

DOI: 10.13382/j.jemi.B2104002

基于嵌入式系统的智能售货柜目标检测算法*

侯维岩¹ 靳东安¹ 王高杰² 王洋² 丁英强¹

(1. 郑州大学 信息工程学院 郑州 450001; 2. 广东顺德创新设计研究院 佛山 528311)

摘要:针对普通商品识别算法在智能售货柜嵌入式系统平台上检测速度慢、识别率低的问题,提出了一种在YOLOv3基础上的改进型商品识别算法DS_YOLOv3。利用k-means++聚类算法得到适应于售货柜中售卖饮料图像数据的先验框;采用深度可分离卷积替换标准卷积,并加入倒置残差模块重构YOLOv3算法,减少了计算复杂度使其能在嵌入式平台实时检测;同时引入CIoU作为边界框回归损失函数,提高目标图像定位精度,实现了对传统YOLOv3算法的改进。在计算机工作站和Jeston Xavier NX嵌入式平台上进行了典型场景下的商品检测实验。实验结果表明,DS_YOLOv3算法mAP达到了96.73%,在Jeston Xavier NX平台上实际检测的速率为20.34 fps,满足了基于嵌入式系统平台的智能售货柜对实时性和商品识别精度的要求。

关键词:商品识别;YOLOv3;k-means++;深度可分离卷积;倒置残差结构;CIoU

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.99

Object detection algorithm of intelligent vending cabinet via embedded system

Hou Weiyan¹ Jin Dongan¹ Wang Gaojie² Wang Yang² Ding Yingqiang¹

(1. School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China;

2. Guangdong Shunde Innovative Design Institute, Foshan 528311, China)

Abstract: In order to solve the problem of slow detection speed and low recognition rate of common commodity recognition algorithms on the intelligent vending cabinet embedded system platform, an improved commodity recognition algorithm DS_YOLOv3 is proposed on the basis of YOLOv3. The traditional YOLOv3 neural network algorithm is improved by obtaining a prior bounding box suited for the image data of beverages sold in the vending cabinet by using k-means++ clustering algorithm, using the deep separable convolution to replace the standard convolution and adding the inverted residual module to reconstruct the YOLOv3 algorithm, which could reduce the computational complexity and enable real-time detection on the embedded platform, and introducing CIoU as the bounding box regression loss function to enhance the accuracy of target image positioning. The commodity testing experiment under typical scenarios is performed on a computer workstation and Jeston Xavier NX embedded platform. The results show that the accuracy of DS_YOLOv3 algorithm reaches 96.73%, and the actual detection rate on the Jeston Xavier NX platform is 20.34 fps, which meet the real-time and commodity detection recognition accuracy requirements of the intelligent vending cabinet based on the embedded system platform.

Keywords: commodity recognition; YOLOv3; k-means++; depth separable convolution; inverted residual structure; CIoU

0 引言

智能售货柜通过识别开门前和开门后采集的图片,判断顾客取走的商品,并对商品进行结算,实现商品零售

的智能化。亚马逊在西雅图推出了 Amazon Go,该商店通过安装大量传感器和摄像头并结合深度学习与移动支付,实现即拿即走的售货方式;阿里巴巴在杭州开设了集餐饮购物于一身的无人售货店淘咖啡,通过生物自主感知系统,精准的识别用户选择的商品;码隆科技以静态视

收稿日期: 2021-01-27 Received Date: 2021-01-27

* 基金项目: 国家自然科学基金重大研究计划(92067106)、广东省科技创新战略专项资金(纵向协同管理方向)(2018FS05020102)、佛山市高质量专利培育(1920025003148)项目资助

觉方法的 RetailAI 智能售货柜也已实现落地工作。但从行业总体发展和用户体验感来看未达到大面积推广的层次,还需要对关键的图像识别目标检测算法加以优化和改进。本文将研究一个基于智能售货柜嵌入式系统平台的、能满足检测精度和实时性要求的目标检测算法,应用于智能售货柜商品的检测对提高用户体验有重要意义。

目前主流的目标检测算法主要分为两大类:1) 双阶段目标检测算法,将物体识别和定位分别完成,优点是识别错误率和漏检率较低,代表算法有 Faster RCNN^[1-2]、R-FCN^[3]和 FPN^[4]等;2) 单阶段目标检测算法,优点是算法识别速度快和准确率高,代表算法如 SSD^[5-6]、RetinaNet^[7]、YOLOv3^[8]等。但上述两类算法计算复杂,推理时间长且对硬件要求高,不适用于部署到嵌入式平台进行物体实时检测。

