

文章编号:1001-7372(2019)06-0034-09

基于 LSTM 网络的驾驶意图识别及 车辆轨迹预测

季学武¹, 费 聪¹, 何祥坤², 刘玉龙¹, 刘亚辉¹

(1. 清华大学 汽车安全与节能国家重点实验室, 北京 100084;

2. 华为技术有限公司 诺亚方舟实验室, 北京 100085)

摘要:自动驾驶汽车需具备预测周围车辆轨迹的能力,以便做出合理的决策规划,提高行驶安全性和乘坐舒适性。运用深度学习方法,设计了一种基于长短时记忆(LSTM)网络的驾驶意图识别及车辆轨迹预测模型,该模型由意图识别模块和轨迹输出模块组成。意图识别模块负责识别驾驶意图,其利用 Softmax 函数计算出驾驶意图分别为向左换道、直线行驶、向右换道的概率;轨迹输出模块由编码器-解码器结构和混合密度网络(MDN)层组成,其中的编码器将历史轨迹信息编码为上下文向量,解码器结合上下文向量和已识别的驾驶意图信息预测未来轨迹;引入 MDN 层的目的是利用概率分布来表示车辆未来位置,而非仅仅预测一条确定的轨迹,以提高预测结果的可靠性和模型的鲁棒性。此外,将被预测车辆及其周围车辆组成的整体视为研究对象,使模型能够理解车-车间的交互式行为,响应交通环境的变化,动态地预测车辆位置。使用基于真实路况信息的 NGSIM (Next Generation SIMulation)数据集对模型进行训练、验证与测试。研究结果表明:与传统的基于模型的方法相比,基于 LSTM 网络的轨迹预测方法在预测长时域轨迹上具有明显的优势,考虑交互式信息的意图识别模块具备更高的预判性和准确率,且基于意图识别的轨迹预测能降低预测轨迹与真实轨迹间的均方根误差,显著提高轨迹预测精度。

关键词:汽车工程;轨迹预测;意图识别;LSTM;交互式行为

中图分类号:U461.91 文献标志码:A DOI:10.19721/j.cnki.1001-7372.2019.06.003

Intention Recognition and Trajectory Prediction for Vehicles Using LSTM Network

JI Xue-wu¹, FEI Cong¹, HE Xiang-kun², LIU Yu-long¹, LIU Ya-hui¹

(1. State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing 100084, China;

2. Noah's Ark Lab, Huawei Technologies, Beijing 100085, China)

Abstract: Autonomous vehicles often need to predict the trajectories of surrounding vehicles for planning and decision making. In this paper, a model for intention recognition and trajectory prediction based on long short-term memory (LSTM) network is proposed. The proposed model comprises an intention recognition module and a trajectory output module. The intention recognition module was employed for identifying the driving intention. The Softmax function was incorporated in the intention recognition module for calculating the probabilities of left lane

收稿日期:2019-03-26

基金项目:国家自然科学基金中国汽车产业创新发展联合基金项目(U1664263);国家自然科学基金项目(51875302);
国家重点研发计划项目(2018YFB1600501)

作者简介:季学武(1964-),男,安徽芜湖人,副教授,博士研究生导师,工学博士,E-mail:jixw@tsinghua.edu.cn.

通讯作者:刘亚辉(1980-),男,河北石家庄人,副教授,博士研究生导师,工学博士,E-mail:liuyahui@tsinghua.edu.cn.

change, lane-keeping, and right lane change. An encoder-decoder structure and a mixture density network (MDN) layer were included in the trajectory output module. The encoder converted the past trajectory information into the context vector. Subsequently, the decoder combined the context vector and the intention recognition information for predicting future trajectories. The MDN layer was employed for representing the future position of a vehicle with its probability distribution rather than with a particular trajectory, which improved the reliability of prediction results and the robustness of the proposed model. Additionally, the compositions of a predicted vehicle and its surroundings were both taken into account, which aided the proposed model in analyzing the interactions among vehicles. Hence, the proposed model can dynamically predict vehicle trajectories according to variations in traffic conditions. The NGSIM data set based on the information of actual road conditions was employed for training, validating, and testing the proposed model. Experimental results indicate that the proposed method based on LSTM network has several advantages over conventional model-based methods with respect to trajectory prediction, especially in a long prediction horizon. Interactive information can ensure that the intention recognition module has high anticipative ability and accuracy. Furthermore, trajectory prediction based on intention recognition can significantly reduce the root-mean-square errors in predicted trajectories with respect to the ground truth, thereby leading to significant improvement in trajectory prediction accuracy.

Key words: automotive engineering; trajectory prediction; intention recognition; long short-term memory; interaction behavior

0 引言

自动驾驶汽车需要具备预测交通场景中周围车辆运动趋势的能力,从而规避危险场景,做出最佳的决策,以保证获得安全、舒适的驾乘体验。同时,轨迹预测也有利于先进驾驶辅助系统的优化,推动更高级别自动驾驶技术的发展。然而,由于道路交通场景的复杂性以及驾驶人驾驶风格的多样性,使得轨迹预测成为一项极具挑战的任务。

轨迹预测方法可以分为2大类:基于模型的轨迹预测和基于数据驱动的轨迹预测^[1]。许多学者采用建立运动学或动力学模型的方法预测车辆轨迹^[2-5]。Barth等^[4]运用卡尔曼滤波和蒙特卡罗方法,在车辆模型的短距离不确定性预测中取得了较高的精度。Houenou等^[3]提出了一种基于恒定横摆角速度和加速度运动模型的预测和多项式轨迹规划相结合的方法,以弥补基于模型方法在长时域预测上的不足。然而,由于车辆轨迹的高度非线性以及驾驶人驾驶习惯的多样性,使得基于模型的轨迹预测方法在长时域预测上难以取得令人满意的精度。

