

基于 YOLOv5 改进算法的桃子瑕疵部位检测

李俊恩¹ 罗志君² 陈幸萍¹

(1. 五邑大学, 广东 江门 529020;

2. 数字江门网络建设有限公司, 广东 江门 529020)

摘要: 桃子瑕疵检测可以在桃果成长期及时发现病害, 在桃果采摘期分出优级果。本文采用轻量级的 YOLOv5s 目标检测模型, 实现对桃子瑕疵部位的检测。数据集由自采图像、公开数据集、网络图片三个渠道获取, 从而使得训练结果具有较好的准确率和跨环境的适应性。最后将训练得出的模型移植到小型设备, 实验结果显示, 模型在推理中仍表现出良好的准确率和处理速度。因此, 本算法能够应用于便携式设备中, 实现对桃子瑕疵部位的实时检测。

关键词: YOLOv5; 目标检测; 人工智能模型

一、引言

近年来随着人们的生活水平、生活质量逐步提高, 居民可支配收入呈现持续增长趋势, 我国水果需求量不断上升, 水果市场规模也稳步提升, 因此果园面积近几年也在不断扩大。我国是世界桃生产第一大国, 在广东粤北山区有成片的鹰嘴桃、九仙桃种植区域。这些地方的桃果种植园通常依山而建, 山坡地带并不利于传统农业管理和大规模机械化生产。产区每年 6 月中至 7 月中将有大量桃果上市, 果园管理、果实采摘等工作都将突发需求大量劳动力, 而由于当前农村年轻劳动力紧缺, 因此桃果种植户对果园智慧化管理的需求就显得越来越迫切。

据前期调研, 桃果的瑕疵部位检测是其中一项重要工作。特别是在桃子成熟期, 果农需要时常关注桃子的健康情况。一是广东地区龙舟雨时期, 刚好也是桃果的快速成长期, 此期间高温高湿的环境, 容易导致大片果树的桃果出现坏点, 需要紧急干预控制, 以保障桃果的产量。二是对于成熟的桃果, 果皮品相好的果实出售价格较高, 因此需要将果皮有瑕疵的桃果检测出来, 从而提高农户的收入。

因此, 本文针对桃果瑕疵部位检测, 提出采用基于人工智能的图像检测算法, 实现对桃子瑕疵部位的自动检测。首先收集桃果数据集, 然后标记其中的瑕疵部位, 将数据集用于智能模型训练, 最后测试后将算法移植到小型设备, 实现对桃果瑕疵部位的实时检测。本算法可应用于无人机巡检、桃果分类装箱, 辅助果农提高果园管理效率。

二、桃子瑕疵检测算法

对于桃子瑕疵部位检测, 也就是桃子好坏分类的问题, 可以采用 YOLOv5 模型进行实现。在这个任务中, 需要训练一个目标检测模型, 使其能够准确地检测出桃子并判断其好坏。

经过 YOLO 系列的多个迭代, YOLOv5 成为目标检测领域经典模型之一, 其从检测精度和端对端的图片检测速度上来说都较为优秀。YOLOv5 提供了四个基准模型, 根据检测任务的复杂度及速度需求可分为 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。其中 YOLOv5s 是最小的 YOLOv5 模型, 速度快, 适用于资源受限的设备和轻量级应用, 因此本文选取 YOLOv5s 模型进行桃子好坏检测任务, 将其作为基准模型进行改进。YOLOv5s 模型主要包括 Input、Backbone、Neck 以及 Head 这四个部分如图 1。

