

基于变分自动编码器的车辆轨迹预测研究

易虹宇, 杨智宇, 杜力

重庆工商大学 制造装备机构设计与控制重点实验室, 重庆 400067

摘要: 针对轨迹预测中车辆与周边车辆、道路几何之间交互关系建模不充分, 以及车辆轨迹多模态建模不完整等一系列问题, 提出了一种基于变分自动编码器的车辆轨迹预测方法。首先, 通过长短时记忆网络从原始数据中提取轨迹数据与车道信息的语义特征; 其次, 引入多头注意力机制, 采用两个单独的注意力模块分别建立车辆与车辆交互模型及车辆与道路交互模型, 能够更好地反映周边车辆与道路几何对车辆轨迹的交互影响, 得到丰富的场景上下文信息; 接着利用变分自动编码器对车辆轨迹多模态建模, 捕捉轨迹预测的随机性质以生成合理的未来轨迹分布; 最后从分布中多次重复采样以生成多条可能的未来轨迹。通过搭建实验平台和使用 Argoverse 自然驾驶数据集进行测试, 改进后的预测方法在平均位移误差和最终位移误差指标下的数值分别为 1.03 和 1.51, 预测精度上相较于其他 3 种预测方法, 分别提升了 45%、46%、32%; 实验结果表明: 预测方法可以有效地改善车辆与周边车辆、道路几何之间交互关系建模不充分, 以及车辆轨迹多模态建模不完整等问题, 预测精度提高, 总体预测性能良好。

关键词: 轨迹预测; 注意力机制; 轨迹多模态; 变分自动编码器

中图分类号: U463.6; TP391 **文献标识码:** A **doi:** 10.16055/j.issn.1672-058X.2024.0002.008

Research on Vehicle Trajectory Prediction Based on Variational Automatic Encoder

YI Hongyu, YANG Zhiyu, DU Li

Key Laboratory of Manufacturing Equipment Organization Design and Control, Chongqing Technology and Business University, Chongqing 400067, China

Abstract: A vehicle trajectory prediction method based on variational automatic coder was proposed to address a series of problems such as inadequate modeling of the interaction between vehicles and surrounding vehicles and road geometry in trajectory prediction, and incomplete multimodal modeling of vehicle trajectories. Firstly, the semantic features of trajectory data and lane information were extracted from the original data through long and short term memory networks. Secondly, a multi-headed attention mechanism was introduced, and two separate attention modules were used to establish the vehicle-vehicle interaction model and vehicle-road interaction model, respectively, which can better reflect the interaction effects of surrounding vehicles and road geometry on vehicle trajectories and obtain rich scene context information. Next, the vehicle trajectory was modelled multimodally by using the variational automatic coder to capture the random nature of trajectory prediction to generate a reasonable future trajectory distribution. Finally, the sampling was repeated several times from the distribution to generate multiple possible future trajectories. By building the experimental platform and testing with the Argoverse naturalistic driving dataset, the improved prediction method yielded values of 1.03 and 1.51 under the average displacement error and final displacement error indicators, respectively, and the prediction

收稿日期: 2022-12-27 **修回日期:** 2023-02-16 **文章编号:** 1672-058X(2024)02-0060-06

基金项目: 重庆市自然科学基金(CSTC2020JCYJ-MSXMX0803); 重庆工商大学研究生科研创新项目(YJSCXX2021-112-100).

作者简介: 易虹宇(1996—), 男, 四川宜宾人, 硕士研究生, 从事车辆轨迹预测研究。

通讯作者: 杜力(1971—), 女, 重庆人, 教授, 从事机构学、机械式无级变速器等研究。Email: dulicq@126.com.

引用格式: 易虹宇, 杨智宇, 杜力. 基于变分自动编码器的车辆轨迹预测研究[J]. 重庆工商大学学报(自然科学版), 2024, 41(2): 60—65.

YI Hongyu, YANG Zhiyu, DU Li. Research on vehicle trajectory prediction based on variational automatic encoder[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2024, 41(2): 60—65.

accuracy has been improved by 45%, 46%, and 32% compared with the other three prediction methods. The experimental results showed that the prediction method can effectively solve the problems of inadequate modeling of the interaction between vehicles and surrounding vehicles and road geometry, and incomplete multi-modal modeling of vehicle trajectories. The prediction accuracy has been improved and the overall prediction performance is good.

