

基于 Faster RCNN 的智能车道路前方车辆检测方法

史凯静 鲍 泓 徐冰心 潘卫国 郑 颖

(北京联合大学 北京市信息工程重点实验室 北京 100101)

摘 要: 使用 Fast RCNN 方法进行特征提取存在耗时较长且检测准确率较低的问题。为此,结合 Faster RCNN 前方车辆检测模型与 3 种不同大小的卷积神经网络,提出一种改进的前方车辆检测方法,研究对比各方法在 3 种交通场景数据库上的前方车辆检测能力。实验结果表明,与深度卷积神经网络方法相比,该方法提高了车辆检测的准确性和鲁棒性,具有一定的泛化能力。

关键词: 智能车; 前方车辆; 深度卷积神经网络; 训练模型; 准确率

中文引用格式: 史凯静,鲍 泓,徐冰心,等. 基于 Faster RCNN 的智能车道路前方车辆检测方法[J]. 计算机工程, 2018, 44(7): 36-41.

英文引用格式: SHI Kaijing, BAO Hong, XU Bingxin, et al. Forward vehicle detection method of intelligent vehicle in road based on Faster RCNN[J]. Computer Engineering, 2018, 44(7): 36-41.

Forward Vehicle Detection Method of Intelligent Vehicle in Road Based on Faster RCNN

SHI Kaijing, BAO Hong, XU Bingxin, PAN Weiguo, ZHENG Ying

(Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China)

【Abstract】 Using Fast RCNN method for feature extraction takes a long time and the detection accuracy is low, this paper proposes an improved forward vehicle detection method that combines the forward vehicle detection model based on Faster RCNN with Convolution Neural Network (CNN) of three different sizes. The forward vehicle detection ability of different methods on three traffic scene databases is researched and compared. Experimental results show that compared with the deep CNN, this method improves the accuracy and robustness of vehicle detection, and has a generalization ability.

【Key words】 intelligent vehicle; forward vehicle; deep Convolution Neural Network (CNN); training model; accuracy rate

DOI: 10.19678/j.issn.1000-3428.0050993

0 概述

前方车辆检测是智能车辅助驾驶系统的重要组成部分,是实现智能车在道路上安全行驶、躲避前方障碍物的关键技术。近年来,进行前方车辆检测时,常用的主要有基于机器学习和深度学习的方法。基于机器学习的前方车辆检测方法,通过提取图像中的车辆目标及其周边背景的特征检测车辆,如支持向量机(Support Vector Machine, SVM)、随机森林或 Adaboost 算法等,但这些算法都难以充分利用图像的信息,并且不易达到较高的准确率。深度学习

作为机器学习新的研究方向,具有较高的检测准确率和较低的误检率和漏检率,是当前前方车辆检测研究的重要方向。

针对提取的特征网络可能影响前方车辆检测准确率以及实时性的问题,本文将 Faster RCNN 前方车辆检测模型与 3 种卷积神经网络 ZF、VGG-CNN_M_1024 和 VGG-16 相结合,提出一种改进的前方车辆检测方法,研究对比不同模型在 BUU-Vehicle、TSD-Max 前方车辆检测数据集和 KITTI 视觉基准库^[1] 3 种交通场景数据库的前方车辆检测能力。

基金项目: 国家自然科学基金“视听觉信息的认知计算”重大研究计划重点支持项目“智能车驾驶脑认知技术、平台与转化研究”(91420202); 北京市教委科研计划项目(KM201811417006)。