研究人员通过使用轻量级网络模型、模型压缩、损失函数优化等方法来降低目标检测算法计算复杂度、提高检测准确率。Google 团队提出的 MobileNetV1^[9],通过深度可分离卷积替换普通卷积,大幅降低了算法计算量和运行开销。Redmon 等^[8]提出了 YOLOv3-tiny,该算法在 YOLOv3 基础上通过简化特征提取网络,在两个不同尺度下对目标进行检测,在牺牲检测精度的情况下提升检测速度。Han 等^[10]通过网络剪枝、共享权重和哈夫曼编码格式压缩网络,可以在 Alexnet^[11]模型上达到 35 倍的压缩比。曹红燕等^[12]通过融合 BN 网络层和卷积网络层,优化网络结构,加快模型训练速度。Zheng 等^[13]提出了新的边界损失框 DIoU loss 和 CIoU loss,并且使用 DIoU-NMS 替换经典 NMS 方法,在 YOLOv3 对比实验上 mAP 提升达 5.91%。

国内学者林付春等^[14]设计了基于机器视觉的无人售货系统,将售货柜采集的图片和视频上传到服务器进行识别;Hu 等^[15]提出了差异检测算法 DiffNet,通过对比售货柜内消费前后图像的差异获得消费信息;魏宏彬等^[16]提出一种改进型 YOLOv3 算法,通过电子秤的摄像头识别蔬菜种类,对蔬菜称重并打印价格。虽然上述方法取得了不错的效果,但需将拍摄的图片或者视频上传到服务器,在服务器通过目标识别算法得到商品信息。当设备增多时,对传输层带宽和服务器要求快速增加。

本文在传统的 YOLOv3 基础上提出了 DS_YOLOv3 算法,利用深度可分离卷积替换标准卷积和倒置残差结构,在不影响精度的前提下,大大减少了计算复杂度;通过 k-means++ 算法^[17]得到数据集的先验框;同时使用 CIoU loss 替换边界框损失函数,克服 YOLOv3 边界损失函数的不足。实验结果表明,DS_YOLOv3 算法在嵌入式平台上能满足智能售货柜对实时检测和检测精度的要求。

1 相关工作

1.1 YOLOv3 网络介绍

YOLOv3 由特征提取网络和多尺度检测网络两部分构成,网络结构如图 1 所示。特征提取网络借鉴 ResNet^[18]和 SqueezeNet^[19]而提出的 Darknet53 网络,其分类性能与 ResNet-101 和 ResNet-152 接近,但检测速度更快;YOLOv3 多尺度检测网络使用了特征金字塔 (FPN) 结构,采用双线性插值的上采样方法重构并融合特征,在 3 个不同尺度的特征图上分别进行检测,提升对小目标的检测效果。

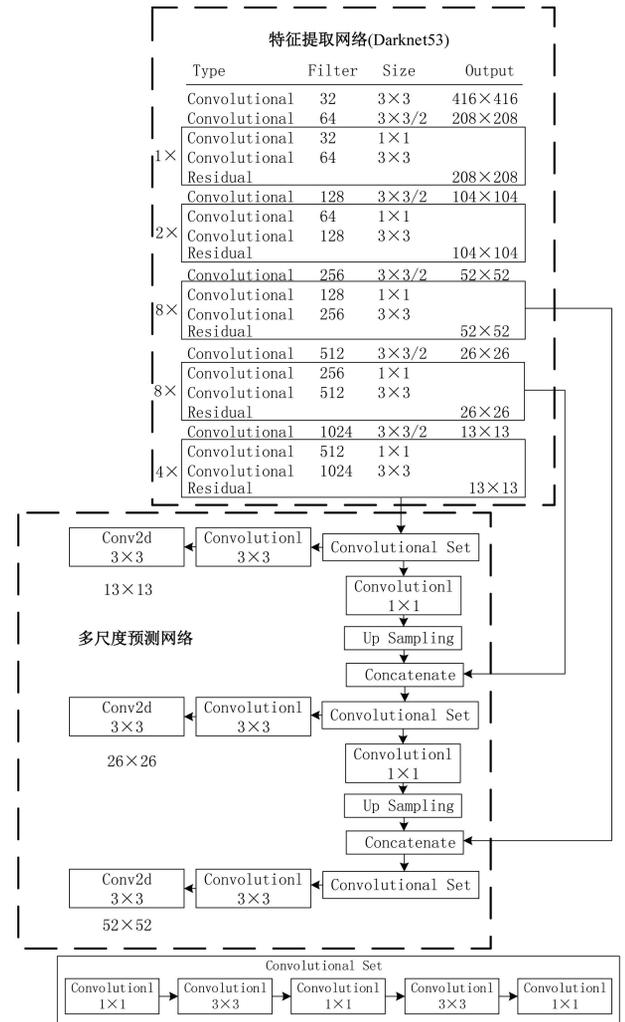


图 1 YOLOv3 网络结构

Fig. 1 YOLOv3 network structure

YOLOv3 算法思想是先通过 Darknet-53 特征提取网络从输入图像中得到 $S \times S$ 的网格单元,该网络负责检测落在该网格的目标,接着对每个网格内的目标通过 3 个先验框进行预测,输出为 $S \times S \times 3 \times (5 + C)$,其中 S 表

