

基于改进 SSD 的青瓜检测算法^{*}

曾 乾¹ 李 博²

(1. 电子科技大学机械与电气工程学院 成都 611731; 2. 电子科技大学中山学院机电工程学院 中山 528402)

摘 要:针对复杂近色背景下青瓜目标识别率低、定位效果不佳等问题,提出一种基于 SSD 的循环融合特征增强(CFFE-SSD)目标检测模型。首先,对 SSD 的前 4 个有效特征层进行循环特征融合,使低层特征层和高层特征层的信息得到有效利用;其次,针对青瓜目标的特殊长宽比以及重叠现象,使用 K-means 算法改进先验框的默认尺寸以及长宽比,提出以 DIoU-NMS 替换普通 NMS;最后,将 ECA 注意力机制引入循环特征融合模块,增强网络特征提取能力。实验结果表明,改进 CFFE-SSD 模型 AP@0.5 达到了 96.63%,提升了 4.61%;AP@0.75 达到了 89.02%,提升了 7.14%,检测速度达到 144 fps,边框回归精度更高,能有效满足青瓜自动采摘的需求。

关键词:青瓜检测;图像识别;深度学习;卷积神经网络;目标检测

中图分类号: TP183 **文献标识码:** A **国家标准学科分类代码:** 520

Cucumber detection algorithm based on improved SSD

Zeng Qian¹ Li Bo²

(1. College of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; 2. College of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan 528402, China)

Abstract: Aiming at the problems of low target recognition rate and poor localization effect of cucumber in complex near-color background, a cyclic fusion feature enhanced based on SSD (CFFE-SSD) target detection model is proposed. First, cyclic feature fusion is performed on the first four effective feature layers of SSD, so that the information of low-level feature layers and high-level feature layers can be effectively used. Secondly, in the view of special aspect ratio and overlapping phenomenon of the cucumber target, the K-means algorithm is used to improve the default size and aspect ratio of the a priori frame, and meanwhile, DIoU-NMS is proposed to replace ordinary NMS. Finally, the efficient channel attention module is introduced into the cyclic feature fusion module to enhance the feature extraction ability of the network. The experimental results show that the AP@0.5 of the improved CFFE-SSD model proposed in this paper reaches 96.63%, an increase of 4.61%, and the AP@0.75 reaches 89.02%, an increase of 7.14%. The detection speed reaches 144 fps, and the frame regression accuracy is higher. Effectively meet the needs of automatic cucumber picking.

Keywords: cucumber detection; image identification; deep learning; convolutional neural network; target detection

0 引 言

青瓜果实的目标检测是实现青瓜自动化采摘的关键步骤,青瓜检测的效果将直接影响青瓜采摘的结果。青瓜

与其他农作物相比有其独特的特点,不同于番茄、苹果等,成熟青瓜的颜色与青瓜叶、藤的颜色非常相近,而且青瓜叶片宽大,容易对果实造成遮挡,这给青瓜的识别检测带来很大挑战^[1]。过去研究人员主要采取利用光谱特性或

收稿日期:2023-02-15

^{*} 基金项目:2021年广东省教育厅普通高校重点领域专项(421N34)、2020年广东省普通高校创新团队项目—机器人与智能装备团队(2020KCXTD035)项目资助

者基于人工特征的传统图像处理方式对青瓜进行识别, Henten 等^[2]采用近红外传感器识别黄瓜果实。Zhang 等^[3]以 RGB 颜色空间中的 B 分量和 HIS 颜色空间中的 S1 分量对青瓜实现粗分割,再利用形态学方法对二值分割图进行滤波处理、最后对连通区的形状和纹理特征进行分析实现青瓜识别。王海青等^[4]利用脉冲耦合神经网络(pulse coupled neural network, PCNN)对青瓜与背景实现粗分割后提取分割图中连通区域的几何形状特征和纹理特征,然后采用支持向量机(support vector machine, SVM)进行特征分类实现青瓜的识别。Bao 等^[5]采用多模板匹配算法进行青瓜的识别定位,但由于模板库的限制对不同距离的青瓜识别效果较差。以上方法的识别准确率较低,对复杂的自然环境不能很好地适应。

近年来随着计算机硬件性能的大幅提升,基于卷积神经网络(convolutional neural network, CNN)^[6]的目标检测模型快速发展并且在各个领域得到广泛应用。深度学习目标检测模型鲁棒性强,对目标有着更强的泛化能力,相比传统机器视觉更适用于复杂环境下的果蔬检测。目标检测模型分为以区域卷积神经网络(region-CNN, RCNN)^[7-9]系列为代表的二阶段模型和以单阶段多框检测器(single shot multibox detector, SSD)^[10]、YOLO(you only look once)^[11]为代表的一阶段模型。二阶段模型对目标有着较好的检测定位效果,但模型参数量较大,检测速度较慢,达不到实时检测的要求;一阶段模型实现了端对端的检测,免去了候选框的提取步骤,在不过多损失分类精度的情况下有着更快的检测速度。吴华昕^[12]通过使用 I-RELIEF^[13]选择的特征通过 LeNet5^[14]神经网络提取青瓜图像特征,然后再将训练特征输入到支持向量机进行特征分类实现对青瓜的识别,在无遮挡情况下对青瓜达到较为准确的识别检测。刘晓洋^[15]通过改进能够实现像素级分