作者简介: 史凯静(1992—),女,硕士研究生,主研方向为深度学习、车辆检测;鲍 泓,教授、博士生导师;徐冰心、潘卫国,讲师;郑 颖,硕士研究生。

收稿日期: 2018-01-29 修回日期: 2018-04-13 E-mail: 1247678671@qq.com

1 前方车辆检测研究现状

常见的基于深度学习的目标检测方法有 CNN、RCNN、MS-CNN、YOLO、Fast RCNN、DCNN 等。文献[2]总结了深度学习的多层非线性函数关系,使得人们能够更合理地对视觉信息进行建模,从而更好地理解图像和视频。文献[3]提出基于区域的卷积神经网络(Region-based Convolution Neural Network, RCNN)^[4],在提取的候选区域基础上,通过使用卷积神经网络(Convolution Neural Network, CNN)从图像中提取可能含有车辆目标的区域特征来检测并识别目标。该方法虽然提高了识别率,但是计算复杂,难以满足实时检测需求。文献[5]使用了AICNN车辆检测方法并提出最大映射高度方法。该方法具有抗噪抗干扰能力,在一定程度上能够准确检测出车辆目标,但适用于场景单一的图像,泛化能力差。文献[6]首先预处理车底候选区域,其次运用卷积神经网络模型检测真正的车辆目标。虽然该方法能够很好地区分车底阴影和非车底阴影干扰,有效地提高车辆检测的准确率和可靠性,但是对超过60 m的小目标算法失效。文献[7]提出一种基于快速区域卷积神经网络(Fast RCNN^[8-9])的场景图像车辆目标发现方法。不足的是该方法需要大量有效样本进行构建,前期样本候选区域^[10]提取过程较为费时。在区域建议网络(Regional Proposal Network, RPN)^[11]中,深度学习的方法结合了选择性搜索(Selective Search)方法和基于区域的卷积神经网络算法的特点,从而得以提升前方车辆的检测技术。RCNN在计算上时间消耗很大,选择性搜索方法要慢一个数量级,在CPU应用中大约每个图像需要耗时2 s。此后这类消耗通过建议框之间共享卷积^[13]而大大降低。文献[14]提出了统一多尺度的卷积神经网络(Multi Scale-Convolution Neural Network, MS-CNN)^[15],将网络分成建议网络和检测网络2个子网络进行车辆检测。通过设计不同的输出层对应不同尺度的目标检测器和去卷积层,并且使用特征的上采样代替输入图像的上采样,可以大大减少内存占用,提高运行速度。该方法在内存和计算方面有了很大的提升,但计算仍然复杂。文献[16-18]为了解决区域建议框提取的效率问题,使用一个深度网络进行区域建议的计算,并采用共享卷积特征的检测网络进行训练。通过这种方式,区域建议计算时间大大减少,利用Faster RCNN模型对大量复杂场景下的车辆训练样本进行学习,能够更准确、稳定地检测出车辆。

2 前方车辆检测方法

首先,输入采集的前方车辆检测数据集图像,将整张前方车辆图片分别输入不同大小的卷积神经网络(ZF、VGG_CNN_M_1024、VGG-16),对感兴趣区域进行特征提取。用区域建议网络(RPN)生成建议窗口,对每张前方车辆图片生成300个建议窗口。把建议窗口映射到CNN的最后一层卷积特征图上。然后,通过RoI池化层使每个RoI生成固定尺寸的特征图。最后,利用探测分类概率和探测边框回归对分类概率和边框进行回归联合训练。Faster RCNN方法整体流程框架如图1所示。

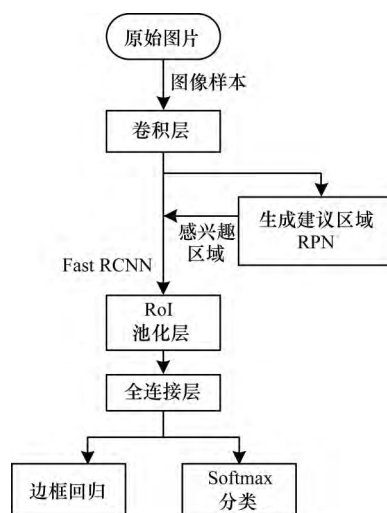


图1 Faster RCNN方法框架流程

本文在进行前方车辆图像目标标注时,使用统一的标注规则。标注框没有重叠,对车辆目标小于 8×8 像素的图像不进行标注。在标注目标完成后,每帧标注图片保存成一个XML文件。

3 前方车辆检测模型的构建

3.1 模型的预训练

使用目标检测的ImageNet分类预训练的模型进行训练网络的初始化,得到前方车辆检测训练模型的初始参数和权重。

迭代训练10 000次得到前方车辆检测粗模型,以此为基础模型继续进行训练,迭代训练110 000次的检测效果较佳。

3.2 前方车辆检测模型的训练

将基于Faster RCNN的模型与ZF小型网络、VGG_CNN_M_1024中型网络、VGG-16大型网络3种不同网络层次结构的卷积神经网络相结合进行迭代训练。通过数次实验对比,择优选择最佳的模型作为前方车辆检测网络的模型。

在Faster RCNN网络中首先进行候选区域的提

取,然后使用区域卷积神经网络对候选区域进行处理判断。这2个网络通过部分层共用和权值共享,简化了模型,提高了训练速度。利用探测分类概率和探测边框回归对分类概率和边框回归进行联合训练,生成建议网络如图2所示。

