

自动驾驶场景中增强深度学习的时空特征提取方法

敬 颀 陈 潭 杜文丽 刘志康 尹 皓

(四川大学软件学院 成都 610065)

摘 要 自动驾驶是当下的热点研究方向,同时交通拥堵也是国内常年存在的社会问题。在未来,交通拥堵很大概率会出现在自动驾驶车辆和人为驾驶车辆共存的道路上。考虑到多种可能会影响自动驾驶的因素,在已有学说的基础上进行实验。为了提升整体交通的运行效率,在保证安全的情况下,所有自动驾驶车辆应当尽可能进行高速行驶,以提升道路效率,从而解决交通拥堵的问题。通过使用二维平面表示道路,将二维信息堆叠形成三维数据以及混合神经网络结构的不同方法来解决这一问题,并利用深度神经网络从中提取出所需的时空特征来进行车辆控制,从而使车辆做出较优的响应。最后,我们利用增强学习的方法来搭建并训练该系统,完成神经网络结构效果的测试。

关键词 深度学习,增强学习,自动驾驶,卷积神经网络,循环神经网络

中图分类号 TP18 文献标识码 A

Spatio-temporal Features Extraction of Traffic Based on Deep Neural Network

JING Jie CHEN Tan DU Wen-li LIU Zhi-kang YIN Hao

(College of Software Engineering, Sichuan University, Chengdu 610065, China)

Abstract Autopilot is a hot research direction, and traffic congestion is a perennial social problem in China. In the future, traffic congestion is likely to occur on the road where self-driving vehicle and artificial driving vehicle coexist. On the basis of existing theories, this paper considered a variety of factors that may affect autopilot, including different speeds and neural networks. In order to improve the overall traffic efficiency, on the premise of maintaining safety, all self-driving vehicles should be as fast as possible to improve road efficiency and fundamentally solve traffic congestion. The feature extraction problem of this special case is different from the feature extraction of the image data, so the method of representing the road in the two-dimensional plane is used to process the three-dimensional data formed by the stacking of two-dimensional information and hybrid neural network. The spatial and temporal features are extracted by using the depth neural network, so that the vehicle can make a better response. Finally, the system design is combining with reinforcement learning so that it can be trained, and thus the effect of neural networks can be tested.

Keywords Deep learning, Reinforcement learning, Autonomous vehicle, Convolutional neural networks, Recurrent neural networks

1 引言

近年来,随着经济的发展,全球交通里程数、汽车拥有量都在持续增加。在中国,交通运输行业的发展也十分的迅速,根据统计,2017年,全国汽车公路总里程达到了477.35万km,而全国拥有公路营运权利的汽车数量共计1450.22万辆^[1]。随之而来的交通拥堵及交通事故也与日俱增。2017年,中国因交通事故死亡的人数约6.3万。据世界卫生组织统计,全球每年有130多万人丧生于道路交通事故,而其中约90%的交通事故是由人为原因引起的。交通的效率与安全问题,日益成为公众关注的焦点。

自动驾驶汽车定位安全、智能且高效,它利用多种车载传感器来感知车辆周围环境,并利用算法分析环境,从而控制车辆的方向盘角度和速度变化,好的自动驾驶算法倾向于提高交通的安全与效率^[2]。

神经网络,是一种利用多层非线性变化处理的网络结构,

而且也是表征学习的通用框架,从深蓝到Alpha Go,各种神经网络结构都在不断的完善和发展。深度学习算法在近些年不断地取得突破。

安全、智能、高效是自动驾驶技术发展的重要指标,深度学习的特性有助于实现自动驾驶软件部分的路况检测及路径决策系统。

从社会层面来看,自动驾驶可以大幅减小驾驶压力,一定程度地提高安全性,同时避免拥堵并降低污染,还有利于提高社会公众的出行体验;从产业发展来看,自动驾驶将成为深度学习、物联网、云计算和大数据技术融合发展的宠儿,它会改善传统的工作环境和工作模式,推动产业转型,提高生产效率。据美国电气和电子工程师协会(IEEE)预测,截至2040年,全球75%的汽车都将配备自动驾驶系统。

基于自动驾驶的发展潜力,我们大胆预测,将来的道路可能会存在自动驾驶车辆与人为驾驶车辆混合的场景^[3],因此,我们希望通过建立一个虚拟系统来模拟此种特殊场景,并利

