

文章编号: 2095-2163 (2021) 02-0080-05

中图分类号: TP391.4

文献标志码: A

基于 TensorFlow Lite 平台的实时目标检测

林向会, 谢本亮

(贵州大学 大数据与信息工程学院, 贵阳 550025)

摘要: 目标检测是计算视觉的重要研究方向之一, 尤其是基于移动设备平台实现快速精准的目标检测功能是非常有必要的。为了能够在移动设备上实现实时目标检测, 本文提出一种基于 Raspberry Pi 4B 硬件平台, 采用 TensorFlow Lite 开发环境, 加载 MobileNet-SSD 网络结构算法的方案。方案采用的 MobileNet 卷积神经网络和 SSD 卷积神经网络结合的方法具有检测速度快、占用内存少等优点。同时, 本文还对 MobileNet-SSD 网络结构算法进行了微小优化。该方案通过在公开的数据集上进行测试, 对 MobileNet-SSD 网络结构算法和其改良算法进行了结果比较, 结果表明其改良算法的检测速度有所提高, 同时其检测精度几乎保持不变, 在检测精准度和检测速度上都有良好表现, 表明该方案具有较高的应用价值。

关键词: 目标检测; 计算机视觉; 网络结构; TensorFlow Lite; SSD; MobileNet; MobileNet-SSD

Real time object detection based on TensorFlow Lite platform

LIN Xianghui, XIE Benliang

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] Target detection is one of the important research directions of computational vision. Especially, it is necessary to achieve fast and accurate target detection based on mobile device platforms. This paper proposes a solution based on the Raspberry Pi 4B hardware environment which uses the TensorFlow Lite development platform and loads the MobileNet-SSD network structure algorithm. The combination of MobileNet convolutional neural network and SSD convolutional neural network used in the scheme has the advantages of fast detection speed and less memory. At the same time, this paper also optimizes the MobileNet-SSD network structure algorithm. The program compares the results of the MobileNet-SSD network structure algorithm with its improved algorithm through experiments on public data sets, and the results show that the detection speed of the improved algorithm has increased, while the detection accuracy remains almost unchanged. Both accuracy and detection speed have good performance, indicating that the program has high application value.

[Key words] target detection; computer vision; network structure; TensorFlow Lite; SSD; MobileNet; MobileNet-SSD

0 引言

随着计算机技术的发展, 目标检测已经广泛应用于生活的各个领域。在交通领域, 目标检测可以通过对行人、车辆、交通灯和道路上的其他物体的检测来实现对交通管理的协助^[1]。医疗领域中, 图像的病理检测和识别也可以通过目标检测来完成^[2]。其他领域, 如遥感图片^[3]、智能监控^[4]和面部识别等等也都应用了目标检测技术。

当前的通用目标检测算法一般基于单阶段过程和两阶段过程。其中, 基于单阶段过程目标检测算法直接生成目标边界框和概率分类, 无需生成候选框区, 代表性的算法主要有 YOLO^[5]、YOLO9000^[6]、SSD^[7]等。基于两阶段目标检测算法要先生成候选区域, 再对候选区域进行分类和调整, 代表性算法主

要有 R-CNN^[8]、Fast R-CNN^[9]等。文献[5]提出的 YOLO 算法实现了端到端的检测效果并提高了检测率, 但由于其粗粒度的运算特征, 检测精度较差。文献[7]提出的 SSD 算法实现了对整个图像的回归检测, 检测速率高。但该算法应用于小物体检测时, 检测精度和准确性较低。文献[10]提出 Faster R-CNN 算法实现了端到端的检测, 该算法的精度高于传统的 R-CNN 算法, 但检测速度较慢, 无法满足实时性检测的要求。此外, 文献[11]提出了一种水平连接的要索金字塔网络 (FPN), 该网络采用多尺度特征映射和自上而下的结构, 在小物体目标检测时的性能优于 YOLO 算法。SSD 算法借鉴了 YOLO 单元的思想 and Faster R-CNN 的锚定机制, 即直接预测对象类别和边界框的多对象检测法。文献[12]中的 YOLOv3 算法引入了一种多尺度预测方法来提高

