Context-Aware Intention and Trajectory Prediction for Urban Driving Environment

Malika Meghjani[1,2(✉)], Shashwat Verma[2], You Hong Eng[2], Qi Heng Ho[2,3], Daniela Rus[4], and Marcelo H. Ang Jr.[3]

[1] Singapore University of Technology and Design, Singapore, Singapore
malika_meghjani@sutd.edu.sg
[2] Singapore-MIT Alliance for Research and Technology, Singapore, Singapore
{shashwat,youhong,qiheng}@smart.mit.edu
[3] National University of Singapore, Singapore, Singapore
mpeangh@nus.edu.sg
[4] Massachusetts Institute of Technology, Cambridge, USA
rus@csail.mit.edu

Abstract. This paper addresses intention and trajectory prediction of exo-vehicles in an urban driving environment. Urban environments pose challenging scenarios for self-driving cars, specifically pertaining to traffic light detection, negotiating paths at the intersections and sometimes even overtaking illegally parked cars in narrow streets. This complex task of autonomously driving while considering anomalous situations make urban driving conditions unique when compared to highway driving. In order to overcome these challenges, we propose to use road contextual information to predict driving intentions and trajectories of surrounding vehicles. The intention prediction is obtained using a recurrent neural network and the trajectory is predicted using a polynomial model fitting of the past lateral and longitudinal components of the vehicle poses and road contextual information. The integrated process of intention and trajectory prediction is performed in real-time by deploying and testing on a self-driving car in a real urban environment.

1 Introduction

Urban driving scenarios involve unique challenges when compared to highway driving due to complexity of the environment and corresponding road-rules. Given the vast range of road-rules that may or may not be applicable to anomalous urban driving situations, it becomes crucial to use road contextual information for autonomous and intelligent driving. An example of the anomalous situation is illegally parked car in a single lane on a bi-directional street which

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-33950-0_30) contains supplementary material, which is available to authorized users.

J. Xiao et al. (Eds.): ISER 2018, SPAR 11, pp. 339–349, 2020.
https://doi.org/10.1007/978-3-030-33950-0_30

requires the ego vehicle to overtake from the opposite side of the road. In such scenarios, using road contextual information such as number and direction of lanes, lane width and distance to the upcoming intersection, can help in human-like decision making. Hence, we propose to use contextual information for intention and trajectory prediction of surrounding vehicles to plan a path for the ego-vehicle.

Our system comprises of five main modules as illustrated in Fig. 1. Specifically, these include:

- **Perception** module comprising of vision-based obstacle detection and classification system along with LiDAR-based point cloud clustering for identifying the road region.
- **Sensor fusion and tracking** module performs data association between the vision and LiDAR systems and tracks the vehicles in the global map frame.
- **Behavior analysis and prediction** module converts the vehicle tracks from map frame to lane coordinate system and uses road contextual information for behavior prediction.
- **Decision making and planning** module uses the predicted behavior of exo-vehicles for decision making of the ego-vehicle. This decision is then translated into a path using the planner.
- **Speed and steering control** module is a low-level controller used to execute the planned path.

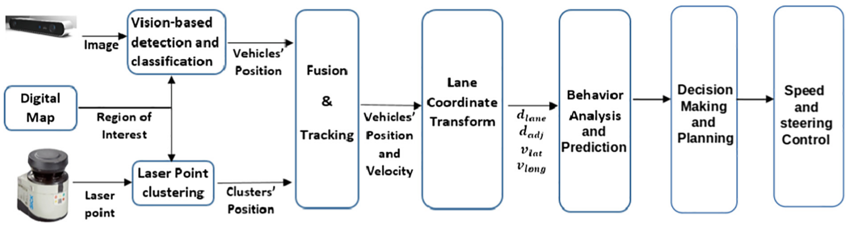


Fig. 1. System overview comprising of five modules: perception, sensor fusion and tracking, behavior analysis and prediction, decision making and planning, speed and steering control.

