

基于树莓派与 YOLOv5-Lite 模型的日常垃圾识别系统

王嘉炜 宁勇强 卓赛龙

江西理工大学 能源与机械工程学院, 江西 南昌 330013

[摘要] 利用深度学习实现视觉检测技术对日常生活垃圾的识别, 从而识别种类繁多的传统生活垃圾有重要意义。把在 PC 端上训练好的 YOLOv5s 与 YOLOv5-Lite 目标检测模型分别部署在搭载树莓派系统平台上, 并在此平台上搭建深度学习环境, 构建垃圾识别系统。在不同光照条件下, 对这两个模型进行分析对比, 实验结果表明, 在识别准确率相差 1.5% 的情况下, YOLOv5-Lite 模型相对于原 YOLOv5s 模型, 网络参数量下降了 80.86%, 模型内存大小下降了 75.52%, 检测速度提高 85.94%。综上所述, 文中提出的基于树莓派与 YOLOv5-Lite 的垃圾识别系统兼顾了准确性好、稳定性好、成本低等综合优点。

[关键词] 垃圾识别; YOLOv5-Lite; 模型部署; 树莓派

DOI: 10.33142/sca.v8i3.15785

中图分类号: TP368.2

文献标识码: A

Daily Garbage Recognition System Based on Raspberry Pi and YOLOv5-Lite Model

WANG Jiawei, NING Yongqiang, ZHUO Sailong

College of Energy and Mechanical Engineering, Jiangxi University of Science and Technology, Nanchang, Jiangxi, 330013, China

Abstract: It is of great significance to use deep learning to implement visual detection technology for the recognition of daily household waste, in order to identify a wide variety of traditional household waste. Deploy the YOLOv5s and YOLOv5-Lite object detection models trained on the PC side separately on the Raspberry Pi system platform, and build a deep learning environment on this platform to construct a garbage recognition system. Under different lighting conditions, the two models were analyzed and compared. The experimental results showed that, with a 1.5% difference in recognition accuracy, the YOLOv5-Lite model had a 80.86% decrease in network parameters, a 75.52% decrease in model memory size, and an 85.94% increase in detection speed compared to the original YOLOv5s model. In summary, the garbage recognition system proposed in the article based on Raspberry Pi and YOLOv5-Lite combines the comprehensive advantages of good accuracy, stability, and low cost.

Keywords: garbage recognition; YOLOv5-Lite; model deployment; Raspberry Pi

引言

随着经济的迅速发展;我国城镇人口数量急剧增加;城市生活垃圾也日益增多;而垃圾分类则是处理垃圾公害的最佳解决方案^[1]。当下的关注热点在于将垃圾进行无害化、资源化处理;因此合理有效地对垃圾进行回收利用具有一定的意义;对垃圾进行有效地分类是处理垃圾回收问题的前提^[2]。近年来;随着深度学习技术的不断发展;目标检测算法在垃圾识别领域得到了广泛的应用。

国内外已有通过基于深度学习的机器视觉进行垃圾检测和分类方面的研究。Mittal G^[3]等人研究的基于 GarbNet 识别算法, 准确率达到了 87.69%, 但并未实现垃圾的分类; Ying Liu 等人^[4]提出的 YOLOv2 检测算法, 使用 MobileNetV1^[5]网络代替原本的卷积模块, 使识别的准确率达到 89.1%, 同样未细分垃圾的种类; Ying Wang 等人^[6]提出用 ResNet 作为特征提取网络, 该方法通过将城市自然环境与垃圾的图片相结合, 从而获取更多的数据集样本, 提高了算法的稳定性和准确性, 但该算法对设备的计算性能的要求较高。

本文研究的 YOLOv5-Lite, 作为 YOLOv5s 的轻量级版本, 在保证检测精度的同时, 进一步降低了模型的复杂度

和计算量, 可以在配置较低, 计算性能较弱的设备上部署。本文获取不同光照条件下日常生活垃圾的图像, 基于 YOLOv5-Lite 深度学习算法模型进行垃圾识别, 以期满足垃圾识别实时性和准确性的要求