近年来,基于数据驱动的轨迹预测方法受到越来越多学者的关注。与许多机器学习应用一样,该方法采用分类算法和回归算法2种思路处理轨迹预

测问题。分类算法旨在识别出驾驶人的意图,一般将轨迹划分为向左换道、向右换道和直线行驶等类别。刘志强等^[6]采用隐马尔可夫模型和支持向量机的方法识别驾驶意图;也有许多学者致力于采用循环神经网络进行意图识别^[7-9]。然而,分类算法最主要的意义在于能提前预判驾驶人换道意图,但无法得到更明确的轨迹信息;回归算法则将轨迹预测表述为一个回归问题,输入历史轨迹并直接获得未来轨迹。李建平等^[10]提出隐马尔可夫模型和高斯混合模型预测车辆轨迹。此外,动态贝叶斯网络^[11]和随机森林^[12]方法也被广泛用以处理此类问题。回归算法的缺陷是当驾驶意图并不明确时,其输出的结果倾向于不同意图轨迹的平均值,这显然不符合实际场景。且无论是基于模型的预测还是基于数据驱动的预测,交通场景的复杂性和驾驶人驾驶风格的多样性,使得轨迹预测中充满不确定性,所以对运动轨迹的预测也应该是概率而非确定性的。

随着长短期记忆(Long Short-term Memory, LSTM)网络在语音识别、机器翻译等领域取得一系列重大突破,其在应对时序问题时展现出强大的信息挖掘能力和深度表征能力。轨迹序列亦是一种时间序列,因此使用LSTM网络处理轨迹预测问题受到越来越多学者的关注。Duan等^[13-14]利用LSTM

成功实现了行人轨迹的预测问题;不少学者尝试将 LSTM 网络应用于车辆轨迹预测^[15-16]。受此启发,本文使用 LSTM 网络分别构建意图识别模块和轨迹输出模块,其中轨迹输出模块采用编码器-解码器结构,使其能够读取和生成任意长度的轨迹序列;并在此基础上进行改进,使模型能在意图识别的基础上预测轨迹。本文主要研究内容如下:

(1)基于意图识别的预测:模型分为意图识别模块和轨迹输出模块。意图识别模块采用 LSTM 网络,由 Sotfmax 层输出对驾驶意图的识别向量,同时设置合适的阈值,避免模型始终做出过于“保守”的预测;轨迹输出模块使用 2 个 LSTM 网络分别构建编码器和解码器。编码器对历史轨迹进行编码,编码向量结合驾驶意图信息共同输入解码器,使解码器能做出基于意图识别的预测。

(2)考虑交互信息:现实场景中相关车辆组成一个相互依赖的整体,各自的行为影响彼此的决策。本文提出的模型考虑车辆间交互作用对车辆运动的影响,将被预测车辆信息与其周围车辆信息作为模型的输入,使模型从车辆间交互式信息中学习到多车行为的依赖关系,响应交通环境变化,动态地调整预测的轨迹。

(3)轨迹概率分布:现实交通场景中各车辆的行为具有多样性,输出确定性的轨迹往往是不可靠的,甚至会误导决策系统做出错误的规划。因此,本文通过混合密度网络(Mixture Density Networks, MDN)层将概率分布中的参数作为全连接网络的

输出,通过训练确定二元高斯分布的参数,描述预测轨迹的概率分布,并体现轨迹预测的不确定性。

1 模型框架

本文提出的预测模型由意图识别模块和轨迹输出模块组成,如图 1 所示。基于 LSTM 网络构建意图识别模块,将车辆的历史轨迹信息作为输入,利用 Softmax 函数计算出驾驶意图分别为向左换道、直线行驶、向右换道的概率。设 I 为模型的输入量; $C=(c_1, c_2, c_3)$,为意图识别模块输出的意图类别向量, c_1, c_2, c_3 分别代表向左换道、直线行驶、向右换道 3 种意图类别; Ω 为各个意图类别概率组成的向量, $\omega_i (i=1, 2, 3)$ 分别代表向左换道、直线行驶、向右换道的概率,则意图识别模块的输出为

$$\omega_i = P(c_i | I), \Omega = (\omega_1, \omega_2, \omega_3) \quad (1)$$

轨迹输出模块由编码器、解码器和 MDN 层组成,编码器将输入的历史轨迹信息编码成上下文向量(Context),上下文向量与意图识别模块输出的意图识别向量 C 相结合,构成轨迹编码向量 r 作为解码器的输入。解码器与 MDN 层相连,使模型在基于意图识别的基础上,预测未来轨迹的概率分布

$$P(O|I) = \sum P_{\pi, \mu, \sigma}(O_i | c_i, I) P(c_i | I) \quad (2)$$

式中: O 为模型输出量; $P_{\pi, \mu, \sigma}(O_i | c_i, I)$ 为解码器输出的轨迹预测信息经由 MDN 层处理后,得到的基于不同驾驶意图的轨迹概率分布。