基金项目: 贵州省半导体功率器件教育部工程研究中心开放基金项目 (ERCMEKFJJ2019-(06))。

作者简介: 林向会 (1994-), 女, 硕士研究生, 主要研究方向: 电路与系统; 谢本亮 (1978-), 男, 博士, 副教授, 主要研究方向: 数字图像处理、电路与系统。

通讯作者: 谢本亮 Email: ic_blxie@gzu.edu.cn

收稿日期: 2020-10-14

小型物体的检测效果。由于 FPN 和 YOLOv3 都有大量参数,即使得检测速度较慢,并且 2 种算法模型的操作都需要高性能 PC 设备支持。

尽管运用在 PC 终端上的算法具有很高的运算效率,但由于 PC 设备的体积大,灵活性差,其应用范围受到了限制。为了达到目标检测算法要用到移动设备的目的,期望目标检测算法的网络不仅要求准确,而且要有很快的检测速度。近年来,研究人员已经提出了数种用于公共目标识别的高效移动神经网络。例如,文献[13]针对移动和嵌入式视觉应用提出了一种称为 MobileNet 的有效模型,文献[14]提出基于 YOLOv3 改进的算法来实现车辆的实时检测,文献[15]提出基于 SSD 算法来实现车辆识别检测。文献[16]提出了 SqueezeNet 模型,这是一种 CNN 架构,其参数比 AlexNet 少 50 倍,并在 ImageNet 上保持 AlexNet 级别的准确性,该模型的参数可存储在 FPGA 内,从而实现移动设备的目标检测。

虽然当前的目标检测技术达到了高精度和速度,但这是在高性能的实验平台上实现的,对于在有限的计算速度和内存资源的移动设备上来说,目标检测的性能下降严重。因此,本文基于 Raspberry Pi 平台,选择 MobileNet 卷积神经网络和 SSD 卷积神经网络,即 MobileNet-SSD 网络结构算法来实现实时目标检测。实验结果表明基于 Raspberry Pi 4B 平台的 MobileNet-SSD 能实时进行目标检测,且检测速度快、精度较高。

1 硬件平台

树莓派是一款微型电脑,相对于普通 PC 机来说,其优点就在于该微型电脑具备普通 PC 机的所有基本功能,体积和重量却远小于普通 PC 机,被称为世界上最小的台式机。基于 Raspberry Pi 4B 的运动目标检测算法研究的意义在于,能够克服 PC 机体积和重量过大的缺点,使对运动目标的检测上更具有便携性。相比于 Raspberry Pi 3B+,Raspberry Pi 4B 的特点为:更快的 CPU 和 GPU,更多、更快的 RAM;支持 USB 3.0 的更高总线速度,还允许板载以太网端口支持真正的千兆位连接(125 MBps);双 micro HDMI 端口连接,并支持 4 K 输出。

TensorFlow Lite 是 TensorFlow 的一个轻量级版本,针对移动设备做了很多优化和改造^[15]。TensorFlow Lite 主要面向移动设备和嵌入式设备,其最大特点为:轻量级、跨平台、快速。TensorFlow Lite 的主要组成部分为:模型转换器、解释执行器、算子库、硬件加速代理。

2 目标检测算法分析

2.1 SSD 算法结构

SSD 算法结构如图 1 所示。SSD 方法基于前馈卷积网络。该网络会生成固定大小的边界框集合,对这些框中存在的对象类实例进行评分,并执行非最大抑制算法以生成最终检测结果。