1 系统方案

垃圾识别系统工作原理: 首先在 PC 端上搭建模型训练测试的环境。然后收集各种类别的垃圾图像, 并使用标注工具对图像进行标注, 生成包含垃圾类别和边界框信息的 XML 或 CSV 文件。再使用 YOLOv5 模型和 YOLOv5-Lite 模型开始训练过程, 输入训练集数据和配置文件, 输出训练好的模型权重文件。使用验证集数据对训练好的模型进行验证, 评估模型的识别精度和性能。最后将在 PC 端训练好的 YOLOv5-Lite 模型与 YOLOv5s 模型分别移植到树莓派上测试, 由连接到树莓派上的摄像头读取垃圾的测试图像, 并显示检测的效果。系统集成流程如图 1 所示。

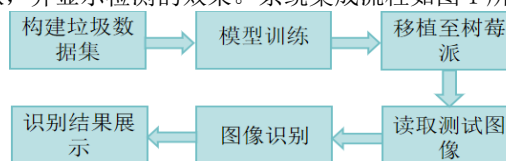


图 1 垃圾检测系统集成图

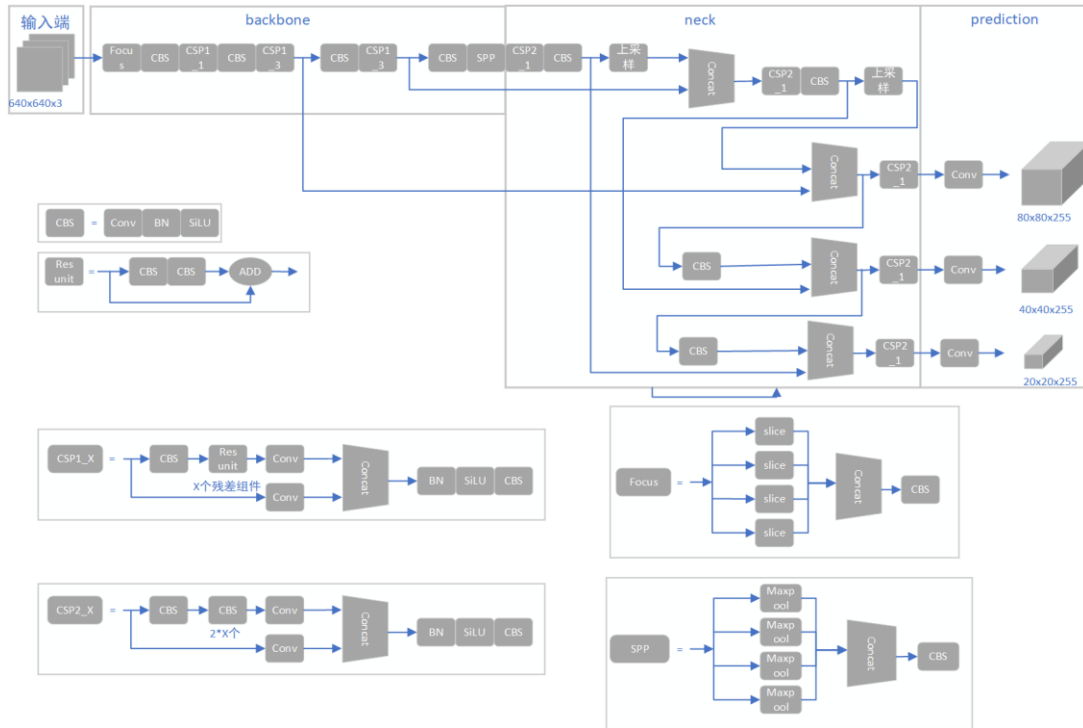


图 2 YOLOv5s 网络结构

2 YOLOv5 算法简介

相对于 R-CNN^[7]、Fast R-CNN^[8]、Faster R-CNN^[9]等 two-stage 的目标检测算法,该两阶段算法在检测速度上难以满足实际的需求,随后 Redmin^[10]等人提出了 YOLO, YOLO 系列是非常经典的 one-stage^[11]目标检测算法极大提高了目标检测速度。本次实验选择 YOLOv5s^[12]网络模型作为基础进行改进。如图 2 为 YOLOv5s 的网络结构。