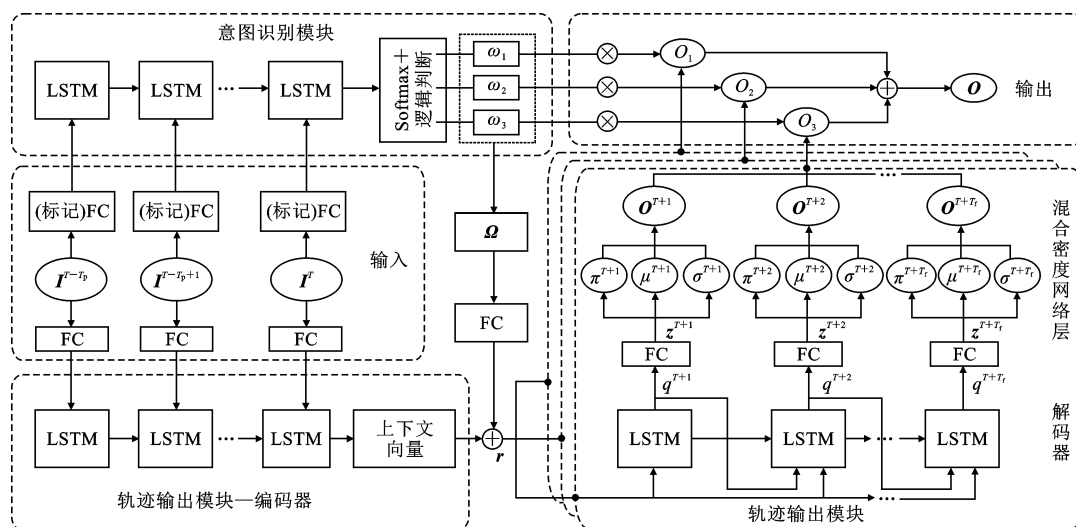


图 1 模型整体结构

Fig. 1 Overall Structure of Model

现实场景中,车辆的行为是与周围车辆交互作用的结果。为了让模型理解这种交互式行为,输入

信息包括被预测车辆历史轨迹信息及环境信息,输入量表示为

$$\mathbf{I}^{(t)} = [\mathbf{S}_e^{(t)}, \mathbf{E}^{(t)}] \quad t = (T - T_p, \dots, T - 1, T) \quad (3)$$

式中: $\mathbf{S}_e^{(t)}$ 为被预测车辆的历史信息; $\mathbf{E}^{(t)}$ 为环境信息; T_p 为历史时域(反映输入轨迹的长度)。

具体而言,被预测车辆的状态信息包括 $\mathbf{S}_e^{(t)} = (x^{(t)}, y^{(t)}, v_e^{(t)})$, 其中, $x^{(t)}$ 为被预测车辆侧向坐标; $y^{(t)}$ 为被预测车辆纵向坐标; $v_e^{(t)}$ 为被预测车辆的绝对速度。

环境信息由被预测车辆左前、正前、右前、左后、正后、右后邻近车辆(分别用 1, 2, 3, 4, 5, 6 表示, 位置关系如图 2 所示)的历史轨迹信息和被预测车辆的 2 个标志位组成, $\mathbf{E}^{(t)} = (\mathbf{S}_{s1}^{(t)}, \mathbf{S}_{s2}^{(t)}, \mathbf{S}_{s3}^{(t)}, \mathbf{S}_{s4}^{(t)}, \mathbf{S}_{s5}^{(t)}, \mathbf{S}_{s6}^{(t)}, \mathbf{M}_r^{(t)}, \mathbf{M}_l^{(t)})$ 。周车状态信息又包括其位置和速度信息, $\mathbf{S}_{si}^{(t)} = (\Delta x_i^{(t)}, \Delta y_i^{(t)}, v_i^{(t)})$, 其中, $\Delta x_i^{(t)}$ 为 i 位置车辆与被预测车辆的侧向相对距离; $\Delta y_i^{(t)}$ 为 i 位置车辆与被预测车辆的纵向相对距离; $v_i^{(t)}$ 为 i 位置车辆的绝对速度; $\mathbf{M}_r^{(t)}$ 为右车道标志位(若被预测车辆存在右车道, 则为 1, 否则为 0); $\mathbf{M}_l^{(t)}$ 为左车道标志位(若被预测车辆存在左车道, 则为 1, 否则为 0)。

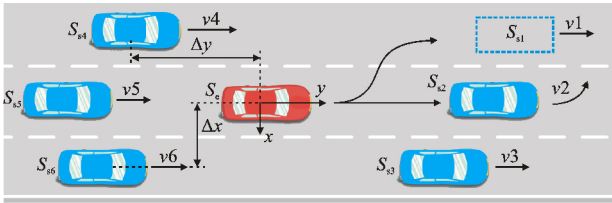


图 2 被预测车辆及其周围车辆标识示意

Fig. 2 Schematic of Predicted Vehicle and Its Surroundings

如果规定范围内 i 位置车辆不存在, 则将该位置车辆信息设置为 $\Delta y_i^{(t)} = +\infty$, $v_i^{(t)} = v_e^{(t)}$ 。经由 MDN 层输出的是预测位置的概率分布, 令 T_f 表示预测时域, \mathbf{X}, \mathbf{Y} 分别表示侧向和纵向坐标序列, 则输出量表示为

$$\mathbf{O}^{(t)} = [\mathbf{X}^{(t)}, \mathbf{Y}^{(t)}] \quad t = (T + 1, T + 2, \dots, T + T_f) \quad (4)$$

2 意图识别模块

意图识别模块需要根据交互式信息理解车辆运行规律, 并对车辆的行驶意图进行预判, 其结构框架如图 1 所示。整个模块基于 LSTM 网络搭建。输入量 $\mathbf{I}^{(t)}$ 通过具有 128 个神经元, 激活函数为 ReLU 的全连接层(Fully Connected Layers, FC)。经处理后的量随即输入 LSTM 循环体。在每一时刻 t , LSTM 单元体读取当前时刻的输入 $\mathbf{I}^{(t)}$ 和上一时刻以前历史轨迹信息的隐藏状态 $\mathbf{h}^{(t-1)}$, 以此更新当前时刻的隐藏状态 $\mathbf{h}^{(t)}$, 即 $\mathbf{h}^{(t)} = f(\mathbf{h}^{(t-1)}, \mathbf{I}^{(t)})$ 。LSTM 通过这种方式学习历史轨迹序列中的规律,

经过 Softmax 层后, 输出概率矩阵 $\boldsymbol{\Omega} = (\omega_1, \omega_2, \omega_3)$, 其中 $\omega_1, \omega_2, \omega_3$ 即驾驶意图分别为左换道、直行和右换道 3 种类别的概率 $P(c_i | \mathbf{I})$ 。然而, 若意图识别模块始终输出 3 种类别的概率, 模型将会始终预测出 2 类甚至 3 类轨迹分布, 这样的预测显然不够“自信”。为了解决这一矛盾, 需要对输出的概率做一逻辑判断: 规定向左、向右换道的确信阈值为 80%, 直线行驶的确信阈值为 70%, 当某一类型的假设意图大于对应的确信阈值, 则认定这就是正确的类型, 故将该类别的概率调整为 100%, 其余两类别的概率为 0, 此时 $\boldsymbol{\Omega}$ 成为 One-hot 向量; 除此之外输出三类别原有的概率。

