

基于深度学习的地铁客流实时监测

燕玲

(宁波市轨道交通集团有限公司运营分公司, 浙江 宁波 315000)

摘要: 提出了一种基于深度学习的视频分析方法, 用于高精度实时客流计数。其中, SSD 算法被应用于目标检测, 以提高检测精度的同时提高其性能。为解决 SSD 中特征提取网络参数数量过大以及大计算量的问题, 对 SSD 算法的特征提取网络进行更改, 最终在满足实时性需求的情况下, 算法模型的 mAP 达到 87.9%。

关键词: 深度学习; 客流统计; SSD

中图分类号: TP391.41 **文献标识码:** A

DOI: 10.16248/j.cnki.11-3723/u.2020.27.004

0 引言

地铁列车数据信息化是智能交通系统中不可或缺的一部分。地铁服务质量中的基本评价参数之一是客流量, 该参数在地铁的日常运营、发展规划、安全问题中具有重要作用。现今, 因为缺乏良好的客流监测方法, 在对人流组织方面有所欠缺, 引起行人乘客的安全形势严峻、进行换乘时不方便、服务质量欠佳等问题。目前的视频监控平台无法做到实时对异常情况进行提醒和预警, 因此不能在异常情况发生的早期进行缓解操作阻止时间的恶化。针对目前这一现状, 提出基于深度学习的目标检测算法的地铁客流实时检测的方案。

1 客流检测方法介绍

传统的客流监测方法包括人工客流统计、红外线感应客流统计、三辊闸设备客流统计。其中, 人工客流统计的准确性受人的主观因素影响大、人力成本投入大等因素使得该方法没有被广泛应用。三辊闸客流统计方法虽然能做到准确地统计出人流数量并且投入成本低, 但是严重影响客流的快速通过性, 可能导致更严重的拥堵情况发生。红外线感应统计使用的设备体积小巧, 不影响人流正常通过, 但是若多人同时经过就会出现漏数的问题。发展至此, 仍然未能简便、精准地解决客流的统计问题。

深度学习是含有神经元层数较多的神经网络统称。通过把神经网络层数增加, 使神经网络层数加深, 以此增强神经网络的特征提取能力与筛选能力。图像领域是深度学习技术发展最成熟的领域之一, 在物体识别、目标检测中已经达到了人类水平。目标检测主要可分为 one-stage 以及 two-stage 两个类别。前者在速度上远优于后者, 但在精度上稍逊于后者。考虑到客流检测的实时性, 确定选用 one-stage 类算法。在 one-stage 的目标检测算法中, SSD^[1]以及 YOLO^[2-5]系列算法均能够在满足一定精度的条件下使目标检测速度有极大的提高。虽然 YOLO 算法能够达到实时的要求, 但是其有两个缺陷: 容易导致目标漏检, 并在物体尺

度变化大的情况下泛化性能下降。因此, SSD 在性能上比 YOLO 系列算法更胜一筹。

2 本文算法

SSD 算法在目标检测中具有优异性能的原因是: 首先, 融合了 YOLO 系列目标检测算法中的回归思路和 Faster-RCNN 算法中的生成 Anchor 机制, 利用全图不同位置的不同尺度区域特征进行回归操作, 既具有 YOLO 系列算法速度快的优点, 也确保窗口预测与 Faster-RCNN 预测得一样精确; 其次, 在图像特征图上采用卷积操作来对 Default Bounding Boxes 的类别、坐标偏移等进行预测。与此同时为进一步提升算法的准确率, SSD 算法在各个不同尺度的特征图上进行预测操作。SSD 算法的主干网络是 VGG16^[4]卷积神经网络。VGGNet 证明了神经网络的性能在一定程度上是受网络深度的制约。伴随隐藏层的增加, 模型的非线性拟合能力逐步增强使得准确率明显提升。同时, VGGNet 采用多个小的卷积核代替一个大的卷积核的做法, 使得参数量较之前大幅减少。但是 VGGNet 仍然具有约 140M 的参数量, 对内存的需求较大, 在追求实时性的场景中表现欠佳。

为满足在一定精度的前提下提高算法的实时性要求, 本文使用了 MobileNet^[6-8]作为主干网络, 使用 SSD 作为检测框架。同时, 为进一步减少计算次数, 降低系统整体检测的运行时间并且防止检测丢失的情况, 保障检测性能。在检测之后加入了 KCF^[7]目标跟踪, 以达到提高系统运行速度, 进一步保证了客流检测的实时性, 同时有利于降低整个系统的运行功耗。