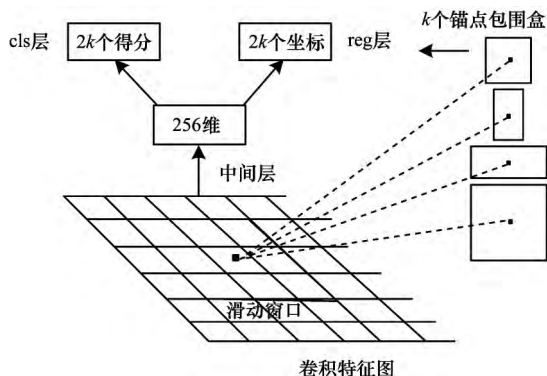


图2 区域建议网络 RPN 的生成过程示意图

在进行边框回归时,假设使用矩形框 F 表示图像的前景锚点,矩形框 G 表示图像中车辆目标的真正位置。将候选窗口的向量用窗口中心点和宽高表示成 (m, n, w, h) 。边框回归的过程就是找到一个回归值 G' 很大程度上接近于真实值 G ,使得 $f(F_m, F_n, F_w, F_h) = (G'_m, G'_n, G'_w, G'_h)$ 映射关系成立。

在训练 RPN 时,给每个锚点分配一个二进制的标签。分配正标签给2类锚点:1)与某个地面实况(Ground Truth, GT)包围盒有最高的交集并集之比(Intersection-over-Union, IoU)(不高于0.7)重叠的锚点;2)与任意 GT 包围盒有(IoU 高于0.7)交叠的锚点。分配负标签给所有 GT 包围盒(IoU 低于0.3)的锚点。

在这些理论基础上,将多任务损失目标函数最小化。将前方车辆图像的损失函数定义为:

$$L = (\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*) \quad (1)$$

对于回归损失,用式(2)计算:

$$L_{\text{reg}}(t_i, t_i^*) = R(t_i - t_i^*) \quad (2)$$

其中 R 是式(3)中定义的具有鲁棒性的损失函数。

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{其他} \end{cases} \quad (3)$$

有些 RPN 建议框和其他建议框大量重叠,为了减少冗余,对其采用非极大值抑制(Non Maximum Suppression, NMS)进行处理。

使用 Faster RCNN 端到端模型进行训练,不需要复杂的特征工程,实现简单快速,在进行训练时 RPN 和 Fast RCNN 两个部分共享卷积特征。通过实验结果分析得出 Faster RCNN 与 VGG-16 的结合在前方车辆检测中能够取得较好的结果。

4 测试过程

4.1 实验平台

本文的实验以北京联合大学研发的“京龙一号”智能车为平台。使用单目摄像头进行真实道路交通场景图像的采集。测试过程使用前方车辆测试集40个,每组平均70帧。智能车实验平台如图3所示。



图3 “京龙一号”智能车实验平台

4.2 实验数据集

实验使用的数据集组成成分2个部分:1)西安交通大学人工智能与机器人研究所提供的 TSD-Max 交通场景数据库;2)北京联合大学智能车队在测试时采集的真实道路交通场景 BUU-Vehicle 数据集。

总的数据集包含约16 000张车辆图片,共6种车型:轿车、越野车、轻型载货汽车、轻型客车、大型客车、重型载货汽车。图片尺寸分为2种:1 280 像素×1 024 像素和1 000 像素×1 000 像素。选取15 000张数据集作为训练验证集以及测试集,将1 000张数据集作为增量测试集,以便研究车辆检测的泛化能力。数据集的图像示例如图4~图6所示。



图4 TSD-Max 数据集车辆类型



图 5 TSD-Max 交通场景数据库



图 6 真实道路交通场景 BUU-Vehicle 数据集

5 实验结果分析

本次实验采用的硬件平台是 Dell precision 深度学习工作站,操作系统为 Ubuntu16.04 平台,编程环境基于 Python,内存为 32 GB, GPU 为 NVIDIA GeForce 1080Ti,在基于深度学习框架 Caffe 下进行实验。将 BUU-Vehicle 随机选择数据集和 TSD-Max 交通场景数据库并集的 80% 作为训练验证集,20% 作为测试集;在训练验证集中,又重新按照相同比例,将数据集分配成训练集和验证集。按照设计思路分别进行实验,通过迭代训练调优,得到最佳的网络模型。

实验评价标准中前方车辆检测任务的评测指标为:

$$recall = \frac{m_{TP}}{m_{TP} + m_{FN}} \quad (4)$$

$$precision = \frac{m_{TP}}{m_{TP} + m_{FP}} \quad (5)$$

$$F - measure = \frac{2precision \times recall}{precision + recall} \quad (6)$$

其中 $precision$ 为查准率, $recall$ 为查全率, $F - measure$ 为 $precision$ 和 $recall$ 的调和平均值, m_{TP} 为前方车辆的正检总数, m_{FP} 为前方车辆的误检总数, m_{FN} 为前方车辆的漏检总数。

