Context-aware behaviour prediction for autonomous driving: a deep learning approach

Syama R.

Department of Computer Science and Engineering, National Institute of Technology, Tiruchirappalli, India and Department of Computer Science and Engineering, SCT College of Engineering, Thiruvananthapuram, India, and

Mala C.

Department of Computer Science and Engineering, National Institute of Technology, Tiruchirappalli, India

Abstract

Purpose – This paper aims to predict the behaviour of the vehicles in a mixed driving scenario. This proposes a deep learning model to predict lane-changing scenarios in highways incorporating current and historical information and contextual features. The interactions among the vehicles are modelled using longshort-term memory (LSTM).

Design/methodology/approach – Predicting the surrounding vehicles' behaviour is crucial in any Advanced Driver Assistance Systems (ADAS). To make a decision, any prediction models available in the literature consider the present and previous observations of the surrounding vehicles. These existing models failed to consider the contextual features such as traffic density that also affect the behaviour of the vehicles. To forecast the appropriate driving behaviour, a better context-aware learning method should be able to consider a distinct goal for each situation is more significant. Considering this, a deep learning-based model is proposed to predict the lane changing behaviours using past and current information of the vehicle and contextual features. The interactions among vehicles are modeled using an LSTM encoder-decoder. The different lane-changing behaviours of the vehicles are predicted and validated with the benchmarked data set NGSIM and the open data set Level 5.

Findings – The lane change behaviour prediction in ADAS is gaining popularity as it is crucial for safe travel in a mixed driving environment. This paper shows the prediction of maneuvers with a prediction window of 5 s using NGSIM and Level 5 data sets. The proposed method gives a prediction accuracy of 97% on average for all lane-change maneuvers for both the data sets.

Originality/value – This research presents a strategy for predicting autonomous vehicle behaviour based on contextual features. The paper focuses on deep learning techniques to assist the ADAS.

Keywords Autonomous driving, Behaviour prediction, Context-aware feature, Lane change scenario, Long short term memory (LSTM)

Paper type Research paper

1. Introduction

With the recent developments in Artificial Intelligence, the vehicles have brought many services to human lives including advanced driver assistance systems (ADAS). Lane change is the most common driving behaviour and lot of accidents have been reported globally due to improper lane changes. The lane change of the neighbouring vehicles has to be anticipated in advance since it involves the lateral displacement of the vehicle which may

International Journal of Pervasive Computing and Communications Vol. 19 No. 4, 2023 pp. 477-490 © Emerald Publishing Limited 1742-7371 DOI [10.1108/IJPCC-10-2021-0275](http://dx.doi.org/10.1108/IJPCC-10-2021-0275)

Context-aware behaviour prediction

477

Received 30 October 2021 Revised 9 January 2022 Accepted 20 January 2022

cause accidents. The driver decides whether to perform a lane change or not by checking the advantages of the lane change. Thus the autonomous vehicle technology research has focused more on reducing road accidents that are caused by improper lane changes and thus improving safety. Behaviour prediction is the major challenge in autonomous driving in a mixed driving scenario where the autonomous vehicles coexist with human-driven vehicles. The vehicles must accurately anticipate the behaviour of the different traffic agents to provide safety. In a mixed traffic environment, the agents include autonomous vehicles, human-driven vehicles and pedestrians. The major factors influencing a traffic agent's behaviour in an environment are other agents' behaviour, road geometry, traffic rules and contextual features. There are several practical challenges to the behaviour prediction problem. Some of the significant challenges are lack of computational resources, limited data from the onboard sensors and computational cost.

Generally, autonomous vehicles must perform several processing steps to automate their movement through the complex traffic environment. The first step is to sense the current environment and to represent it properly. From this representation, the current situation is anticipated. Based on this, a plan of the trajectory is generated. This trajectory is executed as the final step and re-planning may be done if the behaviour changes.

The behaviour prediction of surrounding vehicles involves the understanding of the surrounding environment and foresees future movement. This supports the decisionmaking and provides safety. Several earlier research have looked into predicting lane change behaviour (Mozaffari and Al-Jarrah, 2020). Lane change behaviour of the target vehicle depends on several factors such as the position of the vehicles in the environment, traffic flow in the lanes etc. The behaviour of the vehicles in a traffic environment is highly uncertain. The environment under consideration continuously changes. Based on the current situation, the behaviour has to be predicted. The context-aware behaviour prediction is more challenging. The contextual features that may be considered include weather, daytime and traffic density (Yoon *et al.*[, 2020\)](#page-13-0). The major challenge in behaviour prediction is modelling the interactions among vehicles in real time in a mixed dynamic environment. In the mixed driving scenario, the autonomous vehicles must cooperate with the humandriven vehicles. The existing studies fail to model the spatial and temporal features of the target vehicle and the neighbouring vehicles.