示网格数量,3 表示先验框的数量,5 表示坐标信息和置信度, C 表示类别数。最后通过非极大值抑制算法确定最终的目标位置和类别。

1.2 MobileNetV2 网络

MobileNetV2^[20] 网络是对 MobileNetV1 网络改进,在深度可分离卷积的基础上提出了倒置残差结构作为模型的特征提取网络。其中深度可分离卷积由深度卷积和点卷积组成。不同于常规的卷积,深度卷积将上一层多通道特征图拆分成单通道,各通道内部分别进行深度卷积操作,再使用点卷积对特征图通道信息融合。如图 2 所示。输入样本为 $D_i \times D_i \times M$,卷积核为 $D_k \times D_k \times M$,且步长为 1,标准卷积的参数量为 $D_k \times D_k \times M \times N$,深度可分离卷积的参数量为 $D_k \times D_k \times M$ 与 $M \times N$ 的求和。使用 3×3 的卷积核,参数量可以降到原来的 $1/9 \sim 1/8$,从而在检测精度没有明显变化的情况下,提高检测速率。

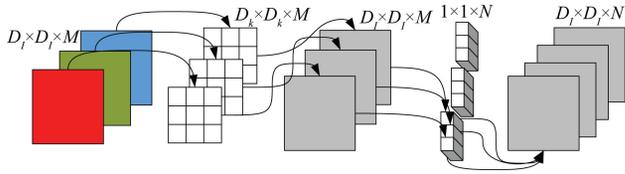


图 2 深度卷积与点卷积

Fig. 2 Deep convolution and point convolution

倒置残差结构如图 3 所示。该结构在 ResNet 的基础上进行改进,使用先扩张后降维的思想,目的是提升通道数,获得更多的特征。当 $stride = 1$ 时,如图 3(a)所示。首先通过 1×1 点卷积将通道扩大为原来的 6 倍,然后使用深度 3×3 深度卷积特征提取,再通过点卷积缩小为原来的通道;当 $stride = 2$ 时,如图 3(b)所示。由于卷积步长不同导致输入输出尺寸不同,因此不使用输入和输出拼接的方式。

2 DS_YOLOv3 算法

2.1 基于 k-means++ 聚类算法的先验框

YOLOv3 借鉴 Faster RCNN 中先验框的思想,使用 k-means 聚类算法^[21]对数据集的目标进行聚类,得到 9 组固定宽高的先验框,并且采用平均交并比 (average intersection over union, AvgIoU) 作为目标聚类的量度,聚类的 AvgIoU 目标函数可表示为:

$$f = \operatorname{argmax} \frac{\sum_{i=1}^k \sum_{j=1}^{n_k} I_{IoU}(B, C)}{n} \quad (1)$$

式中: B 表示目标真值; C 表示簇中心; n_k 表示第 k 个聚

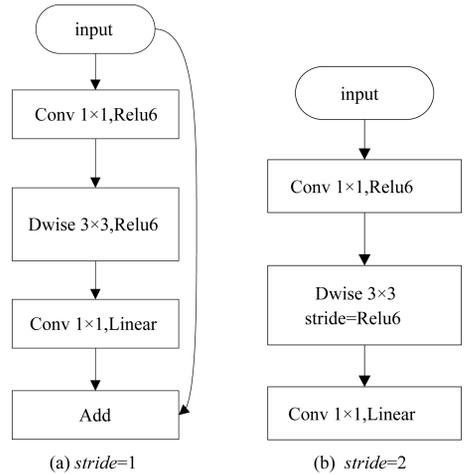


图 3 倒置残差结构

Fig. 3 Inverted residual structure

类中心中样本个数; n 表示样本总数; $I_{IoU}(B, C)$ 表示簇的中心框和聚类框的交并比。

k-means 算法易受随机选取的 k 个聚类中心的影响,初始聚类中心选择不好,算法会陷入局部最优解,影响算法检测效果。为了获得更稳定的先验框,本文采用 k-means++ 聚类算法对商品数据集进行聚类分析,该算法采用初始聚类中心之间的相互距离尽可能远的策略,使得聚类结果不受随机选取初始聚类中心的影响,聚类得到的先验框更贴近商品数据集的目标框。