表 1 的前方车辆检测数据集包含约 16 000 张车辆图片。从中可以看出,由于车辆目标比较清晰、标注完善,基于 Faster RCNN 与较大网络 VGG-16 结合模型方法的前方车辆检测准确率高于其余 2 种小中型网络结合模型方法,前方车辆的检测率为 92.1%。由此得出,卷积神经网络模型越大,检测效果越好。

表 1 TSD-Max 和 BUU-Vehicle 数据集实验结果 %

| 方法 | 准确率 |
|------------------------------|------|
| Faster RCNN + ZF | 86.6 |
| Faster RCNN + VGG_CNN_M_1024 | 90.1 |
| Faster RCNN + VGG-16 | 92.1 |

表 2 为在标准库 KITTI 上 3 种实验方法的准确率对比情况。从表 2 可以看出,在 KITTI 检测标准库上评估本文方法,改进的 Faster RCNN 的检测准确率优于文献 [19] 使用的深度卷积神经网络 (DCNN) 方法,检测率可达 98.9%。同时使用文献 [9] 提出的将 Selective Search 结合 Fast RCNN 的方法在 KITTI 标准库上测试,前方车辆检测率远低于本文方法,实验结果表明提出的 Faster RCNN 结合 VGG-16 的改进方法对于车辆检测问题具有有效性,并提高了准确率。

表 2 在标准库 KITTI 上实验方法的结果对比 %

| 方法 | 准确率 |
|------------|------|
| 文献 [9] 方法 | 80.9 |
| 文献 [19] 方法 | 98.1 |
| 本文方法 | 98.9 |

表 3 为 2 800 帧 TSD-Max 交通场景数据库测试集检测结果。从表 3 可以看出,本文方法在改进了 Faster RCNN 后,检测所用时间优于未改变的 Faster RCNN 的检测效率和 Fast RCNN 的检测效率。使用有共享的卷积计算 RPN 并且建议框由 2 000 个减少至 300 个。这是一种简洁有效的解决方案,使用了 GPU 进行加速,使得训练时间大大减少。在测试时,通过共享卷积计算建议框给检测网络带来的耗时很小。实验结果表明,本文方法有效地提高了车辆检测的实时性。

表 3 不同方法检测相同数据集车辆图像的耗时对比 s

| 方法 | 总检测时间 | 平均单帧检测时间 |
|------------|---------|----------|
| 文献 [9] 方法 | 5 124.0 | 1.830 |
| 文献 [17] 方法 | 820.4 | 0.293 |
| 本文方法 | 800.8 | 0.286 |

在不同交通环境场景下,采用本文方法在 TSD-Max 测试数据集中的部分检测结果如图 7 所示。



(a)分岔路口



(b)城区道路



(c)高速公路



(d)部分遮挡



(e)光照强烈



(f)背景复杂

图 7 TSD-Max 测试数据集中的检测结果

由图 7(a)可见,在分岔路口处,使用本文方法能够准确地检测出不同行驶角度的车辆;由图 7(b)和图 7(c)可见,在城区道路和高速公路上对于道路场景比较单一的情况,本文方法的车辆检测效果最佳;由图 7(d)可见,即使对于邻车道存在部分遮挡(出现 1/2 以上)的车辆,本文方法仍然能够准确检测出这些车辆;由图 7(e)和图 7(f)可见,对于拍摄的光照强烈以及复杂背景的路段,本文方法能够有效去除光照和背景干扰,准确检测前方车辆,漏检和误检情况低。

由实验结果可以验证在不同的交通环境场景下,本文方法获得了较佳的检测识别效果。

6 结束语

针对在前方车辆检测中特征提取耗时较长和检测效率不佳的问题,本文提出一种基于 Faster RCNN 的前方车辆检测方法,并与 3 种不同卷积神经网络相结合的策略进行对比实验。通过实验验证了该方法有助于提高车辆检测的准确性和鲁棒性。最后将网络修改为多分类识别网络,以验证本文所提方法的强扩展性。下一步将围绕前方车辆的漏检情况和样本标注的问题进行改进,以提高车辆检测召回率和准确率,从而全面提升前方车辆的检测能力。