In our initial work [1], we focused on the perception, sensor fusion and tracking modules along with a naive behavior analysis and prediction method. The behavior prediction system was rule-based and could only predict binary outcome of changing or keeping lanes for a single leading vehicle.

In this paper, we focus on predicting intentions as well as trajectories of multiple exo-vehicles. Specifically, the predicted intentions help us distinguish between left and right lane changes for surrounding vehicles based on the learned driving behaviors. The driving intent is predicted using a recurrent neural network (RNN) instead of a rule-based method since, RNNs can capture subtle urban driving behaviors which otherwise cannot be independently represented by road-rules. On the contrary, RNNs are not very promising for trajectory prediction as they require a huge range of data to model various driving patterns. So, we predict the trajectory of the target vehicle based on its short-term tracked motion, the predicted intention and the road-contextual information. This allows us to predict trajectories without explicitly learning a range of driving patterns a-priori. Hence, our proposed context-aware prediction method can provide an early intention and trajectory prediction with reasonably acceptable time horizon in the future.

2 Related Work

There is an extensive amount of literature related to intention and trajectory prediction for motion planning of ego-vehicle [2]. Specifically, in [3] the authors propose a method adapted based on structural RNN to predict lane change intention of exo-vehicles. Unlike our approach, their method is specifically applicable for highway driving and does not predict the trajectories of the vehicles. Another method based on RNN is proposed in [4] for trajectory prediction, also for highway driving. Though the reported results are encouraging, this method requires a large amount of labeled data to model a wide-range of possible trajectories. Ding et al. [5] also propose a neural network based model for trajectory prediction. They address the data dependency issue by using a simulator and recording implicit driving behaviors of human subjects while they controlled the simulated vehicles. In our work, we use a neural network as well but to only predict the intention of the surrounding vehicles. The trajectory is predicted based on past vehicle states. We also use road contextual information which reduces the need of huge amount of data for learning intentions.

A real-time trajectory planning approach is also proposed by Li et al. in [6]. They present a hierarchical motion planning framework comprising of a high-level behavior planner to generate a coarse reference path and a low-level trajectory planner providing locally feasible candidate paths. The candidate paths are then evaluated using an objective function to select the optimal collision-free, smooth and dynamically feasible path. Their proposed approach, unlike our method, does not account for the intentions of the surrounding vehicles and instead passively plans paths in the available collision-free space. In [7,8] the human driving intentions and its uncertainties are taken into account for motion planning of the ego-vehicle. However, this work particularly considers the decision making task of navigating at the intersections which does not require any trajectory prediction.

In this paper, we aim to combine intention and trajectory prediction methods for driving behavior analysis in an urban environment. Similar to our goal, Galceran et al. [9] propose an integrated approach of inference and decision making for autonomous driving. They use hand engineered driving behaviors for model fitting trajectories and detecting change point for most likely behavior prediction. In contrast, our proposed approach does not use predefined trajectories but instead uses RNN for intention prediction and an online polynomial model fitting algorithm for trajectory prediction which is motivated by [10].

3 System Overview

The goal of this paper is to provide real-time prediction of intentions and trajectories of surrounding vehicles such that we can accordingly plan path for the ego-vehicle. Given the individual trajectory of each vehicle i, from the sensor fusion and tracking module, for recent time steps $\tau_i = t_1, ..., t_n$, we use it as input to intention and trajectory prediction algorithms. The intention prediction is obtained using a RNN. Specifically, we propose to use a RNN comprising of a single long-short term memory (LSTM) unit to represent and predict the driving intention I_i of the target vehicle for one-time step in the future. The LSTM outputs a probability distribution over the set of intentions which in our case is lane keeping, right lane change and left lane change. The trajectory of the target vehicle is then predicted using a polynomial curve which is obtained using the vehicle's initial and expected final states. The expected final state is inferred based on the intention predicted by LSTM.

4 Technical Approach

In order to validate our proposed approach we train and test our system based on the data collected from the ego-vehicle. This allows us to have access to the ground truth information which is independent of any perception errors. Our proposed approach has much less dependency on the range of training data as we enhance it with the road contextual information.

