

文章编号: 2095-2163(2021)11-0084-05

中图分类号: TP391.4

文献标志码: A

基于 Lambda 网络和 LSTM 的车辆轨迹预测

谭星祥¹, 谭琛凯²

(1 江苏省丹阳中等专业学校 机电部, 江苏 丹阳 212300; 2 江苏大学 汽车工程研究院, 江苏 镇江 212000)

摘要: 驾驶员在驾驶时,会对周围车辆的轨迹进行预测,并且采取一个安全的行为模式,从而减少道路事故的发生。出于这个想法,研究人员基于深度学习的模型,对车辆的轨迹进行预测,从而提升自动驾驶车辆的安全性。通过不断分析获得的驾驶数据,深度学习可以得到其它汽车的行为模式并进行预测。相较于过去的方案,基于深度学习方案在轨迹预测上具有优势,本文基于 transformer 模型的 LSTM(长短期记忆网络)编码隐藏参数,反映周围车辆的驾驶意图,再通过 Lambda 层获得这些车辆对预测车辆的影响,最后通过 LSTM 解码获得预测的轨迹。相较于以往方案,应用 Lambda 层的方案计算速度更快,同时实验表明相较于以往方案轨迹预测精度有提升。

关键词: 轨迹预测; 自动驾驶汽车; 深度学习; 神经网络

Vehicles trajectory prediction based on LSTM and lambda network

TAN Xingxiang¹, TAN Chenkai²

(1 Department of Mechanical and Electrical Engineering, Jiangsu Danyang Secondary Professional School, Danyang Jiangsu, 212300, China; 2 Automotive Engineering Research Institute, Jiangsu University, Zhenjiang Jiangsu, 212000, China)

[Abstract] When driving, drivers can predict the trajectory of the surrounding vehicles and adopt a safe behavior pattern, thus reducing the occurrence of road accidents. With this in mind, the researchers used models based on deep learning to predict vehicle trajectory to improve the safety of autonomous vehicles. By constantly analyzing driving data, deep learning can predict the behavior patterns of other cars. Compared with previous schemes, deep learning schemes have advantages in trajectory prediction. In this paper, LSTM (Long and short-term memory network) codes hidden parameters based on Transformer model to reflect the driving intention of surrounding vehicles, and then the influence of these vehicles on predicted vehicles is obtained through Lambda layer. Finally, the predicted trajectory is obtained through LSTM decoding. Compared with the previous schemes, the scheme using Lambda layer has a faster calculation speed, and the experiment shows that the trajectory prediction accuracy is improved compared with the previous schemes.

[Key words] trajectory prediction; autonomous vehicles; deep learning; neural network

0 引言

自 2012 年 AlexNet 夺得 ImageNet 图像识别比赛冠军以来,深度学习迎来了爆发式发展。其中递归神经网络(RNN)^[1]在序列预测问题中取得了很大的进展,而车辆轨迹预测可视为一种时间序列回归或预测的问题,通过过去位置及车辆的相互影响来预测车辆的未来轨迹。车辆轨迹预测有助于车辆了解未来情况,从而降低可能发生事故的风险^[2]。

车辆轨迹预测中包含以下单位:目标车辆是方案在场景中感兴趣并且行为预测的车辆;邻居车辆是目标车辆一定范围附近的汽车,可能会对目标车辆的行为存在影响。不同的研究会将对车辆轨迹预测进行不同的建模假设。2013 年 Lefèvre 等人^[3]将车

辆轨迹预测分为:基于物理定律、基于多模态(车辆预期机动性)、基于交互感知这 3 种。近年来,深度学习技术在自然语言处理中取得成功,数据驱动的方式被广泛应用于轨迹预测中。本文主要采用数据驱动的方式,基于交互感知对车辆轨迹进行预测。

1 相关工作

各种基于深度学习的车辆轨迹预测的区别主要体现在学习车辆行为模式的方案不同。目前,比较常用的方案是:通过车辆未来的多重可能性预测结果(多模态)与车辆之间的相互影响预测结果(交互感知)。