割的掩膜区域卷积神经网络(mask region convolution neural network, Mask-RCNN)^[16]网络实现了对青瓜的检测,使其检测的定位精度得到提升,但网络计算量过大,不利于移植到移动设备,也不利于实现实时检测。目前将深度学习应用在果蔬目标检测的研究虽然有很多,如番茄、苹果、柑橘、荔枝、芒果等,这些水果大都与环境颜色对比度高,且形状多为近似球型,相比于青瓜更易于检测,因此研究利用深度学习对复杂背景下的青瓜检测具有重要的实际意义和应用价值。

综上所述,针对青瓜目标的特点和采摘的应用要求,模型的检测速度要求不低于 100 fps,以保证整个采摘过程的效率,考虑后期的移动端植入,模型规模不宜大于 100 MB,自动采摘虽然仍需要人工补充采摘,但要求完成大部分采摘工作,因此模型检测精度(mAP@0.75)要求尽可能高,至少达到 85%以上。基于以上目标,本文提出一种基于 SSD 网络的改进青瓜检测模型,引入循环特征融合模块对 SSD 网络的前面 4 个较大的特征图特征融合,使低层特征利用高层的语义信息提升效果,同时高层特征获得低层特征中更多的位置信息。针对青瓜长宽比较大的特点,通过 K-means 聚类方法对统计的自制青瓜数据集中青瓜目标的长宽比进行聚类,根据聚类结果调整先验框的长宽比,并且将 IoU-NMS 替换成 DIoU-NMS,使预测框和真实框更好地匹配,提升检测和定位精度。在循环特征融合模块重要位置加入通道注意力机制模块,使网络具有跨通道信息获取能力。提升网络的特征提取能力。

1 SSD 目标检测模型

SSD 是一种基于卷积神经网络的单阶段目标检测模型,其基本结构如图 1 所示。

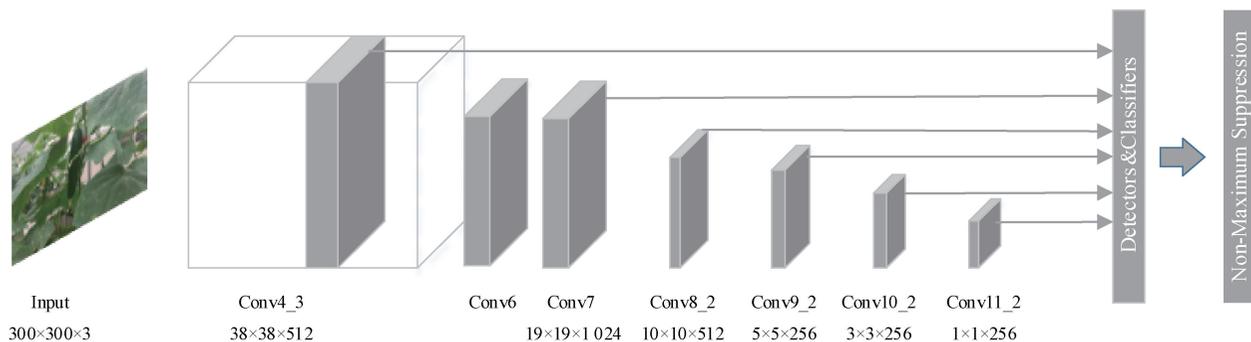


图 1 SSD 模型结构

其使用 VGG16 网络^[17]作为骨干特征提取网络,将 VGG16 的 FC6 和 FC7 层替换成卷积层 Conv6、Conv7,并在后面新增 Conv8、Conv9、Conv10、Conv11 卷积层,然后以 Conv4_3、FC7、Conv8_2、Conv9_2、Conv10_2、Conv11_2 特征层作为有效特征层并对其进行下一步处理来获取预测结果,其相应特征图的尺寸分别为 38×38、19×19、

10×10、5×5、3×3、1×1,分别来处理不同尺度的目标。对上述获取到的每一个有效特征层,都进行一次 num_anchors×4 和一次 num_anchors×num_classes 的卷积,其中 num_anchors 指对应特征层每一个特征点所拥有的先验框数量,num_classes 为预测目标的种类。上述 6 个特征层的每个特征点对应的先验框数量分别为 4、6、6、6、4、

4,其中 Conv4_3 层的默认框尺寸为 30,其他默认框尺寸由式(1)计算可得。

$$S_k = S_{\min} + \frac{S_{\max} - S_{\min}}{m - 1}(k - 1), k \in [1, m] \quad (1)$$

式中: m 为用于预测的特征图个数; k 代表层数; S_{\min} 取 0.2; S_{\max} 取 0.9。候选框的长宽比设定为 $\{1, 2, 3, 1/2, 1/3\}$ 。SSD 的这种先验框机制利用了浅层特征图较小的感受野和丰富的位置信息检测小尺寸目标,能够较准确定位目标位置,但浅层特征图语义信息表达能力弱,不利于确定目标种类。利用深层特征图较大的感受野和丰富的语义信息检测大尺寸目标,能够较好的对其进行分类,但

