas in layer house based on improved particle swarm optimization BP neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 327–335. (in Chinese)

[28] 张伏,陈自均,鲍若飞,等. 基于改进型 YOLO v4-LITE 轻量级神经网络的密集圣女果识别[J]. 农业工程学报, 2021, 37(16): 270–278.
ZHANG Fu, CHEN Zijun, BAO Ruofei, et al. Recognition of dense cherry tomatoes based on improved YOLO v4-LITE lightweight neural network[J]. Transactions of the CSAE, 2021, 37(16): 270–278. (in Chinese)

[29] 何雨,田军委,张震,等. YOLO v5 目标检测的轻量化研究[J]. 计算机工程与应用, 2023, 59(1): 92–99.
HE Yu, TIAN Junwei, ZHANG Zhen, et al. Lightweight research of YOLO v5 target detection[J]. Computer Engineering and Applications, 2023, 59(1): 92–99. (in Chinese)

[30] 丁奇安,刘龙申,陈佳,等. 基于 Jetson Nano + YOLO v5 的哺乳期仔猪目标检测[J]. 农业机械学报, 2022, 53(3): 277–284.
DING Qi'an, LIU Longshen, CHEN Jia, et al. Object detection of suckling piglets based on Jetson Nano and YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 277–284. (in Chinese)