4.1 Road Context Information

We obtained the road contextual information from OpenStreetMap [11] and the lane center information was later modified according to our autonomous vehicle dynamics. An overlay of the road geometry information illustrating the lane center and direction is presented in Fig. 2 (left) along with an instance of ego-vehicle's path while it is changing lanes (right). The lane change path indicated in red is referenced against the lane center marking represented in yellow.

4.2 Intention Prediction

We use urban driving data collected from our ego-vehicle for training and testing the intention prediction module of our system. The data comprises of vehicle's state and road contextual information which are the input features for the LSTM network. Specifically, these features include:

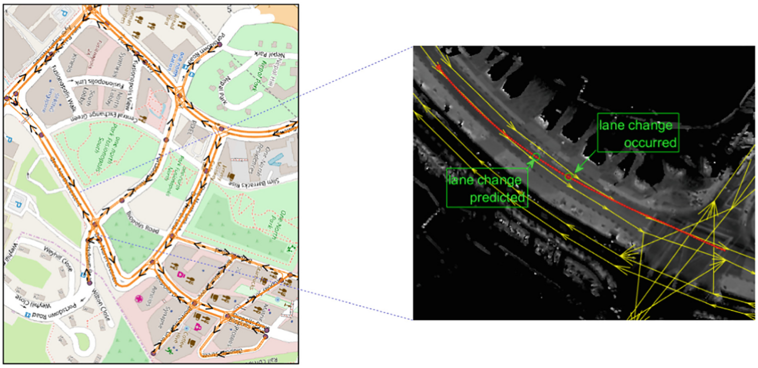


Fig. 2. The digital map overlaid with the lane center in yellow (left). A close-up of the digital map illustrating lane-change instance in red (right).

- δx: Change in lateral pose compared to previous time step
- δy: Change in longitudinal pose compared to previous time step
- ll: If left lane exists (Boolean 0/1)
- rl: If right lane exists (Boolean 0/1)
- d_{center}: Distance from the center of the lane. This value ranges from -1.5 m to 1.5 m for an average lane width of 3 m.

The LSTM network is trained on sample sequences of 1 second comprising of 4 samples. The network outputs predictions in real-time for 1 time step i.e., 0.25 s in future. The LSTM output is converted into a probability distribution over three classes: lane keeping, right lane change and left lane change.

4.3 Trajectory Prediction

Once the intention of the vehicle is known, the next critical step is to predict its trajectory in order to plan path for the ego-vehicle. The trajectory that a vehicle will take in a particular instant is stochastic in nature. First of all, different drivers might have different driving behaviors. For example, some drivers are more aggressive while others are more conservative. This affects how fast the lane change is performed and how the vehicle slows down when approaching an intersection. Second, even for the same driver, those behaviors might change depending on the situation of the road. However, given the road structure, intention of the driver and the past poses of the target vehicle, it is possible to reasonably predict its most likely trajectory.

Given the road contextual information, we know the adjacent lanes that the exo-vehicle can potentially transit to and thus we can generate possible hypotheses accordingly. For example, if the exo-vehicle travels on the left lane of a two lane road, it either continues on its current lane or makes a right lane change. In other words, left lane change is not possible and there are only two hypotheses (lane keeping and right lane change). For each hypothesis, wewill model the lateral component by a 5th-order polynomial and longitudinal component by a 4th-order polynomial as suggested in [10]. The polynomial is solved by providing the initial and final states of the target vehicle. The lateral displacement $d(t)$ is of the form:

$$d(t) = c_5t^5 + c_4t^4 + c_3t^3 + c_2t^2 + c_1t + c_0 \tag{1}$$

where $c_{i,i=\{0,1,2,3,4,5\}}$ are coefficients. Given a start time, $t_0 = 0$, a predefined end time $t_1 = 5$, initial and final states, the coefficients are solved using (2).