多重可能性预测方案^[4-5]是输出多种预测结果,并与真实结果进行比较。通过误差反向传播,可

作者简介: 谭星祥(1969-),男,学士,中学高级教师,主要研究方向:汽车轨迹预测;谭琛凯(1994-),男,博士研究生,主要研究方向:汽车轨迹预测。

通讯作者: 谭星祥 Email:826494517@qq.com

收稿日期: 2021-08-31

以实现训练多模态结果的需求。多重可能性预测方案通常采取换道、加速、减速、红绿灯停止等动作,相较于一般的模型,多重可能性预测方案可以防止预测结果趋近于温和单一,且多模态的模型在负对数似然损失表现更好。

车辆之间的相互影响预测,可以通过神经网络来进行学习。目前主流的方法包括:卷积神经网络、图神经网络、attention 等方案。早先学者提出可以使用全联接层学习车辆之间的相互关系。而卷积神经网络^[5-6]相较于只使用全联接层,在全联接层之后加入了卷积层(学习局部有效信息)和池化层(压缩数据和参数的量,减小过拟合),可以更好地学习车辆之间的关系。但是,卷积神经网络仍存在一些不足之处。由于卷积神经网络是通过卷积核来扫描网格,因此网格中相距较远的车辆关系需要深层的卷积网络才能学习到关系。

图神经网络是一种新颖的解决方案,将行驶场景中的车辆及其相互作用可视为一个图,其中节点是车辆,边表示车辆之间的相互作用。利用这种表示,图形神经网络(GNNs)^[7-8],是解决序列问题的一种好方法,允许可变大小的输入,可以用来预测车辆的行为。有学者使用图注意网络(Graph Attention network)^[9]对行人数据进行了预测^[10-11]。除了使用 GAT 来捕捉相互作用(在每个时间步长中),还使用 LSTM 来记录时间相关性,在预测准确性和计算效率上有一定的改进。除了神经网络,注意力机制也是一种学习车辆之间相互关系的方案。使用 attention 有选择地提取对目标车辆轨迹影响最大的细节^[12-13],从而实现车辆在网格中的相互影响,但 attention 中的参数矩阵会占用较多内存与计算。在注意力机制的基础上,有学者提出线性注意力机制^[14-15],为注意力机制的高内存使用提供了一种可扩展的方法。

2 轨迹预测

2.1 数据载入

以 NGSIM 数据集为例,通过提取数据集中需要的几列数据,再将数据划分为神经网络需要的数据集、验证集、测试集。

根据当前时刻(*obs*),可以获得目标车辆的过去轨迹 $X_V = \{x_V^1, x_V^2, \dots, x_V^{obs}\}$,同时获得目标车辆未来的真实轨迹。其中预测长度为 *pred*, $Y_V = \{Y_V^{obs+1}, Y_V^{obs+2}, \dots, Y_V^{obs+pred}\}$ 。通过网格获得邻居车辆的过去轨迹 $X_i = \{x_i^1, x_i^2, \dots, x_i^{obs}\}, i \in [1, n]$ 。输入过去轨

迹为 X_V 和场景中所有 X_i ,并通过神经网络进行训练,可以得到输出 $Y_V = \{Y_V^{obs+1}, Y_V^{obs+2}, \dots, Y_V^{obs+pred}\}$,通过误差反向传播得到训练的神经网络。训练模型一般分为 3 步:编码,学习车辆行为模式,解码。训练模型如图 1 所示。