用中央决策系统控制道路中所有自动驾驶的车辆,以追求交通效率的最大化。

2 相关工作

2.1 自动驾驶现状调研

自动驾驶系统需要硬件与软件配合,它的典型工作流程如下:首先,通过相机、激光、雷达和声纳对车辆周围的静态和动态环境进行360°的精确鲁棒感知。现阶段,在障碍物的检测识别中,解决方案大多是使用深度网络来提高传感器融合数据的精准程度,从而为后续的感知、定位、规划和决策提供依据。

硬件部分主要为各种传感器原件,包括:全球定位系统和惯性测量单元、感知传感器等。以Google的Waymo(如图1所示)为例,传感器包括LiDAR-激光雷达系统、视觉系统、雷达系统、补充传感器等。

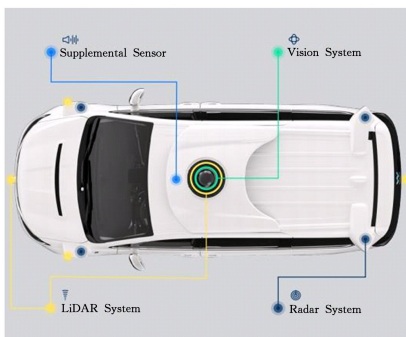


图1 Waymo传感器示意图^[4]

软件系统主要包括:感知、行为预测和规划系统。软件技术主要包括:多传感器信息融合,决策方案及规划技术,车辆控制决策算法等。以决策及规划技术为例,不同的自动驾驶系统会选择不同的方案,如Comma.ai采用End-to-End方案^[5],以摄像头的图像作为输入,直接输出车辆的速度值和方向变化,中间用数学模型来拟合逼近最优驾驶策略。此类模型中,深度学习模型是最常用的。

Waymo的规划器通过感知和行为预测,为汽车规划路径,植入防御性驾驶行为以保证安全。同时,Waymo借助经验数据优化驾驶模式。

2.2 论文实现方法

在针对现有文献进行研究之后^[6-8],我们发现大量的研究未对整体交通进行优化,本文针对这个问题提出解决方案。在一般的自动驾驶场景中,深度学习将贯穿整个交通的调度,对道路中所有自动驾驶的车辆进行控制,最终提升整体交通的效率。

基于上述场景,我们需要以下3个模块:道路模块,汽车模拟模块,中央调度系统模块。汽车模拟模块中,部分车为自动驾驶车辆,受中央调度,其余为人为驾驶车辆,不受中央系统控制,但人为驾驶车辆的移动也可以被检测到。道路信息模拟系统负责搜集车辆信息,然后传输至中央调度系统。中央调度系统中包含所有算法,主要有增强学习^[9-11]和深度学习^[12-13]。我们首先需要设计的是汽车模拟模块以及道路模块,同时提供可视化界面^[14],最后再来完善和实现机器学习相关算法。

2.2.1 数学建模分析

车,作为最小实体数据单元,需要提供的是其速度信息。

道路模拟系统负责搜集其位置信息。而我们关注的是最终提交给中央系统的数据。在需要做全局决策的情况下,中央系统的数据需要包含当前时刻该路段所有车的位置以及速度信息。

能包含整个交通数据的数据结构中,首先想到是使用俯瞰图的二维数组,但很显然这种数组不能满足我们的需求,因为得到的图像只能看出当前时刻的车辆位置信息,并且不包括车的速度信息。

为了简化信息,我们将整个道路划成若干小块,其中小块的长为道路一车道的长度,宽约为1m,每个车都会占据大约4个小块。对于车的位置信息,准备将道路上有车的地方以1表示,无车以0表示。如此一来,交由中央系统的数据便包含了简化的当前路段所有车辆位置信息。

接下来,我们开始引入对交通至关重要的信息——汽车的速度。通过量化速度的方法,以速度的量化值来代替1表示有车,无车仍然以0代表。这样,我们的数据中便包含了所有需要的信息。

2.2.2 中央调度系统的设计

中央调度系统通过接收从道路系统传来的当前时刻的道路信息,反馈给车辆该执行的操作。在常规自动驾驶中,反馈的信息包含方向盘角度和加减速^[15]。此处我们还需要做适当简化,分为加速、减速、左转、右转、无操作5个操作。

我们可以将中央系统的工作理解为对传入数据做特征提取,通过提取到的特征做相应判别,再将结果返回,控制车辆运行。此系统的决策准确决定了整个交通运转速度的快慢。而决策准确依赖于特征提取的优劣。