$$\begin{bmatrix} t_0^5 & t_0^4 & t_0^3 & t_0^2 & t_0 & 1 \\ t_1^5 & t_1^4 & t_1^3 & t_1^2 & t_1 & 1 \\ 5t_0^4 & 4t_0^3 & 3t_0^2 & 2t_0 & 1 & 0 \\ 5t_1^4 & 4t_1^3 & 3t_1^2 & 2t_1 & 1 & 0 \\ 20t_0^3 & 12t_0^2 & 6t_0 & 2 & 0 & 0 \\ 20t_1^3 & 12t_1^2 & 6t_1 & 2 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} c_5 \\ c_4 \\ c_3 \\ c_2 \\ c_1 \\ c_0 \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ \dot{d}_0 \\ \dot{d}_1 \\ \ddot{d}_0 \\ \ddot{d}_1 \end{bmatrix} \tag{2}$$

The initial state is used to determine d_0, $\dot{d}_0$ and $\ddot{d}_0$. The final displacement, d_1 is chosen depending upon the hypothesis whereas, vehicle's final lateral speed and acceleration $\dot{d}_1$ and $\ddot{d}_1$ are both zero. The longitudinal displacement $s(t)$ is of the form:

$$s(t) = c_4t^4 + c_3t^3 + c_2t^2 + c_1t + c_0 \tag{3}$$

Similar to (1), the coefficients c_i are solved using (4), given start time, $t_0 = 0$, a predefined end time $t_1 = 5$, initial and final states.

$$\begin{bmatrix} t_0^4 & t_0^3 & t_0^2 & t_0 & 1 \\ 4t_0^3 & 3t_0^2 & 2t_0 & 1 & 0 \\ 4t_1^3 & 3t_1^2 & 2t_1 & 1 & 0 \\ 12t_0^2 & 6t_0 & 2 & 0 & 0 \\ 12t_1^2 & 6t_1 & 2 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} c_4 \\ c_3 \\ c_2 \\ c_1 \\ c_0 \end{bmatrix} = \begin{bmatrix} s_0 \\ \dot{s}_0 \\ \dot{s}_1 \\ \ddot{s}_0 \\ \ddot{s}_1 \end{bmatrix} \tag{4}$$

The initial state is used to determine s_0, $\dot{s}_0$ and $\ddot{s}_0$. The final displacement, s_1 is unknown and is not required in the equation. The vehicle's longitudinal speed and acceleration are chosen to be $\dot{s}_1 = \dot{s}_0 + \ddot{s}_0 * t$ and $\ddot{s}_1 = \ddot{s}_0$, respectively.

5 Experimental Results

We train and test our prediction algorithms using our self-driving car in real urban environment. Specifically, we use the circuit represented in Fig. 2 which is situated in One-North region of Singapore city. The training set was collected for ≈8 km and tested for 2.3 km over the same path. The perception, sensor fusion and tracking modules are used from our previous work [1]. In this section, we exclusively discuss the intention and trajectory prediction results for the ego-vehicle from the behavior analysis and prediction module of our system as illustrated in Fig. 1. The test on ego-vehicle was specifically chosen to validate our proposed prediction algorithms without any influence of the perception and ground-truth data labeling errors.

5.1 Intention Prediction

We compare intention prediction results from LSTM with our rule-based method from [1]. For the rule-based method, we considered the distance to lane center and lateral velocity of the vehicle to decide on the probability of lane change. It is important to note that in our previous work we did not distinguish between left and right lane changes. However, LSTM based intention prediction can uniquely classify the two lane changes, but we combined their probabilities for a fair comparison to the rule-based method. We compare and report the precision-recall values for the two methods in Table 1. The threshold values for the two methods were independently selected based on the ROC curve shown in Fig. 3.

Table 1. Comparison of rule-based method with our proposed LSTM based approach for intent prediction.

	Class	Precision	Recall
LSTM	Lane keep	0.97	0.80
	Lane change	0.34	0.80
Rule-based	Lane keep	0.93	0.75
	Lane change	0.24	0.57

It is evident from Table 1, that LSTM based method provides significantly better precision recall rate. In addition to these performance metrics, we also calculated the average predicted time to lane change (TTLC) for the test data. The TTLC is calculated as the time difference between when a prediction is made and when the vehicle actually crossed the dividing line between the two lanes. The average TTLC provided by rule-based method is 3.25 s and using LSTM it is 3.69 s (≈14% better).