为了避免出现信息瓶颈和过拟合现象, 选择 LSTM 单元体的隐藏单元个数为 128, 采用 4 个循环层堆叠, 构建深层循环神经网络结构。不同层循环体之间的 Dropout 比率为 0.2。由于意图识别模块是一个分类器, 数据预处理环节需要对输入量 $\mathbf{I}^{(t)}$ 附上相应的类别标记 $\mathbf{C} = \{\text{向左换道, 直线行驶, 向右换道}\}$ 。具体操作于 4.1 节介绍。意图识别模块选用分类交叉熵作为损失函数, 采用 Adam 优化器, 学习率 $\alpha = 0.0005$, 衰减率为 0.9。

3 轨迹输出模块

3.1 编码器与解码器

轨迹输出模块采用编码器-解码器结构^[17], 由 2 个独立的 LSTM 网络循环体组成(图 1)。编码器负责将被预测车辆的历史轨迹信息及其周围车辆的历史轨迹信息编码为一个固定长度的上下文向量(Context), 该向量包含了编码器对历史轨迹特征的理解与记忆, 解码器负责从这个上下文向量中提取重要的信息, 预测车辆未来的位置。

为了使解码器输出基于意图识别的轨迹预测 $P(\mathbf{O} | \mathbf{I}) = \sum P_{\pi, \mu, \sigma}(\mathbf{O}_i | c_i, \mathbf{I}) P(c_i | \mathbf{I})$, 将编码器输出的上下文向量与意图识别模块输出的类别向量 $\boldsymbol{\Omega}$ (经过全连接层后) 相结合, 得到轨迹编码向量 \mathbf{r} 。解码器根据上一时刻输出向量 $\mathbf{q}^{(t-1)}$ 、上一时刻的隐藏状态 $\mathbf{h}^{(t-1)}$ 和轨迹编码向量 \mathbf{r} , 更新当前时刻的隐藏状态 $\mathbf{h}^{(t)}$, 即 $\mathbf{h}^{(t)} = f(\mathbf{h}^{(t-1)}, \mathbf{q}^{(t-1)}, \mathbf{r})$ 。最后依据当前时刻的隐藏状态 $\mathbf{h}^{(t)}$ 、解码器上一时刻输出的向量 $\mathbf{q}^{(t-1)}$ 和轨迹编码向量 \mathbf{r} , 更新当前时刻的解码器输出向量 $\mathbf{q}^{(t)} = g(\mathbf{h}^{(t)}, \mathbf{q}^{(t-1)}, \mathbf{r})$ 。解码器每一时刻输出的轨迹坐标向量 \mathbf{q} , 需经过一个全连接层(FC)调整维度得到向量 \mathbf{z} , 使 \mathbf{z} 的维度与 MDN 匹配。经过以上处理, 式(2)可写成

$$P(\mathbf{O}|\mathbf{I}) = \sum \omega_i P_{\pi, \mu, \sigma}(\mathbf{O}_i | \mathbf{z}) \quad (5)$$

3.2 混合密度网络层

MDN 模型最早由 Christopher Bishop 于 1990 年代提出^[18], 该方法能预测输出的整体概率分布, 并能计算预测结果的置信因子。其核心思想是, 对于每一个输入, 用一组核函数的线性组合去逼近输出, 利用神经网络的框架表达核函数的分布参数, 并进行迭代优化。本文选择 K 个高斯函数的组合作为 MDN 的核函数, 则 MDN 层输出的轨迹分布概率为

$$P_{\pi, \mu, \sigma}(\mathbf{O}_i | \mathbf{z}) = \sum_{k=1}^K \pi_k(\mathbf{z}) \Phi[\mathbf{O}_i | \mu_k(\mathbf{z}), \Sigma_k(\mathbf{z})] \quad (6)$$

$$\sum_{k=1}^K \pi_k(\mathbf{z}) = 1$$

$$\Phi[\mathbf{O}_i | \mu_k(\mathbf{z}), \Sigma_k(\mathbf{z})] = \frac{1}{(2\pi)^{n/2} |\Sigma_k(\mathbf{z})|^{1/2}} \cdot \exp\left\{-\frac{1}{2}[\mathbf{O}_i - \mu_k(\mathbf{z})]^T \Sigma_k^{-1}(\mathbf{z}) [\mathbf{O}_i - \mu_k(\mathbf{z})]\right\} \quad (7)$$

$$\mu_k(\mathbf{z}) = \begin{bmatrix} \mu_{kx} \\ \mu_{ky} \end{bmatrix}, \Sigma_k(\mathbf{z}) = \begin{bmatrix} \sigma_{kx}^2 & \rho_k \sigma_{kx} \sigma_{ky} \\ \rho_k \sigma_{kx} \sigma_{ky} & \sigma_{ky}^2 \end{bmatrix} \quad (8)$$

$$\rho_k \in [-1, 1]$$

式中: π_k 为分配系数; Φ 为二元高斯函数; μ_k 为均值; Σ_k 为协方差矩阵; σ_k 为标准差; ρ_k 为相关系数。混合密度网络中 3 个可变参数 $\pi_k(\mathbf{z})$, $\mu_k(\mathbf{z})$, $\Sigma_k(\mathbf{z})$ 为输入 \mathbf{z} 的函数; k 为第 k 个高斯函数混合。