特征提取的优劣受两个因素的影响,一是训练数据,二是网络结构本身。接下来我们将就这两个方面进行深入研究。

3 详细方法和技术描述

3.1 训练方法

训练过程应遵循以下两个原则:1)训练结果应当能提升整体交通速度;2)不需要手动打标签来做算法设计。因此,对于这种需要不断提升运行结果的算法,我们联想到了增强学习。在不需要标签时,则使用无监督训练方式。

3.1.1 增强学习

根据前文所述,我们需要依据当前时刻的状态来对最优决策做预测,需要使用的增强学习方法的公式为:

$$Q'(S,A) = (1-\alpha) \cdot Q(S,A) + \alpha \times [R + \gamma \cdot \max_A(Q',A)] \quad (1)$$

其中, Q 为效益函数, S 为当前状态, A 为汽车采取的动作, α 为学习速率, R 为环境回馈值。但是按照目前的设计,还不能直接使用该公式。首先,我们需要确定回馈,即 R ,用于衡量系统决策的优劣。根据2.4节中的讨论,最终目的是使交通系统达到更高的运行效率,显然这个回馈量应该由速度决定。而增强学习记忆对决策未来的影响则需要进一步处理。至此我们可以确定:

- 1)神经网络的输出采取5种动作的倾向,用百分数表示;
- 2)增强学习则倾向于采用交通效率效益影响最优的操作。

因此,我们可以以神经网络输出最优操作的对应值来代替增强学习中的记忆收益。但为了确保公式中的当前收益和预估收益保持在一个数量级,我们需要对速度值进行量化处理。由于预期的收益对应为小数(0到1之间),为了避免收

益累加不断增大,我们给予的反馈值应包含负数。我们假设道路限速为 80km/h,其中车速为 20km/h 时表示比较不满意的情况。因此,量化方法为:

$$R=(v-20)\div 60 \quad (2)$$

同时,可得:

$$Q'(S,A)=(1-\infty)\cdot Q(S,A)+\infty\times[(v-20)\div 60+\gamma\cdot\max_A(f(S),A)] \quad (3)$$

使用 q-learning 后我们很自然地发现,其完美地解决了回馈量一些时刻的滞后效应。如在该场景中,自动驾驶车辆向右转后,虽然避开了前方车辆,但是当前时刻的速度并没有得到很好的提升,在使用增强学习后,几个时刻的速度都得到了较大的提升,可以自然地后面的效果提升归功于前面操作的回馈部分。

现在有一处很明显的问题,那就是我们的输入,很显然,判断当前道路信息绝对不能用单个瞬时时刻的信息,当前瞬时状态必须要和前几个时刻一起结合才能判断出应该做的操作。针对这个问题,我们有两种解决办法:1)在每次传入的时候,通过与之前数据的拼接,将几个时刻的信息重叠在一起,进行传入处理;2)利用神经网络结构,传入数据仍然只有当前瞬时的数据,但是网络运算受之前过程的影响,已经存在了满足这种需求的网络模型——循环神经网络,但单纯的循环神经网络结构并不能满足我们的需求。针对这种问题,我们将在下一节深入讨论,并且通过实验来验证两种方法的正确性。

至此,我们已经成功引入增强学习来实现交通速率更高的需求。此时,传入中央系统的信息增加了当前时刻系统所控制自动驾驶车辆的速度,用以做增强学习的回馈。

3.1.2 无监督训练

在上一节根据增强学习设计的算法中,我们确定了神经网络的输出是回归类型。我们用采取值最大的动作的对应值与当前所得反馈值得到新的反馈值,新反馈值和旧反馈值之间的差就是神经网络训练时使用的 BP 算法入口。

$$\Delta y=Q'(S,A)-Q(S,A) \quad (4)$$

由(3)式可得:

$$\Delta y=\infty\times[(v-20)\div 60+\gamma\cdot\max_{value}(f(S))]-\infty\cdot Q(S,A) \quad (5)$$

得到需要修正的差值,我们依据 BP 算法便可训练神经网络,完成无监督训练。

3.2 网络结构

传入的数据是基于对俯瞰图的转化,很显然,这种数据单个时刻的内容在空间上具有很大的关联性。卷积神经网络在对空间上有关联性的信息做特征提取时有很大的优势。但如 3.1 节中所述,必须解决提取时间轴上信息提取的问题。