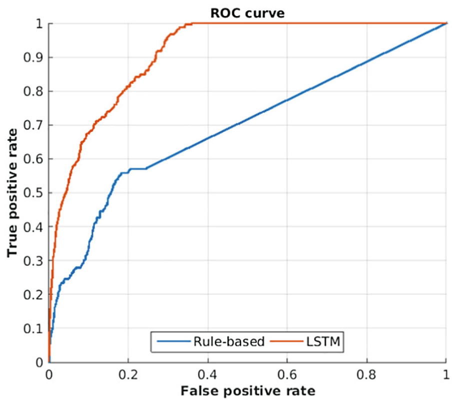
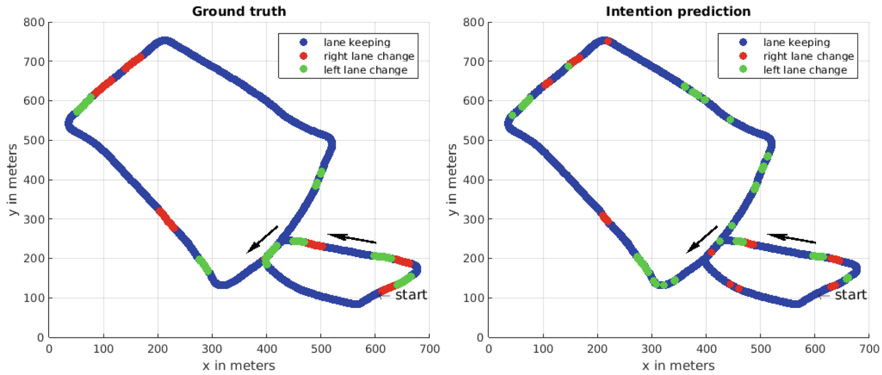


Fig. 3. ROC curve to compare rule-based method and LSTM.

The intention prediction of the ego-vehicle using LSTM for three classes (lane keeping, right lane change and left lane change) is presented in Fig. 4. The ground-truth information is represented in Fig. 4a which is obtained from the indicator values of the ego-vehicle. The corresponding predicted values from LSTM are illustrated in Fig. 4b.



(a) Ground truth

(b) Predicted intention

Fig. 4. Intention prediction of ego-vehicle using LSTM.

Note that there are some delays in prediction due to the fact that the indicator signals are given even before the intended maneuver is started. We also observed that slight off-centering of the vehicle from the lane center often triggers a lane change prediction, resulting in a false positive. This particular outcome can be tuned by varying the threshold for the probability of lane change based on the ROC from Fig. 3. Similar results were observed for exo-vehicle and an instance is presented at the following link: https://youtu.be/1GZMFLk5bk4.

5.2 Trajectory Prediction

An example of the ego-vehicle performing lane change on a two lane road is shown in Fig. 5. The road boundaries and the road separator are indicated using grey solid and dashed line, respectively. The positions marked as stars are the known vehicle positions in the last 5 sampling times. The goal is to predict the likely trajectories from the last sampling point. Given the road structure, we know that the vehicle could either keep lane or lane change. These two trajectories are indicated as dotted purple line and solid orange line for lane keeping and lane changing, respectively. The duration of prediction is 5 s into the future at sampling interval of 0.25 s. A time varying instance of the predicted vehicle trajectories for the two intents is presented at the following link: https://youtu.be/9-8kmSBtMEo.

