

计算机视觉和深度学习的自动驾驶方法分析

马石梓

比亚迪精密制造有限公司 广东深圳 518049

摘要: 随着深度学习算法的进一步提升,它在机器视觉和地图导航等应用领域中都获得了更加巨大的成功。针对传统方式无法解决问题或有效性较差的问题,深度学习能够利用其强大的特征学习能力、了解问题的规律,获得了较为理想的有效性。并以此为基准,研究本文了面向计算机视觉的自动驾驶中的环境目标和运动轨迹辨识问题,研究了深度学习技术在自动驾驶应用领域中的发展现状,并建立了SceneNet等深度网络模式进行场景中的环境目标辨识,以及利用基于视觉特征的场景语义分割构建了自动驾驶的环境特征进行目标的运动轨迹辨识问题,并对环境目标辨识与运动轨迹预测问题开展了试验,从而证明了所提供的方式的有效性。

关键词: 深度学习; 计算机视觉; 自动驾驶; 卷积式神经网络

引言

近年来,由于人工智能技术的迅速发展,中国传统车辆制造业通过与现代计算机技术的结合在研发自动驾驶技术方面也获得了很大发展。该行业的多家大企业在该领域的研究过程中都投入了巨大的资金投入。谷歌、丰田、百度和BYD等国内外公司都发布了令人满意的自动驾驶实验产品。随着自动驾驶科技的进展与完善,预计其将走向实际应用阶段,并将在全世界范围内广泛应用。借助自动驾驶,人类就可以做到自主出行,而不必担忧酒驾、超速、无证行驶和疲倦行驶等道路犯罪。有鉴于此,依靠机械视觉和深度学习的自动驾驶技术将有着更广阔的科学和商务应用前景。由于自动驾驶技术是未来车辆驾驶的必然发展方向,汽车公司和网络企业都很希望在这项崭新的技术应用领域中占得先机。汽车行业是一项特别的技术领域,由于对乘客的安全保护最关键,不允许出现意外,所以对车辆安全性的要求也非常严格。所以,企业在研发自动驾驶技术时,往往对传感器和方法的精确度和安全性都要求相当高。但是,由于自动驾驶车辆的主要消费者仍然是普通用户,所以必须减少成本。因为传感器、计算设备的精确度和传感器的生产成本都是正相关的,这将使得传感器的售价居高不下,并无法研制出大众所可以承受的高产品价格。这个问题在过去很难克服。但是现代计算机视觉下学习技术的高度精准使得在很多重要应用领域中,如任务、决策以及传感器的广泛应用使得研究自动驾驶系统变成了可能。

一、自动驾驶关键技术

作为我国自动驾驶技术的重点业务和核心技术之一,目标识别主要涉及对道路与规避行人的目标辨识。目标

识别系统通过电脑视觉监控交通状况,并通过实时视频自动识别目标,提供实时自主行驶的基础,如起步、停止、转弯、加速与减速。因为交通条件异常复杂,依靠传统目标检测的自主行驶技术的有效性无法显著增强。目前的自主行驶技术通常依赖于经过改进的雷达控制系统来补偿,这样就大大增加了系统的运营成本。但是由于现代科技水平的日益发达,卷积式神经网络已经能够直接认识并解析现实路面上的车辆。而通过一段时间的正确行驶,能够在现实路面条件下掌握一定的行驶技术,而不需要再顾及具体的路面要求以及不同的目标,这样就能够大大提高自主行驶算法的性能。

1. 广义的自动驾驶技术

目前,大部分的自主驾车技术都是通过视频进行分析的。由于在汽车移动场景中所拍摄的视频序列涵盖了不同的兴趣目标,如步行、车辆、道路、障碍物、背景等中的各类目标,因此必须确定在测试图像中的兴趣类目,并提供了车辆控制理论作为决策的基础。标志的确定与表示是一个很关键的过程,涉及对目标图形的编码方法。理想的识别方法必须可以满足以上所有干扰因素,如大小、外观、闭塞、复杂背景等。传统的驾驶方式对路线和车道辨识、机动车监测,以及对行人检查系统识别效果一般。

2. 基于深度学习的自动驾驶技术

基于视频分析的检测和目标辨识技术开始由传统功能过渡为深度学习。这种更加传统和直观的功能在这个阶段得到了非常好的结果。近年来,深度学习技术已经极大地提高了从多个状态中检测和辨识目标的效能,甚至超越了人们在这方面的能力。利用扫描窗口监视和发现目标极大地提高了确定多个目标的有效性。但是,采用扫描