3.2.1 CNN 时间轴堆叠模型

CNN 常被运用于二维或者三维图像的信息提取,而在当前情况可以认为数据是三维的,因为我们引入了时间轴上的信息。首先需要思考,怎样构建传入的数据结构。在中央系统中存储一段时间的单个时刻信息,算法将多层信息重叠起来再传递给神经网络。

由于模拟系统的数据已经做了简化,每个车大概占据 3~4 个小格,相对来说不需要太过复杂的模型。输入的信息

大小为 5(车道) * 50(前后小格数) * 4(时刻数)。我们设计了隐层为 4 层,其中前 3 层为卷积神经网络,每层后面跟着 relu 激活层以及池化层,最后一层为全连接神经网络,输出层为全连接回归层。效果将在实验部分给出。

CNN 时间轴堆叠模型的简单结构示意图如图 2 所示。

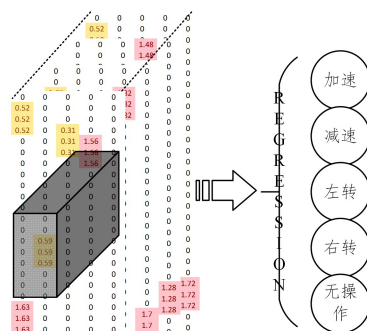


图 2 CNN 时间轴堆叠模型

3.2.2 RNN, CNN 混合神经网络模型

循环神经网络是一种用于处理序列数据的神经网络^[16]。而时间序列数据就是一种典型的序列数据,一般是指在连续的时间点上收集到的一串数据,这类数据反映了某一事物、现象等随时间变化的状态或程度。自然,循环神经网络满足我们的应用需求。

相对于上一小节对多个时刻的信息重叠进行处理,目前单个时刻需要处理的数据量缩小到了 1/5。出于对运算量的考虑,将卷积层减少为一层,隐层最后一层仍然使用全连接。混合神经网络模型的结构图如图 3 所示。

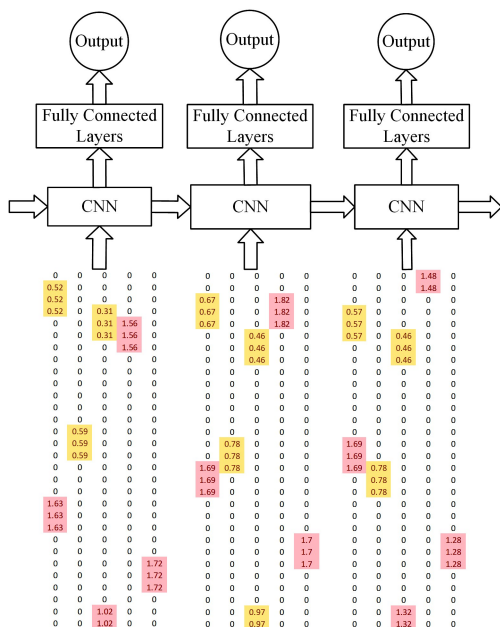


图 3 CNN&RNN 混合神经网络模型

4 实验验证及分析

4.1 实验目的

首先,需要验证系统的正确性。在增强学习的使用中,并未按照一般增强学习的应用使用专门的数据结构去记录历史

回馈值,而是使用了网络的输出作为历史回馈。这种操作有一定的合理性,但是并未验证,如果系统正确地运行起来,车辆速度将有很大的提升,从而也就能证明增强学习的作用及正确性了。

其次,需要验证两种神经网络结构的效果。根据3.1节的论述,网络的特征提取能力最终会反馈到整体交通的速度上,通过多次测试求平均值来比较网络结构的优劣。

最后,我们也关注不同结构的运行效率。因为在模拟场景方面我们做了很多数据上的简化,但实际应用中车的数量以及道路环境的复杂性会明显加大运算量,也就需要考虑不同网络结构的具体运行效率。

4.2 实验设计

基于构想的场景,假设人为驾驶的 vehicle 处于一种在不影响安全的情况下尽可能快速前进并且可能变道的一种状况。道路中的每一个自动驾驶车辆都受中央系统的统一调配。

为了对比两种网络结构的效果,我们设计了对比实验来比较不同神经网络结构在搭建的系统下的运行效果,同时也引入普通全连接神经网络以及随机模型做对比。最后,统计随机模型、全连接神经网络、CNN&数据堆叠以及 CNN 与 RNN 混合神经网络结构的训练收敛速度、运算资源消耗以及交通平均速度。

