

基于深度学习的智能车辆轨迹预测综述*

杨荣森¹ 张国宗²

(1.重庆交通大学机电与车辆工程学院,重庆 400074;2.四川轻化工大学机械工程学院,宜宾 643022)

【欢迎引用】杨荣森,张国宗.基于深度学习的智能车辆轨迹预测综述[J].汽车文摘,2024(2):1-9.

【Cite this paper】YANG R M, ZHANG G Z. An Overview on Intelligent Vehicle Trajectory Prediction Based on Deep Learning [J]. Automotive Digest (Chinese), 2024(2): 1-9.

【摘要】为了全面了解基于深度学习的智能车辆轨迹预测研究方法和现状,通过对现有文献进行分析和总结,分析了基于深度学习的轨迹预测模型的输入表示、输出类型和预测方法。结果表明,基于深度学习的轨迹预测方法在长时域、多模态运动及车路交互场景中能够取得优异表现。

关键词:自动驾驶;车辆轨迹预测;深度学习;复杂行车场景

中图分类号:U463 文献标志码:A DOI:10.19822/j.cnki.1671-6329.20230085

An Overview on Intelligent Vehicle Trajectory Prediction Based on Deep Learning

Yang Rongmiao¹, Zhang Guozong²

(1.School of Mechatronics and Vehicle Engineering, Chongqing Jiaotong University, Chongqing 400074; 2.School of Mechanical Engineering, Sichuan University of Science & Engineering, Yibing 643002)

【Abstract】In order to fully understand the research methods and current status of deep learning-based trajectory prediction of intelligent vehicles, through the analysis and summary of existing literature, the input representation, output types, and prediction methods of deep learning-based trajectory prediction models are analyzed. The results show that trajectory prediction methods based on deep learning demonstrate outstanding performance in long-term, multi-modal motion and vehicle-road interaction scenarios.

Key words: Autonomous driving, Vehicle trajectory prediction, Deep learning, Complex driving scenarios

0 引言

计算机视觉、传感器信号处理、硬件设计等领域的最新技术进展,使自动驾驶技术从“可能”、“可行”发展到“商用”状态。然而,自动驾驶车辆应对复杂行驶场景和紧急交通任务的能力仍存在局限性^[1]。因此,提高自动驾驶汽车智能系统的性能非常重要。预测智能体(Agent)未来时刻的位置并根据预测结果做出合理的规划是自动驾驶车辆和高级监控系统(包括智能交通或者智能城市)的关键任务。理解交通智能体的行为是智能系统与人类共存和互动的一项关键技术,其中涉及表征、感知和运动分析等方面,而预测

在运动分析中起着重要的作用。预测车辆运动主要面临3个重要挑战:

(1)在复杂的多车行驶场景下,车辆的运动状态会受其他车辆影响;

(2)车辆运动呈现多模态特征,即给定相同的车辆运动历史,会输出多个不同的合理轨迹;

(3)在缺失道路几何形状和交通规则约束的情况下,特定场景训练的模型泛化性较差。

本文针对复杂行车场景中车辆轨迹预测问题,综述了近几年的研究成果,并根据输入表示、轨迹输出类型、预测方法进行综述与分析,基于分析结果总结智能车辆轨迹预测领域发展现状,并对未来趋势进行展望。

*基金项目:研究生创新基金项目(Y2023082)。

1 智能车辆轨迹预测算法综述

由于深度学习的方法能高效地提取时间与空间交互关系并能更好地进行长时域预测^[1],现有研究方法主要是基于深度学习的方法来解决轨迹预测问题。因此,本文综述了近几年基于深度学习的车辆轨迹预测研究成果,从输入表示、输出类型和预测方法方面进行总结及探讨。

1.1 输入表示

常用的轨迹预测算法的输入表示见表1。

表1 轨迹预测算法的输入表示

| 算法分类 | 参考文献 | 算法描述 |
|--------|--------|----------------------------|
| 车辆运动状态 | 基于物理因素 | [3] 车辆运动学或动力学参数作为输入 |
| | 单车因素 | [4-5] 单一车的轨迹历史序列作为输入 |
| | 车辆交互 | [6-7] 目标和邻近车的轨迹历史序列作为输入 |
| 地图表示 | 栅格化 | [8-9] 将地图及车辆简化为鸟瞰图的表示方法 |
| | 矢量化 | [10-13] 将地图及车辆表示为图结构的表示方法 |
| 其他方法 | 原始数据 | [14-15] 多任务共享算力处理原始感知输入的数据 |