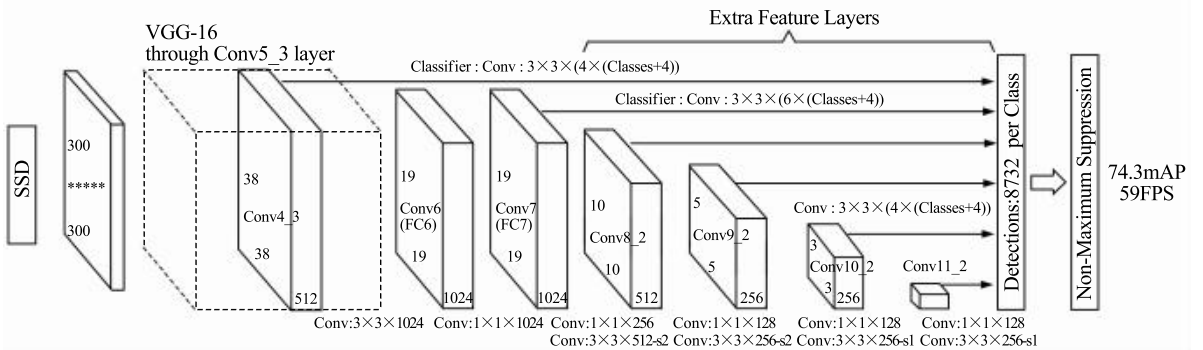


图 1 SSD 算法结构

Fig. 1 SSD algorithm structure

SSD 算法采用的基础网络结构为 VGG16,将 VGG16 最后 2 个连接层 fc6 和 fc7 改为卷积层,再增加 4 个卷积层 conv6, conv7, conv8 和 conv9 来构造网络结构。SSD 算法默认框的个数为 6 个,在对原始图形进行映射时,对应的默认框的大小也不相同,计算相应框大小的数学公式为:

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

其中, S_k 表示先验框大小相对于图片的比例; k 表示特征图的层数; m 表示当前特征图是第几层; S_{\max} 和 S_{\min} 表示默认框和输入图像比率的最大值和

最小值。

SSD 算法采用多任务的损失函数,该函数由 2 部分组成:检测目标类别的分类损失函数和检测目标位置的位置损失函数。函数表达式可写为:

$$L(x, c, l, g) = \frac{1}{N}(L_{con}(x, c) + aL_{Loc}(x, l, g)). \quad (2)$$

其中, N 表示匹配的默认框的数量; x 表示匹配框是否属于类别 p ; g 是真实值; c 表示所选框目标属于类别 p 的置信度。

2.2 MobileNet 网络结构

MobileNet 模型基于深度级可分离卷积构建的网络结构,可将标准卷积拆分为 2 种形式:深度卷积和逐点卷积(1×1 卷积核)^[15]。和标准卷积不同,标准卷积的卷积核是对应所有的输入通道上,而深度卷积为一个卷积核,且只对应一个输入通道。标准卷积和深度卷积过程示意图如图 2 所示。

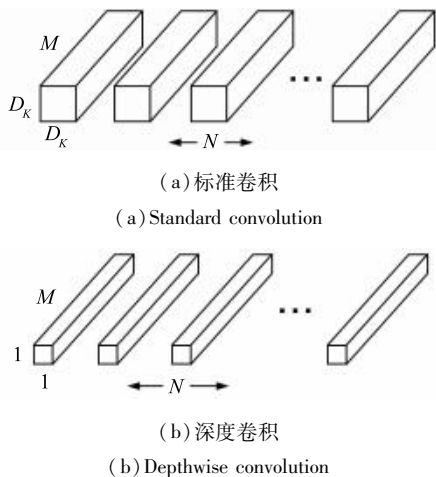


图 2 2 种卷积过程

Fig. 2 Two convolution processes

图 2 中, M 表示输入通道的数量, N 表示输出通道数量, D_k 表示卷积核, $D_k \times D_k$ 表示输出通道数量 N 内核大小。深度可分离卷积先是使用深度卷积对不同输入通道进行卷积,此后再使用点卷积将上一步的输出进行处理。这样一来,卷积分解过程即能有效减少计算量和模型尺寸。