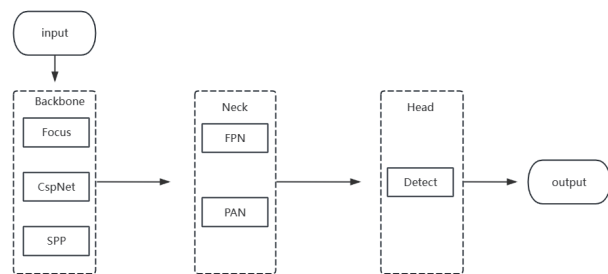


图 1 YOLOv5s 模型结构图

在 YOLOv5 的 input 中, 对输入图像进行尺寸调整、归一化、色彩空间转换等, 会对数据集进行数据增强, 例如: Mosaic, 将四张图片拼成一张图片; Copy paste, 将部分目标随机粘贴在图片中, 前提是数据要有 segments 数据才行, 即每个目标的实例分割信息; Random affine, 随机进行仿射变换, 但根据配置文件里的超参数发现只使用了 Scale 和 Translation 即缩放和平移; MixUp, 将两张图片按照一定的透明度融合在一起, 较大的模型才会使用到 MixUp, 而且每次只有 10% 的概率会使用到; Augment HSV, 随机调整色度、饱和度以及明度; Random horizontal flip, 随机水平翻转。通过以上操作, 提高网络模型的鲁棒性。

在经过预处理过后, 图像输入到 Backbone 进行特征提取, 其主干网络包括 Focus 模块、CSP 模块以及 SPP 模块。Focus 模块通过将输入图像分成不同的块, 然后将每个块进行卷积操作, 从而提取局部特征, 用于减少计算量和提高特征提取效率。CSPNet 结构通过将浅层的特征图在通道维度一分为二, 一部分经由特征提取模块向后传播, 另一部分则经过跨阶段层次结构直接与特征提取模块的输出进行合并。SPP 模块通过引入多尺度的空间金字塔池化操作, 从而在不同尺度上提取特征, 并且保持了特征图的尺寸不变, 更好地适应不同大小的目标, 提高了模型的泛化能力和鲁棒性。

Neck 部分负责对特征图进行多尺度特征融合, 并把这些特征传递给预测层。其采用 FPN+PAN 的结构, FPN 是自顶向下的, 将高层的特征信息通过上采样的方式进行传递融合, 得到进行预测的特征图, 而特征金字塔则自底向上传达强定位特征, 进一步提高特征提取的能力。

Head 是用来对特征金字塔进行目标检测的模块,主要包括三个部分: Anchors, 提前定义不同比例和尺寸的锚框来适应不同大小的目标; Classification, 采用全连接层加 Softmax 函数对每个检测框进行分类; Regression, 对每个检测框进行回归, 得到其位置和大小。

三、模型训练与实验测试

(一) 实验平台

本文实验在 Windows 10 环境下使用 Python 语言, 基于深度学习框架 PyTorch 1.7. 版本进行, 硬件配置为 NVIDIA GeForce RTX 4070 Ti GPU, Intel (R) Core (TM) i7-12700K CPU。

(二) 实验细节

本文实验在模型训练中, 将训练 epochs 设定为 100, batch-size 为 16, 优化器选择 SGD, 初始学习率为 0.01。

(三) 数据集准备

本文实验使用的数据集由 3 个部分组成。第一部分来自线下桃园实拍, 总共 1129 张, 第二部分来自于公开数据集, peach-classification, 这是一个区分桃子大小的数据, 其有 4 个大小的桃子, 从每个尺寸中抽取 400 张桃子, 将其作为好桃子数据集的一部分, 总共 1600 张; 第三部分数据集使用自建数据集, 图像来自于互联网, 通过爬虫等方式, 获取了 450 张图像。这三部分数据集加起来一共 3179 张图像, 使用 lableImg 标注图像中的桃子目标框, 抽取 429 张作为测试集, 剩余图片以 8: 2 的比例将其分为训练集 2200 张, 验证集 550 张。

(四) 实验结果

在本次实验中, 总共检测了 429 张桃子图片, 效果如图 1, 处理一张图像的速度为 1.1 毫秒, 推理 (inference) 的速度为 16.5 毫秒, NMS (Non-Maximum Suppression, 非极大值抑制) 的速度为 1.3 毫秒, 因此分类一张图片的时间为 18.9 毫秒, 满足实时检测的要求, 而准确率也能高达 99.76%。