选择 MDN 层由 $K=6$ 个二元高斯函数的线性组合构成, 每个二元高斯分布有 5 个参数, 加一个分配系数 π_k , 所以 \mathbf{z} 的维度为 36。 π_k 经由 Softmax 层确保分配系数总和为 1, 且各项均大于 0, 同时指数运算保证 σ_k 为正^[18]

$$\pi_k = \exp(z_k) \left[\sum_{i=1}^6 \exp(z_i) \right]^{-1} \quad (9)$$

$$\mu_k = z_{7-18}, \rho_k = z_{19-24}, \sigma_k = \exp(z_{25-36})$$

MDN 层中每个输入都对输出一个不同的概率分布。设有 T 个训练样本 $(\mathbf{I}^{(t)}, \mathbf{O}^{(t)})$, $t=1, 2, \dots, T$, 其极大似然函数 L 为

$$L = \prod_{i=1}^T \left[\sum_{i=1}^3 \omega_i^{(t)} P_{\pi, \mu, \sigma}(\mathbf{O}_i^{(t)} | \mathbf{z}) \right] \quad (10)$$

取极大似然函数的最小负对数作为轨迹输出模块训练的代价函数, 其表达式为

$$J = - \sum_{i=1}^T \log \left[\sum_{i=1}^3 \omega_i^{(t)} P_{\pi, \mu, \sigma}(\mathbf{O}_i^{(t)} | \mathbf{z}) \right] \quad (11)$$

式中: J 为整个轨迹输出模块的整体代价函数, 以使得各个意图类型的轨迹与真实轨迹尽量接近。

轨迹输出模块的编码器和解码器中, LSTM 循环体都采用和意图识别模块相似的结构, 由 4 层 LSTM 堆叠, 各层间 Dropout 率为 0.2, 每层隐藏单

元个数为 128。采用 Adam 优化器, 学习率 $\alpha = 0.001$, 衰减率为 0.9。

4 试验与分析

4.1 数据预处理

本文选取 NGSIM^[19] 数据集中 I-80 路段和 US-101 路段的数据用于训练和测试模型。该数据集源于美国联邦公路管理局发起的 Next Generation Simulation (NGSIM) 计划, 采样频率为 10 Hz, 记录了包括车辆坐标、速度、加速度、车辆类型和车道编号等信息。为了减小运算成本, 试验数据的采样频率定为 5 Hz (每隔一点取一次样), 研究路段如图 3 所示。由于原始数据存在一定的误差和噪声, 尤其是速度信号抖动明显, 因此先采用对称指数移动平均滤波算法^[20]对原始数据进行滤波处理。

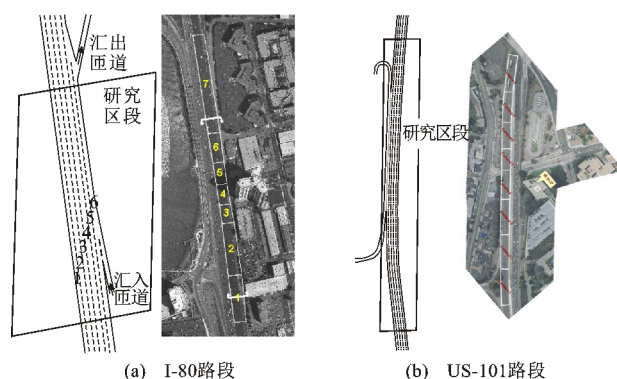


图 3 研究路段

Fig. 3 Study Area

对于意图识别模块, 需要将提取的轨迹片段分成向左换道、向右换道和直线行驶 3 类, 并附上相应标记。本文采用的分类依据为: 首先求出车辆轨迹和车道线的交点, 将其定义为换道点, 由车辆位置参数 (x, y) 计算出车辆的航向角 $\theta = \arctan\left(\frac{x^{(t)} - x^{(t-3)}}{y^{(t)} - y^{(t-3)}}\right)$; 然后从换道点向时间轴反方向遍历每个采样点的航向角 θ , 若轨迹序列连续 3 个采样点的 $|\theta| \leq \theta_s$ (换道起始点航向角阈值), 则将自第 1 次达到阈值 θ_s 的位置定位为换道起点; 最后用类似的方法判断 $|\theta| \leq \theta_e$ (换道终止点航向角阈值), 确定换道终点。连续三点确认是为了避免噪声引起的误判。换道起点和换道终点之间的点都被定义为换道过程点, 如图 4 所示。

同时, 采用滑动时窗的方法提取指定长度的序列^[21], 时间步长设置为 1 s, 即每次更新时向前 10 个采样节点。设截取序列的长度为 n 个采样点, 则相邻提取 2 个序列中有 $(n-10)$ 个节点的信息是一

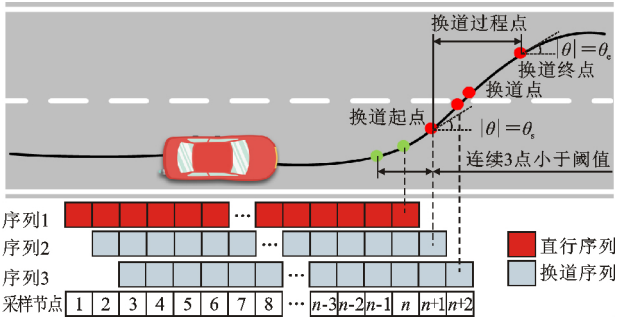


图 4 轨迹分类依据

Fig. 4 Basis for Trajectory Classification

样的。采用滑动时窗的方式能最大限度地利用数据。采样频率为 10 Hz,设输入序列的历史时域为 T_p ,则序列长度 $n=10T_p$ 。若提取轨迹序列最后一个采样节点为换道过程点,则该序列被标记为换道序列,否则为直行序列。左、右换道序列的判别依据序列始末位置的横向坐标确定。对所有轨迹采取以上方式提取轨迹片段。

由于直线行驶工况远多于换道工况,所以提取的序列中直线行驶类别远多于向左换道和向右换道类别。故从 3 种类型数据集中各随机选取 50 000 个序列(共 150 000 个序列)作为整个数据集,将其中 80% 作为训练集、20% 作为测试集。最后,需要对所有提取的数据做标准化处理,以便于神经网络进行训练。