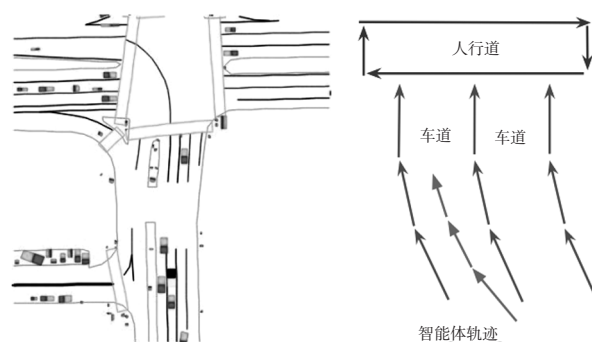
1.1.1 车辆运动状态

传统基于模型驱动轨迹预测方法常以速度、加速度等作为输入构建车辆运动学或动力学模型来实现单一车辆短时域的轨迹预测。Abbas等^[3]使用5种卡尔曼滤波模型对车辆的5种运动状态进行预测,但该方法仅限于短时域预测,并难以考虑周围车辆的影响。文献[4]用车辆的横纵向位置坐标、速度和航向角的轨迹历史序列预测车辆的机动行为,但忽略了车辆间的交互感知。文献[5]指出受车辆传感器限制和障碍物遮挡,车辆不能有效地获取周围车辆的运动状态,使用单一车辆的运动状态信息作为输入的深度学习模型进行轨迹预测。为了研究车辆间相互作用关系,学者们在轨迹预测研究中输入目标车辆与周围车辆的运动状态信息。在文献[6-7]中将车辆和邻近车辆的运动状态历史作为输入来预测车辆未来的轨迹,研究表明车辆在合理感知范围内能准确地获取周围车辆的运动状态信息。

1.1.2 地图表示

考虑环境条件和交通因素能够更好地预测车辆的运动状态,学者们将高清(High Definition, HD)地图信息和车辆轨迹信息融合来精确预测车辆运

动状态。高清地图主要有2种表示方法,包括栅格化地图和矢量化地图,如图1所示。



(a) 栅格化地图表示

(b) 矢量化地图表示

图1 地图表示方法^[11]

栅格化地图将高清地图简化为鸟瞰图(Bird's Eye View, BEV)。该方法将驾驶场景中动态交通参与者和静态环境等因素表示为多边形和线条集合,并保留交通参与者的形状、位置和道路信息集合。文献[8]使用CoverNet来提取栅格化地图的语义特征。文献[9]使用卷积神经网络(Convolutional Neural Network, CNN)在栅格化地图中提取车辆的运动状态特征和交互表示。栅格化的方法可解释性好,语义信息丰富全面,可融入更多的地图信息(红绿灯),但该方法感受野受限,栅格化处理过程中会丢失信息,连续的卷积计算效率较低。

栅格化表示交互作用局限,为了更好地构建车辆间、车辆与地图间的相互作用关系,一些学者将交通参与者和静态环境表征为矢量化图结构。文献[10]认为矢量化地图能够更好地用于轨迹预测。文献[11]将驾驶场景中动静对象及交互关系表征为节点和折线进行交互。文献[12]使用矢量化地图来表征车辆与对应的车道信息。文献[13]使用矢量化方法构造了复杂的车辆、拓扑结构和车道的相互依赖关系。矢量化地图具有较强的交互关系建模能力,高效的图数据结构有更少的模型参数,但存在一定信息损失。

1.1.3 其他方法

在一些研究方法中,将原始数据输入到轨迹预测模型中,输入数据包含了所有周围环境的可用信息,但输入数据量较大,需要更多的计算负载。为了缓解计算负载,可使用不同任务共享计算资源。文献[14]以数据驱动的方法实现自动驾驶系统中目标检测、跟踪和预测。文献[15]利用激光雷达数据作为输入,使用同样的方法来完成任务功能。该模型使用了简化的BEV处理,同时将3D点云数据输入到深

度模型中。

1.2 输出类型

预测模型的轨迹生成方法主要分为机动意图、单模态以及多模态3个类型,如表2所示。

1.2.1 机动意图

机动意图可以是预测模型的最终输出,也可以是预测模型的一个输出分支(通过预测车辆机动意图来协助轨迹预测)。刘志强等^[16]使用基于隐马尔可夫和支持向量机的方法来识别驾驶意图。Ding等^[17]使用意图预测模型预测高速场景下车辆的变道和保持车道行为。季学武等^[18]利用长短期记忆网络(Long Short Term Memory, LSTM)识别驾驶意图后进行车辆轨迹预测,如图2所示。但基于机动意图的方法存在意图识别误差,会导致预测结果偏差过大,甚至造成车辆驾驶安全事故。

1.2.2 单模态

预测模型输出单个或多个交通参与者的未来轨迹。文献[19–20]采用随每个时间步产生对应位置的轨迹生成方法。文献[21]使用深度学习的方法学习3

次多项式的3个参数表示未来的轨迹。文献[22]解码模块生成每个时间步对应位置的二元高斯分布的均值、标准差和相关系数,并讨论了单模态的轨迹生成方法收敛趋于所有模态的平均值,该方法表示复杂的车辆运动状态比较局限。

表2 轨迹预测算法的输出类型

| 算法分类 | 参考文献 | 算法描述 |
|------|---------|-------------------|
| 机动意图 | [16] | 输出驾驶意图 |
| | [17–18] | 输出驾驶意图及轨迹预测 |
| 单模态 | [19–20] | 输出随时间步对应的位置 |
| | [21] | 输出基于深度学习的3次多项式轨迹 |
| | [22] | 输出随时间步对应位置的二维高斯分布 |
| | [23] | 基于生成对抗网络的轨迹预测 |
| 多模态 | [24] | 基于条件变分自编码的轨迹预测 |
| | [25] | 基于归一化流的轨迹预测 |
| | [26] | 基于预测终点条件的锚点的轨迹预测 |
| | [9] | 基于原轨迹条件的锚点的轨迹预测 |
| | [27] | 基于网格的轨迹预测 |
| | [6] | 基于双变量高斯的轨迹预测 |