Considering this, a behaviour prediction scheme is presented here which uses deep learning techniques to anticipate the lane-change behaviour of the vehicle in a highway incorporating contextual features.

The following is how the rest of the paper is organized: The available frameworks and approaches for prediction are discussed in Section 2. The descriptions of the suggested approach of behaviour prediction are included in Section 3. The suggested scheme's evaluation results are discussed in Section 4. Section 5 contains the conclusion, which is followed by a list of references.

2. Related works

Several studies on behaviour analysis of vehicles have been published. A study of vehicle monitoring, behaviour prediction and analysis is provided in [\(Shirazi and Morris, 2017\)](#page-13-1). Several studies have been undergone in different literature (Mozaffari and Al-Jarrah, 2020) to identify and predict the surrounding agents' behaviour in the traffic environment in highway scenarios. The studies on this area can be divided into different categories in [Lefevre](#page-12-0) *et al.* (2014). Some literature concentrates on approaches that are based on the laws of Physics (Mozaffari and Al-Jarrah, 2020). Some literature anticipates the behaviour based

IIPCC 19,4

on their intended manoeuver [\(Mahajan](#page-13-2) *et al.*, 2020) and others considers the interaction between different agents in the environment [\(Schulz](#page-13-3) *et al.*, 2018, [2019\)](#page-13-4).

In [Mofan Zhou](#page-13-5) et al. (2017), Zyner et al. [\(2017\)](#page-13-6) and Worrall et al. (2020), vehicles' past behaviour is used to anticipate how they would behave at intersections. But these methods fail to consider the behaviour of the nearby vehicles. The target vehicle alone is the consideration. Few studies focused on the deep learning approaches that considered the inputs and the presence of other vehicles on the environment [\(Deo and Trivedi, 2018;](#page-12-1) [Patel](#page-13-7) et al.[, 2018](#page-13-7)).

Context-aware computing is the process of gathering data from numerous sensors and making decisions based on the present context's requirements without the need for human participation. It describes an application's capacity to obtain current information about its surroundings and makes decisions based on that information. Few studies on the contextual features are included in Gite et al. [\(2021\).](#page-12-2)

The existing literature has failed to study the target vehicle's lane change behaviour in a mixed scenario, considering the surrounding vehicles and contextual information. This paper presents the behaviour prediction method using Long Short-Term Memory (LSTM), considering the contextual properties. The basic LSTMs used in Alahi et al. [\(2016\)](#page-12-3), [Wiest](#page-13-8) et al. [\(2012\),](#page-13-8) [Hermes](#page-12-4) et al. (2009) suffer from two problems: the basic LSTM model failed to instantaneously model the spatial interactions of vehicle and trajectories and sometimes they suffer from vanishing gradient problem [\(Heaton](#page-12-5) *et al.*, 2018). But this problem can be solved using LSTM with some structural modifications to well represent the spatial interactions as time series. This new LSTM model is used to model the trajectories and spatial interactions between vehicles (Jain *et al.*[, 2016](#page-12-6)).

Considering these, an encoder-decoder LSTM model is proposed to predict the lane change behaviour of the vehicles in an autonomous driving scenario. The motivation of using encoder-decoder LSTM model is that it is well suited for time-series data, it can identify and accurately predict behaviours in time-series data, and it is capable of automatically extracting the features from massive long time series data.

This study makes substantial contributions in the following areas:

- Apply a deep-learning based approach to generate a model to foresee the maneuvers and generate upcoming positions of the vehicle in a time window of 5s.
- Discuss the influence of contextual knowledge on the behaviour prediction considering traffic density and show that this improves prediction capabilities.
- The proposed model is trained and tested using the benchmarked data set, NGSIM, which is the most realistic public data set available for traffic data and the Level 5 open data set.

3. Methodology

This paper presents a maneuver-based prediction model that uses LSTM to predict the human-driven vehicles' long-term driving behaviour in a mixed driving environment considering the contextual information. Keep Lane, Change Left and Change Right are the basic maneuvers. The scenario under consideration is shown in [Figure 1](#page-3-0).