4.2 意图识别模块性能分析

意图识别模块(IRM)的性能直接影响到轨迹预测质量的高低,为了测试其性能,本文以支持向量机(SVM)模型^[22]为基准模型,对比了两者在查准率(分类正确的正样本个数占分类器判定为正样本的样本个数的比例)、召回率(分类正确的正样本个数占真正的正样本个数的比例)、F1-分数(查准率与召回率的调和平均值)和准确率(分类正确的样本个数占总样本个数的比例)上的表现。由于意图识别模块输出的是 3 个类别的概率,而对应的标记只有一个正确的类别,故规定输出概率最大的类别为预测类别。用测试集中的数据对意图识别模块进行测试,结果见表 1、2。

表 1 意图识别混淆矩阵

真实意图		向左换道		直线行驶		向右换道	
		IRM		SVM		IRM	
		IRM	SVM	IRM	SVM	IRM	SVM
预测意图	向左换道	8 841	8 328	719	876	3	19
	直线行驶	1 153	1 656	8 589	8 280	1 198	1 632
	向右换道	6	16	692	844	8 799	8 349

表 2 意图识别性能度量

Table 2 Performance Measurement of Intention Recognition Module

评估指标	精确率		召回率		F1-分数		准确率	
	IRM	SVM	IRM	SVM	IRM	SVM	IRM	SVM
意图	向左换道	0.924 5	0.90 3	0.884 1	0.832 8	0.903 8	0.886 5	0.874 3
	直线行驶	0.785 1	0.715 8	0.858 9	0.828 0	0.820 3	0.767 8	
	向右换道	0.926 5	0.906 6	0.879 9	0.834 9	0.902 6	0.869 3	

输入轨迹的长度为 $T_p=3\text{ s}$,由表 2 可以看到,本文提出的意图识别模块的各项指标全面优于传统 SVM 分类器,意图识别模块的 3 类轨迹的识别召回率和整体准确率都达到 0.85 以上,说明意图识别模块具备良好的意图识别能力;左、右换道行为的识别精确率接近,高达 0.92,高于直线行驶意图的识别准确率,分析原因,可能是有小部分数据存在较大抖动,导致意图识别模块容易将其误判为换道类型,且任一(向右/向左)换道类别的轨迹若存在误判,很少会识别成相反(向左/向右)方向的类型,而是易误判为直线行驶,这也导致直线行驶类别精确率低于另外 2 个类别。

此外,意图识别模块还需要尽早识别出换道意图。定义预判时间为车辆从当前位置(序列最末采样时刻)行驶至换道点的时间间隔,从测试集的换道序列中选择预判时间分别为-3,-2.5,-2,-1.5,-1,-0.5,0 s 的序列作为输入,比较不同预判时间下的识别精确率,见图 5。由图 5 可以看出,随着被预测车辆驶向换道点,意图识别精确率逐渐提升。当车辆在换道前 1.5 s,就能较为准确地识别出换道意图,精确率达到 85% 以上。在车辆行驶至换道点精确率高达 95%。查阅相关文献可知,高速公路环境换道时间一般在 3.5~6.5 s 之间,平均 5 s 可以实现一次完整的换道过程^[23]。假设换道点前和换道点后各占 2.5 s 时间,则意图识别模块在驾驶人执行换道动作之前就能做出预判。原因在于,意图识别模块接收的信息除了有被预测车辆状态信息外,还有周围车辆的状态信息,所以模型能站在被预测车辆的角度“观察”路况,“理解”交通环境的变化规律,并作出合理预判。

4.3 完整模型性能分析

选择均方根误差(Root Mean Square Error, RMSE)作为完整模型预测结果的衡量标准,对比以下几个模型在预测时域 T_f 为 5 s 内的均方根误差,

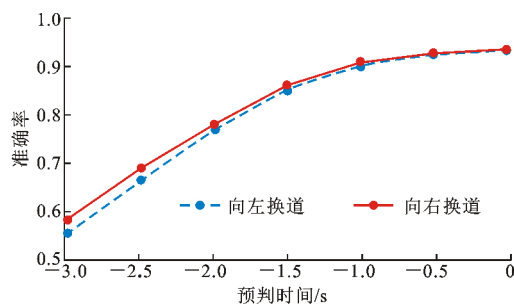


图5 预判时间与识别准确率关系

Fig. 5 Relationship Between Predicting Time and Recognition Accuracy

以检验模型的有效性。

(1)CTRA: 使用参考文献[2]中介绍的基于CTRA模型的轨迹预测方法。

(2)XY_LSTM: 使用LSTM结构, 仅输入被预测车辆的位置信息。

(3)E_LSTM: 在XY_LSTM的基础上, 增加被预测车辆的速度信息。

(4)S_LSTM: 在E_LSTM的基础上, 增加周边车辆的位置和速度信息。

(5)I_LSTM: 本文提出的完整模型, 包括意图识别模块和轨迹输出模块。

因本模型输出的是基于不同意图分类的位置概率分布, 为了便于RMSE的计算以及模型对比的公平, 取每一时刻概率最大类别(向左换道、直线行驶、向右换道)对应的高斯分布的均值作为最终的输出。试验结果见表3和图6。

表3 模型均方根误差对比

Table 3 RMSE Comparison of Models

模型	预测时域/s				
	1	2	3	4	5
CTRA	0.236 7	0.833 7	2.273 2	5.072 1	9.621 4
XY_LSTM	0.307 2	1.303 3	2.970 1	4.789 1	7.891 4
E_LSTM	0.280 0	1.064 7	2.436 7	4.368 7	6.860 7
S_LSTM	0.163 3	0.694 7	1.607 5	2.949 7	4.836 3
I_LSTM	0.125 7	0.531 2	1.120 4	2.180 4	3.427 4