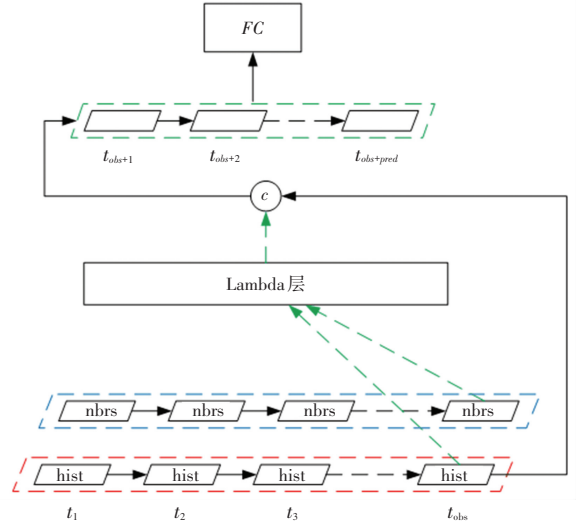


图 1 模型总体架构

Fig. 1 The architecture of the model

图中,目标车辆和邻居车辆经过相同的嵌入及 LSTM-Encoder; C 将网格内邻居车辆对目标车辆的交互影响与目标车辆的隐藏状态进行融合;通过 LSTM-Decoder 递推出未来时刻的隐藏状态;最后通过全连接层(FC)得到目标车辆预测的未来轨迹。

2.2 训练模型

2.2.1 编码

编码是实现车辆轨迹预测模型的重要一步,除了从数据集中提取的车速、车辆种类、速度、加速度等信息,车辆的行驶轨迹还受到道路结构、交通标识、交通规则等先验知识的约束^[2],而编码层能很好地提取这些数据的隐藏信息。首先,通过全连接层来形成嵌入向量,可通过 LSTM^[16]编码器得到该车的隐藏状态向量 h_V 与周围汽车的隐藏状态向量 h_i 。

$$h_V^t = \text{LSTM}(h_V^{t-1}; \psi(x_V^t; W_{emb}); W_{encoder}) \quad (1)$$

$$h_i^t = \text{LSTM}(h_i^{t-1}; \psi(x_i^t; W_{emb}); W_{encoder}) \quad (2)$$

式中, $\psi(\cdot)$ 代表全连接层; LSTM 代表长短期记忆编码器; W_{emb} 、 $W_{encoder}$ 分别是全连接和长短期记忆网络的嵌入参数。根据上一层隐藏状态 h_V^{t-1} 、 h_i^{t-1} 与 t 时刻的轨迹,嵌入 $\psi(x_V^t; W_{emb})$ 计算得到 t 时刻的隐藏状态。

2.2.2 学习车辆行为模式

本文采用基于 lambda 层^[17]的方法来学习车辆

之间的交互影响。 λ 是 Google 在 2021 年提出的一种新的 self-attention 框架,用于捕获输入和结构化上下文信息之间的长期交互。

在计算车辆交互作用之前,需要将周围车辆进行嵌入,与 Deo N^[5]等人的方法一样,将目标车辆周围 3X13 的所有邻居车辆嵌入网格,所有周围车辆隐藏状态 h_i 嵌入后,得到嵌入的隐藏状态 H 。

如图 2 所示,首先将最后一个时刻 (*obs*) 的隐藏状态 H 与 h_V^{obs} 输入 λ 模块。 λ 层通过一维卷积来计算 K 和 Q ,然后使用 *softmax* 函数将 K 进行归一化,计算得到目标函数 λ_c 。具体计算公式为:

$$\lambda_c = \text{softmax}[\text{Conv}(H, W_K)] \otimes \text{Conv}(h_i^{obs}, W_V) \quad (3)$$

式中, Conv 代表一维卷积, W_K, W_V 分别代表一维卷积的参数。通过公式(3)可得到 λ 层中周围车辆对目标车辆的交互影响系数 λ_c , 这是一个线性函数,可提高模型的计算效率。应用函数 λ_c 可以获得比 attention 更高的计算效率。计算出车辆与邻居车辆隐藏状态之间交互影响。

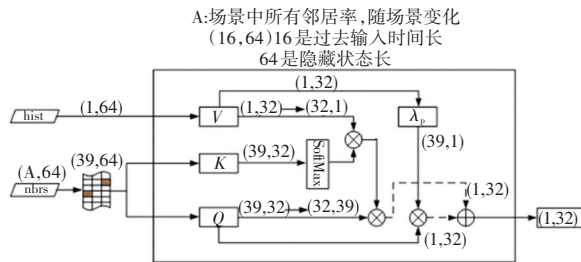


图2 λ 层学习车辆交互

Fig. 2 λ layer learns vehicles' interaction

将所有相邻车辆的交互影响,通过学习到的线性投影得到对目标车辆的影响。具体计算公式为:

$$J_c = \text{Conv}(H, W_Q) \otimes \lambda_c \quad (4)$$

式中, Conv 代表一维卷积; W_Q 分别代表一维卷积的参数;模型获得了邻居车辆对目标车辆的交互影响 J_c 。

λ_c 计算了车辆之间的交互影响对目标车辆的影响,复杂的模型效果在某些情况里表现较差。应用简单的函数映射可能是更有效的方案, λ_p 计算了目标函数映射,计算公式如下:

$$\lambda_p = W_{\lambda_p} \otimes \text{Conv}(h_i^{obs}, W_V) \quad (5)$$

式中, Conv 代表一维卷积, W_{λ_p} 分别参数矩阵,经过该公式得到了自车隐藏状态的简单函数映射,再经过公式(6)计算出简单映射与邻居车辆的交互。

$$J_p = \text{Conv}(H, W_Q) \otimes \lambda_p \quad (6)$$

最后,将 $J_c \oplus J_p = J$ 计算出邻居车辆对目标车辆总的交互影响。模型采用 λ 层将可用的车辆交互总结为固定大小的线性函数,可直接应用于相关 Query 计算。由此产生的车辆轨迹预测网络,计算效率高,并以较小的内存成本捕获长距离依赖关系,使其能够应用于大型结构化输入,同时具有较好的扩展性。

2.2.3 解码

解码同样采用 LSTM 解码器。输入目标车辆的编码轨迹与学习到的多种预测结果或车辆之间的相互影响进行解码。LSTM 解码器会生成时间步长 T 的目标车辆未来位置估计分布的预测参数。通过解码,可以得到 $Y_V = \{Y_V^{obs+1}, Y_V^{obs+2}, \dots, Y_V^{obs+pred}\}$, 然后通过损失函数训练模型。通常采取负对数似然函数与均方根误差函数作为损失函数。

在解码之前,连接车辆交互影响与 LSTM 编码最后一个时刻的隐藏状态 h_V^{obs} 。

$$z = \text{concat}(J, h_V^{obs}) \quad (7)$$

既包含邻居车辆对目标车辆的交互影响,又包含目标车辆自身运动信息。根据隐藏状态 z , 通过递归神经网络,对目标车辆未来的隐藏状态进行预测,最后通过全连接,将未来的隐藏状态转换成车辆未来的轨迹。具体公式为:

$$Y_V = \psi[LSTM(z^{t-1}; W_z); W_{dec}] \quad (8)$$

式中, z^{t-1} 是 $t-1$ 时刻解码器的隐藏状态向量。根据 z^{t-1} 和隐藏状态 LSTM 得到 z^t ,再根据 z^t 全连接得到车辆未来轨迹坐标。

2.3 损失函数

在本文中使用均方根误差 (RMSE) 来评价该模型的性能。度量预测轨迹和真实轨迹之间的差异公式为:

$$L_{RMSE} = \sqrt{\frac{1}{pred} \sum_{t=obs+1}^{obs+pred} [(x_V^t - x_V^t)^2 + (y_V^t - y_V^t)^2]} \quad (9)$$

3 实验

3.1 数据集

深度学习是机器学习中一种基于对数据进行表征学习的方法,因此研究车辆轨迹预测时可选取公开的数据集,方便其他研究人员复现成果。NGSIM (Next Generation Simulation) 数据集是美国智能交通系统联合计划办公室提供的美国高速公路行车数据,广泛应用于车辆轨迹预测。包含实验时间段内的所有车辆,根据时间和车辆编号保存数据。

3.2 预测精度与结果分析

模型在 NGSIM 数据集上的实验结果见表 1。本文模型输入为 5 个参数, 分别是车辆横纵坐标、速度、加速度、车辆种类(汽车、摩托车、卡车等)。

表格中对比的模型分别为:

CV^[18]: 恒速模型使用匀速递归贝叶斯滤波器, 用于短时间范围内(< 2 s)的轨迹预测。

GAIL-GRU^[19]: 生成对抗性模仿学习模型使用 GRU。GRU 是 LSTM 网络的变体, 较 LSTM 网络的结构更加简单, 具有更少的参数。

V-LSTM: 该方案是采取典型的 LSTM 编码-解码模型, 历史轨迹作为 LSTM 输入, LSTM 解码作为输出。

Social-LSTM^[6]: 在 LSTM 编码-解码器模型基础上, 应用社交池层, 学习车辆之间的交互影响。

CS-LSTM^[5]: 利用卷积在社交池中提取特征, 同时在模型输出中使用多模态预测车辆未来的多重可能性。

NLS-LSTM^[20]: 基于一个 LSTM 编码-解码器, 非本地社交池用于捕捉车辆之间的互动, 并采用非局部多头注意机制对互动进行了总结。

MATS: 将背景场景特征和车辆历史轨迹特征连接成一个张量, 再用生成对抗网络(GAN)模块生成未来的轨迹。

表 1 中显示了所有对比模型的 RMSE 性能, 相

较于数据驱动的方法传统的模型 CV 性能, 其它数据驱动的方案性能更好、更有效。由于 V-LSTM、S-LSTM 只考虑车辆的动力性能, 利用 LSTM 或变体来预测车辆轨迹, 而未考虑车辆之间的交互影响, 因而预测表现不佳。CS-LSTM、NLS-LSTM 使用卷积或注意力机制学习车辆的交互影响, 邻居车辆的信息有助于车辆轨迹预测。CS-LSTM、NLS-LSTM、MATS 均考虑了相邻车辆的特征, 并通过卷积神经网络进行处理。此外, MATS 还使用生成对抗网络来提升模型性能。其模型相较于过去的模型性能更好。具体优势体现在:

(1) 模型考虑车辆速度、加速度、车辆类型对轨迹预测的影响, 在预测时加入这些元素可提升模型性能。

(2) 模型生成 lambda 层的目标函数 λ , 共享 λ 可以加强模型计算速度。

4 结束语

本文提出了一种结合 LSTM 编码-解码器与 Lambda 层的车辆轨迹预测方案。应用 Lambda 层共享的 λ 参数, 为车辆轨迹应用加快了计算速度。同时实验结果证明, 预测精度相较于过去的方案有一定的提升。本文概述了一种深度学习轨迹预测方案, 不仅为计算机科学研究人员和专业人员, 也为处理轨迹的更广泛的领域提供了参考。

表 1 模型均方根误差性能对比

Tab. 1 RMSE of Several models on NGSIM dataset

数据集	预测范围/s	CV	GAIL-GRU	V-LSTM	S-LSTM	CS-LSTM	NLS-LSTM	MATS	Lambda-LSTM
NGSIM	1	0.73	0.69	0.68	0.65	0.61	0.56	0.66	0.46
	2	1.78	1.51	1.65	1.31	1.27	1.22	1.34	1.08
	3	3.13	2.55	2.91	2.16	2.09	2.02	2.08	1.81
	4	4.78	3.65	4.46	3.25	3.10	3.03	2.97	2.74
	5	6.68	4.71	6.27	4.55	4.37	4.30	4.13	3.92

参考文献

- [1] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. arXiv preprint arXiv:1406.1078, 2014.
- [2] 王少博. 动态场景下基于交互性预测的自动驾驶汽车轨迹规划方法研究[D]. 合肥: 中国科学技术大学, 2020.
- [3] LEFÈVRE S, VASQUEZ D, LAUGIER C. A survey on motion prediction and risk assessment for intelligent vehicles [J]. ROBOMECH journal, 2014, 1(1): 1-14.
- [4] DEO N, RANGESH A, TRIVEDI M M. How would surround vehicles move? a unified framework for maneuver classification and motion prediction [J]. IEEE Transactions on Intelligent

Vehicles, 2018, 3(2): 129-140.

- [5] DEO N, TRIVEDI M M. Convolutional social pooling for vehicle trajectory prediction[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018: 1468-1476.
- [6] ALAHI A, GOEL K, RAMANATHAN V, et al. Social lstm: Human trajectory prediction in crowded spaces[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 961-971.
- [7] GORI M, MONFARDINI G, SCARSELLI F. A new model for learning in graph domains [C]//Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005: 729-734.

(下转第 91 页)