由于 YOLOv3 的 9 组先验框是由 COCO 数据集聚类产生,该数据集和本文商品数据集目标偏差较大,因此需要重新进行聚类得到新的先验框。使用 k-means++ 对商品数据集进行聚类分析,得到先验框数量 k 与 AvgIoU 之间的关系如图 4 所示。

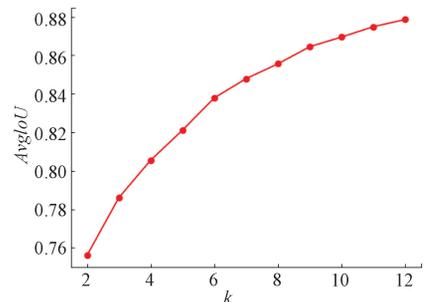


图 4 先验框数量与平均交并比的关系

Fig. 4 The relationship between the number of prior frames and the average intersection union ratio

随着 k 值的增大, AvgIoU 的值也不断增大并趋于平稳。先验框选择越多算法的精确度相应越高,但会增加算法的复杂度。当 k 值大于 9 时曲线增长速度变慢,故选择 9 个先验框分到 3 个不同尺度的特征图中,其中大

先验框分配给小尺寸特征图,小先验框分配给大尺寸特征图。聚类结果如图 5 所示。通过 k-means++ 算法对本文数据集进行聚类得到 9 个先验框,尺寸分别为 (38, 70)、(55, 101)、(67, 111)、(85, 95)、(59, 152)、(83, 126)、(106, 108)、(78, 178)、(100, 158)。

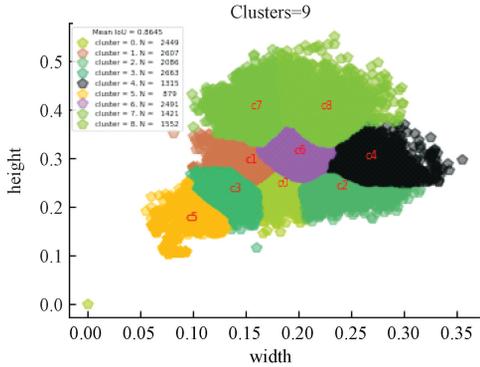


图 5 聚类可视化
Fig. 5 Cluster visualization

2.2 使用 CIoU 作为边界损失函数

交并比 (intersection over union, IoU) 用来衡量两个边界框重叠的相对大小,是目标检测领域评价目标检测器性能的一个重要指标,预测框与真实框的重叠度越高, IoU 值越大。IoU 的定义如下:

$$IoU = \frac{P \cap G}{P \cup G} \quad (2)$$

式中: P 表示目标预测框; G 表示目标真实框。

IoU 可以表示两个边框的距离,且具有非负性、尺度不变性,因此经常被用来作为损失函数。但当预测框和真实框不重叠时, IoU 值为 0, 导致优化损失函数时梯度为 0, 网络的权重和阈值无法更新, 网络模型不能收敛到全局最优。如图 6 所示, 图 6 (a)、(b) 中 IoU 值都为 0, 真实框和预测框的相对位置却无法用 IoU 表示。

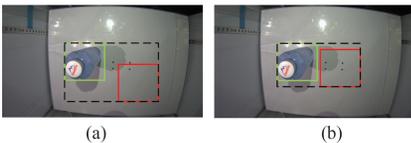


图 6 IoU 为 0 的不同场景

Fig. 6 Different scenarios where IoU is 0

为了弥补 IoU 的不足,本文引入 CIoU^[12] 作为算法的边界框回归损失函数。CIoU 不仅将真实框与预测框之间的距离、重叠率及尺度作为评判标准,并且加入惩罚项,使得目标回归精度更准确,loss 值收敛更快,不会出现 IoU 在训练过程中出现的发散问题。CIoU 定义定义如下:

$$CIoU = IoU - \frac{\rho^2(b, \rho^{gt})}{c^2} - av \quad (3)$$

式中: b 表示预测框中心; b^{gt} 表示真实框中心; ρ 是两个中心点的欧氏距离; c^2 表示能够包含预测框和真实框的最小闭包区域的面积,如图 6 黑色虚线框所示; α 为权重函数; v 是用来度量长宽的相似比。参数 α 与 v 的定义如下:

$$\begin{cases} a = \frac{\beta}{1 - IoU + \beta} \\ v = \frac{1}{\pi^2} (\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h})^2 \end{cases} \quad (4)$$