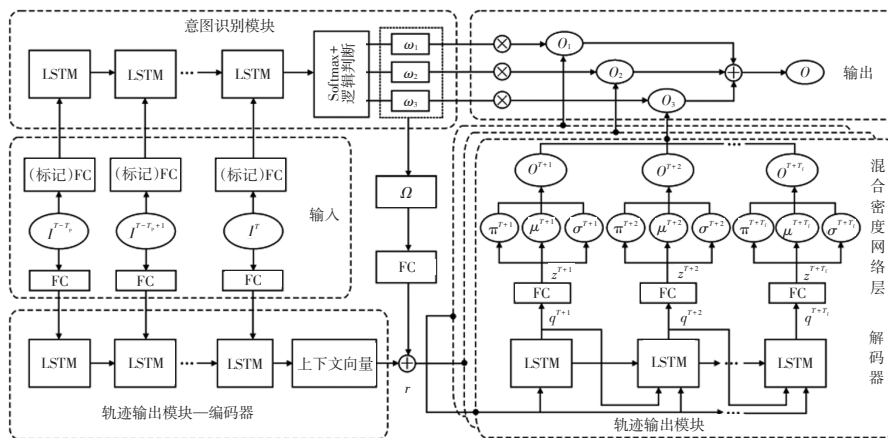
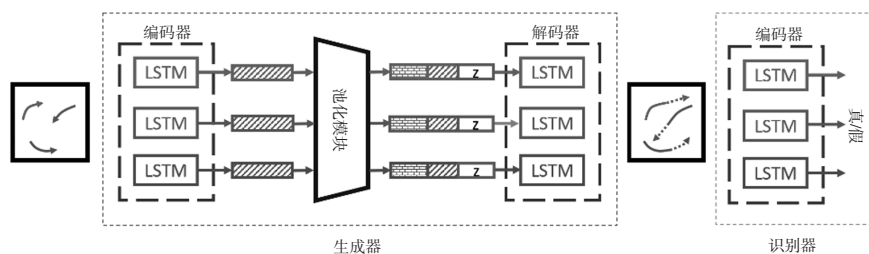
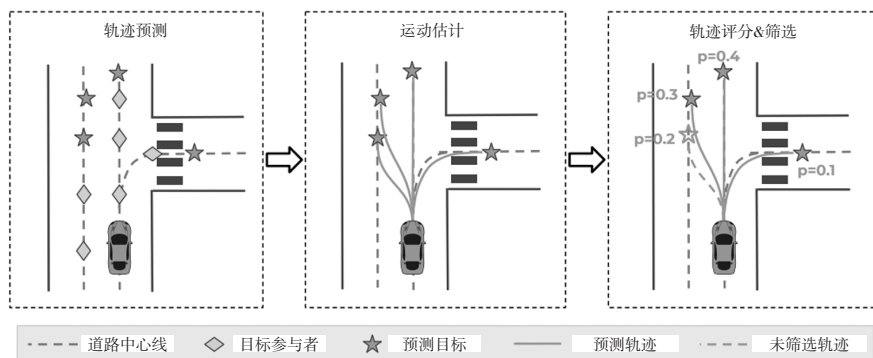


图2 机动意图及轨迹预测的网络结构^[18]

1.2.3 多模态

多模态预测方法是指在预测车辆未来运动轨迹时,考虑到车辆可能存在不同的运动模式,并预测每种模式下的轨迹。为了更好地获取预测轨迹,Guptade等^[23]提出一种生成对抗网络,通过学习鉴别器来区分预测轨迹是否正确,即设计一种对抗损失函数来判定预测的好坏,如图3所示。Kingma等^[24]提出了条件变分自动编码器,采用神经网络参数化对变分推理中的分布,提高了模型的生成能力。基于生成对抗网络和条件变分自编码器的方法采用隐式分布建模而难以训练,基于归一化流模型能够显示学习参数分布,该方法可以通过可逆变换将复杂的分布变换为易处理的形式。Bhattacharyya等^[25]通过基于Haar小波轨迹块自回

归该模型进行拆分耦合学习预测分布。Zhao等^[26]使用基于预测终点条件的锚点方法,该方法通过采样车道中心线的位置来预测车辆的终点位置,但应考虑障碍物导致无法到达终点的情况,如图4所示。Chai等^[9]使用原轨迹条件的锚点方法,该方法对轨迹集中进行聚类分析来获得多模态未来分布,但应对复杂的场景时该方法比较局限。Li等^[27]使用基于网格的方法,搜索具有top K个累积对数概率的轨迹。基于网格的方法可以高度兼容具有高级训练策略的场景,并且适应于长时域预测,其缺点是需要大量卷积计算和对地图分辨率要求较高。Deo等^[6]使用双变量高斯表示输出,在每个时间步的位置生成双变量高斯分布,通过负对数似然损失函数来获取最高可能性的预测分布。

图3 生成对抗网络结构^[23]图4 运用锚点的多模态方法^[26]

1.3 预测方法

将现有研究所使用的预测模型分为递归神经网络(Recurrent Neural Network, RNN)、卷积神经网络(Convolutional Neural Network, CNN)、递归和卷积神经混合网络(RNN&CNN)、图神经网络(Graph Neural Network, GNN)和注意力机制(Attention),如表3所示。