MobileNet 中,只有第一层是普通卷积层外,其他都是深度可分离卷积层。而且除了最后的全连接层没有 ReLU 层外,每个网络层后都有 BN 层和 ReLU 非线性激活层。网络最后的全连接层将续接 softmax 层进行分类。网络的下采样是使用带步长的卷积层来实现的。而在全连接层前,使用了全局池化将张量的尺寸变为 1×1 。文献[13]通过比较普通卷积和深度可分解卷积的计算量得出,由于

MobileNet 使用了 3×3 的卷积,其计算量减少了接近 9 倍。同时准确率几乎保持不变,下降得很少。

2.3 MobileNet-SSD 算法及其优化

MobileNet-SSD 算法是结合了 SSD 算法理论和 MobileNets 网络结构的一种算法结构,相对其它算法来说,MobileNet-SSD 算法能大幅减少计算量,同时其检测速度和精度几乎不变。因此,MobileNet-SSD 算法能够在嵌入式设备上使用。

MobileNet-SSD 算法结构是将 SSD 算法结构里的 VGG16 网络去掉,而后采用 MobileNet 网络。MobileNet-SSD 算法结构大致可分为 2 类。一类是输入层,本文实验中的输入层为 300×300 像素值矩阵;另一类为卷积层,这类卷积层一共有 22 层。前 14 个卷积层 conv0 ~ conv13 和 MobileNet 算法结构的卷积层相同,而于其后增加的 8 个卷积层则用来对网络中进行检测。

MobileNes 算法结构由于采用了深度可分离卷积,其网络模型进行了压缩、网络尺寸得到了减小。但是若要应用于嵌入式的平台的实时检测,其网络结构仍需进行优化。在对 MobileNets 算法结构的前 14 个卷积层分析后得出:对于 Conv0 ~ Conv3 卷积层,深度可分离卷积层和普通卷积层两者计算耗费时长都较多;对于 Conv4 ~ Conv13 卷积层,深度可分离卷积层和普通卷积层计算时间计算耗时有差异,前者计算耗时较少,而后者计算耗时较多。因此,本文通过调整 Conv0 ~ Conv13 卷积层的结构,采用 ReLU6 函数,来实现目标的实时检测。

3 平台搭建与检测分析

3.1 树莓派视觉系统搭建

树莓派视觉系统包括树莓派 4B 主板、IO 扩展板、500 万像素摄像头、PCA 9685 舵机驱动板、16 G 内存卡等配件。在树莓派系统里实现目标检测,需要搭建树莓派桌面,首先使用远程 SSH 工具连接树莓派,接着在树莓派上安装 Tight VNC 包,最后运行远程桌面程序 VNC-Viewer。在 VNC Server 栏中,输入树莓派的 IP 地址,按 connect 连接并输入树莓派用户名和密码后就可以看到树莓派桌面。

在此基础上将进行 TensorFlow Lite 环境搭建^[17],具体步骤为:

- (1)更新 Raspberry Pi。
- (2)下载及保存库并创建虚拟环境。
- (3)安装 TensorFlow Lite 和 OpenCV。

3.2 TensorFlow-Lite 运行及实时目标检测

首先,通过关闭所有不使用的应用程序来释放内存和处理能力;接着,检查摄像头是否正确连线,确保摄像头可以正常使用。将自己在 PC 端训练优化的模型通过 FileZilla 客户端软件,打开名为“TFLite_model”的文件夹,其中包含 detect. tflite 和 labelmap. txt 文件,复制到树莓派桌面文件 home/pi/tflite1 目录。接着,通过从 home/pi/tflite1 目录中发出以下命令来运行实时网络摄像头检测脚本。TFLite_detection_webcam. py 脚本可与 USB 摄像头一起使用。输入命令时要注意字母大小写和文件名不要输错,否则会提示不能打开文件夹,无法运行程序。