Keywords: trajectory prediction; attention mechanism; multimodal trajectory; variational automatic encoder

1 引言

轨迹预测是自动驾驶系统中最重要的任务之一,准确预测车辆的未来轨迹对于智能汽车高效、安全的规划未来路径起至关重要的作用。轨迹预测的关键在于正确表示车与车之间的交互、车与场景的交互且同时考虑多个可能的未来轨迹,这是由于车辆未来轨迹会受到驾驶场景中周边车辆的影响以及道路几何的约束,同时车辆未来轨迹具备多模态性质,在相同的历史条件下,车辆的未来驾驶行为依旧有多种不同的可能,从而产生不同的未来轨迹。因此,如何有效地利用车辆的历史轨迹数据结合驾驶场景中的交互信息来预测多个可信的未来轨迹成为轨迹预测的核心挑战之一。

早期的轨迹预测方法主要关注交互建模和场景表示,传统方法如社会力模型^[1]使用人工制定的特征和规则来模拟车辆之间的互动以及道路对车辆的约束,然而,车辆之间的隐式交互有时并不能被显式的表达。随着近年来机器学习及其相关技术的快速发展,尤其是深度学习与神经网络的大量应用,推进了轨迹预测领域的发展,一些学者开始利用深度学习^[2]的方法来隐式的建模车辆与车辆之间的交互以及驾驶场景对车辆的约束,通过这样方法来表示驾驶场景中道路几何以及周边车辆对车辆未来轨迹的影响。在此基础上 Kim 等^[3]提出了一种使用递归神经网络(RNN)的车辆轨迹预测方法,在占用网格图上生成车辆未来位置的概率,但是 RNN 网络由于梯度爆炸或消失的问题,无法学习到轨迹数据之间的长期依赖关系,因此无法用于长期轨迹预测;为此,Deo 等^[4]将 RNN 网络的变体长短时记忆网络(LSTM)用于轨迹预测任务以解决长期依赖问题,然而 LSTM 网络存在无法并行计算,计算耗时,同时对交互关系建模并不充分等缺陷;Alahi 等^[5]引入社交池化模块用以更好的建模交互关系;Mercat 等^[6]采用了注意力机制通过并行计算解决了长期依赖的问题。然而这些方法未能有效的整合不同的交互信息,同时只预测了未来多种可能轨迹中的一种,忽略了未来轨迹的多模态性质。

由于车辆未来行为的不确定性,在相同车辆运动

历史下,存在不止一种可能的未来行为。因此,车辆未来轨迹自然的呈现出多模态分布。目前有许多关于多模态建模问题的研究,早期的方法通过学习高斯混合模型的参数来解决多模态回归任务。然而,由于在高维空间中操作时的数值不稳定性,在实践中往往难以训练。为了克服这个问题,研究人员提出变分自动编码器^[7](VAE)和生成式对抗网络^[8](GAN)等深度生成模型,这些深度生成模型使用潜在变量,从数据中学习场景上下文信息,生成合理的未来轨迹分布,最后从该分布中多次采样以生成多个可能的车辆未来轨迹。综上所述,轨迹预测算法的最新趋势是在关注交互建模和场景表示的基础上预测车辆多个可能的未来轨迹。为此,提出一种基于变分自动编码器的轨迹预测方法,引入长短时记忆网络和多头注意力机制。长短时记忆网络有效的提取了轨迹数据时间维度的特征,并通过隐藏状态保留;多头注意力机制能够更好地结合轨迹数据时间维度的特征对车辆与周边车辆、道路几何之间交互关系建模,同时还可以有效的整合这两种不同的交互信息;通过整合后的交互信息,变分自动编码器生成更合理的未来轨迹分布。改进后的轨迹预测方法有效的改善了交互关系建模不充分、车辆轨迹多模态建模不完整等问题,提高了预测精度,总体性能良好。

2 相关工作

2.1 场景表示

场景表示是轨迹预测的一个重要方面。通常,准确的场景表示可以更好地建模车辆与场景的交互关系。场景上下文信息由二维简单鸟瞰图(BEV)表示的地图以及地图上交通参与者的过去轨迹组成,包含了丰富的空间信息。根据对地图信息处理方式的不同,目前有两种类型的场景上下文表示方法:光栅化和矢量化。基于光栅化的方法将高清地图的元素和历史轨迹光栅化为图像,接着采用卷积神经网络(CNN)提取特征并执行坐标预测。IntentNet^[9]开发了一个多任务模型从光栅化地图中提取特征。MultiPath^[10]使用场景 CNN 提取中级特征,这些特征在自上而下的场景表示上编码交通参与者的状态及其交互。然而,这些基于