CIoU 边界框损失函数定义为:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, \rho^{gt})}{c^2} + av \quad (5)$$

2.3 DS_YOLOv3 算法模型的建立

YOLOv3 算法计算复杂度高且对硬件资源依赖较高,部署在嵌入式平台检测速度慢。针对这一问题,本文构建轻量级目标检测网络,使用 MobileNetV2 主干网络替换 Darknet53 网络,采用点卷积和深度卷积逆向残差结构来构造多尺度预测网络。

本文提出的 DS_YOLOv3 网络结构如图 7 所示。其中 DBL 为卷积+Batch Normalization+relu6,是该网络的最小组件,PDP 为 $stride = 2$ 时的倒置残差块;而 IRLB 为 $stride = 1$ 时带有 shortcut 的倒置残差块。图 7 中红色虚线是 MobileNetv2 主干网络,提取商品图片特征信息。主干网络之后是特征交互层,分为 3 个尺度。每个尺度内,通过卷积核的方式实现局部特征交互,输出 13×13 、 26×26 、 52×52 的特征图,本文采用 9 种饮料制作的数据集,故每个特征图包含 42 个通道,在此基础上对目标进行分类和位置回归。

3 实验结果与分析

3.1 数据集与实验设置

为了验证改进算法的有效性,在售货柜内部采集图片,共收集 9 类市场上常见的饮料,数据集类别分别是可乐、雪碧、芬达、脉动、尖叫、百岁山、营养快线、茶 π、红牛。本数据集共 8 512 张图片,包含 17 463 个样本框,使用 MRLabeler 软件对样本进行标注,建立一个完善的商品数据集。其中 6 808 张作为训练集和验证集 1 704 张作为测试集,商品数据集目标类别及样本框数量分布如图 8 所示。

本文使用 Darknet 框架搭建网络,在服务器进行训练,嵌入式平台部署。服务器的配置为 i5-9300H 处理器、16G RAM、RTX1660Ti 显卡、Ubuntu 18.04 系统,安装了 CUDA 和 cuDNN。售货嵌入式开发板搭载 ARM A57