位置信息表达能力不如浅层特征。

2 改进 SSD 模型

SSD 模型利用多尺度特征图进行多尺度预测,从而改善了 Faster-RCNN 检测速度慢的问题,但同时由于浅层特征图包含的语义信息较少,在对青瓜这种与背景相似度较高的目标时容易产生漏检和误检现象,并且青瓜目标存在较多重叠、被遮挡的情况。针对以上问题,本文提出一种基于改进 SSD 的循环融合特征增强(cyclic fusion feature enhanced-SSD, CFFE-SSD)的目标检测算法,其结构如图 2 所示。

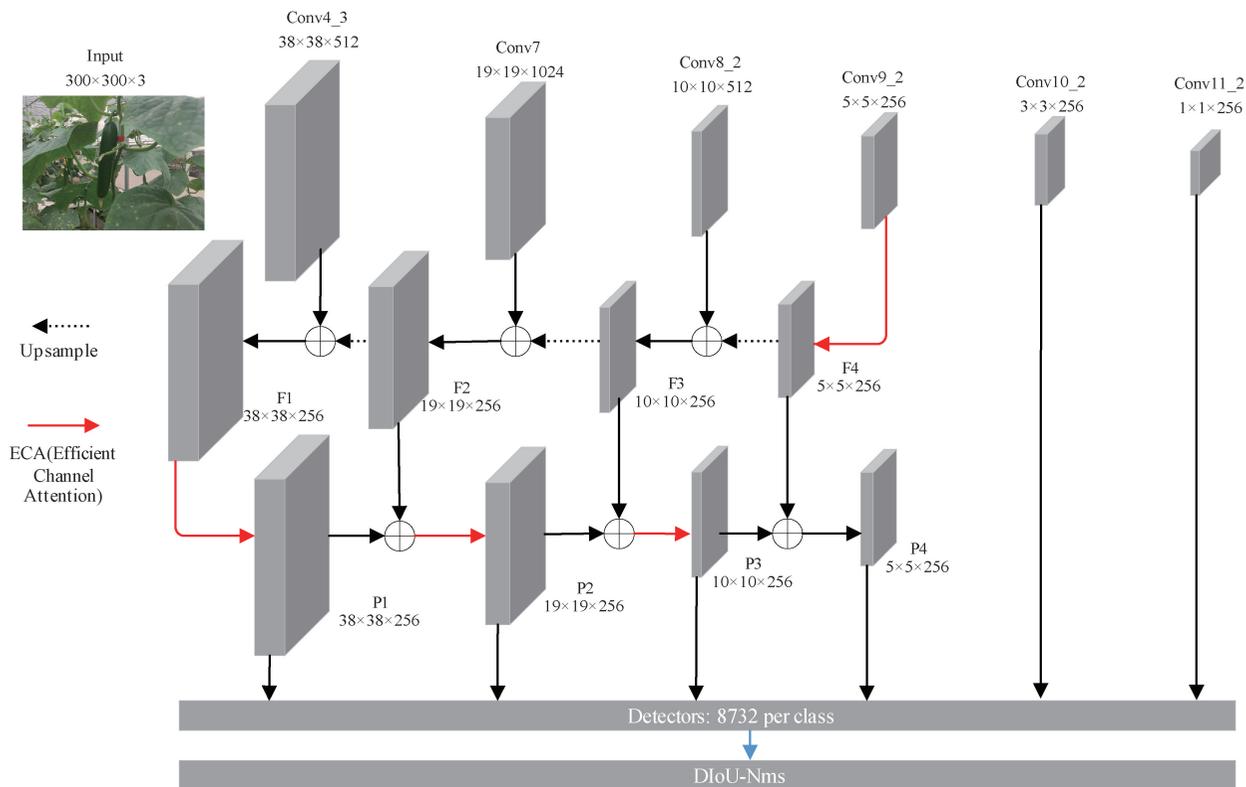


图 2 CFFE-SSD 模型结构

考虑到青瓜目标在大棚环境中与背景相似度高这一特点,如何区分青瓜与背景特征从而提升检测效果成为研究重点。因此本文算法借鉴了特征金字塔网络^[18](feature pyramid networks, FPN)和 PANet^[19](path aggregation network)网络的思想,通过一条 Up-bottom Path 和一条 Bottom-up Path 循环融合浅层特征图和深层特征图,利用深层语义信息来提升浅层的效果,同时浅层到深层的通道提高浅层信息的利用率,使不同深度的特征层语义和定位信息表达能力得到提升;另外在循环特征融合模块的关键位置嵌入 ECA (efficient channel attention) 模块,提升网络的跨通道信息获取能力,让网络自适应地注意重要的物体;以 DIoU-NMS^[20] 替换掉 IoU-NMS 实现更好的定位效果且一定程度解决漏检问题。除此之外针对青瓜