2D CNN 的方法由于渲染时耗费的 cpu 资源过大导致在提取图结构地图的特征方面效率低下,同时在渲染后会丢失掉车道间的连通性,车道之间的依赖关系等,因此,赵行等^[11]提出矢量化的方法采用向量表示场景上下文,通过分层表示来捕捉道路网络的复杂拓扑以及语义实体的空间位置,矢量化的表示方式使计算效率提高,并保留了地图中更多的信息。

与现有工作不同,为了模拟道路几何对车辆轨迹的约束,基于车辆驾驶方向总是与车道中心线相切的假设,首先将高清地图中每条车道的中心线表示为一系列等距二维坐标点,接着由 LSTM 网络编码,编码后的信息输入注意力模块建立车辆与道路交互模型。这样的处理方式可以计算车道级别的场景上下文信息,使交互关系建模更充分。

2.2 多模态轨迹预测

由于车辆未来轨迹的多模态性质,预测车辆多个可能的未来轨迹是轨迹预测任务中不可或缺的一部分。基于概率热图的方法^[12]专注于输出概率热图以表示车辆轨迹的未来分布。HOME^[13]在输出概率热图的基础上设计了用于优化的确定性采样算法。然而基于概率热图的方法尽管简单灵活,却无法描述变量之间的关系。此外,基于 Transformer^[14]设计了一种基于区域的培训策略,以捕获特定的轨迹模式。DenseTNT^[15]从密集的目标候选轨迹中输出一组轨迹。TNT^[16]首先对车道上密集的候选目标进行采样,并生成以高分目标为条件的轨迹。基于深度生成网络的方法,直接对车辆未来轨迹分布建模,接着从建模的分布中采样固定数量的轨迹序列,并根据其似然性进行排序。这些方法虽然需要更多的样本数量,但是能更加完整的对车辆轨迹的多模态性质建模。

利用变分自动编码器进行多模态轨迹预测任务,根据丰富的场景上下文信息,生成合理的轨迹分布,并采样固定数量的轨迹。与其他工作不同的是,结合了车辆与车辆的交互信息以及车辆与道路的交互信息,为生成更合理的未来轨迹分布提供了更多样的信息。

3 轨迹预测方法

3.1 问题描述

假设在真实驾驶场景中有 N 辆车,目的是生成目标车辆 $\{V_i\}_{i=1}^N$ 的轨迹分布 $p(Y_i | X_i, E_i)$ 。其中 $X_i = P_{ih}(x, y)$ 表示目标车辆 V_i 在 t 时刻过去 h 个时间步的历史位置, $Y_i = P_{if}(x, y)$ 表示目标车辆 V_i 在 t 时刻之

后 f 个时间步的未来位置, E_i 表示与目标车辆 V_i 相关的场景上下文信息。对于 E_i , 使用的是场景中周边车辆的历史位置信息 X_j 以及 t 时刻目标车辆 V_i 所在的车道信息 L_f , 其中 $L_f = (1, 2, \dots, f)$ 表示车道中心线上 f 个等距坐标点。每个位置信息都由车辆当前位置和方向表示。

3.2 预测网络整体结构

如图 1 所示,基于变分自动编码器的轨迹预测网络整体结构,由 3 个不同的模块组成。首先是一个特征提取模块,从轨迹数据和车道信息中提取语义特征,通过 LSTM 最后一个隐藏状态编码,目的是保留轨迹数据时间维度的特征;接着引入了注意力机制,分别由两个不同的注意力模块从社会维度学习复杂的交互关系,包括车辆与周边车辆的交互关系以及车辆与道路的交互关系,同时对其数量和排序保持不变,输出对应的特征向量;接着将两种不同的特征向量结合经 LSTM 网络编码后的车辆历史位置向量输入先验网络,未来位置编码输入编码器,通过潜变量 Z 以及解码器生成符合场景上下文的轨迹分布。