The human-driven vehicle whose trajectory has to be predicted is represented as the target vehicle (TV). The vehicles in the environment influence the trajectory of every other vehicle; but not all of them produce a significant impact on the other vehicle. The vehicles that impact the target vehicle's future behaviour can be termed as Nearby Vehicles (NV) and all others are termed as Distant Vehicles (DV). Here, a threshold distance-based scheme is

Context-aware behaviour prediction

479

used to classify vehicles into NV and DV. If the vehicles' longitudinal distance is less than the threshold distance, they are added to the set NV. The interactions of a maximum of nine such vehicles that surround the target vehicle are taken. All the other interactions are considered least significant and are neglected. [Figure 2](#page-3-1) shows the scenario under consideration. IJPCC

> Here, NV9 is included since sometimes the vehicle just in front of TV may not be enough to determine the future traffic condition. The future movement of NV1 always depends on the change in the positions of NV9.

The aim is to model TV's future positions within a time window t in the longitudinal and lateral positions (x_t, y_t) . The historical trajectories of TV and NV_i can be represented as x_t^t and x_i^t respectively. The input to the model is the historical trajectories of TV and NV_is and outputs the predicted trajectory of TV. The interactions of all the vehicles of interest are identified and the information is added to the data set.

The vehicle trajectories can be represented as displacements. The inputs and output to the system is $(Dai *et al.*, 2019)$ $(Dai *et al.*, 2019)$ $(Dai *et al.*, 2019)$:

$$
x_T^t = \Delta X_T^{(t-th+\Delta t)}, \Delta Y_T^{(t-th+\Delta t)} \dots \Delta X_T^{(t)}, \Delta Y_T^{(t)}
$$
(1)

$$
y_T^t = \Delta X_T^{(t+\Delta t)}, \Delta Y_T^{(t+\Delta t)} \dots \Delta X_T^{(t+tp)}, \Delta Y_T^{(t+tp)}
$$
\n(2)

 $[\Delta X, \Delta Y]$: trajectory, which is represented as $O = (O_1, O_2, \ldots O_t)$ X, Y: Lateral and Longitudinal positions

Figure 1. The lane-change scenario in a straight highway

19,4

Number of historical time NH =
$$
^{th}/_{\Delta t}
$$
 (4)

The following characteristics of the target vehicle are taken into account:

- The target vehicle's lateral and longitudinal positions represented as x_T^t and y_T^t , respectively.
- Lateral and longitudinal velocity V_{xTV} and V_{vTV} , respectively.

Each nearby vehicle of interest has the following features.

- The lateral velocity, V_{xSVi}
- The longitudinal velocity V_{ySVi}
- Distance from the target vehicle X_i^t and Y_i^t

The contextual features is represented by a vector $F = (f_1, f_2, f_i)$ where each f_i represents the context feature.

In the traffic scenario, the vehicles interact among themselves in real-time. These interactions must be properly modelled to improve the prediction accuracy. The long sequences can be recognized and predicted by LSTMs, but the dependencies between numerous correlated sequences cannot be captured by a single LSTM. So in this proposed work, an encoder-decoder LSTM network together with the contextual features is used to model the interactions among the vehicles in a mixed driving scenario. During each updation step, one LSTM is assigned to the target vehicle and the neighbouring vehicles and extract the spatial and temporal features and these features together with the extracted contextual features are transferred to the higher-level LSTM and analysed the interaction pattern to predict the lane change behaviour.

The interactions between the vehicles are modelled using LSTM models using Encoder-Decoder as shown in [Figure 3.](#page-5-0) The contextual features are represented by a vector F and are passed to the decoder. These models detect the dependencies among the different vehicles and store the information in the memory. The LSTM network, together with the contextual information, can be used to detect the different driving patterns and classify them into three classes such as Keep Lane, Change Left and Change Right. At every prediction window, the historical data and the interaction details are fed to the input of the LSTM. The sigmoid

function is applied as an activation function here. The classification results are then passed into the softmax layer for multiclass classification.

Even if the vehicles' velocity and driving intentions affect the future behaviour of target vehicles, external factors may still influence the vehicle's future trajectory. Traffic density, weather and other contextual factors can all be taken into account when predicting vehicle behaviour. The traffic density is taken in this paper as a feature that affects the subject vehicle's behaviour. To calculate the influence of traffic density on the behaviour of target vehicle, the calculated values of lateral and longitudinal distances are taken. The data are transformed to contain the lane number in increasing order from right to left. Then three Gaussian models are used to model the future positions of the vehicle to generate three classes for lane change. These models are trained using the input features, the lateral velocity and the distance to the current lane. Then a mixture of experts method is used, which predicts the position of the vehicle.