YOLOv5s 算法的网络模型主要由四部分组成:第一部分为输入端,负责接收并预处理图像数据,采用 Mosaic 数据增强技术丰富背景信息,同时执行自适应锚框计算和图像缩放,确保输入图像与模型要求匹配,为后续特征提取和检测任务奠定基础;第二部分为 Backbone,即主干网络,主要用于提取图片中的特征信息,为后续的网络提供输入,在 YOLOv5s 中,Backbone 包含 Focus、CBS、C3、SPP 等多种模块,这些模块通过卷积层的堆叠来提取不同尺度的特征信息;第三部分为 Neck,位于 Backbone 和 Head 之间,用于进一步利用 Backbone 提取的特征,提高模型的鲁棒性,在 YOLOv5s 中,Neck 采用 FPN+PAN 结构,实现了特征的多尺度融合;第四部分为 Prediction,获取网络输出,利用之前提取的特征做出预测。在 YOLOv5s 中, Prediction 使用解耦预测头,将分类任务与回归任务拆分开,使用两条并行分支分别进行预测,提高了模型的检测精度和收敛速度。

针对传统目标检测模型存在的问题, YOLOv5s 进行了多项改进与优化:①使用 EIIOU Loss 替换 CIIOU Loss:加快了预测框回归损失函数的收敛速度,并提高了预测框的

回归预测精度;②使用 CARAFE 上采样算子:替换了原始的最近邻插值上采样模块,减少了上采样过程中特征信息的损失,提高了上采样特征图的质量;③添加小目标检测分支:在原始的三尺度检测层基础上新添加一层检测尺度更小的小目标检测分支,提高了模型对于小目标的检测能力;④通道剪枝:对改进后的模型进行通道剪枝,剪除对于检测效果影响不大的冗余通道,降低模型体积,使得模型更加适用于资源受限条件下的目标检测任务。

3 改进的 YOLOv5-Lite 网络模型结构

3.1 摘除 Focus 层

YOLOv5-Lite 摘除了部分 Focus 层,以避免多次采用 slice 操作导致的计算量增加。Focus 层虽然可以通过分组处理特征图来减少计算量,但在某些情况下可能会增加模型的复杂度和计算负担。因此, YOLOv5-Lite 在设计时选择了摘除部分 Focus 层,以进一步降低模型的计算量。

3.2 使用 ShuffleNetV2

YOLOv5-Lite 的主干网络主要使用了含 Shuffle channel 的 Shuffle block 组成。Shuffle block 是一种轻量化的卷积块,它通过 Shuffle channel 操作来打乱特征图的通道顺序,以增加特征图的多样性,同时减少模型的计算量和参数数量。

3.3 减少使用 C3 Layer

C3 Layer 是 CSPBottleneck 改进版本,它更简单、更快、更轻,在近乎相似的损耗上能取得更好的结果。但频繁使用 C3 Layer 以及通道数较高的 C3 Layer 会占用较多的缓存空间,降低运行速度。

3.4 使用 CIOU 损失函数

传统 YOLOv5-Lite 采用 GIOU 作为损失函数, GIOU 损失函数虽然在一定程度上解决了 IOU 损失函数在边界框不重叠时无法提供有效梯度的问题, 但当两个框之间存在包含关系时, GIOU 会退化为 IOU, 无法准确反映两个框之间的相对位置关系。这种情况下, 即使预测框与真实框之间存在较大的位置偏差, GIOU 的值也可能较高, 从而导致模型训练过程中的误判。且 GIOU 在计算过程中需要考虑两个框的最小外接矩形, 这增加了计算的复杂性, 导致收敛速度慢。本实验采用 CIOU 代替 GIOU 作为损失函数, 它不仅能够更全面地评估预测边界框与真实边界框之间的差异, 使得预测目标框回归的速度更快, 精度更高还能够更快地收敛到最优解^[13], 提高检测精度和泛化能力。本实验采用基于 CIOU 损失函数的 YOLOv5-lite 模型。CIOU 损失函数公式如下:

$$CIOU = \frac{\rho^2(A, B)}{C^2} + \alpha v \quad (1)$$

$$\alpha = \frac{v}{(1 - IOU) + v} \quad (2)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (3)$$

$$IOU = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

$$Loss_{CIOU} = 1 - IOU + \frac{\rho^2(A, B)}{C^2} + \alpha v \quad (5)$$

式中 α 是权重函数; v 用来度量宽高比的一致性; $\rho^2(A, B)$ 表示 A(预测框)与 B(目标框)的中心距离 c 代表 A 与 B 最小并区间对角线距离; IOU 等于 A 和 B 的交集除以

二者的并集。

以上方法使得 YOLOv5-Lite 网络相较于 YOLOv5s 网络更具优势: ①部署灵活: 由于 YOLOv5-Lite 具有较小的模型大小和较快的推理速度, 因此它更容易部署在资源受限的设备上。这使得 YOLOv5-Lite 在物联网、智能家居等领域具有广泛的应用前景。②降低成本: 轻量化设计使得 YOLOv5-Lite 在运行时对计算资源的需求降低, 从而减少了能源消耗和硬件成本。这对于需要大规模部署的应用场景来说尤为重要。③提升 fps: 在相同的硬件条件下, YOLOv5-Lite 的 fps 相较于 YOLOv5s 有显著提升。这使得 YOLOv5-Lite 在实时检测任务中更具优势。

因此, YOLOv5-Lite 网络在原来的 YOLOv5s 网络上的改进主要体现在轻量化设计、效率提升以及保持高精度等方面。这些改进使得 YOLOv5-Lite 在实时检测任务中更具优势, 同时也为它在各种应用场景中的广泛部署提供了可能。本此实验所用的 YOLOv5-Lite-s 模型结构如图 3 所示。

4 模型部署

在树莓派上部署 YOLOv5-Lite 模型是一个涉及硬件、软件、模型转换、系统部署、测试与优化等多个环节的综合任务。

硬件: 本次实验选择的树莓派型号是树莓派 4B, 该型号树莓派能提供足够的计算能力。同时配备了与树莓派兼容的摄像头模块用于捕捉实时图像, 从而完成垃圾识别的任务。

软件: 下载并安装适用于树莓派的操作系统 Raspbian, 并通过烧录工具将镜像写入 SD 卡。随后, 将 SD 卡插入树莓派, 连接显示器、键盘和鼠标等外设, 启动树莓派并完成操作系统的初始化设置。接着, 更新软件源, 安装 Python 3、pip 等必要的软件包, 以及 OpenCV 等计算机视觉库, 用于图像处理和显示。