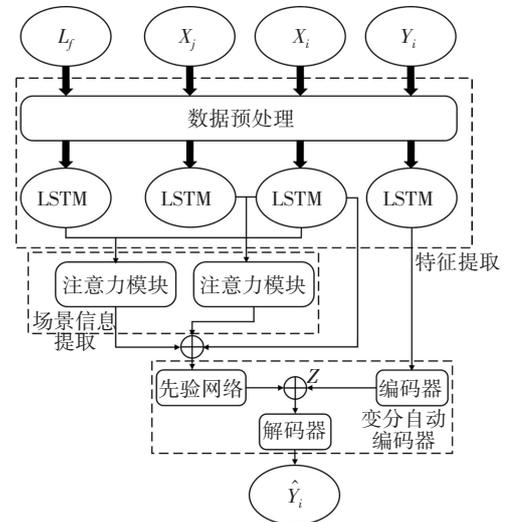


图 1 预测网络整体框架

Fig. 1 Overall framework of the prediction network

3.2.1 特征提取

在真实驾驶场景中,沿着车道行驶的车辆总是尝试与车道的切线向量平行,基于这样的假设,首先对车辆的位置信息和车道信息做对应的数据预处理,车辆的位置数据通过公式计算每个时间步的速度和方向:

$$v = \alpha \sqrt{(p_x^t - p_x^{t-1})^2 + (p_y^t - p_y^{t-1})^2}$$

$$h = \tan^{-1} \left(\frac{p_y^t - p_y^{t-1}}{p_x^t - p_x^{t-1}} \right)$$

其中, p_x^t, p_y^t 表示车辆在 t 时刻的位置, α 表示采样频率。

此外, 车道信息需要计算车道的切线向量 τ_f 和切线向量的方向 d_{τ_f} , 通过公式计算:

$$\tau_f = l_f - l_{f-1}$$

$$d_{\tau_f} = \tan^{-1} \left(\frac{\tau_f, y}{\tau_f, x} \right)$$

其中, l_f 表示 t 时刻车道中心线上第 f 个坐标点。

当数据经过预处理后, 接着通过不同的 LSTM 网络对车辆的位置信息 Y_i, X_i, X_j 以及车道信息 L_f 编码, 并提取相应时间维度的特征信息, LSTM 网络通过上一个时刻的隐藏状态, 和当前时刻的输入来计算当前时刻的隐藏状态, 并向前传播。隐藏状态的计算公式如下:

$$h_t = f(Uh_{t-1} + Wx_t + b)$$

其中, 0 时刻的 $h_0 = 0$, U, W, b 为神经网络的可学习参数, 由网络自行训练, f 表示激活函数, 现采用 ReLU 函数。

因为隐藏状态将每个时刻的输入数据信息向前传播, 最后一个时刻的隐藏状态包含所有的输入数据信息, 所以保留最后一个隐藏状态的向量作为编码结果。

3.2.2 场景信息提取

模块引用多头注意力机制^[17], 学习车辆与驾驶场景中交通参与者的复杂交互。注意力机制是通过一个打分函数来计算输入向量 $\mathbf{X} = [x_1, x_2, \dots, x_n]$ 与查询向量 q 的相关性, 查询向量 q 是一组与特定任务有关的向量, 在文中查询向量为车辆的历史位置信息 X_i 经 LSTM 网络编码后的隐藏状态 X'_i , 输入向量为 X_j 与 L_f 编码后的隐藏状态。使用了两个注意力模块, 分别计算车辆历史位置信息 X_i 与周边车辆的历史位置信息 X_j 以及车道信息 L_f 的相关性(即目标车辆与周边车辆的交互关系以及目标车辆与道路的交互关系)。注意力函数如下:

$$s(\mathbf{x}, \mathbf{q}) = \mathbf{x}^T \mathbf{q}$$

$$\alpha_n = \frac{\exp(s(\mathbf{x}, \mathbf{q}))}{\sum_{n=1}^N \exp(s(\mathbf{x}, \mathbf{q}))}$$

$$att(\mathbf{X}, \mathbf{q}) = \sum_{n=1}^N \alpha_n \mathbf{x}_n$$

其中, \mathbf{X} 为输入向量, $s(\mathbf{x}, \mathbf{q})$ 为打分函数。

多头注意力是将 \mathbf{q} 分为多个查询向量, $\mathbf{q} = [q_1, q_2, \dots, q_n]$, 默认是采用 8 头, 采取并行的方式从输入信息中选取多组相关信息, 每个头部关注输入信息的不同