MobileNet 是一个小巧而高效的卷积神经网络模型, 在保持模型性能的前提下降低模型的大小。MobileNet 借鉴了 Inception^[8-10]模型的思路, 提出了一种深度可分离卷积的基本单元结构。区别于标准卷积, 深度可分离卷积可看作由深度卷积和点卷积共同完成。

深度卷积的操作是 feature map 中的每一个通道都采用不同的卷积核进行卷积操作。点卷积操作是把经过深度卷

收稿日期: 2019-12-30

作者简介: 燕玲 (1980—), 女, 山西运城人, 研究方向为运营管理。

积后的采用 1×1 的传统卷积操作。虽然得到与传统卷积基本一致的效果,但是模型的参数量以及计算量却会大幅减少。一般卷积的计算量如公式 1 所示,深度可分离卷积如公式 2, 3, 4 所示。可得深度可分离卷积与传统卷积的计算量之比约等于卷积核大小的平方分之一,如公式 5 所示。若以卷积核的大小选取 3×3 为例,计算量则变为原来的 $1/9$,大幅减少网络模型的计算工作量,提升网络的运行速度。

$$CountNormal = D_k \times D_k \times D_{out} \times D_{out} \times M \times N \quad (1)$$

$$CountDW = D_k \times D_k \times D_{out} \times D_{out} \times 1 \times M \quad (2)$$

$$CountPW = 1 \times 1 \times D_{out} \times D_{out} \times M \times N \quad (3)$$

$$Countdsc = CountdPW + CountdDW \quad (4)$$

$$Countdsc / CountNormal = \frac{1}{N} + \frac{1}{D_k^2} \approx \frac{1}{D_k^2} \quad (5)$$

式 (1) ~ (5) 中: D_k 为卷积核的大小, D_{out} 为经过卷积运算得到的特征图大小, M 为特征图通道数, N 为卷积核个数。

3 实验结果

本次实验的环境是在 Ubuntu16.04 的系统下进行,采用 caffe 开源框架。GPU 型号为 RTX 2080 Ti、CPU 为 Intel Xeon E5-2440 v2。训练集来源于深圳地铁各站口摄像头的视频数据。将得到的视频数据读取为图片数据并进行打标,制作成 lmdb 格式的数据集,最终获得的效果如图 1 所示。除远距离、过于密集目标检测不出外。基本能满足客流统计的要求,目前该模型的 mAP 达到 87.9%。满足客流检测场景的使用需要。



图 1 利用目标检测进行客流统计实验结果

4 结语

相较于传统的人工客流统计法、三辊闸客流统计法、红外线感应客流统计方法,基于深度学习的目标检测客流统计方法能做到投入成本低、维护成本低,不影响通道的通行性,受外界因素影响少,统计精度高的优点。此方法

在目标检测技术方面具有优势,为解决客流实时监测提供了理论方法,具有较高的可操作性。

参考文献:

- [1] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[J]. Computer Vision-ECCV, 2016.
- [2] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[J]. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [3] Wang Y, Zhang H. Pedestrian detection and counting based on ellipse fitting and object motion continuity for video data analysis[J]. International Conference on Intelligent Computing, 2015.
- [4] 张雅俊, 高陈强, 李佩, 等. 基于卷积神经网络的人流量统计[J]. 重庆邮电大学学报(自然科学版), 2017, 29(2): 265-271.
- [5] 李继秀, 李啸天, 刘子仪. 基于 SSD 卷积神经网络的公交车下车人数统计[J]. 计算机系统应用, 2019, 28(3): 51-58.
- [6] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. ICLR, 2015 (10): 5-8.
- [7] Andrew G. Howard, Menglong Zhu, Bo Chen, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[J]. Computer Vision and Pattern Recognition, 2017(17): 56-58.
- [8] Henriques João F, Caseiro Rui, Martins Pedro, Batista Jorge. High-Speed Tracking with Kernelized Correlation Filters[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(3): 583-596.
- [9] Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions[J]. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
- [10] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[J]. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [11] Arcos-Garcia A, Alvarez-Garcia J A, Soria-Morillo L M. Evaluation of deep neural networks for traffic sign detection systems[J]. Neurocomputing, 2018(316): 332-344.