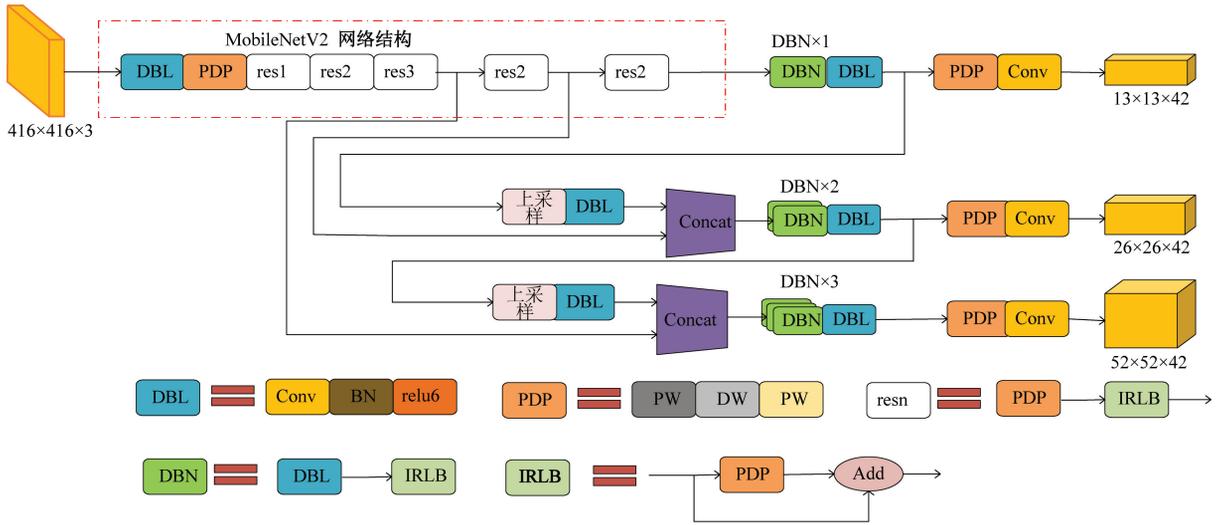


图 7 改进的算法网络结构

Fig. 7 Improved algorithm network structure

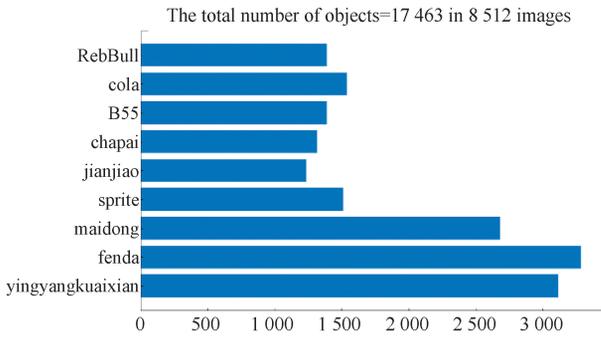


图 8 商品数据集目标类别及样本框数量分布

Fig. 8 Target category and sample box quantity distribution of commodity dataset

处理器、8GRAM、Volta-GV10B 显卡与 Ubuntu 18.04 系统。

为验证本文改进算法的性能,对三种网络模型进行评估。其中将使用深度可分离卷积、倒置残差结构和边界损失函数改进的网络记为 DS_YOLOv3,并与 YOLOv3 与 YOLOv3-tiny 进行对比实验。检测速度、AvgIoU、mAP、计算量作为训练后的模型评价指标。其中 mAP 可由以下计算公式得到:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (7)$$

$$AP = \int_0^1 P(R) dR \quad (8)$$

$$mAP = \frac{\sum_{n=1}^N AP}{N} \quad (9)$$

式中: P 表示精确率; TP 表示正例预测为正例的样本个数; FP 表示负例预测为正例的样本个数; R 表示召回率; FN 表示正例预测为负例的样本个数; AP 表示每一个类别的平均精度; N 表示数据集类别数。检测速度计算公式如下:

$$FPS = \frac{F}{S} \quad (10)$$

式中: F 表示处理图片的数量; S 表示处理 F 张图片所用的时间。

本文 3 种算法以 YOLOv3 论文数据集训练时的参数设置为基础,并对部分参数调整,经过反复实验选取最佳参数。样本共进行 50 000 次迭代训练,其中初始学习率设置为 0.001, momentum 设置为 0.9,表示该训练采用加速梯度下降法,加速模型收敛,减少迭代次数, weight decay 设置为 0.000 5 使得网络权值逐渐衰减,防止模型过拟合,学习率分别在 40 000 和 45 000 减小到之前的 0.1 倍,学习率随着训练次数的增加而衰减,有利于模型训练达到全局最优,每迭代 5 000 次保存一次模型,使用 k-means++ 聚类出的 9 个先验框,将训练后的模型在测试集上测试。

3.2 实验结果与分析

1) 服务器上对比实验

表 1 为 YOLOv3、YOLOv3-tiny 与本文改进算法 DS_YOLOv3 在商品数据集上的实验结果。其中 DS_YOLOv3 获得 52.50 F/s 的检测速度,在测试集上 mAP 达到 96.73%。相比于传统 YOLOv3 网络,DS_YOLOv3 网络在检测速度提升 4 倍,AvgIoU 提升 0.52%,计算复杂度减少 12 倍,mAP 仅仅下降 1.08%。

表 1 网络性能对比

Table 1 Network performance comparison

网络	图片数	检测速度/fps	AvgIoU/%	mAP/%	理论计算量(FLOPs)
YOLOv3	1 740	12. 83	87. 05	97. 81	65. 36B
YOLOv3-tiny	1 740	57. 31	84. 33	92. 86	5. 46B
DS_YOLOv3	1 740	52. 50	87. 57	96. 73	5. 51B



图 9 商品检测效果对比

Fig. 9 Comparison of commodity inspection results

可以看出 DS_YOLOv3 使用深度可分离卷积替换普通卷积和倒置残差结构构造网络模型,大大减少计算量,提高模型检测速度。

相比于 YOLOv3-tiny 算法,DS_YOLOv3 算法检测速度上稍有下降,计算复杂度相近,但检测精度提升 3. 87%,AvgIoU 提升 3. 22%。可以看出 DS_YOLOv3 算法使用新的边界损失函数 CIoU,不仅饮料目标定位更加准确,而且也提高了算法检测精度。因此在检测精度与检测速度的综合考虑下 DS_YOLOv3 适合在嵌入式平台部署。

对改进前后检测结果对比,如图 9 所示。图 9(a)、(b)分别为 3 种算法检测结果。根据两组实验可以看出 YOLOv3 和 DS_YOLOv3 算法均能准确识别商品的类别和位置,识别效果较好,而 YOLOv3-tiny 算法出现了漏检和误检的情况。实验 1 中 YOLOv3-tiny 算法将茶 π (图 9(a)红色箭头)误检为尖叫,由于茶 π 被尖叫遮挡较多,且两种饮料瓶盖特征极为相似,导致误检情况的发生。实验 2 中 YOLOv3-tiny 算法漏检了一瓶芬达(图 9(b)红色箭头),由于目标框定位不准,图中两瓶芬达只被一个目标框所识别,且两个瓶盖都在同一框内,因此导致该目标漏检。