1.3.1 递归神经网络

RNN在应对时序问题时展现出强大的信息处理能力和深度表征能力,因此越来越多的学者使用递归神经网络来处理时间序列问题^[18]。RNN存储了当前时间步的历史信息,其输出由隐状态和输入共同决定。其模型的参数量不会随时间步的增加而增加,但时间步过长,其梯度存在“梯度消失”和“梯度爆炸”的问题。在轨迹预测中常采用LSTM和门控循环单元(Gated Recurrent Unit, GRU)解决这一问题。同时,GRU相比LSTM结构较简化,计算速度更快,适用于数据量较小的任务^[28]。根据现有RNN的轨迹预测研究,可分为单RNN和多RNN模型。

单RNN模型常适用于基于机动和单模态的轨迹预测模型或者其他辅助模型,来实现更复杂的任务,例如车辆交互感知。文献[4]将LSTM作为序列分类器来预测车辆的机动类型。Alché等^[19]将车辆历史运动状态序列输入单层LSTM来预测高速场景车辆的轨迹。为了预测基于机动的轨迹,Ding等^[20]采用LSTM

编码器对车辆运动状态信息编码来预测车辆的机动,并且结合地图信息预测轨迹。

多RNN结构常适用于解决预测问题。Xin等^[5]使用2组LSTM搭建模型,一个LSTM预测车辆未来的目标车道,同时,结合目标车辆的运动状态和预测的目标车辆结果预测车辆未来的轨迹。Dai等^[7]也采用2组LSTM实现车辆的轨迹预测,一组用于目标车辆与所有周围车辆的单个轨迹序列编码,另一组用于目标车辆与周围车辆的相互作用关系的交互建模。Ding等^[17]利用GRU编码器来构建目标车辆与周围车辆之间的相互作用关系,并预测车辆在长时域内的意图,如图5所示。

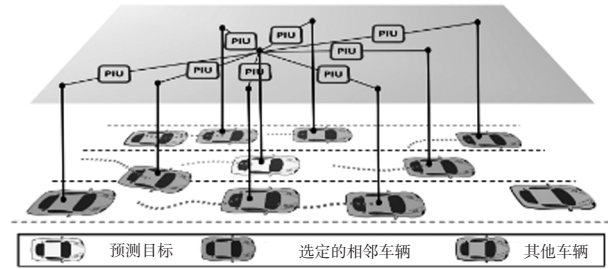
1.3.2 卷积神经网络

随着CNN在计算机视觉等领域快速发展,学者们也将CNN用于解决轨迹预测问题。卷积神经网络组成包括卷积层、激活函数、池化层及全连接层。Nikhil等^[29]认为车辆轨迹序列具有强时间空间连续性,使用CNN优于RNN预测轨迹,其使用卷积层和全连接层实现了车辆未来轨迹的预测。Song等^[30]使用一维时间卷积将车辆运动状态序列嵌入到高维向量进行数据提取。然而,现有轨迹预测研究方法中,CNN常被用于提取鸟瞰图的特征。Phan-Minh等^[31]根据车辆运动状态(速度、加速度和横摆角速度)生成轨迹集,并结合CNN处理的栅格图像,将发生概率最高的轨迹作为车辆未来的轨迹,如图6所示。Luo等^[14]使用三维卷积在时间维上对体素化激光雷达数据的4D表示提取时

间特征,然后使用二维卷积提取空间特征。最后通过2个分支,分别输出当前和未来帧的被检测对象的标记框。

表3 车辆轨迹预测算法

| 算法分类 | 参考文献 | 算法描述 |
|--------------------------|---------------------|-------------------------------------|
| RNN | 单一 RNN | [4] 基于 LSTM 的机动意图识别 |
| | | [19] 基于 LSTM 的单车轨迹预测 |
| | | [20] 基于 LSTM 的车辆机动识别及轨迹预测 |
| | 多 RNN | [5] 基于 2 组 LSTM 的车辆机动识别及轨迹预测 |
| | | [7] 基于 2 组 LSTM 的车辆交互轨迹预测 |
| | | [17] 基于 GRU 的车辆交互意图预测 |
| CNN | CNN | [29] 基于 CNN 的轨迹预测 |
| | 一维 CNN | [30] 基于一维时间卷积的轨迹嵌入 |
| | 鸟瞰图 & CNN | [31] 基于鸟瞰图和 CNN 结合的轨迹预测方法 |
| | 激光雷达数据 & CNN | [14] 基于 4D 激光雷达数据和三维 CNN 结合的轨迹预测方法 |
| RNN & CNN | LSTM & CNN | [32-33] 基于 LSTM 和 CNN 结合的轨迹预测方法 |
| | LSTM & CNN & 鸟瞰图 | [34] 基于 LSTM、CNN 和鸟瞰图结合的轨迹预测方法 |
| | GRU & CNN & 鸟瞰图 | [35] 基于 GRU、CNN 和鸟瞰图结合的轨迹预测方法 |
| | CNN & GRU & LSTM | [36] 基于 CNN、GRU、LSTM 和鸟瞰图结合的轨迹预测方法 |
| Attention | 多头注意力 | [37] 基于多头注意力的轨迹预测 |
| | 多头注意力 & LSTM | [38] 基于多头注意力和 LSTM 结合的轨迹预测方法 |
| | 自注意力 & LSTM | [39] 基于自注意力和 LSTM 结合的轨迹预测方法 |
| | 通道注意力 & Transformer | [41] 基于通道注意力和 Transformer 结合的轨迹预测方法 |
| | Transformer | [42] 基于 Transformer 融合环境因素的轨迹预测 |
| GNN | GCN | [43-44] 基于 GCN 编码的轨迹预测 |
| | | [45-46] 基于谱聚类 GCN 的轨迹预测 |
| | 矢量化地图 & GNN | [10] 车辆及环境因素的向量化表示 |
| | | [11] 基于交互的车辆及环境因素的向量化表示 |
| | | [12] CNN 和 GCN 融合的轨迹预测 |
| | 其他 | [47] 基于 GAT 和 LSTM 结合 |
| [48] 基于有向图 GNN 和 LSTM 结合 | | |
| [50] 基于有向图 GNN | | |