由表3和图6可知, 基于车辆运动学模型的轨迹预测方法(CTRA)在较短预测时域能保证较高的预测精度, 但随着预测时域的增加, 预测误差迅速增加, 长预测时域的可靠性很差。相比之下, 其他4种基于LSTM结构的模型对均方根误差的控制能力较强, 预测时域大于4 s后的均方根误差均小于基于模型的预测方法, 可见LSTM网络结构在处理长时间序列的优越性。

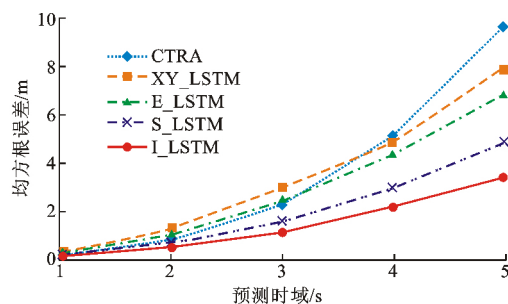


图6 模型均方根误差对比

Fig. 6 Comparison of RMSE of Different Models

同时, 对比4个基于LSTM网络的模型可以看出, 仅在输入轨迹位置信息的基础上, 增加车辆速度信息、周车信息以及驾驶意图信息均能提高模型输出轨迹的准确率; 对比E_LSTM曲线和S_LSTM曲线可知, 周车信息对模型预测性能有提升, 这也体现了本模型采用交互式信息作为输入的有效性; 除此之外, 对比S_LSTM曲线和I_LSTM曲线表明, 意图识别模块能明显改善预测较长轨迹的能力, 尤其是在预测时域较长时, 基于意图识别模块的轨迹预测模块仍然能保持较高的准确率。

4.4 基于场景的分析

在实际应用中, 轨迹预测模块需要能够实时地预测出目标车辆未来的位置分布, 因此本文提出的完整模型会在每个采样节点更新输入的历史轨迹信息, 动态调整最新的预测。为了体现这一动态调整过程, 本文从测试集中选择一段具有代表性的向右换道行驶的完整轨迹序列, 在每一个采样节点, 随机取5个采样点, 并滑动更新输入轨迹序列。其中, 历史时域 $T_p = 3$ s, 预测时域 $T_f = 5$ s。图7展示了车辆接近换道点过程中, 意图识别模块实时输出的3类驾驶意图的概率。图8展示了预判时间分别为-3, -2, -1, 0 s时预测位置的分布, 其能直观地展示历史轨迹、未来实际轨迹和预测轨迹三者之间的关系。

结合图7, 8可知, 随着车辆行驶接近换道点, 模型做出的预测质量越来越高, 且长期预测的置信度仍比较高, 这正是使用LSTM网络处理长时间序列的优势。当被预测车辆处在距换道点3 s时距的位置, 意图识别模块已经能预判到车辆将采取向右换道的行为, 向右换道的概率大于直线行驶的概率, 但没有达到确信阈值80%, 所以直接输出Softmax层得到的三类型概率, 因而轨迹输出模块会预测出直行和向右换道两类轨迹; 当被预测车辆行驶至距离换道点2 s位置时, 识别出向右换道的概率已经超

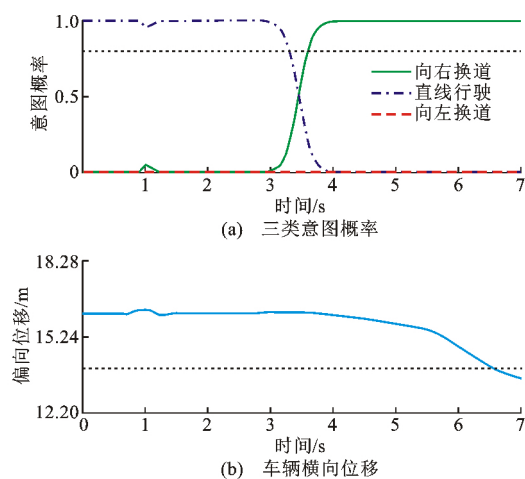


图7 意图识别结果

Fig. 7 Results of Intention Recognition

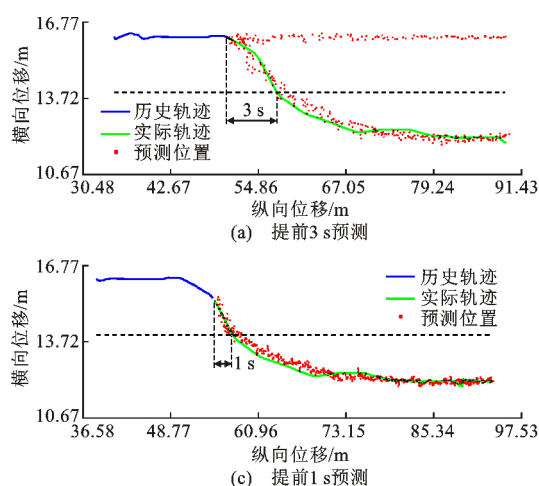


图8 预测轨迹散点

Fig. 8 Scatter Plot of Predicted Trajectory

(2) 由于试验数据存在噪声与误差,所以图中的实际未来轨迹不代表现实中车辆的真实轨迹。而本模型做出的预测轨迹位置更加平顺,反而更接近车辆真实的行驶轨迹。

5 结 语

(1) 本文设计了一种基于 LSTM 网络的意图识别及轨迹预测模型,其中的意图识别模块负责计算出驾驶人向左换道、直线行驶、向右换道的概率,轨迹输出模块在此基础上,通过 MDN 层输出预测轨迹分布;并利用 NGSIM 数据集对模型进行训练与测试,并与其他模型进行对比。试验结果表明,该模型相较于传统的基于模型的方法在长时间预测精度上有明显提高,周围车辆的交互式信息和驾驶意图的预判能有效提高轨迹预测的准确性。

