

基于卷积神经网络的车辆分类与检测技术研究

解 丹¹, 陈立潮², 曹玲玲¹, 张艳丽¹

(1.山西晋中理工学院信创与大数据学院, 山西 晋中 030600;

2.晋中信息学院大数据学院, 山西 晋中 030600)

✉ever_xd@163.com; 2676393759@qq.com; clzzjzc0708@163.com; xcjiaowu@163.com



摘 要: 车辆分类与检测在智能交通系统、道路交通规划、安全预警、无人驾驶等领域发挥着越来越重要的作用。随着图形处理器(Graphics Processing Unit, GPU)运算能力的增强以及数据量的剧增,以卷积神经网络为主的深度学习成为研究热点。基于LeNet、AlexNet、VGG、GoogLeNet、ResNet等卷积神经网络,介绍并分析了6种车辆分类方法、8种车辆检测方法及4个用于评估这些方法的数据集,并对车辆分类及检测的研究方向进行了展望。

关键词: 卷积神经网络; 车辆分类; 车辆检测; 深度学习

中图分类号: TP391 **文献标识码:** A

Research on Vehicle Classification and Detection Technology based on Convolutional Neural Network

XIE Dan¹, CHEN Lichao², CAO Lingling¹, ZHANG Yanli¹

(1.Department of Information Technology Application Innovation and Big Data, Shanxi Jinzhong Institute of Technology, Jinzhong 030600, China;

2.Department of Big Data, Jinzhong College of Information, Jinzhong 030600, China)

✉ever_xd@163.com; 2676393759@qq.com; clzzjzc0708@163.com; xcjiaowu@163.com

Abstract: Vehicle classification and detection plays an increasingly important role in intelligent transportation systems, road traffic planning, safety warning, unmanned driving and so on. Deep learning based on convolutional neural networks has become a research focus with the enhancement of GPU computing power and the increase in data volume. Based on convolutional neural networks such as LeNet, AlexNet, VGG, GoogLeNet, ResNet, this paper introduces and analyzes 6 vehicle classification methods, 8 vehicle detection methods and 4 datasets used to evaluate these methods. In addition, the research direction of vehicle classification and detection is prospected in this paper.

Keywords: convolutional neural network; vehicle classification; vehicle detection; deep learning

1 引言(Introduction)

随着以深度学习为代表的人工智能技术的发展,越来越多的领域开始应用深度学习技术。基于深度学习的车辆分类与检测是当前智慧交通领域的研究热点,其在智能交通系统中处于基础地位,对车辆跟踪、交通管控、无人驾驶等的发展具有重要意义。深度学习主要涉及卷积神经网络(Convolutional Neural Network, CNN)、自编码神经网络、深度置信网络,其中CNN的发展最为迅速。随着CNN

的发展,其在车辆分类的应用被不断优化,例如对LeNet^[1]、AlexNet^[2]、VGG、GoogLeNet等模型的应用。在车辆检测中对CNN的应用大致分为两类:一类是一阶段检测方法,也称为区域提议方法;另一类是两阶段检测方法,也称为无区域提议方法。本文以CNN的发展为主线,介绍并分析基于LeNet、AlexNet等CNN的车辆分类方法、基于R-CNN及YOLO系列的车辆检测方法以及用于评估这些方法的数据集,并对未来的研究方向进行展望。

2 基于CNN的车辆分类(Vehicle classification based on CNN)

车辆分类是对图像中车辆的种类进行划分。传统的车辆分类方法是通过方向梯度直方图、尺度不变特征转换等方法提取车辆特征,并将特征输入支持向量机、迭代器等分类器中,实现对车辆进行分类。这些机器学习方法在提取车辆特征时大多依赖研究人员的专业知识和经验,受主观因素影响较大,并且难以适应天气、光线、角度等的变化,泛化能力较差。随着车辆数量增多,车辆类型也逐步增多,依靠机器学习提取车辆特征,无法较好地对车辆进行分类。深度学习中CNN训练时具有更多的数据及参数,因而其模型特征表达能力更强,性能优于机器学习。