实时目标检测的模块内容有:模型训练、模型转换、目标检测、图像获取、图像处理和结果输出。实时目标检测模块的关系示意图如图 3 所示。在命令编写结束后,系统会初始化几分钟,这时就会出现一个窗口,显示网络摄像头。检测的对象将实时显示边框和标签。如图 4 所示。

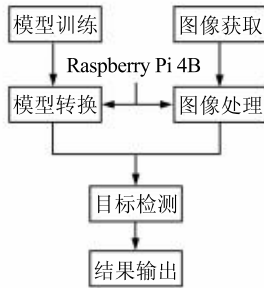


图 3 实时目标检测的模块关系

Fig. 3 Module relationship of real-time target detection

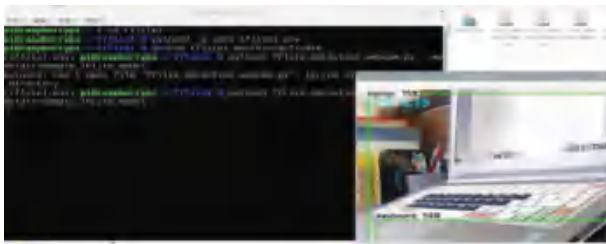


图 4 命令及检测窗口

Fig. 4 Command and detection window

3.3 实时目标检测系统效果及分析

本次实验对数据集分别使用 MobileNet-SSD 算法及其改良算法在 Raspberry Pi 4B 上进行测试,2 种算法测试后的结果见表 1、表 2。通过对比发现,经过优化的 MobileNet-SSD 算法的检测速度有所提高,mAP 下降了 3.728 个百分点,在对精度没有较大减小的情况下,提高了检测速度。MobileNet-SSD

算法及其改良算法各性能的对比见表 3。

表 1 MobileNet-SSD 算法 AP 值

Tab. 1 MobileNet-SSD model AP value

检测目标	AP 值	检测目标	AP 值
Bicycle	84.65	Person	62.82
Bird	84.97	Car	55.79
Bus	83.74	Tv	73.65
Chair	65.94	Train	93.39
Dog	86.14	Laptop	78.29

表 2 MobileNet-SSD 算法改良后的 AP 值

Tab. 2 MobileNet-SSD model optimized AP value

检测目标	AP 值	检测目标	AP 值
Bicycle	82.67	Person	59.32
Bird	83.67	Car	48.89
Bus	77.88	Tv	72.08
Chair	59.12	Train	89.34
Dog	87.44	Laptop	71.69

表 3 算法改良前后各性能的对比

Tab. 3 Comparison of each performance before and after model optimization

	mAP/%	视频检测/FPS
优化前	76.938	6.86
优化后	73.210	7.22

4 结束语

本文研究并实现的是在树莓派上搭建 TensorFlow-Lite 框架并运行转换为 TFLite 模型的目标检测程序,系统设定在 Raspberry Pi 4B 硬件平台上运行,整个系统实现了实时目标检测。经过实验可知检测精度和识别速度较好,表明系统的硬件实现是可行的,可以用在需要移动嵌入式设备进行检测的场景中。

参考文献

- [1] JU Jianguo, XING Jinsheng. Moving object detection based on smoothing three frame difference method fused with RPCA [J]. Multimedia Tools and Applications, 2019, 78(21): 29937-29951.
- [2] HUSSEIN S, KANDEL P, BOLAN C W, et al. Lung and pancreatic tumor characterization in the deep learning era: Novel supervised and unsupervised learning approaches [J]. IEEE Transactions on Medical Imaging, 2019, 38(8): 1777-1787.
- [3] LIU Jing, YANG Shuojin, TIAN Liang, et al. Multi-component fusion network for small object detection in remote sensing images [J]. IEEE Access, 2019(7): 128339-128352.
- [4] XU H, YANG Z, CHEN G, et al. A ground moving target