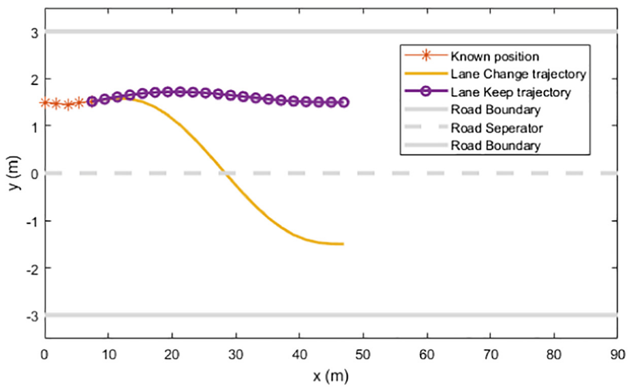


Fig. 5. Two hypothesis trajectories: lane keep and lane change for a vehicle traveling on a two lanes road.

Figure 6 shows the accuracy of the trajectory prediction algorithm for three different look-ahead times: 0.25 s, 1.0 s and 2.0 s. It is important to explain what it means by look-ahead time. If the look-ahead time is 1.0s, for example, we are trying to predict the vehicle's position at $t + 1.0\,\mathrm{s}$ given only the vehicle's position until time t. Since there are two potential trajectories to choose from as shown in Fig. 5, we choose the trajectory based on the driver's intention given by the LSTM network.

For 0.25 s look-ahead, the trajectory prediction is perfectly in-line with the predicted path, indicating that the intention of the driver is inferred correctly for this short look-ahead time. The motion model used to generate the lane keeping and lane changing hypotheses, also managed to predict the position of the vehicle accurately, as vehicles generally maintain their momentum at this short interval.

Similarly for 1.0 s look-ahead, the LSTM network is able to accurately predict the instance of lane change intention of the driver. However, there are some position errors while predicting the trajectory as the motion model needs to predict longer into the future. Lastly, for 2.0 s look ahead, there is a noticeable error in trajectory prediction at the start of the lane change maneuver which is caused due to the delay in predicting the lane change intention. In addition to that, there are errors due to the motion model as there is more uncertainty now to predict position 2.0 s into the future.

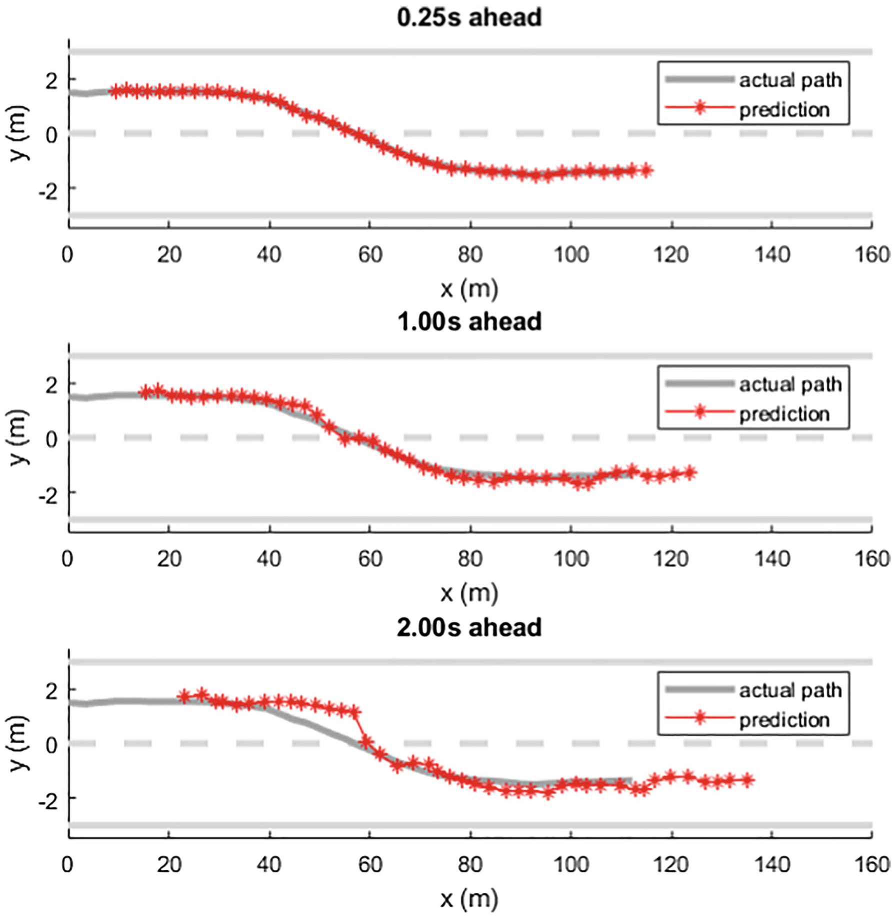


Fig. 6. Predicted trajectory for different look ahead times: 0.25 s, 1 s and 2 s. The thick grey line shows the actual path.