常见的CNN如表1所示,其中LeNet为最早发布的CNN,由两个卷积层和三个全连接层组成,主要用于手写数字的识别。LeNet虽然在小数据集上取得了较好的效果,但是在大量真实数据中的应用效果还有待研究。计算机硬件的发展,使得训练大数据成为可能。2012年,KRIZHEVSKY等^[2]在GPU上运行深度CNN,实现了卷积运算的快速操作,该网络被称为AlexNet。AlexNet和LeNet类似,但AlexNet比LeNet层数更多,它由五个卷积层和三个全连接层组成。随着AlexNet在训练大数据集上的成功,CNN开始逐渐成为研究热点。SIMONYAN等^[3]提出了VGG网络,SZEGEDY等^[4]提出了GoogLeNet。网络越来越深的发展趋势,使得人们开始思考如何使新添加的网络层更有效地提升网络性能。2016年HE等^[5]提出了残差网络(ResNet),之后HUANG等^[6]对ResNet进行改进提出了稠密连接网络(DenseNet)。随着CNN的发展,其越来越多地被应用于车辆分类领域。

表1 常见的CNN

Tab.1 Common CNN

CNN名称	参数个数/ 10^5 个	层数(卷积层+全连接层)/层	错误率/%(Top5)
LeNet	0.6	2+3	—
AlexNet	570	5+2	15.3
VGG	1,340	13+2	6.8
GoogLeNet	60	22	6.7
ResNet	234	49	3.6
DenseNet	180	200	—

曹林等^[7]对LeNet-5进行改进以学习车辆的雷达视频信号特征,从而对大小车型进行分类,为交通场景下车型识别的研究提供了新思路。马传香等^[8]对AlexNet进行改进,提出一种maxout加dropout的六层CNN用于区分空车和重车,模型对34,220张图像进行测试,结果显示分类的准确度、敏感度都有较大的提高。张超^[9]将VGG-16的网络结构进行改进,减少全连接层的个数及改变池化层的大小和性质,在保持对车辆特征提取能力的基础上,提升了网络速度。刘泽康等^[10]将GoogLeNet应用于车辆检测,发现GoogLeNet需要较长的训练时间才能达到较好的分类精度。余烨等^[11]对ResNet进行改进,提出基于残差网络特征重用的深度卷积神经网络模型FR-ResNet用于车辆型号的精细分类,该模型引入多尺度输入和特征图权重学习策略,能防止网络过深而陷入局部最优。陈立潮等^[12]针对细粒度车辆分类准确率低及深度学习不能

较好地关注车辆鉴别性部件特征的问题,在DenseNet中引入处理层和独立组件增强部件信息和高位特征的提取,在斯坦福大学的Stanford Cars数据集上获得95%的准确率,但此方法重点关注局部鉴别性部件信息,对车辆全局特征的提取能力仍有待提高。

可以看出,车辆的分类包括对车辆的车型大小、车辆的类型、车辆的承重等进行分类,针对不同的分类目标可以采用不同的模型。在改进网络模型时,可以改进网络结构、网络参数或引入其他有利于模型学习的策略。此外,车辆所在的场景对于车辆的分类也有很大的影响,场景越复杂,对车辆分类算法的要求越高。

3 基于CNN的车辆检测(Vehicle detection based on CNN)

车辆检测包含车辆识别及定位,是智能交通系统中重要的一部分。在城市交通拥堵路段,车流量大、车辆之间相互遮挡严重,交通环境复杂,对车辆检测提出了更高的要求。CNN未出现之前,对车辆的检测主要有背景差分法、帧差法、光流法等,这些方法存在可移植性差、鲁棒性差及速度慢的问题。随着GPU并行技术和CNN的发展,以CNN为基础的目标检测算法成为主流。2014年,GIRSHICK等^[13]提出了R-CNN用于目标检测,实现了使用少量标注数据训练高质量模型,但该模型占用的内存较大、效率较低。为了提高R-CNN的效率,GIRSHICK等又提出了Fast R-CNN^[14]、Faster R-CNN^[15],使模型能更好地应用于实时检测。基于R-CNN的目标检测属于两阶段检测,其首先生成目标候选区,然后对候选区的目标分类和定位。此外,YOLO、SSD、RetinaNet等为一阶段检测,它们将单个神经网络用于整张图像,并对目标的类别和边界框进行回归,检测速度较快。其中,YOLO算法发展最快,至今已发展至YOLOv5。2016年REDMON等提出了YOLOv1^[16],将目标检测问题处理为分类回归问题。为实现更好、更快的检测,2017年REDMON等又提出YOLOv2^[17]、2018年进一步提出YOLOv3^[18]。2020年BOCHKOVSKIY等融合深度学习的多种优势改进YOLOv3,提出YOLOv4^[19]。为提高网络的多尺度检测能力,曾凯等^[20]对YOLOv5进行研究。随着目标检测算法的发展,各种算法根据车辆目标检测的特点进行改进并得到实际应用。