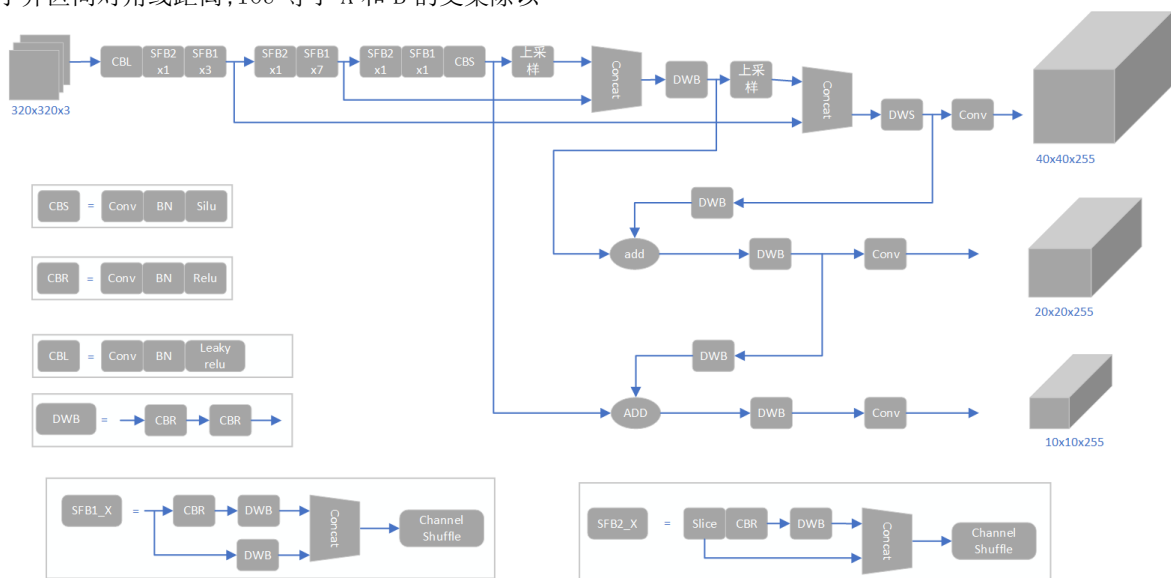


图 3 YOLOv5-Lite-s 网络结构

模型转换与部署环节：首先使用 PyTorch 提供的转换工具将训练好的 YOLOv5-Lite 模型从 PyTorch 格式转换为 ONNX 格式。在转换过程中，需要确保模型的输入和输出格式与树莓派上的代码相匹配。完成模型转换后，将模型文件上传到树莓派上，并编写 Python 脚本加载模型进行推理。包括读取摄像头捕获的图像、对图像进行预处理、将图像输入到模型中、获取模型的输出并进行后处理等步骤。

测试与优化：在不同光照条件下，使用不同的垃圾图像对系统进行测试，验证模型的识别精度和实时性。根据测试结果调整模型的参数和预处理步骤，以提高识别效果。同时，对代码进行优化，减少不必要的计算和资源消耗。

通过这一系列步骤，可以成功在树莓派上部署 YOLOv5-Lite 模型，构建一个高效、准确的日常垃圾识别系统。

5 实验设置与结果分析

5.1 数据集

训练集是模型训练过程中使用的图像数据^[14]。本次实验所用数据集共 8366 张，本实验以比例 8:2 划分训练集和测试集，即训练集 6693 张，测试集 1673 张。

5.2 计算机实验运行环境

本文计算机使用的显卡型号为 NVIDIA GEFORC RTX 3070、CPU 为 Intel Core i9-10900K，2.50 GHz，Python 版本为 3.8，基于 Pytorch 深度学习框架。

5.3 模型训练

本文训练所用的初始权重分别为 yolov5s.pt 和 v5lite-s.pt，其中 Batch_size 设为 128，迭代周期 epochs 设为 200，置信度阈值 0.45，IOU 阈值为 0.5，通过训练模型，分别得到了两个最好的权重文件。

5.4 评估指标

为评估两种模型的检测效果，本文使用参数量、模型占用内存、检测速度以及识别精确度作为评价指标。

①参数量：参数量是指模型中所有权重和偏置的总数，是衡量模型大小和复杂度的重要标准。参数量越小，模型越轻量，对计算资源的需求越低，但可能在性能上有所牺牲。反之，参数量越大，模型可能具备更强的表示能力，但也可能导致过拟合和更高的计算成本。

②模型占用内存：模型占用内存是指在训练或推理过程中，模型所占用的物理内存大小。这包括模型参数、中间计算结果以及必要的缓存等。内存占用的大小直接影响模型的运行效率和可扩展性。在训练大型模型时，内存占用过高可能导致训练速度下降甚至训练失败。

③检测速度：检测速度是指模型处理每张图片所需的时间。检测速度越快，模型在实时应用中的表现越好。在目标检测任务中，快速准确的检测对于实际应用至关重要。

④识别精确度：识别精确度是指模型正确识别目标类别的能力。识别精确度越高，模型的性能越好，能够更准

确地识别出目标类别。在目标检测任务中，高识别精确度是实现高精度定位的前提。

5.5 实验结果与检测效果

通过对训练好的两种模型进行测试，表 1 为在树莓派上测试所得的网络对比结果。

表 1 网络对比实验结果

网络模型	精确度	检测速度	网络参数	模型内存大小
YOLOv5-Lite	86.7%	0.45s	1634792	3.5MB
YOLOv5s	88.2%	3.2s	8539572	14.3MB