目标特殊的长宽比,利用 K-means 算法对先验框的长宽比和尺寸进行优化,使其更适应青瓜目标。

2.1 循环特征融合

现有 SSD 模型的多尺度检测中,采用的 VGG16 骨干网络不深,在保证网络模型小的同时也使其精度收到一定影响,本文提出的循环特征融合模块在不加深网络卷积层数的情况下使得原来的深层特征层和浅层特征层能够互相利用到彼此的特征信息,在小幅增加网络模型体积的情况下大大提升网络精度。本文的特征融合策略为在 Up-bottom Path 中,高层特征图通过最近邻域插值实现上采样从而保证与低层特征图维度相同,如图 2 所示,相对于双线性插值其算法简单且保留了初始特征层的像素信息,符合特征融合提高信息利用率的策略,插值方式如图 3 所示。

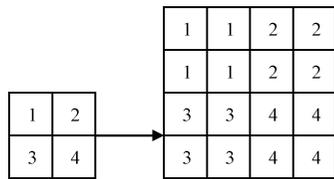


图3 最近邻域插值

现有的特征融合方式主要有 Concat 和 Add 两种, Concat 融合在通道维度上对原有的特征图进行拼接, 输出特征图的高和宽与输入特征图相同, 融合后通道数为待融合特征图通道数的和。Add 融合则通过对对应特征点相加的方式进行融合, 在不改变通道数的情况下增加特征层单个通道所包含的特征信息。两种融合方式如图 4 所示。两种融合方式均要求连接的特征层拥有相同的长宽维度, Add 融合要求连接的特征层有相同的通道数。Add 不会引入额外的卷积参数, 计算量比 Concat 小, 本文选择 Add 特征融合方式。

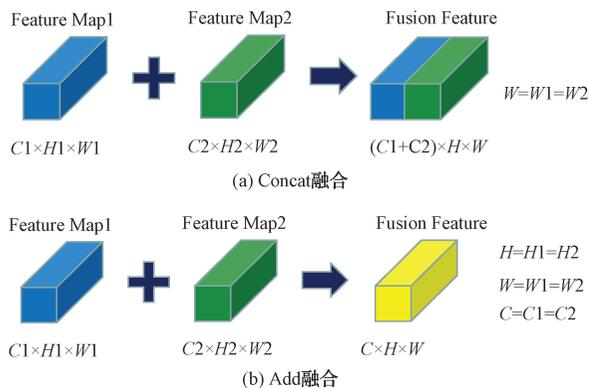


图4 特征融合方式

在 Bottom-up Path 中的浅层特征图和两条路径经过特征融合后得到的特征图还需要 3×3 的卷积。普通 3×3 卷积计算量大, 本文将普通 3×3 卷积替换成深度可分离卷积^[21-22], 将其卷积分为逐通道卷积和逐点卷积, 减少计算量, 加快检测速度。以 $W_F \times W_F$ 表示输入特征图的尺寸, 以 $W_K \times W_K$ 代表卷积核的尺寸, 以 M, N 分别表示输入和输出的通道数, 且一般卷积核和特征图的长和宽相等, 普通卷积与深度可分离卷积计算量对比如式(2)、(3)所示。

普通卷积为:

$$W_K \times H_K \times M \times N \times W_F \times W_F \quad (2)$$

深度可分离卷积为:

$$W_K \times H_K \times 1 \times M \times W_F \times H_F + 1 \times 1 \times M \times N \times W_F \times H_F \quad (3)$$

深度可分离卷积在达到与普通卷积相同效果的同时, 使得参数量能够减少至原来的 $1/9 \sim 1/8$, 降低资源消耗从而加快模型的计算速度。

2.2 先验框改进

根据式(1)计算可得 SSD 模型的先验框 6 个有效特

征层的最小默认框尺寸为 30、60、111、162、213、264, 最大默认框尺寸为 60、111、162、213、264、315。这些尺寸是针对大型数据集中各类物体的平均尺寸设计的, 用于多类物体检测具有较好的鲁棒性, 但对单类物体的目标检测不具有专一性。针对青瓜目标, 其平均尺寸相对大型数据集中物体较小, 因此将 S_{\min} 设置为 0.14, 调整后 6 个有效特征层的最小默认框尺寸 21、45、99、153、207、261, 最大默认框尺寸为 45、99、153、207、261、315。除此之外, 对于青瓜目标, 因其近似圆柱体的形状导致其在图像中拥有的长宽比有一定特殊性, 因此本文首先统计了自制青瓜数据集中共计 28 348 个青瓜目标的长宽比, 如图 5 所示。

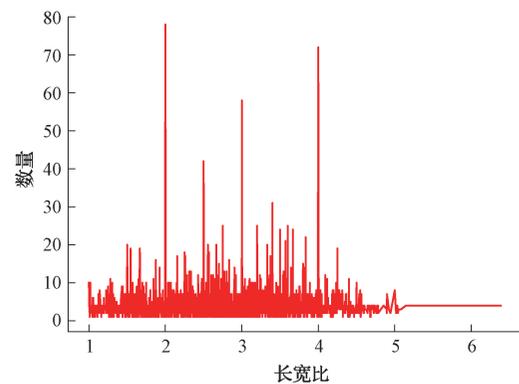


图5 青瓜目标长宽比

然后用 K-means 聚类^[23-24]确定先验框的高宽比和尺度, 根据得到的聚类结果调整 CFFE-SSD 模型的长宽比设定, 以此提升先验框和真实框的匹配效果, 提高模型的检测效果, 聚类结果如图 6 所示。