刘杰^[21]采用R-CNN对常见的车辆事故场景及车身关键部位进行训练,并结合空间辅助定位检测事故车辆,本质上是对特殊场景的车辆进行检测。张琦等^[22]改进Fast R-CNN,在原网络上引入内置的子网络和尺度感知加权层,实现不同空间尺度的车辆检测。王林等^[23]针对远近场景中车辆尺寸存在较大差异造成Faster R-CNN误检、漏检的问题,在训练Faster R-CNN时引入多尺度训练策略,针对数据集中目标区域远小于背景区域、负样本空间大的问题,在训练时引入难负样本挖掘的策略,两个策略的引入,使Faster R-CNN在检测KITTI数据集的精确度提高了约8%,但该模型训练时对硬件的要求较高。LU等^[24]对三个公共航空影像数据集进行处理,集成一个适合YOLO训练的航空影像数据集,使训练好的YOLO模型能更好地检测小目标、旋转目标及紧凑目标车

辆;张成标等^[25]将DenseNet、多尺度层引入YOLOv2,并设计Kelu激活函数检测不同形态的车辆,提高了车辆检测的准确率和精度;张富凯等^[26]改进YOLOv3,提出单阶段深度神经网络(DF-YOLOv3),将DenseNet的特征图与多尺度预测网络的特征图融合,得到更丰富的车辆特征,以较好地检测不同尺度的车辆目标。为解决车辆检测中的错检、漏检问题,於积荣等^[27]将YOLOv4的主干网络替换为CBAM-DenseNet-BC, SPPNet替换为ASPPNet,并引入CBAM模块,张漪等^[28]将注意力机制SE模块引入YOLOv5,并使用了Focal Loss焦点损失,经过改进的YOLOv5可以较好地检测遮挡车辆和小目标车辆。

对车辆检测的研究致力于解决对不同尺度车辆的检测以及错检、漏检问题。针对不同尺度车辆的检测,大多是在原网络上增加不同大小的卷积层,从而得到不同大小的特征图。对错检、漏检问题的解决大多采用引入优势策略或加深网络层数的方法。除此之外,车辆检测的难点除了车辆尺度、车辆类型的问题,还有车辆所在的场景问题,以及对特殊天气的车辆检测,仍然有待深入研究。

4 车辆分类与检测质量评估数据集(Quality assessment dataset of vehicle classification and detection)

4.1 KITTI数据集

KITTI数据集由德国卡尔斯鲁厄理工学院及丰田美国技术研究院制作,可通过<http://www.cvlibs.net/datasets/kitti>网址下载,主要用于车辆的检测与跟踪。该数据集中有激光雷达数据、标注数据、图像数据及标定校准数据。其中,图像的分辨率为 1392×512 ,包括市区、乡村、校园及高速公路等真实场景,如图1所示。数据集的标注采用激光雷达坐标系,共有Car(小汽车)、Van(厢式货车)、Truck(卡车)、Pedestrian(行人)、Person(sitting)(坐着的人)、Cyclist(骑自行车的人)、Tram(有轨电车)和Misc(其他)八类,有各种程度的遮挡及截断。

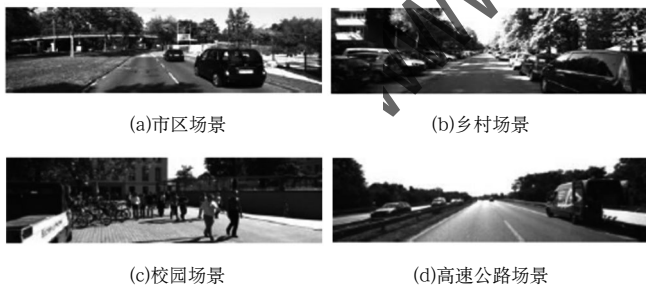


图1 KITTI数据集

Fig.1 KITTI dataset

4.2 BIT-Vehicle数据集

BIT-Vehicle数据集由北京理工大学制作,主要用于车辆的分类与检测,来源于<http://iitlab.bit.edu.cn/mcislav/vehicledb/>网站。该数据集中有9,850张车辆图像,来自两个不同地方的摄像机,图像的分辨率分别为 1600×1200 和 1920×1080 ,包含光照条件、尺度、车辆颜色和视角的变化。部分数据集如图2所示,车辆被分为Bus(公共汽车)、Microbus(中型客车)、Minivan(小型货车)、Sedan(轿车)、SUV(越野车)、Truck(卡车)五类,每类车辆的数量分别为