注:PIU(path Information Unit)为路径信息单元

图5 运用GRU的车辆交互表示^[17]

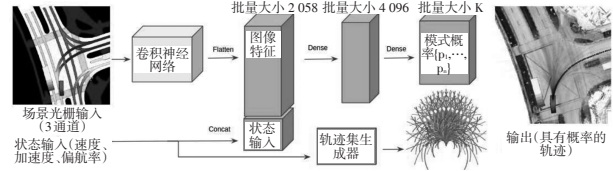


图6 运用CNN的网络结构^[31]

1.3.3 卷积和循环神经网络

卷积和循环神经网络混合的方法使用RNN在时间序列信息中提取时间特征,并且使用CNN提取空间特征。因此,学者们使用CNN和RNN混合的方法挖掘车辆运动状态的时间和空间的特征。Alahi等^[32]将LSTM和全连接池化层结合,将空间范围内LSTM的特征池化到以目标车辆为中心的网格。Deo等^[33]认为卷积池化层相比全连接池化层能更好地提取空间特征,其在合理感知范围内给每辆车一个LSTM编码器提取车辆的时序特征,然后馈送到卷积池化层中提取局部空间交互特征,进而输入到基于机动的LSTM解码器来输出车辆未来的轨迹分布。Schreiber等^[34]使用CNN提取简化鸟瞰图中车辆在不同时间步的周围环境特征,然后输入到基于编码-解码框架的LSTM提取输入数据的时间特征,进而将解码器输出馈入到反CNN来生成输出图像,这些图像表示车辆的周围环境在未来的演变。Lee等^[35]使用基于编码-解码的GRU生成轨迹分布,并输入到解码器GRU中进行细化和排序。同时,后一个模块还接收使用CNN提取的场景特征。Zhu等^[36]使用基于CNN、GRU和LSTM混合网络,该方法由3部分组成:基于池化和GRU混合网络提取全局时空交互特征、基于CNN网络提取环境特征以及基于编码-解码的LSTM预测网络。

1.3.4 注意力机制

注意力机制能在复杂的数据中快速高效地关注到高价值信息。随着注意力机制在视觉、自然语言处理等领域快速发展及应用,学者们将注意力机制用来解决轨迹预测问题。在视觉处理中,现有的注意力机制主要分为通道域注意力、空间域注意力及混合域注

意力。在自然语言处理中,注意力机制主要包含自注意力和多头注意力。文献[37]使用基于多头注意力机制赋予车道和车辆注意力权重,并输出未来轨迹的分布。Messaoud等^[38]赋予场景中每辆车一个LSTM编码器,并利用多头注意力池来挖掘车辆之间的相互作用关系。Lin等^[39]将注意力和LSTM编码器结合起来提取车辆运动状态的时间序列特征,并赋予每个时间步注意力权重。同时,利用注意力机制挖掘空间特征,并赋予周围车辆空间交互注意力权重,提高了模型的可解释性。Vaswani等^[40]提出了Transformer模型,该模型使用了单头注意力机制和多头注意力机制搭建来完成序列机器翻译任务。该方法可以并行计算序列数据,RNN会受序列计算的约束。Zhao等^[41]使用基于通道-空间注意力机制的方法,利用通道注意力机制在输入数据的通道维度上表征车辆间的全局交互,并利用Transformer提取空间特征并输出车辆未来的轨迹预测,分析了通道和空间注意力串并行结构和模型在不同交互场景密度预测的性能,如图7所示。现有研究中使用Transformer将交通参与者与环境之间进行交互性建模。Liu等^[42]使用Transformer将环境信息融合到轨迹序列来预测未来轨迹。由此可见,基于Transformer模型在处理时间序列数据方面具有优势。

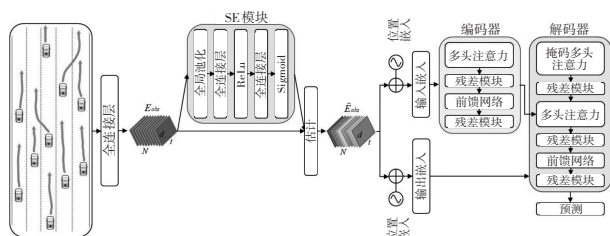


图7 通道注意力和Transformer结合的网络结构^[41]