参考文献

- [1] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: the KIT dataset [J]. International Journal of Robotics Research 2013, 32(11): 1231-1237.
- [2] 郑胤, 陈权崎, 章毓晋. 深度学习及其在目标和行为识别中的新进展[J]. 中国图象图形学报, 2014, 19(2): 175-184.
- [3] TANG Yong, ZHANG Congzhe, GU Renshu, et al. Vehicle detection and recognition for intelligent traffic surveillance system [J]. Multimedia Tools and Applications 2017, 76(4): 5817-5832.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. USA: IEEE Press 2014: 580-587.
- [5] 邹冲, 蔡敦波, 赵娜, 等. 基于 SVM-LeNet 模型融合的行人检测算法[J]. 计算机工程, 2017, 43(5): 169-173.
- [6] ZHANG Chong, LI Hongsheng, WANG Xiaogang, et al. Cross-scene crowd counting via deep convolutional neural networks [C]//Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Washington D. C. USA: IEEE Press 2015: 833-841.

- [7] 李琳辉,伦智梅,连静,等.基于卷积神经网络的道路车辆检测方法[J].吉林大学学报(工学版),2017,47(2):384-391.
- [8] 曹诗雨,刘跃虎,李辛昭.基于Fast R-CNN的车辆目标检测[J].中国图象图形学报,2017,22(5):671-677.
- [9] GIRSHICK R. Fast R-CNN [C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2015: 1440-1448.
- [10] WANG X, SHRIVASTAVA A, GUPTA A. A Fast-RCNN: Hard positive generation via adversary for object detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 3039-3048.
- [11] 韩凯,张红英,王远,等.一种基于Faster R-CNN的车辆检测算法[J].西南科技大学学报(自然科学版),2017,32(4):65-70.
- [12] UIJLINGS J R, SANDE K E, GEVERS T, et al. Selective search for object recognition [J]. International Journal of Computer Vision, 2013, 104(2): 154-171.
- [13] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 37(9): 1904-1916.
- [14] CAI Zhaowei, FAN Quanfu, FERIS R S, et al. A unified multi-scale deep convolutional neural network for fast object detection [M]. Berlin, Germany: Springer, 2016.
- [15] 姚相坤,万里红,霍宏,等.基于多结构卷积神经网络的高分遥感影像飞机目标检测[J].计算机工程,2017,43(1):259-267.
- [16] 桑军,郭沛,项志立,等. Faster-RCNN 的车型识别分析[J].重庆大学学报(自然科学版),2017,40(7):32-36.
- [17] REN Shaoqing, HE Kaiming, 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, 2015, 39(6): 1137-1149.
- [18] 王盛玉,曾碧卿,胡翩翩.基于卷积神经网络参数优化的中文情感分析[J].计算机工程,2017,43(8):200-207.
- [19] 蔡英凤,王海,陈龙,等.采用视觉显著性和深度卷积网络的鲁棒视觉车辆识别算法[J].江苏大学学报(自然科学版),2015,36(3):331-336.

编辑 顾逸斐

(上接第35页)

- [7] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of go without human knowledge [J]. Nature, 2017, 550(7676): 354-359.
- [8] LILICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning [J]. Computer Science, 2015, 8(6).
- [9] RĂDULESCU R, VRANCIU P, NOWÉ A. Analysing congestion problems in multi-agent reinforcement learning [C]//Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems. [S. l.]: International Foundation for Autonomous Agents and Multiagent Systems, 2017: 1705-1707.
- [10] 赵冬斌,邵坤,朱圆恒,等.深度强化学习综述:兼论计算机围棋的发展[J].控制理论与应用,2016,33(6):701-717.
- [11] 张德兆,王建强,刘佳熙,等.加速度连续型自适应巡航控制模式切换策略[J].清华大学学报(自然科学版),2010,50(8):1277-1281.
- [12] 凌滨,宋梦实.汽车自适应巡航系统车距控制策略研究[J].计算机仿真,2017,34(11):143-148.
- [13] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing Atari with deep reinforcement learning [EB/OL]. [2017-12-20]. <https://arxiv.org/pdf/1312.5602v1.pdf>.
- [14] SERNA C G, RUICHEK Y. Dynamic speed adaptation for path tracking based on curvature information and speed limits [J]. Sensors, 2017, 17(6): 1383.
- [15] THRUN S, MONTEMERLO M, DAHLKAMP H, et al. Stanley: the robot that won the DARPA grand challenge [J]. Journal of Field Robotics, 2006, 23(9): 661-692.
- [16] HAN X, BAO H, XUAN Z, et al. A predictive control algorithm based on driving behavior data model [C]//Proceedings of International Conference on Computational Intelligence and Security. Washington D. C., USA: IEEE Computer Society, 2017: 390-394.

编辑 金胡考