558辆、883辆、476辆、5,922辆、1,392辆、822辆,共有10,053辆车。

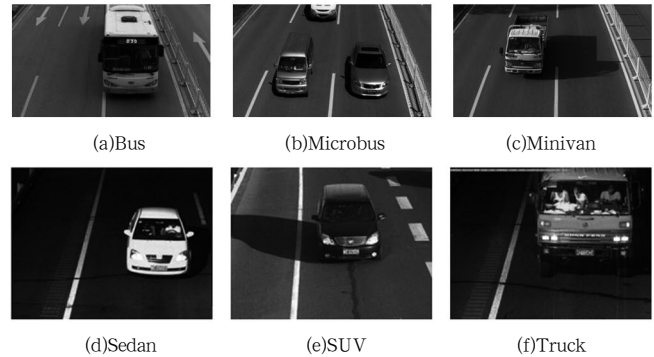


图2 BIT-Vehicle数据集

Fig.2 BIT-Vehicle dataset

4.3 Stanford Cars数据集

Stanford Cars数据集由斯坦福大学整理,可通过http://ai.stanford.edu/~jkrause/cars/car_dataset.html下载,主要用于车辆的细粒度分类。该数据集的类别主要根据汽车的牌子、车型及年份进行划分,共包含16,185张图像,图像中的车辆共有196类,如图3所示。数据集划分为训练集和测试集,训练集中有8,144张图像,测试集中有8,041张图像,在训练集和测试集中每个类别的数量大致相同。



图3 Stanford Cars数据集

Fig.3 Stanford Cars dataset

4.4 UA-DETRAC数据集

UA-DETRAC数据集来源于<http://detrac-db.rit.albany.edu/>,主要拍摄于北京和天津的道路、过街天桥等24个不同的地点,存在多云、夜间、晴天和雨天四种天气状况,视频每秒25帧,每帧图像的分辨率为 960×540 ,如图4所示。该数据集中手动标注了8,250辆车辆,共有121万标记的对象外框。车辆分为轿车、公共汽车、厢式货车和其他四类。此数据集主要用于多目标检测和多目标跟踪的训练和测试。

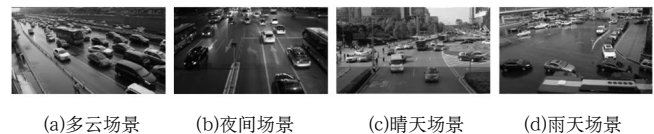


图4 UA-DETRAC数据集

Fig.4 UA-DETRAC dataset

5 结论(Conclusion)

车辆分类与检测对于智能交通系统的发展极为重要。随着GPU并行技术的发展及数据量的成倍增长,深度学习得到了极好的发展。在深度学习中CNN的表现较好,CNN在车辆分类和检测中具有许多的应用,在车辆分类中有LeNet、AlexNet、VGG、GoogLeNet等模型的应用,在检测中有R-CNN、YOLO等模型的应用。这些模型的效果常在

KITTI、BIT-Vehicle、Stanford Cars、UA-DETRAC等数据集上进行测试。如今, 车辆分类和检测模型在这些数据集上的测试效果有了极大的提高, 但对于特殊天气、特殊场景的车辆分类和检测仍有待深入研究。

参考文献(References)

