

基于 Faster R-CNN 和增量学习的车辆目标检测^①



张子颖, 王 敏

(河海大学 计算机与信息学院, 南京 210098)
通讯作者: 张子颖, E-mail: 137718791@qq.com

摘 要: 随着深度学习的研究热潮, 近年来对车辆目标的检测逐步从机器学习方法转变为深度学习方法. 目前, 大多数深度学习对车辆目标的检测都存在不同程度的错检漏检问题. 针对车辆目标检测中存在的小目标的错检漏检、截断式待检目标的漏检和重叠遮挡待检目标的漏检等问题, 提出一种基于增量学习数据集的车辆目标检测方法, 该方法与 Faster R-CNN 算法结合对车辆目标实现检测和分类. 在实验的最后, 分别从主观判断和客观检测数据两个方面, 对比了车辆目标检测中未使用增量学习方法和使用增量学习方法对实验结果的影响. 实验证明, 使用基于增量学习和 Faster R-CNN 的车辆目标检测方法在主观判断方面对错检漏检的目标有明显地改善效果, 从客观数据分析, 使用该方法与未使用增量学习方法相比, VGG16 网络 mAP 值提升了 4%, ResNet101 网络 mAP 值提升了 6%.
关键词: 深度学习; 机器学习; 增量学习; Faster R-CNN 算法; 目标检测

引用格式: 张子颖, 王敏. 基于 Faster R-CNN 和增量学习的车辆目标检测. 计算机系统应用, 2020, 29(2): 181-186. <http://www.c-s-a.org.cn/1003-3254/7259.html>

Vehicle Detection Based on Faster R-CNN and Incremental Learning

ZHANG Zi-Ying, WANG Min

(College of Computer and Information, Hohai University, Nanjing 210098, China)

Abstract: With the research boom of deep learning, vehicle target detection has gradually changed from machine learning to deep learning in recent years. At present, most of the deep learning methods have different degrees of error detection and omission in vehicle target detection. The vehicle detection method based on incremental learning dataset is proposed to solve the problems of small targets error detection and the omission of truncated and overlapping targets. This method is combined with faster R-CNN algorithm to detect and classify vehicle targets. At the end of the experiment, the influence of with/without incremental learning method on the experimental results was compared from two aspects of subjective judgment and objective test data. The result show that the vehicle detection method based on incremental learning and faster R-CNN has significantly improved the performance in subjective judgment of the missed targets. Objective data also show that the VGG16 network mAP value is increased by 4% and the ResNet101 network mAP value is increased by 6% compared with the incremental learning method.

Key words: deep learning; machine learning; incremental learning; faster R-CNN algorithm; target detection

1 引言

随着科学技术的发展, 智能交通系统应运而生. 智能车辆作为智能交通系统的重要组成部分, 成为了研

究人员研究的热点问题. 智能车辆包括车辆自动驾驶技术和车辆目标检测技术, 其中车辆目标检测技术是智能车辆领域和智能交通系统的重要应用, 它在提高

^① 收稿时间: 2019-07-02; 修改时间: 2019-07-23; 采用时间: 2019-07-26; csa 在线出版时间: 2020-01-16

车辆自动驾驶技术安全性方面起着关键的作用。国内外研究人员对车辆目标检测提出许多不同的算法,这些算法主要分为机器学习和深度学习两类。张全发等使用梯度方向直方图(HOG)结合支持向量机(SVM)分类器进行训练,实现了工程车辆图像检测。但是该方法对于车辆边缘的定位不太理想^[1]。Zhang YZ等改进了多维 Haar 特征和 Adaboost 算法,实现了级联分类器训练,从而实现了可靠的车辆检测。与传统的算法相比,它可以有效提高车辆检测的准确率和检测效率^[2]。Qiu QJ等利用 Adaboost 分类器对车辆类别进行分类^[3]。上述机器学习算法,大多都是基于人工设计的特征,这些特征都有自己应用的局限性,算法泛化能力和鲁棒性不太理想^[4],为改善这些算法的弊端,基于深度学习的卷积神经网络应运而生。如今,卷积神经网络已广泛应用于目标检测。其中,基于区域的卷积神经网络发展非常迅速,包括 R-CNN、Fast R-CNN、Faster R-CNN、MS-CNN。R-CNN 采用 Selective Search 方法得到物体候选框,然后送入卷积神经网络提取特征,最后通过 SVM 分类器获得最终的检测结果,该方法效率低,计算复杂。Fast R-CNN^[5]采用兴趣区域(ROI)池化层得到候选框,大大提升了检测效率。但提取候选框需要大量时间,因此降低了整体的检测效率。Faster R-CNN^[6]设计了一个区域建议网络(RPN)来生成区域建议框,采用共享卷积层特性来提高速度和性能,但它对小目标的检测效果不太理想^[7]。MS-CNN^[8]由提议子网和检测子网组成,该算法精度明显提高,大大节省了内存和计算量,但它需要高设备性能的支持。