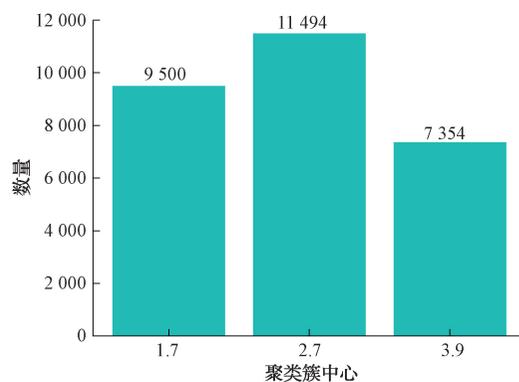


图6 K-means 聚类结果

根据聚类结果, 将原始 SSD 模型先验框长宽比设定 $\{1, 2, 3, 1/2, 1/3\}$ 调整为 $\{1, 2.7, 3.5, 10/27, 2/7\}$, 调整后的默认框尺寸及长宽比设定与真实框更接近从而实现更好的匹配, 加速模型收敛速度, 提升目标的检测效果。

在传统 SSD 算法中, 通常利用 NMS 算法去除冗余候选框, 普通的 IoU 对预测框与真实框的重合度不敏感, 因为其只考虑了两者的交并比。若两个青瓜目标距离较近甚至有重叠情况, 经过 NMS 处理后只留下一个检测框, 而

DIoU 将真实框与预测框之间的距离, IoU 和尺度都纳入了考虑范围, 其示意图如图 7 所示, 计算方法如式(4)所示。

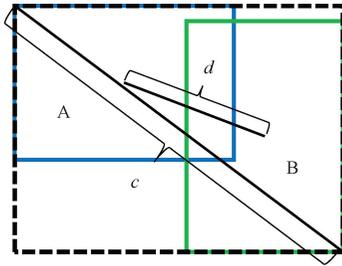


图 7 DIoU 示意图

其中 A 为真实框, 对应式(4)中的 A, B 为预测框, 对应式(4)中的 B, c 为预测框与真实框最小外接区域黑色框的对角线距离, d 为预测框与真实框的中心点的欧氏距离。

$$DIoU = \frac{(A \cap B)}{(A \cup B)} - \frac{d^2}{c^2} \quad (4)$$

式中: $A \cap B$ 表示真实框与预测框重叠区域的面积; $A \cup B$ 表示真实框与预测框最小外接区域的面积。式(4)去掉后面的项即退化为 IoU。相对于普通 IoU, DIoU 使得目标框的回归变得更稳定, 因此用 DIoU-NMS 替换原来的 IoU-NMS, 能够实现更好的定位效果, 同时一定程度降低有重叠目标出现漏检或误检的情况。

2.3 注意力机制 ECA 模块

在目标检测中, 更希望卷积神经网络去关注应该注意的地方, 即目标物体所在的地方。注意力机制就是实现网络自适应注意的一种方式, 本文使用的 ECA 模块^[25]是一种通道注意力机制, 通过跨通道交互策略, 在几乎不增加计算量情况下能使网络对复杂背景下的青瓜目标赋予更多的关注度, 有效提高模型的检测性能。该模块首先对输入的特征图 $F \in R^{H \times W \times C}$ 进行全平均池化, 然后通过一维卷积得到通道间的相关性参数 l_i , 其中卷积核大小 k 由通道数 C 自适应计算得到, 计算公式为:

$$W_i = \sigma(l_i) \quad (5)$$

$$k = \psi(C) = \left\lfloor \frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (6)$$

其中, $\gamma=2, b=1$ 。最后将通道注意力权重 W_i 与原始特征 F 相乘得到输出特征信息 $F' \in R^{H \times W \times C}$ 。其结构如图 8 所示。

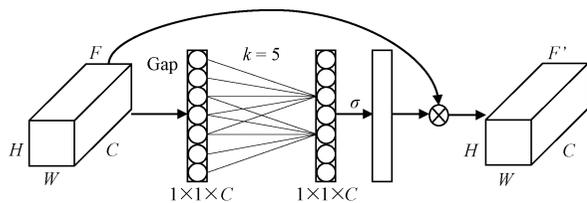


图 8 通道注意力模块

3 实验与分析

3.1 实验数据集和测试平台

青瓜数据集采集于广东省珠海市金湾区, 采集设备为 Intel RealSense D435 相机, 总共采集有效图像 2 000 张, 为考虑后续采摘应用, 采集的图像包含多种角度, 多种方位和亮度等不同条件, 保证样本多样性, 提高泛化能力。