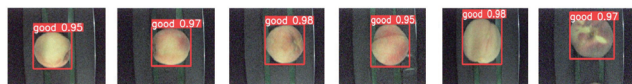


图 2 公开数据集测试结果



图 3 网络桃子图片测试结果

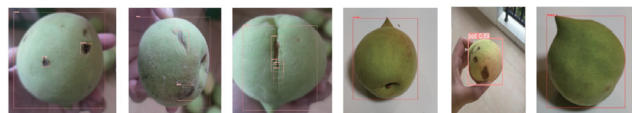


图 4 线下实地采集图片测试结果

四、模型部署应用

为了方便进行桃果瑕疵检测的现场应用, 本文将训练的人工智能模型移植到小型设备如图 5, 使得检测设备功耗得到降低, 并且更加便于携带使用。第一种小型设备采用的 CPU 为 Intel (R)

Celeron (R) CPU N3050 @ 1.60GHZ, 内存为 4GB, 操作系统为 Ubuntu 20.04。将模型部署至该设备, 处理每张图片的时间为 0.44s。

此外, 本文还将算法移植到高通 SG865 开发板, 该板提供神经网络加速的模块, 实测处理每张图片的时间为 30ms, 比小主机的处理速度有一个数量级的提升。也显示出本文所提方案的实时有效性。

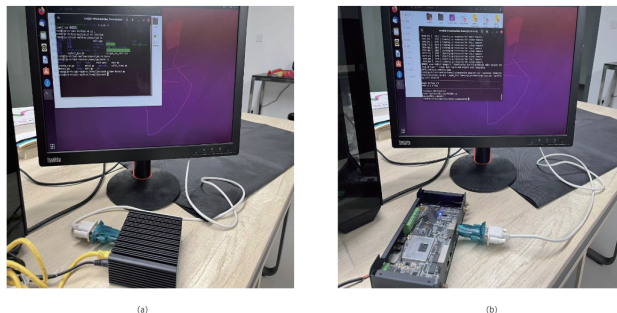


图 5 (a) 基于 x86 架构的小型主机 (b) 基于 ARM 架构的高通 SG865 开发板

五、结论

本文采用 YOLOv5s 实现轻量级的桃子瑕疵部位检测算法, 通过选取多样性的数据集, 让训练得到的模型在保持速度快、高精度的同时, 能够具有较强的鲁棒性, 从而能够在复杂的环境中稳定地工作。实验表明, 该模型的复杂度较低, 即使移植到工作频率较低的小型设备中, 仍然具有较快的速度, 并能保持推理的准确度。下一步, 一方面可以考虑继续优化检测算法, 另一方面可以继续将算法的模型移植到更低功耗的嵌入式设备中, 采用锂电池供电, 从而设计为小型便携式设备, 更加适合嵌入安装到其他功能模块。

参考文献:

- [1] 王婉婷, 姜国龙, 褚云飞, 等. 从 RCNN 到 YOLO 系列的物体检测系统综述 [J]. 齐鲁工业大学学报, 2021, 35 (05): 9-16.
 - [2] 赵永强, 饶元, 董世鹏, 等. 深度学习目标检测方法综述 [J]. 中国图象图形学报, 2020, 25 (04): 629-654.
 - [3] 许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述 [J]. 计算机工程与应用, 2021, 57 (08): 10-25.
 - [4] 朱晓慧, 钱丽萍, 傅伟. 图像数据增强技术研究综述 [J]. 软件导刊, 2021, 20 (05): 230-236.
 - [5] 韦小龙, 王方田, 刘超, 等. 基于 CSPNet-YOLOv7 目标检测算法的煤矸图像识别模型 [J]. 煤炭科学技术, 2024 (5): 1-13.
 - [6] 杨波, 何金平, 张立娜. 基于改进 SPP-x 的 YOLOv5 神经网络水稻叶片病害识别检测 [J]. 中国农机化学报, 2023, 44 (09): 190-197.
 - [7] 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+201.
 - [8] 陈科圻, 朱志亮, 邓小明, 等. 多尺度目标检测的深度学习研究综述 [J]. 软件学报, 2021, 32 (04): 1201-1227.
- 基金项目: 江门市科技计划基础科学项目 (项目编号: 2022JC01026)