(2) 本文模型是基于高速公路直线路段的行车

过 80%,所以意图识别模块输出的左换道、直行、右换道的概率分别为 0,0,100%,轨迹输出模块只输出一种类型的位置分布。从图 8(a)、(b)还能看出,预测超前时间超过 2 s 时预测的轨迹稍有滞后,预测位置的方差较大,确信度较低;被预测车辆继续行驶过程中,预测的置信度逐渐提高,预测的位置分布逐渐变得集中,更接近真实的轨迹。除此之外,有 2 处细节值得留意:

(1) 在距离换道点 3 s 时距的位置,仅从图 8(a)中历史轨迹几乎看不出有转向的意图,然而本模型仍然能做出向右换道的预测。模型接收的是被预测车辆和其周围车辆的状态信息,相当于处在被预测车的角度“观察”路况,推断意图,收到更多的信息,做出更加准确的预测。

数据进行训练的,因此适用范围有限。今后的研究工作将尝试拓展模型应用范围,使之能识别包括转向、起停等类别,使模型能应用于诸如十字路口等工况。

参考文献:

References:

- [1] 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.
- [2] SCHUBERT R, ADAM C, OBST M, et al. Empirical Evaluation of Vehicular Models for Ego Motion Estimation [C] // IEEE. 2011 IEEE Intelligent Vehicles Symposium (IV). New York: IEEE, 2011: 534-539.
- [3] HOUENOU A, BONNIFAIT P, CHERFAOUI V,

- et al. Vehicle Trajectory Prediction Based on Motion Model and Maneuver Recognition [C] // IEEE. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. New York: IEEE, 2013: 4363-4369.
- [4] BARTH A, FRANKE U. Where Will the Oncoming Vehicle Be the Next Second? [C] // IEEE. 2008 IEEE Intelligent Vehicles Symposium. New York: IEEE, 2008: 1068-1073.
- [5] TOLEDO-MOREO R, ZAMORA-IZQUIERDO M A. IMM-based Lane-change Prediction in Highways with Low-cost GPS/INS [J]. IEEE Transactions on Intelligent Transportation Systems, 2009, 10 (1): 180-185.
- [6] 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.
- LIU Zhi-qiang, WU Xue-gang, NI Jie, et al. Driving Intention Recognition Based on HMM and SVM Cascade Algorithm [J]. Automotive Engineering, 2018, 40 (7): 858-864.
- [7] ZYNER A, WORRALL S, WARD J, et al. Long Short Term Memory for Driver Intent Prediction [C] // IEEE. 2017 IEEE Intelligent Vehicles Symposium (IV). New York: IEEE, 2017: 1484-1489.
- [8] KHOSROSHAHI A, OHN-BAR E, TRIVEDI M M. Surround Vehicles Trajectory Analysis with Recurrent Neural Networks [C] // IEEE. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). New York: IEEE, 2016: 2267-2272.
- [9] PHILLIPS D J, WHEELER T A, KOCHENDERFER M J. Generalizable Intention Prediction of Human Drivers at Intersections [C] // IEEE. 2017 IEEE Intelligent Vehicles Symposium (IV). New York: IEEE, 2017: 1665-1670.
- [10] 李建平. 面向智能驾驶的交通车辆运动预测方法研究[D]. 长春: 吉林大学, 2018.
- LI Jian-ping. Research on Vehicle Motion Prediction Method for Intelligent Driving [D]. Changchun: Jilin University, 2018.
- [11] SCHREIER M, WILLERT V, ADAMY J. BAYESIAN, Maneuver-based, Long-term Trajectory Prediction and Criticality Assessment for Driver Assistance Systems [C] // IEEE. 17th International IEEE Conference on Intelligent Transportation Systems (ITSC). New York: IEEE, 2014: 334-341.
- [12] SCHLECHTRIEMEN J, WIRTHMUELLER F, WEDEL A, et al. When Will it Change the Lane? A Probabilistic Regression Approach for Rarely Occurring Events [C] // IEEE. 2015 IEEE Intelligent Vehicles Symposium (IV). New York: IEEE, 2015: 1373-1379.
- [13] DUAN Y, LV Y, WANG F Y. Travel Time Prediction with LSTM Neural Network [C] // IEEE. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). New York: IEEE, 2016: 1053-1058.
- [14] ALAHI A, GOEL K, RAMANATHAN V, et al. Social LSTM: Human Trajectory Prediction in Crowded Spaces [C] // IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 961-971.
- [15] 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.
- [16] PARK S H, KIM B D, KANG C M, et al. Sequence-to-sequence Prediction of Vehicle Trajectory Via LSTM Encoder-decoder Architecture [C] // IEEE. 2018 IEEE Intelligent Vehicles Symposium (IV). New York: IEEE, 2018: 1672-1678.
- [17] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning Phrase Representations Using RNN Encoder-decoder for Statistical Machine Translation [EB/OL]. [2019-01-04]. <https://arxiv.org/pdf/1406.1078.pdf>.
- [18] BISHOP C M. Mixture Density Networks [R]. Birmingham: Aston University, 1994.
- [19] FEDERAL HIGHWAY ADMINISTRATION. NG-SIM-Next Generation Simulation [EB/OL]. <https://ops.fhwa.dot.gov/trafficanalysisistools/ngsim.htm>.
- [20] THIEMANN C, TREIBER M, KESTING A. Estimating Acceleration and Lane-changing Dynamics from Next Generation Simulation Trajectory Data [J]. Transportation Research Record, 2008 (2088): 90-101.
- [21] ALTCHÉ F, DE LA FORTELLE A. An LSTM Network for Highway Trajectory Prediction [C] // IEEE. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). New York: IEEE, 2017: 353-359.
- [22] CORTES C, VAPNIK V. Support-vector Networks [J]. Machine Learning, 1995, 20 (3): 273-297.
- [23] TIJERINA L, GARROTT W R, STOLTZFUS D, et al. Eye Glance Behavior of Van and Passenger Car Drivers During Lane Change Decision Phase [J]. Transportation Research Record, 2005 (1937): 37-43.