4. Results and discussion

The benchmark trajectory data set NGSIM US-101 is used to train and test our model. The US-101 data set was collected from U.S. Highway 101 in Los Angeles, CA and is the largest public natural driving data set available. This is a real-world vehicle trajectory data set, which includes 45 minutes of data. This data set captured the trajectories of more than 6,000 unique vehicles at 10 Hz. This also includes the different lane-changing scenarios. So this data set is well suited for the scenario discussed here. The data frame in the data set consists

of several vehicle parameters such as position, velocity, yaw rate etc (LI *et al.*). The aerial view of the US 101 study area is as shown in [Figure 4](#page-6-0) ([Punzo](#page-13-9) et al., 2011). Context-aware behaviour

The details of different types of trajectories in NGSIM data set is shown in [Table 1](#page-6-1).

But this data set is highly unbalanced. The right amount of data must be used for training LSTM models since both insufficient and excess data can affect the training. To balance the data set, downsampling and oversampling are done. Each trajectory with line change can be cropped so that only the data part that contains the lane change only is stored. Random sampling is of the trajectories having no lane change also is done to keep it balanced. The sampled data set is used for training. Overfitting and underfitting is major problem while using LSTM. A random set of 80% sampled trajectories is used for training and gives a suitable model. Then another set of 20% trajectories is chosen for testing.

[Figure 5](#page-7-0) and [6](#page-7-1) show the lateral distance and lateral velocities of the vehicle.

The model is trained using LSTM. The different interactions I_i of the nearby vehicles NV_i have a different impact on the target vehicle. Therefore each interaction is trained independently. All the models are trained at a rate of 0.001. Then the different LSTMs are combined and trained. At each training step, a batch of 50 trajectories each of length

> Vehicle Study **Vehicle Trajectory Study Area** 503 meters (1.650 feet) Eastbound/

Figure 4. Aerial view of the US-101 study area

Data set	No. of vehicles	Total No. of lane changes	Mandatory lane changes (MLC)	Discretionary lane-change	Table 1. Trajectory statistics of US highway-101 data set
NGSIM US Highway -101	2169	199		128	

Northbound I-80

483

prediction

50 (i.e. 5 s) is chosen. The randomly chosen parameters are used for training. Adam optimizer with default values for α and β are used.

To show the future movement, the lane-change trajectory is selected randomly and observes the result. Different snapshots are shown in [Figure 7.](#page-8-0)

The loss function of each individual trajectory can be calculated as follows:

$$
loss = \sum_{i=1}^{Np} (Y_i)^T Y_i
$$
 (3)

The MSE loss is calculated as a function:

$$
Loss = \frac{1}{N} \sum loss \tag{4}
$$

N: Total number of trajectories used.

Figure 6. Lateral velocity in a lane change scenario

Note: (a) Predicted trajectory at T=0s; (b) predicted trajectory at T=1; (c) predicted trajectory at T=2s; (d) predicted trajectory at T=4s

Context-aware behaviour prediction

485

The prediction window chosen here is $5s$ since on average, the lane change process takes $4 \sim 5$ s.

The mean and the standard deviation of velocity changes in each prediction window is calculated to analyse the causes of trajectory prediction error. Then the mean absolute deviation of position changes is plotted and shown in [Figure 8.](#page-9-0)

Here, it is shown that the position changes are almost the same in different prediction windows. Thus when the prediction window is increased, the absolute deviation increases and error gets accumulated. This reduces the prediction accuracy.

The model is trained and tested using the open data set, Level 5 also. The snapshot of the data set is shown in [Figure 9.](#page-9-1) This data set is a motion data set which consists of object trajectories and the 3 D maps which is collected from Palo Alto, CA [\(Houston](#page-12-8) et al., 2020). This is a real-world vehicle trajectory data set which includes over 20 million frames which encodes the precise positions and the position and motions of nearby vehicles, cyclists and pedestrians. The objects are labelled into three classes: vehicles, pedestrians and cyclists. This data set can be used for behaviour prediction as it includes different driving behaviours such as lane changes, unprotected turns, merges and intersections and is

Figure 9. Snapshot of Level 5 Data set

recorded at various environments that represent congestion, transition between uncongested and congested conditions, etc. Context-aware

[Tables 2](#page-10-0) and [3](#page-10-1) compare the prediction accuracy obtained with the proposed method and LSTM. The proposed method outperforms the LSTM method in all the lane-changing classes for both the data sets.