部分。公式如下:

$$att(\mathbf{X}, \mathbf{q}) = att(\mathbf{X}, q_1) \oplus \dots \oplus att(\mathbf{X}, q_n)$$

3.2.3 变分自动编码器

假设轨迹数据服从一个未知的轨迹分布 $p(Y_i | X_i, E_i)$, 由于轨迹数据之间存在复杂的依赖关系, 直接建模车辆未来轨迹分布显得非常困难, 所以引入了一个不可观测的隐变量 z 来简化模型, 这样就将问题转换为两个局部的条件概率 $p(Y_i | z, X_i, E_i)$ 和 $p(z | Y_i, X_i, E_i)$ 。

通过编码器网络生成隐变量 z 的分布 $p(z | Y_i, X_i, E_i)$, 假设 $p(z | Y_i, X_i, E_i)$ 服从高斯分布, 即 $p(z | Y_i, X_i, E_i) = N(z; \mu_b, \sigma_b^2)$ 。其中 μ_b, σ_b^2 是高斯分布的均值和方差, 通过下方的公式采用一个多层感知机来估计均值和方差。

$$\mu_b, \sigma_b = MLP([Y'_i; X'_i; E'_i])$$

$Y'_i; X'_i; E'_i$ 表示对应数据经前两个模块编码后的输出, 分号表示串联操作。

由于未来位置在预测阶段是不可知的, 所以加入一个先验网络, 仅利用车辆的历史位置和场景上下文信息生成一个同样服从高斯分布先验概率分布 $p_x(z | X_i, E_i) = N(z; \mu_x, \sigma_x^2)$ 。其中均值和方差通过下方的公式进行估计:

$$\mu_x, \sigma_x = MLP_{p_x}([X'_i; E'_i])$$

隐变量 z 在模型训练期间利用重构参数的方法从 (μ_b, σ_b) 中采样, 在预测阶段从 (μ_x, σ_x) 中采样。

解码器网络通过得到的隐变量 z 来得到条件概率 $p(Y_i | z, X_i, E_i)$, $p(Y_i | z, X_i, E_i)$ 同样服从高斯分布:

$$p(Y_i | z, X_i, E_i) = N(z; \mu_j, \sigma_j^2)$$

接着根据条件概率 $p(Y_i | z, X_i, E_i)$, 采样生成估计位置 \hat{y}_i^t 。与一般的编码器-解码器网络不同的是, 变分自动编码器的输出为概率分布的参数, 而不是确定的编码。

最后通过 LSTM 网络来生成未来轨迹预测, 如下所示:

$$q_i^t = MLP_q(\hat{y}_i^t)$$

$$h_i^{t+1} = LSTM([q_i^t; E_i; z], h_i^t)$$

$$\hat{y}_i^{t+1} = MLP_f(h_i^{t+1})$$

在 $t=0$ 时刻时, 将位置估计 \hat{y}_i^0 和 LSTM 网络的隐藏向量 h_i^0 分别初始化为最后观测的位置和零向量。最终通过多次随机采样生成 K 条可能的未来轨迹, K 为预先设定好的常数。

3.3 损失函数

本文是基于变分自动编码器模型通过隐变量生成轨迹分布进行多模态轨迹预测任务。因为轨迹分布的部分变量是不可观测的,所以引入一个隐变量 z ,将轨迹分布的联合概率密度函数 $p(Y_i|X_i, E_i)$ 通过以下公式分解:

$$p(Y_i|X_i, E_i) = p(Y_i|z, X_i, E_i)p(z|Y_i, X_i, E_i)$$

其中, $p(z|Y_i, X_i, E_i)$ 为潜在变量 z 先验分布的概率密度函数, $p(Y_i|z, X_i, E_i)$ 为已知 z 时 Y_i 的条件概率密度函数,均服从正态分布,所以通过最大化似然来估计均值和方差,通过优化以下改进 ELBO^[18] 目标:

$$L_{\text{ELBO}} = -E_{z \sim q_b} [\log p(Y_i|z, X_i, E_i)] + KL(p(z|Y_i, X_i, E_i) || p(z|X_i, E_i))$$