在经过左右翻转, 上下翻转, 高斯噪声等数据增强方式将数据集扩充至 8 000 张, 使用图像标注工具 Labelimg 进行手动标注, 并且忽略图像中遮挡超过 80% 的青瓜目标, 对于被叶片截断进能看见首尾部分的目标, 按照一个目标进行标注。按照 8:1:1 的比例将其随机划分为训练集、验证集和测试集数据集样本如图 9 所示。



图 9 青瓜数据集

实验所用到的硬件平台和软件系统如表 1 所示。

表 1 实验硬件平台和软件系统

硬件/软件	型号
处理器	AMD Ryzen 9 5900X@3.70 GHz
显卡	NVIDIA GeForce RTX3090(24 GB)
内存	64 GB
操作系统	Windows 10
运行平台	Python 3.7
运行框架	Pytorch 1.7.1(torchvision 0.8.2)
CUDA/CUDNN	11.1 / 8.0.5

3.2 训练策略

在训练模型时, 首先采用 Pytorch 官方提供的 COCO 数据集的训练权重作为预训练权重, 防止主干特征提取网络的权值太过随机, 特征提取效果不明显。共训练了 500 个 epoch, 训练 loss 和验证 loss 均得到收敛, 如图 10 所示。

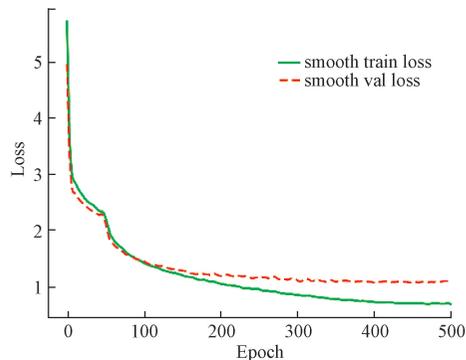


图 10 训练过程训练集和验证集 loss 变化曲线

其中实线为训练集,虚线为验证集。训练前 50 个 epoch 对主干网络进行冻结,批大小(batch size)设置为 32,采用的优化器为 SGD 优化器,初始最大学习率(learning rate)设置为 0.001,最小学习率设置为最大学习率的 0.01 倍,学习率调整方式采用余弦退火算法,优化器动量参数 momentum 设置为 0.937,权重衰减(weight decay)设置为 0.0005。

3.3 评价指标

实验采用的评价指标有平均精度(average precision, AP)衡量模型检测效果,考虑到后续采摘任务,除了考量 AP@0.5,另加入 AP@0.75 来评价边框回归的准确度提升效果,AP@0.x 表示置信度阈值 IoU 为 0.x,该值越高表明模型的精确度越高。帧率衡量模型检测速度,模型大小来衡量模型的轻量级别。

AP 表示在不同召回率(recall, R)下精确率(precision, P)即预测率-召回率(P-R)曲线与坐标轴围成的面积,如式(7)所示。

$$AP = \int_0^1 P(R) dR \quad (7)$$

式中:P 为查准率即正确分类的正样本占所有正样本的比例;R 为查全率即预测到的正样本占真正样本的比例。P 和 R 公式为:

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

式中:TP 为正确分类的正样本数量;FP 为错误分类的正样本数量;FN 表示将正样本错误分类为负样本的数量。由各类的 AP 可计算出平均精度均值(mAP),公式为:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (10)$$

式中:N 为预测的类别数;AP_i 为第 i 个类别对应的 AP 值,由于目标种类只包含青瓜这一类,因此 mAP 等于平均精度。

3.4 检测模型对比

为验证本文模型的有效性,在相同实验条件下对当前常见目标检测模型进行对比。为控制变量,将各个模型的输入图像分辨率设置(resize)在 300×300 左右,由于 YOLO 系列网络分辨率必须为 32 的倍数,因此 YOLO 系列网络输入图像分辨率设为 320×320,而 Faster RCNN 网络模型则由于限制规则将图像分辨率吧设为 600×600。

对比结果如表 2 所示,与经典目标检测算法 Faster RCNN、YOLOv3、YOLOv4 相比。本文模型的 AP@0.5 值和 AP@0.75 值均是所有模型中最高的,相对于原 SSD 模型分别提高了 4.61%和 7.14%,说明无论在检测精度还是边框回归精度上较原网络都有很大提升。虽然与 YOLOv5l 的结果比较,仅相差 0.01%和 0.05%,但其模型大小仅为对方的 1/2,且检测速度将近对方的 3 倍。从

检测速度上看,CFFE-SSD 的 144 fps 虽然比原模型的 157 fps 有所降低,但幅度很小,依旧远超其他模型,达到其他模型的 2~3 倍,能够实现较好的实时检测效果。从模型大小上看,改进网络模型相比原 SSD 只增加了 2 MB,相对于其他网络模型有更小的体积,更有利于后续在移动端设备上的使用。

表 2 不同模型的对比结果

模型	主干网络	AP@0.5/%	AP@0.75/%	帧率 /fps	模型大小/MB
Faster-RCNN	VGG16	91.23	81.93	24	521
YOLOv3	DarkNet53	94.66	86.51	73	235
YOLOv4	CSPDarkNet53	92.74	82.65	53	244
YOLOv5l	CSPDarkNet53	96.62	88.97	52	178
SSD	VGG16	92.02	81.88	157	91
CFFE-SSD	VGG16	96.63	89.02	144	93