The performance of prediction in both set of data can be compared. [Table 4](#page-10-2) shows the confusion matrix values for prediction accuracies, precision values, recall values and F1 values are shown.

[Figure 10](#page-11-0) and [Figure 11](#page-11-1) show the ROC curves for both the data sets.

The training time of the model with different data sets is also evaluated. [Table 5](#page-12-9) shows the approximate training time for each data set. From the table, it is obvious that the training time rapidly increases in Level 5 data set.

Now, the effect of contextual information on prediction accuracy has to be studied. The traffic density, T can be calculated as:

$$
T = \frac{N}{km.N_L} \tag{5}
$$

487

behaviour prediction

The impact of traffic density on lateral position error can be shown in [Figure 12](#page-12-10) for different traffic scenarios. It shows that the lateral prediction error during lane-keep reduces as the traffic density increases. The error values fluctuate and then almost remain constant with increasing traffic density.

5. Conclusion

A deep learning approach based on LSTMs, which considers current and past behaviours and contextual features is proposed to predict the various driving behaviours in a highway driving scenario. The benchmarked data set NGSIM US-101 and Level 5 are used to validate the results. The method gives a prediction accuracy of 97% on average for all the driving maneuvers for both the data sets. The effect of traffic density on the driving maneuvers is also investigated.

Figure 11. ROC Curve for different behaviours in Level 5 data for proposed model

References

- Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L. and Savarese, S. (2016), "Social LSTM: human trajectory prediction in crowded spaces", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 961-971.
- Dai, S., Li, L. and Li, Z. (2019), "Modeling vehicle interactions via modified LSTM models for trajectory prediction", IEEE Access, Vol. 7, pp. 38287-38296.
- Deo, N. and Trivedi, M. (2018), "Multi-Modal trajectory prediction of surrounding vehicles with maneuver based LSTMs", IEEE Intelligent Vehicles Symposium (IV), pp. 1179-1184.
- Gite, S., Kotecha, K. and Ghinea, G. (2021), "Context–aware assistive driving: an overview of techniques for mitigating the risks of driver in real-time driving environment", International Journal of Pervasive Computing and Communications.
- Heaton, J., Goodfellow, I., Bengio, Y. and Courville, A. (2018), "Deep learning", *Genetic Programming* and Evolvable Machines, Vol. 19 Nos 1/2, pp. 305-307.
- Hermes, C., Wohler, C., Schenk, K. and Kummert, F. (2009), "Long-term vehicle motion prediction", IEEE Intelligent Vehicles Symposium, Xi'an, China, pp. 652-657.
- Houston, J. Zuidhof, G. Bergamini, L. Ye, Y. Chen, L. Jain, A. Omari, S. Iglovikov, V. and Ondruska, P. (2020),"One thousand and one hours: self-driving motion prediction dataset", arXiv:2006.14480
- Jain, A., Zamir, A.R., Savarese, S. and Saxena, A. (2016), "Structural-RNN: deep learning on spatiotemporal graphs", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 5308-5317.
- Lefevre, S., Vasquez, D. and Laugier, C. (2014), "A survey on motion prediction and risk assessment for intelligent vehicles", Robomech Journal, Vol. 1 No. 1.

- Zhou, M., Qu, X. and Li, X. (2017), "A recurrent neural network based microscopic car following model
to predict traffic oscillation", *Transportation Research Part C: Emerging Technologies*, Vol. 84, pp. 245-264.
- Zyner, A., Worrall, S., Ward, J. and Nebot, E. (2017), "Long short term memory for driver intent prediction", IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, USA, 2017, pp. 1484-1489.

Further reading

- Mozaffari, S., Al-Jarrah, O.Y., Dianati, M., Jennings, P. and Mouzakitis, A. (2020), "Deep learning-based vehicle behavior prediction for autonomous driving applications: a review", IEEE Transactions on Intelligent Transportation Systems, Vol. 23 No. 1, pp. 33-47.
- Zyner, A., Worrall, S. and Nebot, E. (2020), "Naturalistic driver intention and path prediction using recurrent neural networks", IEEE Transactions on Intelligent Transportation Systems, Vol. 21 No. 4, pp. 1584-1594.

Corresponding author

Syama R. can be contacted at: $\frac{\text{syma}}{\text{intra}}$.com

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm Or contact us for further details: permissions@emeraldinsight.com