1.3.5 图神经网络

图神经网络在轨迹预测研究中将驾驶场景中的交通参与者表示为一个节点,将交通参与者间的关系表示为一条边,所有的点和边集合形成一个图。图神经网络与现有研究中RNN和CNN的方法相比,能够提取非欧几里得空间数据特征。在复杂的驾驶场景中,每个场景都可以视作一个不规则的图,每个图的节点由交通参与者的运动状态特征表示或者静态环境信息表示。每个图的节点间存在相互关联的边,能够用来表示场景中交通参与者的相互依赖关系。因此,图神经网络适用于解决基于交互的车辆轨迹预测问题。

图卷积网络(Graph Convolutional Network, GCN)

将传统的CNN扩展到图数据卷积处理,能够将图中的节点特征构成的邻域结构映射获取数据特征信息。Li等^[43]提出了一种基于GCN和LSTM混合网络的轨迹预测模型,用几个图卷积块编码输入轨迹数据和基于编码-解码的LSTM模型输出车辆未来的轨迹预测。在驾驶场景中考虑车辆间的距离来定义交互关系,该方法相比CNN和RNN的轨迹预测有较大的改进,但缺乏泛化性。使用固定图和动态图优化了原有方法预测不同类型的交通参与者的轨迹^[44]。同时,由于GCN模型存在一个典型问题—过平滑,可使用跳过连接(类似残差连接)来缓解这个问题,最后将LSTM替换成GRU来预测车辆未来的轨迹。

为了考虑交通参与者的机动,一些学者采用基于谱聚类的图卷积网络。Chandra等^[45]提出一种基于谱聚类GCN的车辆轨迹预测方法,该方法可以提取交通参与者的低级信息(未来轨迹)和高级信息(交通参与者行为机动),其使用2层基于GCN和LSTM混合网络,分别用来预测交通参与者的空间坐标和建模交通参与者的空间交互来预测其超速等行为,但其使用单一行为特征进行预测导致模型的泛化性差和计算量过大。Zhao等^[46]使用基于频谱GCN预测轨迹,该方法考虑了周围车辆的交互能够共享场景内所有车辆的状态信息,符合车辆间社会交互。

使用GNN表示矢量化地图。Ziegler等^[10]率先将矢量化地图应用于轨迹预测,将车辆及车道进行地图拓扑来提取出未来沿车道的轨迹,但其忽略了车辆间交互。Gao等^[11]使用层次图神经网络VectorNet,该方法首先利用向量表示交通参与者以及静态环境的空间局域性,同时建模所有对象的高阶相互作用。Liang等^[12]使用基于CNN和GCN混合网络,该方法使用CNN提取车辆特征并用GCN提取矢量地图中车道特征来预测轨迹,如图8所示。

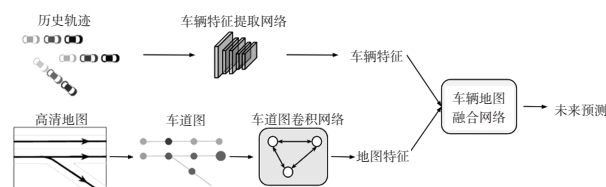


图8 运用GCN的网络结构^[12]

一些研究也使用其他图神经网络来实现轨迹预测。图注意力网络(Graph Attention Network, GAT)通过注意力机制对邻接节点进行聚合来对不同邻接节点的自适应校核,该方法能够提高图神经网络特征提取能力^[47]。Huang等^[48]使用图注意力网络预测轨迹,该

方法利用 LSTM 编码器编码交通参与者的运动状态信息,然后将编码信息输入到 GAT 块计算每个交通参与者的注意力权重。最后,该模型使用 LSTM 解码器来生成预测轨迹。图神经网络根据边的方向性可以划分为有向图和无向图。Zhang 等^[49]使用社会图网络轨迹预测。该方法根据实时位置和速度方向动态构建有向图来有效地获取交通参与者的运动状态。使用基于社会图和 LSTM 构成的网络表示社会交互,利用时间随机方法序列学习社会交互不确定性形成先验模型,对先验模型进行采样,使用 LSTM 解码器来预测轨迹。Salzmann 等^[50]提出图结构递归网络来生成动态可行的未来轨迹,该方法使用有向时空图表示驾驶场景,并紧密地集成到车辆规划控制系统。

2 结束语

为使智能车辆在复杂的交通场景下安全高效地行驶,智能车辆需要能够自主准确地理解交通参与者的相互作用关系和预测其运动状态。传统的预测方法已经无法满足自动驾驶系统对场景的复杂性和车辆行为多样性的需求。本文综述近几年智能车辆的轨迹预测研究方法,通过输入表示、输出类型和预测方法进行分析,对本文的总结及展望如下。

(1)目前大部分预测模型将交通参与者的运动状态和静态环境嵌入,然后使用深度学习等方法对车辆间、车辆与道路间的交互建模,进而利用多模态轨迹生成方法产生车辆未来的轨迹。相比传统预测方法,深度学习能更好地建模时空特征联合交互以及解决长时域预测的问题。

(2)车辆轨迹预测的嵌入方式主要采用栅格化和矢量化。栅格化方法可解释性好,语义信息丰富全面,但感受野受限,存在信息丢失以及连续的卷积计算效率较低。矢量化方法具有较强的交互关系建模能力,高效的图数据结构有更少的模型参数,但存在一定信息损失。针对输入表示的问题也需要提出更好的方法。