目前,国内外研究人员对车辆目标检测已有不错的研究成果。曹诗雨等使用 Fast R-CNN 算法对公交车和小汽车两类图像进行训练和测试,证明了测试样本场景与视觉任务的相关度,以及车辆形变的程度决定了车辆检测的最终效果^[9]。桑军等使用 Faster R-CNN 目标检测模型与 ZF、VGG16、ResNet101 3 种卷积神经网络分别结合的策略,验证了该方案在 BIT-Vehicle 和 CompCars 两个数据集上的泛化能力^[10]。陈超时等提出一种新的车脸定位的方法,该方法通过修改卷积神经网络结构来解决车脸识别在无车牌情况下无法定位的问题,取得良好的检测精度和较快的实时处理效率^[11]。Zhang YJ 等提出使用多线程技术通过并行计算来检测和跟踪车辆^[12]。徐乐先等结合过融合数据分层法和 Timsort 算法实现对停车场停车位的检测^[13]。

本文主要提出一种基于增量学习的车辆目标检测方法,结合国内外目前的研究近况和现有的设备支持,选用 Faster R-CNN 算法与 VGG16^[14]、ResNet101 两种卷积神经网络分别结合的策略,多次测试数据集,最后对比测试结果。实验结果表明,与未使用增量学习方法的车辆目标检测方法相比,识别精度得到了显著提高。

2 基于 Faster R-CNN 和增量学习的车辆目标检测算法

对交通道路上的车辆目标分类检测的困难在于图像中总会存在不完整的待检目标。监控视频和人为拍摄图像中的车辆在大小、姿态上都会有很多的变化。本文使用 Faster R-CNN 网络对车辆进行提取特征并检测目标。

2.1 网络结构

基于 Faster R-CNN 的车辆目标检测网络结构主要由 3 部分构成:卷积层、RPN 层和坐标回归层。其中卷积层用来提取图像特征,特征提取部分与普通 CNN 网络特征提取结构相似,用一串卷积层和 pooling 从原图像中提取 feature map。本文使用 VGG16 和 ResNet101 分别实现车辆目标的特征提取。RPN 层是 Faster R-CNN 提出的全新结构,作用是通过网络训练的方式从 feature map 中获取目标的大致位置,在下一小节中,将对 RPN 结构做详细地介绍。全连接层将接受特征提取所产生的特征向量,所有特征向量将输入全连接层最终将结果共享,产生两个支路,进入两个不同的层中。一个层负责利用 Softmax 回归计算图像中目标类别的概率估算值,另一个层负责输出用来表征每个图像上目标类别检测框坐标的 4 个值(检测框左上角坐标 (x, y) 和右下角坐标 (w, h))。该网络结构图如图 1 所示。

2.2 RPN 网络

区域建议网络 RPN 与传统的 Selective Search 方法的作用相似,都是输出一批矩形候选框,但 RPN 网络更高效^[15]。RPN 的主要功能是根据特征图和输入图像生成候选窗口。整个过程分为两个步骤:(1)将特征图上的每个特征点映射到输入图像的相应位置,以输入图像中的位置为中心绘制一个区域建议框。(2)通过计算候选区域和原标记框之间的 IoU 来获得候选窗口的样本,将候选窗口的样本和特征点用于监督训练。回归层用于确定候选窗口的位置和大小,分类层用于确定候选窗口的前景和背景置信得分。具体过程如图 2 所示。