2) 嵌入式平台上对比实验

在 Nvidia Xavier NX 系统平台上,对传统 YOLOv3 和

本文改进算法 DS_YOLOv3 对测试集分别测试。检测结果如表 2 所示。由表 2 可知,本文算法在嵌入式开发板上检测单张图片耗时 49. 16 ms,比 YOLOv3 快 2 倍,达到 20. 34 fps 的检测速度。

表 2 嵌入式端检测时间对比

Table 2 Comparison of embedded detection time

网络	图片数	检测速度/fps
YOLOv3	1740	10. 53
DS_YOLOv3	1740	20. 34

智能售货柜商品识别实验平台如图 10 所示。图 10(a)为智能售货柜,每层顶部安装一个高清摄像头,通过拍摄开门前后的图像传输到开发板上进行识别;图 10(b)为智能售货柜硬件部分,通过接收每层摄像头拍摄的图像,在 DS_YOLOv3 模型上得到检测结果。

将本文改进算法 DS_YOLOv3 部署到开发板上对商品进行检测,结果如图 11 所示。图 11(a)是初始检测结果,图 11(b)~(d)是顾客取出物品不同时刻的检测结果。时刻 2 比时刻 1 少了 1 瓶脉动,时刻 3 比时刻 2 少了 1 瓶营养快线和 1 瓶尖叫,时刻 4 比时刻 3 少了 1 瓶可乐,因此可以根据不同检测时刻的检测结果计算出顾客的消费金额。可以看出该算法能够对售货柜中饮料进

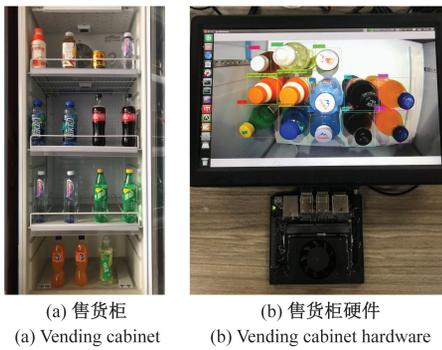


图 10 商品目标检测实验平台

Fig. 10 Commodity target detection experimental platform

行准确分类和定位,且能根据商品变化得到顾客取走的商品,网络模型达到预期效果。

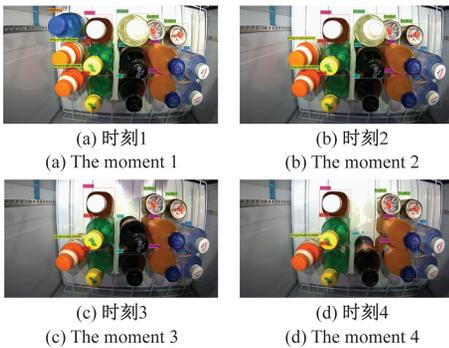


图 11 检测结果

Fig. 11 Test results

4 结 论

本文提出的 DS_YOLOv3 算法,减少了传统 YOLOv3 算法的复杂度,提高了目标检测精度和定位精度,在 Nvidia Xavier NX 嵌入式系统平台上部署,检测速度是原始 YOLOv3 算法的 2 倍左右,mAP 为 96.73%,完全可以满足一个基于嵌入式系统平台对目标检测速度和精度的要求。

下一步的工作重点将继续改进算法模型,研究在减少售货柜摄像头数量的条件下,增加对商品的动态识别,降低售货柜成本。

参考文献

- [1] 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.
- [2] 陈丹,林清泉.基于级联式 Faster RCNN 的三维目标最优抓取方法研究[J]. 仪器仪表学报,2019,40(4):229-237.

CHEN D, LIN Q Q. Research on 3D object optimal grasping method based on cascaded Faster RCNN [J]. Chinese Journal of Scientific Instrument, 2019,40(4):229-237.