6 Conclusion

The results from our intention and trajectory prediction algorithms seem promising without requiring big datasets that are generally needed by the existing deep learning prediction methods. Instead, we use road contextual information which helps us to significantly improve our prediction outcome. We validated our proposed LSTM based intention prediction method which outperforms the rule-based method without explicitly accounting for anomalous scenarios. Our trajectory prediction method has proven to provide real-time results with very low error rates. Lastly, the integrated output of our intention and trajectory prediction algorithms is shown to be applicable for decision making and path planning by the autonomous vehicles in urban driving environments [12].

Acknowledgment. This research was supported by the National Research Foundation, Prime Minister's Office, Singapore, under its CREATE programme, Singapore-MIT Alliance for Research and Technology (SMART) Future Urban Mobility (FM) IRG. We would also like to acknowledge the support of NVIDIA Corporation through its NVAIL program and NUS-NVIDIA MoU.

References

1. Verma, S., Eng, Y.H., Kong, H.X., Anderson, H., Meghjani, M., Leong, W.K., Shen, X., Zhang, C., Ang, M.H., Rus, D.: Vehicle detection, tracking and behavior analysis in urban driving environments using road context. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE (2018)
2. Paden, B., Čáp, M., Yong, S.Z., Yershov, D., Frazzoli, E.: A survey of motion planning and control techniques for self-driving urban vehicles. IEEE Trans. Intell. Veh. **1**(1), 33–55 (2016)
3. Patel, S., Griffin, B., Kusano, K., Corso, J.J.: Predicting future lane changes of other highway vehicles using RNN-based deep models. arXiv preprint arXiv:1801.04340 (2018)
4. Altché, F., De La Fortelle, A.: An LSTM network for highway trajectory prediction. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), pp. 353–359. IEEE (2017)
5. Ding, C., Wang, W., Wang, X., Baumann, M.: A neural network model for driver's lane-changing trajectory prediction in urban traffic flow. Math. Probl. Eng. **2013**, 8 (2013)
6. Li, X., Sun, Z., Cao, D., He, Z., Zhu, Q.: Real-time trajectory planning for autonomous urban driving: framework, algorithms, and verifications. IEEE/ASME Trans. Mechatron. **21**(2), 740–753 (2016)
7. Bandyopadhyay, T., Won, K.S., Frazzoli, E., Hsu, D., Lee, W.S., Rus, D.: Intention-aware motion planning. In: Algorithmic Foundations of Robotics X, pp. 475–491. Springer (2013)
8. Sezer, V., Bandyopadhyay, T., Rus, D., Frazzoli, E., Hsu, D.: Towards autonomous navigation of unsignalized intersections under uncertainty of human driver intent. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3578–3585. IEEE (2015)
9. Galceran, E., Cunningham, A.G., Eustice, R.M., Olson, E.: Multipolicy decision-making for autonomous driving via changepoint-based behavior prediction: theory and experiment. Auton. Robots **41**(6), 1367–1382 (2017)
10. Houenou, A., Bonnifait, P., Cherfaoui, V., Yao, W.: Vehicle trajectory prediction based on motion model and maneuver recognition. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4363–4369. IEEE (2013)
11. OpenStreetMap contributors. Planet dump. https://planet.osm.org, https://www.openstreetmap.org (2017)
12. Meghjani, M., Luo, Y., Ho, Q.H., Cai, P., Verma, S., Rus, D., Hsu, D.: Context and intention aware planning for urban driving. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2019)