(3)车辆预测轨迹生成方法可以分为3种类型:机动意图、单模态和多模态。机动意图产生车辆行为意图来协助预测,但存在意图识别误差,会导致轨迹预测偏差过大。单模态输出单个或多个交通参与者的未来轨迹,但表示复杂的车辆运动状态局限较大以及输出单一可能轨迹会忽略未来的真实轨迹。多模态将输出轨迹表示在合理的轨迹集中,能更好地生成未来可能的轨迹。多模态也是轨迹预测领域未来主要

研究方向。

(4)目前大多数预测方法只考虑车辆的运动状态和环境结构,交通规则可以重塑车辆在驾驶场景中的轨迹,但是很少被作为输入。车辆其他视觉和听觉数据也可以用来预测车辆的轨迹。大部分预测方案仅限于特定的行车场景(十字路口),其方案存在局限性。

参 考 文 献

- [1] ATKINS W S. Research on the Impacts of Connected and Autonomous Vehicles (CAVs) on Traffic Flow[R]. Stage 2: Traffic Modelling and Analysis Technical Report, 2016.
- [2] 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.
- [3] ABBAS M T, JIBRAN M A, AFAQ M, et al. An adaptive approach to vehicle trajectory prediction using multimodel Kalman filter[J]. Transactions on Emerging Telecommunications Technologies, 2020, 31(5): e3734.
- [4] ZYNER A, WORRALL S, NEBOT E. A recurrent neural network solution for predicting driver intention at unsignalized intersections[J]. IEEE Robotics and Automation Letters, 2018, 3(3): 1759-1764.
- [5] XIN L, WANG P, CHAN C Y, et al. Intention-aware long horizon trajectory prediction of surrounding vehicles using dual LSTM networks[C]//2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2018: 1441-1446.
- [6] DEO N, TRIVEDI M M. Multi-modal trajectory prediction of surrounding vehicles with maneuver based lstms[C]//2018 IEEE intelligent vehicles symposium (IV). IEEE, 2018: 1179-1184.
- [7] DAI S, LI L, LI Z. Modeling vehicle interactions via modified LSTM models for trajectory prediction[J]. IEEE Access, 2019, 7: 38287-38296.
- [8] CUI H, NGUYEN T, CHOU F C, et al. Deep kinematic models for kinematically feasible vehicle trajectory predictions [C]//2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020: 10563-10569.
- [9] CHAI Y, SAPP B, BANSAL M, et al. Multipath: Multiple probabilistic anchor trajectory hypotheses for behavior prediction[J/OL]. (2019-10-12) [2024-01-25]. <https://doi.org/10.48550/arxiv.1910.05449>.
- [10] ZIEGLER J, BENDER P, SCHREIBER M, et al. Making bertha drive—an autonomous journey on a historic route [J]. IEEE Intelligent transportation systems magazine, 2014, 6(2): 8-20.

- [11] GAO J, SUN C, ZHAO H, et al. Vectornet: Encoding hd maps and agent dynamics from vectorized representation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 11525–11533.
- [12] LIANG M, YANG B, HU R, et al. Learning lane graph representations for motion forecasting[C]//Computer Vision – ECCV 2020: 16th European Conference. ECCV, 2020: 541–556.
- [13] ZENG W, LIANG M, LIAO R, et al. Lanercnn: Distributed representations for graph-centric motion forecasting[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021: 532–539.
- [14] LUO W, YANG B, URTASUN R. Fast and furious: Real time end-to-end 3d detection, tracking and motion forecasting with a single convolutional net[C]//Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. IEEE, 2018: 3569–3577.
- [15] CASAS S, LUO W, URTASUN R. Intentnet: Learning to predict intention from raw sensor data[C]//Conference on Robot Learning. PMLR, 2018: 947–956.
- [16] 刘志强, 吴雪刚, 倪捷, 等. 基于 HMM 和 SVM 级联算法的驾驶意图识别[J]. 汽车工程, 2018, 40(7): 858–864.
- [17] DING W, CHEN J, SHEN S. Predicting vehicle behaviors over an extended horizon using behavior interaction network[C]//2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019: 8634–8640.
- [18] 季学武, 费聪, 何祥坤, 等. 基于 LSTM 网络的驾驶意图识别及车辆轨迹预测[J]. 中国公路学报, 2019, 32(6): 34–42.
- [19] ALTCHÉ F, DE LA FORTELLE A. An LSTM network for highway trajectory prediction[C]//2017 IEEE 20th international conference on intelligent transportation systems (ITSC). IEEE, 2017: 353–359.
- [20] DING W, SHEN S. Online vehicle trajectory prediction using policy anticipation network and optimization-based context reasoning[C]//2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019: 9610–9616.
- [21] MIN K, KIM D, PARK J, et al. RNN-based path prediction of obstacle vehicles with deep ensemble[J]. IEEE Transactions on Vehicular Technology, 2019, 68(10): 10252–10256.
- [22] MA Y, ZHU X, ZHANG S, et al. Trafficpredict: Trajectory prediction for heterogeneous traffic-agents[C]//Proceedings of the AAAI conference on artificial intelligence. AAAI, 2019, 33(1): 6120–6127.
- [23] GUPTA A, JOHNSON J, FEI-FEI L, et al. Social gan: Socially acceptable trajectories with generative adversarial networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, 2018: 2255–2264.
- [24] KINGMA D P, WELLING M. Auto-encoding variational bayes[J/OL].(2013–12–20) [2024–01–24]. <https://doi.org/10.48550/arXiv.1312.6114>.
- [25] BHATTACHARYYA A, STRAEHLE C N, FRITZ M, et al. Haar wavelet based block autoregressive flows for trajectories[C]//Pattern Recognition: 42nd DAGM German Conference, 2021: 275–288.
- [26] ZHAO H, GAO J, LAN T, et al. Tnt: Target-driven trajectory prediction[C]//Conference on Robot Learning. PMLR, 2021: 895–904.
- [27] LI L, PAGNUCCO M, SONG Y. Graph-based spatial transformer with memory replay for multi-future pedestrian trajectory prediction[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2022: 2231–2241.
- [28] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J/OL]. (2014–12–11) [2024–01–24]. <https://doi.org/10.48550/arXiv.1412.3555>.
- [29] NIKHIL N, TRAN MORRIS B. Convolutional neural network for trajectory prediction[C]//Proceedings of the European Conference on Computer Vision (ECCV) Workshops. ECCV, 2018.
- [30] SONG H, DING W, CHEN Y, et al. Pip: Planning-informed trajectory prediction for autonomous driving[C]//Computer Vision – ECCV 2020: 16th European Conference. ECCV, 2020: 598–614.
- [31] PHAN-MINH T, GRIGORE E C, BOULTON F A, et al. Covernet: Multimodal behavior prediction using trajectory sets [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020: 14074–14083.
- [32] 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. IEEE, 2016: 961–971.
- [33] 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. IEEE, 2018: 1468–1476.