- [3] DAI J, LI Y, HE K, et al. R-FCN: Object detection via region based fully convolutional networks[C]. Advances in Neural Information Processing Systems, 2016(6):379-387.
- [4] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, 2017:2-3.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]. Computer Vision - ECCV 2016:21-37.
- [6] 张睿鑫,齐艺璇,朱秋煜.基于人头检测与跟踪的人群行为异常检测[J]. 电子测量技术,2018,41(24):72-76.
- ZHANG R X, QI Y X, ZHU Q Y. Detection of abnormal crowd behavior based on head detection and tracking[J]. Electronic Measurement Technology, 2018, 41(24):72-76.
- [7] 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.
- [8] REDMON J, FARHADI A. YOLOv3: An incremental improvement[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [9] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:710-729.
- [10] HAN S, MAO H Z, DALLY W J. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding [J]. Fiber, 2015, 56(4):3-7.
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]. Proceedings of the Advances in Neural Information Processing Systems, 2012:1097-1105.
- [12] 曹红燕,沈小林,刘长明,等.改进的 YOLOv3 的红外目标检测算法[J]. 电子测量与仪器学报,2020,34(8):188-194.
- CAO H Y, SHEN X L, LIU CH M, et al. Improved YOLOv3 infrared target detection algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2020,

- 34(8):188-194.
- [13] ZHENG ZH, WANG P, LIU W, et al. Distance-IoU Loss: Faster and better learning for bounding box regression [C]. Proceedings of the 2020 AAAI Conference on Artificial Intelligence, 2020: 12993-13000.
- [14] 林付春, 张荣芬, 何倩倩, 等. 机器视觉中的智能无人零售系统设计[J]. 电子技术应用, 2018, 44(9): 96-98, 103.
LIN F CH, ZHANG R F, HE Q Q, et al. The design of intelligent unmanned retail system in machine vision[J]. Application of Electronic Technology, 2018, 44(9): 96-98, 103.
- [15] HU B, ZHOU N, ZHOU Q, et al. DiffNet: A learning to compare deep network for product recognition[J]. IEEE Access, 2020, DOI: 10.1109/ACCESS.2020.2967090.
- [16] 魏宏彬, 张端金, 杜广明, 等. 基于改进型 YOLO v3 的蔬菜识别算法[J]. 郑州大学学报(工学版), 2020, 41(2): 7-12, 31.
WEI H B, ZHANG D J, DU G M, et al. Vegetable recognition algorithm based on improved YOLOv3[J]. Journal of Zhengzhou University (Engineering Edition), 2020, 41(2): 7-12, 31.
- [17] 舒军, 吴柯. 基于改进 YOLOv3 的航拍目标实时检测方法[J]. 湖北工业大学学报, 2020, 35(1): 21-24.
SHU J, WU K. Research on real time detection method of aerial targets based on improved YOLOv3[J]. Journal of Hubei University of Technology, 2020, 35(1): 21-24.
- [18] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [19] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size [C]. International Conference on Learning Representations, 2017.
- [20] SANDLER M, HOWARD A, ZHU M, et al. Mobilenet2: Inverted residuals and linear bottlenecks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.

- [21] 王署东, 尹柏强, 何怡刚, 等. 高压局部放电定位模型转换求解与改进 K-means 聚类优化方法[J]. 电子测量与仪器学报, 2018, 32(11): 178-186.

WANG SH D, YIN B Q, HE Y G, et al. High-voltage partial discharge positioning model conversion solution and improved K-means clustering optimization method [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(11): 178-186.

作者简介



侯维岩, 1986 年于西安电子科技大学获得学士学位, 1998 年于清华大学获得硕士学位, 2004 年于上海大学获得博士学位, 现为郑州大学教授, 主要研究方向为分布式智能控制、无线通信和数字图像处理。

E-mail: houwy@zzu.edu.cn

Hou Weiyan received his B. Sc. degree from Xidian University in 1986, M. Sc. degree from Tsinghua University in 1998, and Ph. D. degree from Shanghai University in 2004. Now he is professor at Zhengzhou University. His main research interests include distributed intelligent control, wireless communication and digital image processing.



靳东安, 2018 年于河南工程学院获得学士学位, 现为郑州大学硕士研究生, 主要研究方向为深度学习、目标检测。

E-mail: jda@gs.zzu.edu.cn

Jin Dongan received his B. Sc. degree from Henan Institute of Technology in 2018. He is currently a M. Sc. candidate at Zhengzhou University. His main research interests include deep learning and object detection.



丁英强, 2003 年于郑州大学获得学士学位, 2006 年于天津大学获得硕士学位, 2009 年于天津大学博士学位, 现为郑州大学讲师, 主要研究方向为机器学习、计算机视觉、信号处理和工业自动化相关领域。

E-mail: cnpowerfoot@gmail.com

Ding Yingqiang received his B. Sc. degree from Zhengzhou University in 2003, M. Sc. degree from Tianjin University in 2006, and Ph. D. degree from Tianjin University in 2009. Now he is professor at Zhengzhou University. His main research interests include machine learning, computer vision and optimization algorithms.