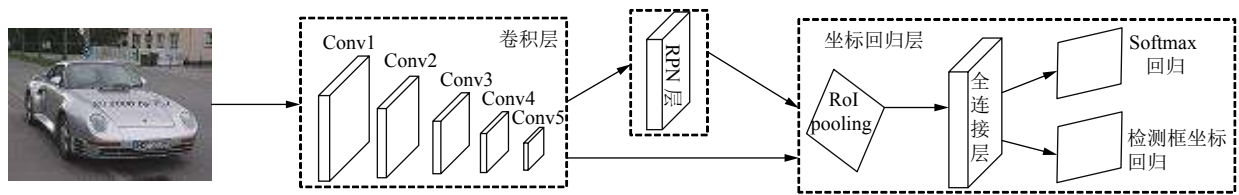


图1 基于 Faster R-CNN 的车辆目标检测网络结构图

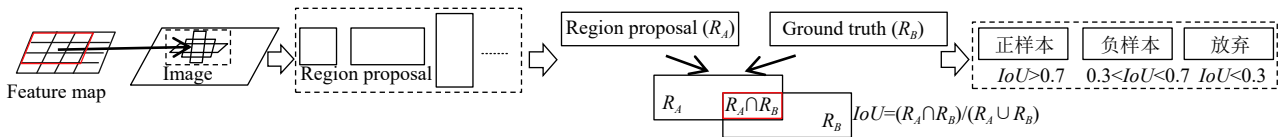


图2 RoI 的生成过程

感兴趣区域 (RoI) 生成过程中, 输入的是图像, 输出是候选区域. 如果 $IoU > 0.7$, 候选区域被视为正样本, $IoU < 0.3$, 候选区域被视为负样本, 当 $0.3 < IoU < 0.7$ 时, 候选区域视为放弃.

监督训练过程输入是图像和标签, 输出是候选区

域的置信度得分和坐标. 回归层用于微调候选区域的位置. 候选区域可以通过分类层获得置信得分, 通过回归层和分类层获取对象和对象的位置, 通过候选区域的置信得分和坐标来获得候选窗口. 具体过程如图3所示.

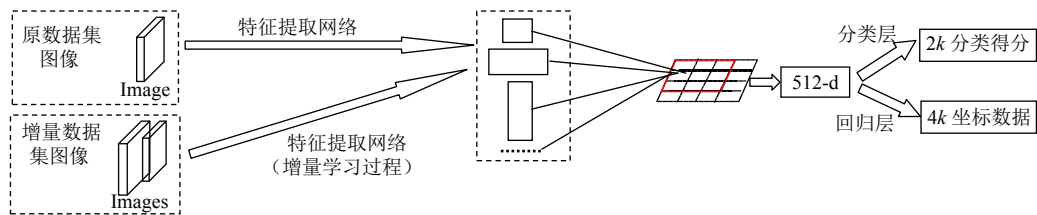


图3 监督训练过程

2.3 增量学习

随着人工智能和机器学习的发展, 出现了很多机器学习算法, 这些算法大都是批量学习模式, 即假设在训练前可一次性获取所有训练样本, 学习已知样本后, 学习过程即终止. 对已学习到的模型, 存在各种缺陷, 若只使用同一样本反复训练, 机器学习到的特征有限, 会限制模型的有效性. 增量学习算法可以渐进地更新知识, 对机器学习增加新的样本知识, 可修正和加强以前的知识, 使更新后的模型更完善.

在本文中, 对 VOC2007 数据集中车辆目标的检测存在两个问题: (1) 车辆目标形变较大, 小目标车辆较多, 使用 VOC2007 数据集训练所得的模型对小目标车辆的检测效果不太理想; (2) VOC2007 数据集中图像背景比较复杂, 机器学习不能更为准确地提取单一车辆目标的特征, 导致待检图像中一些截断式车辆目标被漏检. 为解决这两个问题, 选用了 MIO-TCD 数据集作

为增量数据集对原模型进行重新训练测试, 选择理由主要有两个: (1) 该数据集中小目标车辆较多, 可以很好地学习小目标车辆的形态特征; (2) 该数据集的图像背景较为单一, 可以更加精准地学习车辆目标的特征, 减少了很多区域框内的背景干扰因素.

2.4 算法描述

对道路上车辆目标的检测存在很多问题影响检测效果, 本文以 VOC2007 数据集的检测结果为, 直观分析在 VOC2007 数据集上进行的车辆目标检测效果, 可以发现: (1) 图像中远处小目标出现错检漏检的情况; (2) 对图像边缘截断式待检目标出现漏检; (3) 对待检目标重叠遮盖区域部分出现错检漏检. 为解决这些问题, 本文提出一种基于增量学习的车辆检测方法, 该方法分为两个阶段, 即训练阶段和测试阶段. 在训练阶段, 重新训练在 ImageNet 上预训练后的初始参数的卷积神经网络, 以获得车辆检测模型; 在测试阶段, 将测试

样品输入检测模型以获得测试结果。

为证明本文方法的有效性,在训练阶段首先使用 VOC2007 数据集训练网络得到车辆检测模型 A,并使用 VOC2007 数据集测试模型 A,得出测试结果;其次,选用 MIO-TCD 数据集作为增量学习的数据集,将 MIO-TCD 数据集加入到 VOC2007 数据集中重新训练网络,得到车辆检测模型 B,使用原测试集加 MIO-TCD 数据集中分配出的测试集测试模型 B;最后,对比模型 A 和模型 B 的测试结果.整个方法流程如图 4 所示.