4 实验结果分析

4.1 数据集

Argoverse^[19] 是一个公开的大型真实数据集,包含专用于轨迹预测任务的数据子集 Argoverse Forecasting。数据集提供了训练、验证和测试子集,包含超过 30 万个驾驶场景,每个场景的时间长度为 5 s,观测时间窗口长度为 2 s,预测窗口长度为 3 s,采样频率 10 Hz。主要的目标是通过车辆过去 2 s 的轨迹预测未来 3 s 的轨迹。

4.2 评价指标

采用轨迹预测任务中最常用的两种度量,包括平均位移误差(ADE)和最终位移误差(FDE),前者关注所有预测时间步中预测结果与真实结果的距离,后者仅关注最终位置和真实结果的距离,两者定义如下:

$$ADE(\hat{Y}, Y) = \frac{1}{T} \sum_{t=1}^T ||\hat{p}^t - p^t||_2$$

$$FDE = ||\hat{p}^T - p^T||_2$$

其中, \hat{Y}, Y 分别表示预测轨迹和真实轨迹, \hat{p}^t, p^t 分别表示预测位置的坐标和真实坐标,通过计算预测位置与真实位置之间的误差来判断预测精度。在其余部分 ADE_K 与 FDE_K 表示 K 个生成轨迹中 ADE 与 FDE 的最小值。

4.3 网络参数设置

搭建实验平台的具体参数如下,预测算法操作系统为 Linux18.04,显卡 GPU 是 NVIDIA GeForce RTX 2060,基于 1.10.0 版本的深度学习框架 Pytorch,使用 cuda11.0 作为运算平台。一个自适应矩估计优化器^[20] (Adam)在模型训练过程用于优化,初始学习率为 $5 * 10^{-4}$,batch size 为 8。此外。使用训练数据集中的验证样本本来评估每 4 轮训练后的预测性能,当预测性能高

于之前时,保存模型的网络参数。

4.4 实验结果分析

为了检验多模态轨迹预测对预测精度的影响,根据预先设定好不同的 K 值采样以生成对应数量的预测轨迹,并计算平均位移误差和最终位移误差进行对比。如表 1 所示,当 K 为 1 时输出单条轨迹,等同于单模态的轨迹预测,此时平均位移误差和最终位移误差的值较大,预测精度低,这可能是由于没有考虑车辆未来行为的不确定性。随着采样次数的增加,输出更多的预测轨迹,预测性能有明显提高。

表 1 不同 K 值下预测精度对比

Table 1 Comparison of prediction accuracies under different K values

K	1	5	6	10	12
ADE	1.61	1.11	1.03	0.96	0.86
FDE	3.26	1.63	1.51	1.37	1.14

同时,将轨迹预测方法与 3 种不同的方法相比较,结果如表 2 所示。轨迹预测方法在大多数指标上都是优于现存的一些方法,预测方法相较于 Argoverse 数据集的基准方法在预测精度上提升了 45%,相较于 DESIRE 和 R2P2 的预测方法分别提升了 32%与 46%。

表 2 不同轨迹预测方法预测精度对比

Table 2 Comparison of prediction accuracies of different trajectory prediction methods

预测方法	ADE ₁ /FDE ₁	ADE ₅ /FDE ₅	ADE ₆ /FDE ₆	ADE ₁₂ /FDE ₁₂
Argoverse baseline	2.92/6.45	—	2.08/4.19	—
DESIRE ^[21]	2.38/4.64	1.17/2.06	1.09/1.89	0.9/1.45
R2P2 ^[22]	3.02/5.41	1.49/2.54	1.40/2.35	1.11/1.77
本文方法	1.61/3.26	1.11/1.63	1.03/1.51	0.86/1.14

5 结束语

准确预测车辆未来轨迹是实现安全且高效规划的必要条件,因此也是自动驾驶技术的关键部分。提出了一种基于变分自动编码器的轨迹预测方法,结合多头注意力机制以及 LSTM 网络预测未来轨迹,是一种编码器-解码器结构。方法考虑了驾驶场景中的交互,利用场景上下文信息生成合理的轨迹分布,并在 Argoverse 数据集上进行了实验验证。实验结果表明:轨迹预测方法与其他现有的方法相比,获得了具有竞争力的结果,相较于其他方法预测精度分别提升了 45%、32%以及 46%;同时多模态轨迹预测相较于输出单条轨迹,在预测精度上有明显的提升,但本文选用的采样方法为随机多次重复采样,一