窗口技术的另一项主要缺陷是扫描窗口的尺寸和位置的组合非常繁琐,有时候甚至无法计算。

二、SceneNet深度网络模型

卷积神经网络(CNN)属于人工神经网络,因为其网路模式通常由几个层次组成,也被称为深度卷积神经网络(DCNN)。CNN可被认为是对多层感知网络技术的进一步发展,它利用空间关系减小了待调整参数的数量,从而大大提高了反向传播算法的学习效果。随着输入信息不断地被传递至不同的感觉层次,每一层的数字滤波器都被用于获得最关键的信息特征。而基于局部感觉神经元对图象的基本元素,例如在图像中的边界和角度,通过CNN的信息传输方式能够获得缩放、偏移和翻转等信息特性。

1. 网络结构

卷积式神经网络的典型模型,是由卷积层、池化层和全连接层构成,如图1所示。

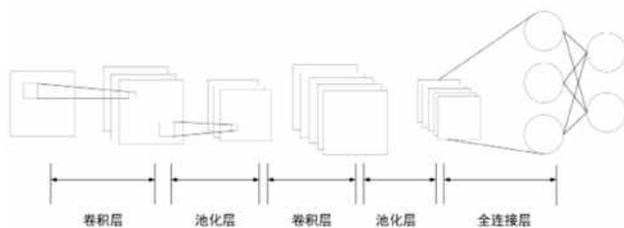


图1 典型卷积神经网络模型

2. SceneNet 场景识别网络

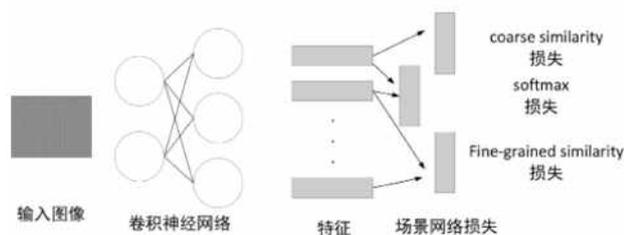


图2 SceneNet 模型

和所有其他类别任务一样,场景识别也是独一无二的,它并不仅仅是类别间的巨大区别,而是有场景类别中的巨大视觉区别。这是有很多错误的分类,同时也是识别场景的最困难所在。关于CNN分析能力的提高,一直有比较多的研究。有二个方式可以提高网络的深度和宽度。例如,VGG-NET网络的架构根据 AlexNet 网络的架构,增加了网络层次结构。而另一个方式则是开发一种全新的网络结构,包括了谷歌的原始架构,以及 ResNet 的所有残差模块化架构。另外,梯度因为网络深度过大而消失,从而导致了学习复杂度的提高。使用随机梯度算法(CNN)基于学习水平(微包子集)的下降,训练效果并没有充分反映网络的深度。所以,研究高效且可识别的损失函数才是提

高CNN有效性的关键问题。并以此为出发点,在本章中介绍了新的损失函数概念,并构建了基于自动感知场景相似率的网络模型,如图二所示。

三、基于深度学习的自动驾驶技术

基于视频分析的目标测量和辨识技术经过了由传统特征,如:HOG、SIFT、Bagofvisualwords和Fisher核矢量,到深度学习的转变过程。HOG所获得的定义保证了图象的几何与光学转换不变性。如Fisher核矢量,可整合各种特性间的维度、缩小时精度损失极小等,正是这些传统直观的特性,在目前阶段实现了良好的应用效果。但是因为目标的类型很多,颜色变化也很大,还有视角的改变等,导致了传统基于特征的目标检测技术出现了很难突破的瓶颈。近年来,由于深度学习的出现,导致在大量多类多状态下目标检测和辨识的特性都得以提高到拟人水准,甚至可能在很多方面都超过人。学习特性是在大量训练数据中自动学习到目标的特性,比传统特性而言,更能刻画目标的本性。学习中有许多经典模型架构,如自动编码器、稀疏编码技术、控制波耳兹曼机、深信度网路、卷积神经网络系统等。当中,采用卷积神经网络系统(Convolutional Neural Network, CNN)的深入学习模式是最经典的模式和热点之中。20世纪60年间,Hubel和Wiesel正在深入研究猫脑皮质中负责局部敏感性和方位选择性的神经元,后认为其独特的网络组织可以更有效地减少传递神经网络的复杂度,进而提出了CNN。于1980年发明的新识别机,是CNN的首次实施网络。在2014年,GiRshick等人发明了R-CNN24J,其思路是把每张图分成大约二千个区块并输入CNN训练,在预选框中使用CNN抽取出固定长度的特征,最后再使用特殊类型的支持向量机器(SVM)来分析。SPPnet也修改了以往采用剪裁的一个图像以其宽度为满 Alexnet 的输入条件,而改为可以采用任何规格大小尺寸图像用作入口。Fast-RCNN在SPPnet的基础上,通过显著性测量方式从原始图形上抽取出预选区,并把每一区坐标系数数据反映在特定图形上,当完成目标测量时,利用ROI池化层选择映射的坐标系数,将部分卷积图象送入分类器中,无需再对每一预先选择区域实施卷积计算,从而极大地提高了测量效率。2015年,Ren等人发明Faster-RCNN,在以前的基础上采用了一种RPN网络技术,可以通过卷积算法一次性获取卷积的特征图形,而Faster-RCNN则是对Fast-RCNN的更进一步加速。在2015年12月的ICCV国际大会上,邹文斌教授在R-CNN iii o的基石上,提交了基于RCNN的多层次结构与显著性目标测量方法旧八I,在MSRA-B,PASCAL-