3 实验分析

3.1 实验环境与数据集

实验选取了 Faster R-CNN+VGG16、Faster R-CNN+ResNet101 对使用增量学习数据集和原数据集进行对比,实验环境:操作系统为 Ubuntu18.04,训练 GPU 为 2 块 1080Ti,内存为 448 GB.

实验所用数据集包括原数据集与增量学习数据集,原数据集使用的是来自 VOC2007 数据集中 bus、bicycle、car、motorbike 等 4 类车辆图像,共计 3599 张带标签图像;增量学习数据集为 MIO-TCD 数据集中随机挑选出 1000 张包含 bus、bicycle、car、motorbike 等 4 类车辆图像,4 类图像平均分布,使用 labelImg 标注工具,将 1000 张车辆图像按照 VOC2007

格式进行标注.

3.2 实验结果及分析

实验首先对原数据集按照训练集与测试集 8:2 的比例随机分配,将训练集输入到网络中通过训练迭代得到车辆目标分类模型,再将测试集输入到训练好的模型进行测试,测试结果如图 5 所示.

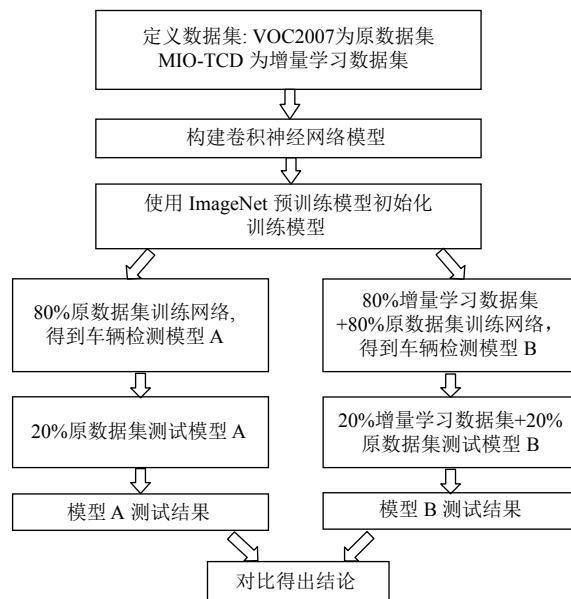


图 4 基于增量学习的车辆目标检测方法流程图



图 5 使用模型 A 测试结果示例图

由图 5 可以看出,使用 Faster R-CNN 网络训练出的车辆分类检测模型去检测车辆目标,检测能力方面体现出非常优秀的效果.对背景清晰单一并且重叠遮盖较少的待检目标,检测区域置信度高达 0.98 以上,但是,在复杂的背景下,如背景干扰目标较多、待检目标互相遮挡重叠、待检目标车体大部分超出图片区域等,此时容易出现漏检错检或检测框置信度较低的情况.

为解决这些问题,我们将增量学习数据集按照训练集与测试集 8:2 的比例随机分配,再将增量学习数据集与原数据集合并,将新的训练集输入到卷积神经网络中重新训练车辆目标分类模型,最后将测试集输入到新的模型进行测试.为直观的区别使用增量学习数据集为车辆目标分类任务带来的改进,我们在原测试结果与新的测试结果中分别挑选出 4 组相同的图像进行对比,对比图如图 6.

由图 6 可以看出,图 6(a)虽检测到汽车遮挡截断的左前方车辆,但区域建议置信度只有 0.946,且对右前方截断式的 car 目标完全漏检,而使用增量学习数据集重新训练并测试后,图 6(e)不仅对左前方遮挡截断式的 car 目标提高了置信度到 0.990,并且也检测到了

右前方 car 目标. 同样, 图 6(b) 中, 对远处小目标有良好的检测效果, 但对图片左右边缘处的截断式 car 目标漏检, 而图 6(f) 对图 6(b) 中漏检的目标做了很好的修正. 对重叠严重的 bicycle 目标图 6(c), 区域建议框标注并不十分准确, 置信度相对较低, 且对远处模糊的目标

出现漏检错检情况, 在图 6(g) 中, bicycle 目标检测框明显变得更为准确, 置信度比原图像的置信度要高, 对远处目标检测也给出了修正结果. 更为明显的图 6(d) 与图 6(h) 对比, 可以明显看出使用增量学习数据集方法对重叠的待检目标有良好的检测效果.