3.5 不同改进点对模型的影响实验

在改进模型 CFFE-SSD 中,作了如下 3 种改进:1) 对前 4 层有效特征层进行循环特征融合,用 F 表示;2) 改进默认先验框,并用 DIoU-NMS 替换普通 NMS 重新设计先验框生成规则,用 A 表示;3) 引入通道注意力机制 ECA 模块,用 E 表示。为了进一步探究不同改进点对模型的影响,本文研究做了一系列消融实验。将上述 3 种改进点进行组合,探究每种组合对模型的影响。实验结果如表 3 所示。实验共包括 6 组,其中组别 1 为原 SSD 模型,组别 2 为单独加入循环特征融合模型。组别 3 为单独改进先验框的模型。其他组别为 3 种改进的组合实验。由于 ECA 模块嵌入在特征融合模块中,因此只在有特征融合模块的组合中加入 ECA 模块的对比。

表 3 不同改进点对模型的影响

组别	改进点	AP@0.5/%	AP@0.75/%
1	—	92.02	81.88
2	F	94.07	84.71
3	A	94.02	85.00
4	F+A	95.54	88.31
5	F+E	95.97	85.96
6	F+A+E	96.63	89.02

第 2 组和第 3 组的结果表明,在单独加入循环特征融合结构和单独改进先验框时,网络的检测精度都有较大提升,其中改进网络结构 AP@0.5 和 AP@0.75 分别提升了 2.05%和 2.83%,说明循环特征融合使模型对近色背景下青瓜的检测进行了改善;而单独改进先验框则分别提升了 2.00%和 3.12%,说明先验框的调整和 DIoU-NMS 对边框回归精度有较好的提升;第 4 组和第 5 组的结果显示,在循环融合特征的基础上改进先验框使 AP@0.75 提

升了6.43%，加入ECA模块使AP@0.5提升了3.95%，进一步说明先验框的改进在边框回归精度上起到更大作用，而ECA模块的注意力机制使模型在检测精度上有更好的效果；第6组在同时加入循环特征融合、先验框改进以及ECA模块时，AP@0.5和AP@0.75分别提升了4.61%和7.14%，同时模型大小从91MB增加到93MB，综上实验结果表明本文模型的改进能在小幅增加模型体积的情况下改善模型的检测精度和边框回归精度，同时保证模型检测速度能够达到应用要求。

3.6 检测效果对比

将测试集中的图像分别输入原SSD模型与CFE-SSD模型进行检测，部分样本的检测效果对比如图11所示。其中，图11(a)~(c)为原SSD模型的检测效果，图11(d)~(f)为改进SSD模型的检测效果。其中

图11(a)中青瓜目标在无遮挡的情况时出现漏检现象；图11(d)则精确检测出未被遮挡和局部被藤蔓遮挡的青瓜目标；图11(b)为逆光角度，青瓜果实颜色和纹理等特征受光照影响，其中重叠青瓜果实出现漏检的现象；在图11(e)中CFE-SSD模型检测出被遮挡青瓜目标；图11(c)中两根青瓜由于重叠被误检为一根青瓜，并且右下角的预测框过大；图11(f)中重叠的青瓜果实被正确检测，并且预测框与青瓜果实边缘呈现相切的效果，相比原模型的预测框回归精度更高，定位更精准。

综上所述，CFE-SSD模型相对于原SSD模型在检测性能和边框回归精度上都提升，特别在边框回归精度方面得到了有效改善，更有利于对青瓜目标的精准空间定位，对近色背景下的绿色果实检测有着很好的泛化能力和鲁棒性。