1500E301,和SOD三个数集上的试验表明,其测量率超过了当时业界最高水准。在该研讨会上,Kontschieder旧引等人提交了在CNN各层输出的特性研究基本上,通过随机森林,在开放的数据集MNIST和Imagenet上,得出了较高的检测效率。

CNN与多通道处理技术结合的方式,在图像识别上也有良好的作用。2011年,PierreSermanet等人采用了多尺度CNN算法,将原始图片和其子取样的卷积成果经过了线性分类器分类后,其在GTSRB数据集上准确率超过了98.97%。在2012年DanCi.Resan等人则主张,利用多通道深度神经网络的方式辨识交通信号。该方案通过将训练图像同步进入的N个深度神经网络同步练习,在得到预测结果时,通过统计同步进入图片的N个深度神经网络预测结果的平均数为最后结论,其预计结论精确度将达到99.46%,超越了传统人工辨识精确度。在2014年,KarenSimonyan等人将连续录像分成了空气流和时间流,通过不同的CNN解决了同一个录像的物质特性和动作特点,并将二者融合完成了行为判断,也就大大提高了辨识的准确性。

在辅助驾驶与自动驾驶领域中,由于需要辨识与预测通过电脑视觉的自动驾驶算法研究与综述的主要目标数量众多,涉及前方车辆、非车辆、行者、路面标志、机动车道本身、机动车道线等,使得基于目标检测和辨识的机器学习计算越来越十分复杂。在自动驾驶在机器人导航中,另一个方式直接透过在视频图像中了解前进路线的角度来选择路线和绕过障碍物,以及YannLecun的工作方式,即进行端到端教学,以进行非车道上的障碍物避让,再透过六阶的CNN学会人的驾驶动作,就能够在跨越视线内范围的时候掌握底层和高级特性,从而免去了人工的校准、矫正、参数调节等,该系统最主要的好处就是对各种不同环境下的非车道环境都具有很好的鲁棒性。上述工作,均是利用深度学习直接把图像反馈到汽车行驶的视角下完成的。在这一思路的直接影响下,在ICCV2015上,由普林斯顿大学学生发明了深入驾车计算I,其计算结构如图一所显示,通过利用深入神经网络直接感知的驾车操控能力(drivingaffordance),不但大大简化了计算的复杂性,同

时也极大地提高了自主驾车的鲁棒性和自动化水平,是在自动驾驶关键技术上的又一项突破。而深入驾车的关键技术,通过使用CNN来进行学习和感知在一段时间的正常行驶过程后,就可以掌握并了解在实际路面状况下的相关行车智能,而无需再通过了解具体的道路状况和各种目标,从而大幅增强了辅助行车算法的性能。

四、总结与展望

自动驾驶技术也是未来车辆智能的研发热点之一。从综述的论文中我们能够看出,采用了中国传统目标测试最有效的HOG特性、SIFT特性、CSS等特性的自动驾驶科技,已经获得了良好成果。但因为现实道路环境非常复杂多变,采用了中国传统目标测试的辅助自动驾驶科技性能上难以获得大幅改善,而目前的自动驾驶技术,由于普遍依赖于先进的雷达控制系统,显著提高了系统实施的成本。在具体的辅助行驶算法中,若对道路特征和目标信息没有整体了解,将很难获得实用化和商用化水平。而借鉴传统自动驾驶技术中的精髓,并借助深度学习研究中的新研究成果,融合传统特征与深度学习特性,并提取更多信息,不失为一种较好的处理方式。设计自动驾驶技术的新算法,进而提高深度驾驶拟人性与实用性水准,是一个有待去进一步探讨的新道路。

参考文献:

- [1] 胡艳,曹钻.基于机器视觉和深度学习的自动驾驶方法研究[J].环境技术,2021,39(3):100-105,110.
- [2] 刘崇宇,陈晓雪,罗灿杰,等.自然场景文本检测与识别的深度学习[J].中国图象图形学报,2021,26(6):1330-1367.
- [3] 秦晓晖.基于计算机视觉的辅助自动驾驶系统的设计与实现[J].办公自动化,2021,26(11):62-64.
- [4] 张新钰,高洪波,赵建辉,等.基于深度学习的自动驾驶技术综述[J].清华大学学报(自然科学版),2018,58(4):438-444.
- [5] 孙学强,谷俊荣,孙泽辉,等.基于计算机视觉的辅助自动驾驶系统的设计与实现[J].电脑爱好者(普及版)(电子刊),2021(3):910-911.