图 6 对截断式和重叠区待检目标测试结果对比图

在主观评价后, 表 1 中使用具体数据对实验结果进行了客观展示. 由于召回率准确率存在单点值局限性的问题, 所以本文采用 AP 值来反映检测结果. AP 值为平均准确率, 是一个能够反应全局性能的指标, 准确率召回率曲线下的面积即为 AP 值. mAP 值为所有不同类 AP 值的平均值, mAP 值越大, 检测效果越好.

表 1 实验结果

模型	方法	Bus	Bicycle	Car	Motorbike	mAP
A	VGG16	0.806	0.808	0.876	0.803	0.823
	ResNet101	0.823	0.853	0.874	0.798	0.837
B	VGG16	0.895	0.816	0.875	0.885	0.868
	ResNet101	0.896	0.895	0.881	0.894	0.891

4 结论与展望

本文提出一种增量学习和 Faster R-CNN 的车辆分类识别的方法. 该方法有效地利用了 Faster R-CNN 算法端到端的网络结构, 能够快速、鲁棒地检测车辆目标. 实验结果表明, 使用 Faster R-CNN 卷积神经网络结构结合增量学习数据集创新性地训练图像, 能够从图像中提取更加全面丰富的目标特征, 使用所得车辆目标分类检测模型测试数据集, 从主观和客观两个角度

来分析, 能够明显看出本文方法存在的优势. 然而, 我们从结果中也发现很多不足的地方, 包括: (1) 对远处的小目标仍会出现漏检错检情况; (2) 在昏暗背景下对目标的检测, 给出的区域建议框置信度较低; (3) 在更为复杂的背景下, 如闹市背景目标重叠度太大、背景干扰目标较多等, 检测效果仍需改进. 因此, 在接下来的研究中, 解决上述问题是未来研究的关键.

参考文献

- 张全发, 蒲宝明, 李天然, 等. 基于 HOG 特征和机器学习的工程车辆检测. 计算机系统应用, 2013, 22(7): 104-107. [doi: 10.3969/j.issn.1003-3254.2013.07.023]
- Zhang YZ, Sun PF, Li JF, *et al.* Real-time vehicle detection in highway based on improved Adaboost and image segmentation. Proceedings of 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems. Shenyang, China. 2015. 2006-2011.
- Qiu QJ, Liu Y, Cai DW. Vehicle detection based on LBP features of the Haar-like characteristics. Proceedings of the 11th World Congress on Intelligent Control and Automation. Shenyang, China. 2014. 1050-1055.
- 武凯利, 全宗和, 张鹏升, 等. 基于深度学习的复杂气象条

- 件下海上船舶识别. 科学技术与工程, 2019, 19(3): 130–135. [doi: [10.3969/j.issn.1671-1815.2019.03.022](https://doi.org/10.3969/j.issn.1671-1815.2019.03.022)]
- 5 Girshick R. Fast R-CNN. Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile. 2015. 1440–1448.
- 6 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
- 7 Li SX, Zhang ZL, Li B. A plane target detection algorithm in remote sensing images based on deep learning network technology. Journal of Physics: Conference Series, 2018, 960: 012025. [doi: [10.1088/1742-6596/960/1/012025](https://doi.org/10.1088/1742-6596/960/1/012025)]
- 8 He KM, Gkioxari G, Dollar P, *et al.* Mask R-CNN. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018. [doi: [10.1109/TPAMI.2018.2844175](https://doi.org/10.1109/TPAMI.2018.2844175)]
- 9 曹诗雨, 刘跃虎, 李辛昭. 基于 Fast R-CNN 的车辆目标检测. 中国图象图形学报, 2017, 22(5): 671–677.
- 10 桑军, 郭沛, 项志立, 等. Faster-RCNN 的车型识别分析. 重庆大学学报, 2017, 40(7): 32–36.
- 11 陈超村, 蔡晓东, 赵勤鲁, 等. 基于多层次特征的高效车脸定位方法. 电视技术, 2017, 41(9–10): 161–165.
- 12 Zhang YJ, Wang J, Yang X. Real-time vehicle detection and tracking in video based on faster R-CNN. Journal of Physics: Conference Series, 2017, 887(1): 012068.
- 13 徐乐先, 陈西江, 班亚, 等. 基于深度学习的车位智能检测方法. 中国激光, 2019, 46(4): 0404013.
- 14 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2015.
- 15 宋焕生, 张向清, 郑宝峰, 等. 基于深度学习方法的复杂场景下车辆目标检测. 计算机应用研究, 2018, 35(4): 1270–1273. [doi: [10.3969/j.issn.1001-3695.2018.04.067](https://doi.org/10.3969/j.issn.1001-3695.2018.04.067)]