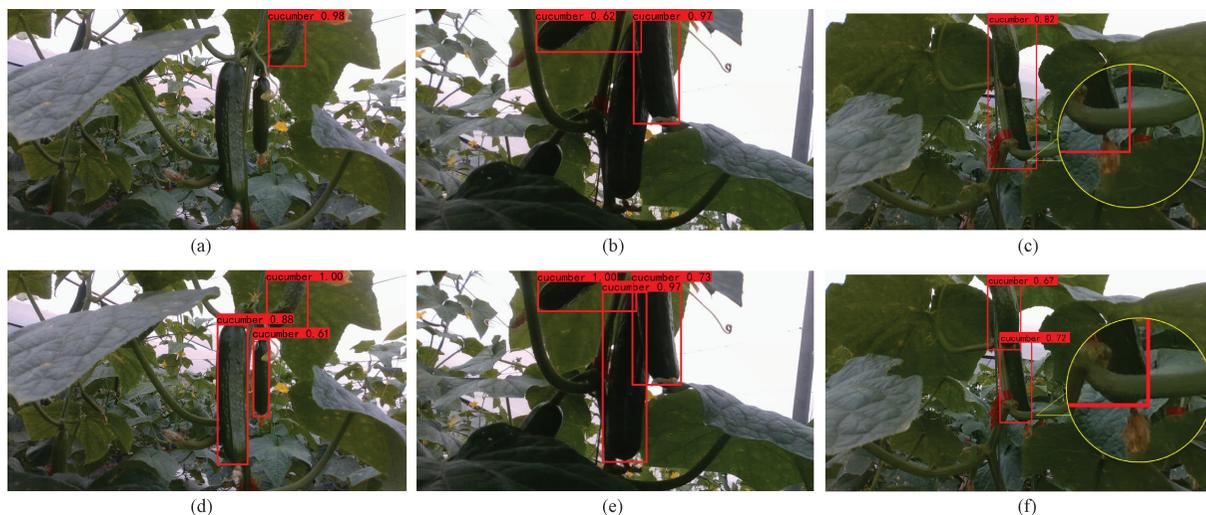


图11 SSD模型与CFE-SSD模型检测效果对比

4 结论

针对复杂近色背景下青瓜目标难以检测这一问题，本文基于SSD模型进行改进，使用了一种新的方式改善了青瓜检测的效果。首先对浅层的有效特征层加入循环特征融合模块，使低层特征图的位置信息信息和高层特征图的语义信息得到有效利用；其次针对青瓜目标的特殊比例以及果实重叠现象多这一现状，改进了模型的先验框尺寸和比例，并用DIoU-NMS替代传统NMS提升模型的检测精度、速度以及边框回归精度；最后加入通道注意力机制ECA模块进一步提升模型检测效果。实验结果表明，改进后的模型有效提升了近色背景下青瓜目标的检测率，改善了重叠青瓜目标误检、漏检的现象，提升了模型的边框回归精度，有利于后续结合双目视觉的青瓜自动采摘研究。

参考文献

[1] ARIMA S, KONDO N. Cucumber harvesting robot

and plant training system[J]. Journal of Robotics and Mechatronics, 1999, 11(3):208-212.

[2] HENTEN E J, HEMMING J, VAN TUIJL B A J, et al. An autonomous robot for harvesting cucumbers in green-house [J]. Autonomous Robots, 2002, 13(3): 241-258.

[3] ZHANG L, YANG Q, XUN Y, et al. Recognition of greenhouse cucumber fruit using computer vision[J]. New Zealand Journal of Agricultural Research, 2007, 50(5): 1293-1298.

[4] 王海青, 姬长英, 顾宝兴, 等. 基于机器视觉和支持向量机的温室黄瓜识别[J]. 农业机械学报, 2012, 43(3): 163-167, 180.

[5] BAO G, CAI S, QI L, et al. Multi-template matching algorithm for cucumber recognition in natural environment[J]. Computers and Electronics in Agriculture, 2016, 127: 754-762.

[6] 王涛, 鲁昌华, 孙怡宁, 等. 多尺度卷积神经网络检测

- 睡眠呼吸暂停[J]. 电子测量与仪器学报, 2021, 35(7):30-35.
- [7] GIRSHICK R, DONAHUE J. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). IEEE Computer Society, 2014: 580-587.
- [8] GIRSHICK R. Fast R-CNN[C]. 2015 IEEE Conference on Computer Vision. IEEE Computer Society, 2015: 1440-1448.
- [9] REN S, HE K, 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, ANGELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Proceedings of the European Conference on Computer Vision. Springer, 2016: 21-37.
- [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 (CVPR). Seattle: IEEE Computer Society, 2016: 779-788.
- [12] 吴华昕, 黄瓜智能采摘机器人关键技术研究[D]. 扬州:扬州大学, 2021.
- [13] SUN Y. Iterative RELIEF for feature weighting: Algorithms, theories, and applications [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6): 1035-1051.
- [14] LE CUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11), 2278-2324.
- [15] 刘晓洋. 多功能果蔬采摘机器人的图像识别算法研究[D]. 镇江:江苏大学, 2020.
- [16] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386-397.
- [17] SIMONVAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]. Computer Vision and Pattern Recognition, 2014.
- [18] LIN T, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [19] 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, 2018: 8759-8768.
- [20] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [21] 蒋晨, 钱永明, 姚兴田, 等. 基于可变形卷积改进 SSD 算法的目标检测方法[J]. 电子测量技术, 2022, 45(16):116-122.
- [22] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[C]. Computer Vision and Pattern Recognition, 2017.
- [23] 向宽, 李松松, 栾明慧, 等. 基于改进 Faster RCNN 的铝材表面缺陷检测方法[J]. 仪器仪表学报, 2021, 42(1):191-198.
- [24] GIRDHAR G L, RAVI K. An efficient distance estimation and centroid selection based on K-means clustering for small and large dataset [J]. International Journal of Advanced Technology and Engineering Exploration (IJATEE), 2020, 7(73): 234-240.
- [25] 张奇, 陈梦蝶, 赵杰. 融合特征金字塔和通道注意力的轻量车辆检测算法[J]. 国外电子测量技术, 2023, 42(1):41-48.

作者简介

曾乾, 硕士研究生, 主要研究方向为机器视觉、深度学习。

E-mail: EthanQ_97@163.com

李博, 硕士, 副教授, 主要研究方向为机器视觉检测、工业自动化。

E-mail:superblee@163.com