4.3 实验结果

实验结果如表1所列,对应柱状图(10次实验的平均值)如图4所示。

表1 实验结果

	随机模型	全连接网络	CNN & 数据堆叠	CNN 与 RNN 混合神经网络
收敛速度/min	—	30.2	56.7	71.5
平均速度效果/(km/h)	23.2	68	75.3	76.8
平均计算资源占用(GPU)/%	6.00	32.75	87.50	100

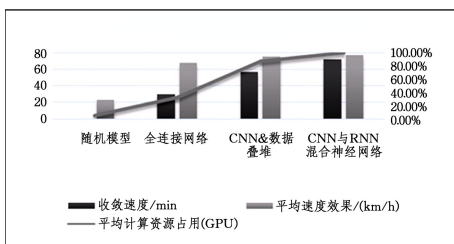


图4 实验结果

4.4 结果分析

通过实验结果,我们可以得到以下几个结论。

(1)相对随机模型,引入增强学习的模型的平均速度均处于高水平值,由此可证得系统设计的正确性。

(2)混合神经网络的方式在多次实验测试中取得最快平均速度,由前文可知,其特征提取能力较 CNN&数据堆叠的方式有一定程度的提升。

(3)计算资源方面,采用 1080Ti,混合神经网络模型几乎稳定地完全占用了所有 GPU 资源,而 CNN 与数据堆叠混合的方式也占用了较高资源,但相对全连接神经网络,平均速度均有很大提升。

实验结果验证了系统设计的正确性,以及混合神经网络

结构在交通时空特征提取上的高效性。

结束语 本文基于自动驾驶与人工驾驶混合路段的情况,结合增强学习方法设计出了一套决策系统,并且提出将 RNN 与 CNN 混合神经网络结合的想法,最终通过实验验证其可行性。

RNN 与 CNN 混合的神经网络结构在数据既包含时域特征又包含空域特征的情况下有着强大的潜力,除自动驾驶外,还在动作识别、视频信号处理等领域有着极其重要的参考意义。

参考文献

- [1] 孟海波,江洪波,汤天波.全球自动驾驶发展现状与趋势(上)[J].华东科技,2014(9):68-70.
- [2] 翁岳暄,多尼米克·希伦布兰德.汽车智能化的道路:智能汽车、自动驾驶汽车安全监管研究[J].科技与法律,2014,6(4):632-655.
- [3] 夏伟,李慧云.基于深度强化学习的自动驾驶策略学习方法[J].集成技术,2017,6(3):29-40.
- [4] WAYMO L L C. Remote assistance for autonomous vehicles in predetermined situations;US,US 9,720,410 B2[P]. 2017.
- [5] BOJARSKI M,DEL TESTA D,DWORAKOWSKI D,et al. End to End Learning for Self-Driving Cars[J]. arXiv:1604.07316, 2016.
- [6] CUI Z,KE R,WANG Y. Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction[J]. arXiv:1801.02143,2018.
- [7] CHEN C,SEFF A,KORNHAUSER A,et al. DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving[C]// IEEE International Conference on Computer Vision. IEEE,2015:2722-2730.
- [8] MA X,DAI Z,HE Z,et al. Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction[J]. Sensors,2017,17(4).
- [9] MNH V,KAVUKCUOGLU K,SILVER D,et al. Playing Atari with Deep Reinforcement Learning[J]. arXiv:1312.5602,2013.
- [10] CHAE H,KANG C M,KIM B D,et al. Autonomous Braking System via Deep Reinforcement Learning [J]. arXiv: 1702.02302,2017.
- [11] LILLICRAP T P,HUNT J J,PRITZEL A,et al. Continuous Control with Deep Reinforcement Learning: US 20170024643 A1[P]. 2017.
- [12] FRIDMAN L,JENIK B,TERWILLIGER J. DeepTraffic: Driving Fast through Dense Traffic with Deep Reinforcement Learning[J]. arXiv:1801.02805,2018.
- [13] 郭丽丽,丁世飞.深度学习研究进展[J].计算机科学,2015,42(5):28-33.
- [14] MEEDS E,HENDRIKS R,FARABY S A,et al. MLitB: Machine Learning in the Browser[J]. arXiv:1412.2432,2015.
- [15] SHAI S,NIR B,AVIAD C,et al. Long-term Planning by Short-term Prediction [J]. arXiv:1602.01580,2016.
- [16] LIPTON Z C,BERKOWITZ J,ELKAN C. A Critical Review of Recurrent Neural Networks for Sequence Learning[J]. arXiv: 1506.00019,2015.