- [1] CUN Y L, BOSER B, DENKER J S, et al. Handwritten digit recognition with a back-propagation network[J]. Advances in neural information processing systems, 1990, 2(02):396-404.
- [2] KRIZHEVSKY A, SUTSKYER I, HINTON G. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25(2):1097-1105.
- [3] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]// SINHA K, MAHARAJ T. Proceedings of the International Conference on Learning Representations. San Diego, USA: Computational and Biological Learning Society, 2015:1-14.
- [4] SZEGEDY C, LIU W, JIA Y. Going deeper with convolutions[C]// MORTENSEN E, MASSON M, DERPANIS K, et al. Proceedings of the Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015:1-9.
- [5] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// MORTENSEN E, MASSON M, DERPANIS K, et al. Proceedings of the Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016:770-778.
- [6] HUANG G, LIU Z, LAUERNS V, et al. Densely connected convolutional networks[C]// MORTENSEN E, MASSON M, DERPANIS K, et al. Proceedings of the Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017:4700-4708.
- [7] 曹林, 李佳, 张鑫怡, 等. 一种基于微波雷达回波信号的车型分类方法[J]. 电讯技术, 2020, 60(05):542-548.
- [8] 马传香, 汪焰杰, 王旭. 基于卷积神经网络的煤炭运载车辆识别[J]. 计算机科学, 2020, 47(S2):219-223.
- [9] 张超. 基于深度学习的铁路货车类型识别研究[D]. 北京: 北京交通大学, 2019.
- [10] 刘泽康, 孙华志, 马春梅, 等. 基于卷积神经网络多特征联合的车辆识别模型[J]. 计算机科学, 2019, 46(S1):254-258.
- [11] 余烨, 傅云翔, 杨昌东, 等. 基于FR-ResNet的车辆型号精细识别研究[J]. 自动化学报, 2021, 47(05):1125-1136.
- [12] 陈立潮, 朝昕, 潘理虎, 等. 基于部件关注DenseNet的细粒度车型识别[J]. 智能系统学报, 2022, 17(02):402-410.
- [13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// FREIFELD O, HAUBERG S, BLACK M J. IEEE conference on computer vision and pattern recognition. Columbus, OH, USA: IEEE Computer Society, 2014:580-587.
- [14] GIRSHICK R. Fast R-CNN[C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2015:1440-1448.
- [15] REN S, HE K, GIRSHICK R. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// KOH Y J, JANG W D, KIM C S. IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE Computer Society, 2016:779-788.
- [17] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]// MORTENSEN E, MASSON M, DERPANIS K, et al. IEEE Conference on Computer Vision & Pattern Recognition. Honolulu: IEEE, 2017:6517-6525.
- [18] LI Y D, HAN Z Q, XU H Y, et al. YOLOv3-Lite: A lightweight crack detection network for aircraft structure based on depthwise separable convolutions[J]. Applied Sciences, 2019, 9(18):3781.
- [19] BOCHKOVSKIY A, WANG C Y, LIAO. YOLOv4: optimal speed and accuracy of object detection[R]. 2020. <https://arxiv.org/pdf/2004.10934v1>.
- [20] 曾凯, 李响, 陈宏君, 等. 引入注意力机制的改进型YOLOv5网络研究[J]. 软件工程, 2023, 26(01):55-58, 54.
- [21] 刘杰. 一种基于RCNN的翻车事故检测算法[J]. 计算机时代, 2019(01):55-58.
- [22] 张琦, 胡广地, 李雨生, 等. 改进Fast-RCNN的双目视觉车辆检测方法[J]. 应用光学, 2018, 39(06):832-838.
- [23] 王林, 张鹤鹤. Faster R-CNN模型在车辆检测中的应用[J]. 计算机应用, 2018, 38(03):666-670.
- [24] LU J, MA C, LI L, et al. A vehicle detection method for aerial image based on YOLO[J]. Journal of Computer and Communications, 2018, 6(11):98-107.
- [25] 张成标, 童宝宏, 程进, 等. 改进的Yolo_v2违章车辆检测方法[J]. 计算机工程与应用, 2020, 56(20):104-110.
- [26] 张富凯, 杨峰, 李策. 基于改进YOLOv3的快速车辆检测方法[J]. 计算机工程与应用, 2019, 55(02):12-20.
- [27] 於积荣, 黄德启, 曾蓉, 等. 基于改进YOLOv4的车型检测算法研究[J]. 激光杂志, 2022, 43(04):52-59.
- [28] 张漪, 张美月. 基于改进YOLOv5的交通监控视频车辆检测方法研究[J]. 内蒙古公路与运输, 2022(02):50-55.

作者简介:

解 丹(1994-), 女, 硕士, 助教. 研究领域: 软件开发, 图像处理与模式识别.

陈立潮(1961-), 男, 博士, 教授. 研究领域: 图像处理与模式识别.

曹玲玲(1983-), 女, 硕士, 讲师. 研究领域: 大数据与人工智能.

张艳丽(1983-), 女, 硕士, 讲师. 研究领域: 人工智能与图像处理.