一定程度上忽略了样本的多样性,如何优化采样方法,这将是未来的工作方向。

参考文献(References):

- [1] HELBING D, MOLNAR P. Social force model for pedestrian dynamics[J]. *Physical Review E*, 1998, 51(5): 4282—4286.
- [2] MOZAFFARI S, AL-JARRAH O, DIANATI M, et al. Deep learning-based vehicle behavior prediction for autonomous driving applications: A review [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 5(2): 1—15.
- [3] KIM B, KANG C M, KIM J. Probabilistic vehicle trajectory prediction over occupancy grid map via recurrent neural network[C]//*Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017: 399—404.
- [4] DEO N, TRIVEDI M. Multi-modal trajectory prediction of surrounding vehicles with maneuver based LSTMs [C]//*Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018: 1179—1184.
- [5] ALAHI A, GOEL K, RAMANATHAN V, et al. Social LSTM: Human trajectory prediction in crowded spaces [C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016: 27—30.
- [6] MERCAT J, GILLES T, EL ZOGHBY N, et al. Multi-head attention for multi-modal joint vehicle motion forecasting[C]//*2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020: 9638—9644.
- [7] MIGUEL M, ARMINGOL J, GARCL F. Vehicles trajectory prediction using recurrent vae network [J]. *IEEE Access*, 2022, 10(4): 32742—32749.
- [8] ZE L, XIAOCI H, WENGUAN C. An improved GAN with transformers for pedestrian trajectory prediction models [J]. *International journal of intelligent systems*, 2022, 37(8): 4417—4436.
- [9] CASAS S, LUO W, URTASUN R. Intentnet: Learning to predict intention from raw sensor data [C]// *Proceedings of the Conference on Robot Learning*, 2018: 947—956.
- [10] CHAI Y, SAPP B, BANSAL M, et al. MultiPath: Multiple probabilistic anchor trajectory hypotheses for behavior prediction [C]//*Proceedings of Machine Learning Research (MLR)*, 2020: 86—99.
- [11] GAO J, SUN C, ZHAO H, et al. VectorNet: Encoding HD maps and agent dynamics from vectorized representation[C]//*proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020: 13—19.
- [12] GILLES T, SABATINI S, TSISHKOU D, et al. GOHOME: Graph-oriented heatmap output for future motion estimation [C]//*Proceedings of the 2022 International Conference on Robotics and Automation (ICRA)*, 2022: 23—27.
- [13] GILLES T, SABATINI S, TSISHKOU D, et al. HOME: Heatmap output for future motion estimation[C]//*Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021: 19—22.
- [14] LIU Y, ZHANG J, FANG L, et al. Multimodal motion prediction with stacked transformers [C]//*Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021: 20—25.
- [15] GU J, SUN C, ZHAO H. DenseTNT: End-to-end trajectory prediction from dense goal sets [C]//*Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021: 10—17.
- [16] ZHAO H, GAO J, LAN T, et al. TNT: Target-driveN trajectory prediction [C]// *Proceedings of the Conference on Robot Learning (CRL)*, 2020: 16—22.
- [17] SEONGJIN C, JIWON K, HWASOO Yeo. Attention-based recurrent neural network for urban vehicle trajectory prediction [J]. *Procedia Computer Science Systems*, 2019, 151(2): 327—334.
- [18] LI L, YAO J, WENLIANG L, et al. Grin: Generative relation and intention network for multi-agent trajectory prediction[J]. *Advances in Neural Information Processing Systems*, 2021, 34(6): 27107—27118.
- [19] CHANG M F, LAMBERT J, SANGKLOY P, et al. Argoverse: 3D tracking and forecasting with rich maps [C]//*Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019: 15—20.
- [20] KUDARI J M, JEBAKUMARI A, Kumar S, et al. Image classifier using the Adam optimizer and the Relu activation function [J]. *International Journal of Advanced Research in Engineering & Technology*, 2021, 12(3): 56—60.
- [21] LEE N, CHOI W, VERNAZA P, et al. DESIRE: Distant future prediction in dynamic scenes with interacting agents [C]//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017: 2165—2174.
- [22] RHINEHART N, KITANI K M, VERNAZA P. r2p2: A Reparameterized pushforward policy for diverse, precise generative path forecasting [C]//*Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 794—81.