- [34] SCHREIBER M, HOERMANN S, DIETMAYER K. Long-term occupancy grid prediction using recurrent neural networks[C]//2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019: 9299–9305.
- [35] LEE N, CHOI W, VERNAZA P, et al. Desire: Distant future prediction in dynamic scenes with interacting agents [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, 2017: 336–345.
- [36] ZHU Y, REN D, FAN M, et al. Robust trajectory forecasting for multiple intelligent agents in dynamic scene[J/OL]. (2020-05-27) [2024-01-24]. <http://doi.org/10.4855/arXiv.2005.13133>.
- [37] KIM H, KIM D, KIM G, et al. Multi-head attention based probabilistic vehicle trajectory prediction[C]//2020 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2020: 1720–1725.
- [38] MESSAOUD K, YAHIAOUI I, VERROUST-BLONDET A, et al. Attention based vehicle trajectory prediction[J]. IEEE Transactions on Intelligent Vehicles, 2020, 6(1): 175–185.
- [39] LIN L, LI W, BI H, et al. Vehicle trajectory prediction using LSTMs with spatial-temporal attention mechanisms [J]. IEEE Intelligent Transportation Systems Magazine, 2021, 14(2): 197–208.
- [40] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//ICASSP 2021 – 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021.
- [41] ZHAO J, LI X, XUE Q, et al. Spatial-channel transformer network for trajectory prediction on the traffic scenes[J/OL]. (2021-01-27) [2024-01-24]. <https://doi.org/10.48550/arXiv.2101.11472>.
- [42] LIU Y, ZHANG J, FANG L, et al. Multimodal motion prediction with stacked transformers[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2021: 7577–7586.
- [43] LI X, YING X, CHUAH M C. Grip: Graph-based interaction-aware trajectory prediction[C]//2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE, 2019: 3960–3966.
- [44] LI X, YING X, CHUAH M C. Grip++: Enhanced graph-based interaction-aware trajectory prediction for autonomous driving[J/OL]. (2019-07-17) [2024-01-24]. <http://doi.org/10.4855/arXiv:1907.07792>.
- [45] CHANDRA R, GUAN T, PANUGANTI S, et al. Forecasting trajectory and behavior of road-agents using spectral clustering in graph-lstms[J]. IEEE Robotics and Automation Letters, 2020, 5(3): 4882–4890.
- [46] ZHAO Z, FANG H, JIN Z, et al. Gisnet: Graph-based information sharing network for vehicle trajectory prediction [C]//2020 International Joint Conference on Neural Networks (IJCNN). IEEE, 2020: 1–7.
- [47] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph attention networks[J/OL]. (2017-10-30) [2023-12-30]. <http://doi.org/10.48550/arXiv.1710.10903>.
- [48] HUANG Y, BI H, LI Z, et al. Stgat: Modeling spatial-temporal interactions for human trajectory prediction[C]//Proceedings of the IEEE/CVF international conference on computer vision. IEEE, 2019: 6272–6281.
- [49] ZHANG L, SHE Q, GUO P. Stochastic trajectory prediction with social graph network[J/OL]. (2022-05-25) [2024-01-24]. <http://doi.org/10.48550/arXiv:1907.10233>.
- [50] SALZMANN T, IVANOVIC B, CHAKRAVARTY P, et al. Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data[C]//Computer Vision—ECCV 2020: 16th European Conference. ECCV, 2020: 683–700.

(责任编辑 明慧)

【作者简介】

杨荣森(1998—),男,重庆交通大学,硕士研究生,研究方向为自动驾驶系统预测与规划。

E-mail: yrmcqjtu@163.com