实验结果表明，在不同光照条件下改进 YOLOv5-Lite 网络与 YOLOv5s 网络相比，其识别精确度和 YOLOv5s 数值非常相近，网络参数量下降了 80.86%，模型内存大小下降 75.52%，检测速度提高了 85.94%。因此 YOLOv5-Lite 综合性能优势明显，成功满足了实验需求。最终实现的检测效果图如图 4、图 5 所示。



图 4 强光拍摄



图 5 弱光拍摄

6 结束语

本文提出了一种基于树莓派和 YOLOv5-Lite 垃圾检测系统，通过将训练好的 YOLOv5s 和 YOLOv5-lite 模型部署到树莓派平台上，搭建一个垃圾检测系统。通过实验分析，在不同光照条件下，基于轻量化的 YOLOv5-Lite 网络模型的垃圾识别系统优势更加明显。在识别精确度差别很小的情况下，推理速度方面，YOLOv5-Lite 模型凭借其轻量级的网络结构和高效的计算优化，实现了更快的检测速度。这对于实时垃圾检测系统尤为重要，能够显著提升系统的响应速度和用户体验。此外，YOLOv5-Lite 模型的参数量和内存占用均较低，更适合在资源受限的嵌入式设备或边缘设备上部署，降低了系统的硬件成本。

基金项目：江西理工大学大学生创新创业训练资助项目“智能垃圾拾取小车”（S202210407044）。

[参考文献]

- [1]魏潇潇,王小铭,李蕾,等.1979-2015 年中国城市生活垃圾产生和处理时空特征[J].中国环境科学,2018,38(10):3833-3834.
- [2]梁建胜,温贺平.基于深度学习的视频关键帧提取与视频检索[J].控制工程,2019,26(5):965-970.
- [3]Mittal G,Yagnik K B,Garg M,et al.Spotgarbage:smart—phone app to detect garbage using deep learning[C].Heidelberg,GERMANY:Assoc Comp Machinery,2016.
- [4]Liu Y,Ge Z,Lv G,et al.Research on automatic garbage detection system based on deep learning and narrowband internet of things[J].Journal of Physics:Conference Series,2018(1):1069-1076.
- [5]Howard A,Zhu M,Chen B,et al.MobileNets:efficient convolutional neural networks for mobile vision applications [C].Hawaii:Computer Vision and Pattern Recognition,2017.
- [6]WANG Ying,XU Zhang.Autonomous garbage detection for intelligent urbanmanagement[C].Shanghai:MATEC Web of Conferences,2018.
- [7]TAO X,ZHANG D P,WANG Z H.Detection of power line insulator defects using aerial images analyzed with convolutional neural networks[J].IEEE Transactions on Systems Man and Cybernetics Systems,2020,50(4):1486-1498.
- [8]Girshick R.Fast R-CNN[C].IEEE:Proceedings of 2015 IEEE International Conference on Computer Vision.Santiago,2015.
- [9]REN S,HE K,GIRSHICK R.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]REDMAN J,DIVVALAS,GIRSHICK R,et al.You only look once: unified, realtime object detection[C].IEEE:2016 IEEE Conference on Computer Vision and Pattern Recognition,2016.
- [11]刘芳,韩笑.基于多尺度深度学习的自适应航拍目标检测[J].航空学报,2020,7(5):1-13.
- [12]Glenn Jocher. YOLOv5: The Leader in Realtime Object Detection[EB/OL]. [2023-03].
- [13]武张静,刘敏,史禹龙,等.语音示教自主巡航智能垃圾分类机器人的研究与设计[J].科技创新与应用,2020(32):43-47.
- [14]Gu S,Ding L.A complex-valued VGG network based deep learning algorithm for image recognition[C].Wangzhou:Proceedings of the Ninth International Conference on Intelligent Control and Information Processing,2018.

作者简介:王嘉炜(2003—),男,汉族,江西上饶人,本科在读,江西理工大学能源与机械工程学院,研究方向:机器视觉、智能控制;宁勇强(1992—),男,汉族,江西吉安人,硕士,讲师,江西理工大学能源与机械工程学院,研究方向:无线通信、机器学习。