Inferring Human Driver Intent in Partial Deployment of Connected Autonomous Vehicles: the Lane Change Case

Jonghwan Na, Hojeong Lee, and Hyogon Kim

Abstract—Co-existence between autonomous vehicles (AVs) and human-driven vehicles expected in the next few decades poses a problem for AVs to infer human drivers' intents nearby and cope with them safely and efficiently. To address this issue, we develop a light-weight deep learning model for a connected autonomous vehicle (CAV) to infer intents in a safety-critical case of lane changes made by human-driven vehicles. Through experiments with the real trajectory dataset NGSIM, we show that a simple Multi-Layer Perceptron (MLP) model can predict lane change events with high accuracy comparable with more sophisticated models. The model is intentionally designed to work with the simplest 3-vehicle topology to foster real-time execution on the resource-constrained computing platforms on AVs. Still, the model achieves 85% accuracy over 5 to 8 seconds prediction horizons so that AVs can have enough time to prepare for an upcoming lane change event.

Index Terms—Connected Autonomous Vehicles, partial deployment, V2X, human driver, intent, lane change, deep learning

I. INTRODUCTION

TUMAN-driven vehicles will continue to exist alongside autonomous vehicles (AVs) for the next few decades [1]–[3]. Because of the long transition period, AVs must be designed to be able to infer human drivers' intents for the sake of safer and more efficient driving. Unfortunately, it is a difficult task - even human drivers themselves sometimes have a hard time figuring out what other drivers plan to do [4]. In this paper, we consider the case of lane change, a highly critical maneuver, intended by human drivers. The significance of correctly inferring the lane change intent is evident in the statistics. Lane change accounts for nearly 5% of crashes and 7% of crash fatalities in the road [5], [6]. This paper investigates how AVs can predict the maneuver early on, if both AVs and human-driven vehicles are equipped with vehicle-to-everything (V2X) communication capability to exchange their kinematics information with each other. In the rest of the paper, we will refer to the former as connected autonomous vehicles (CAVs) and the latter as connected vehicles (CVs).

Note that CVs do not, and in fact cannot, convey through V2X communication the intent of their human driver *per se*. Although "Day 3" scenarios address intent broadcast [7], [8], it pertains to CAVs, not CVs. Therefore, the human driver's intent in CVs can only be inferred from the kinematics information such as position and speed that is periodically broadcast from the CVs through V2X in Basic Safety Message

(BSM) [9] or Cooperative Awareness Message (CAM) [10]. A CAV infers through the collected kinematics information from the set of vehicles that can be involved in the lane change maneuver whether a human-driven CV will make the move within a prediction time horizon. If so, the CAV will be able to prepare for it, for instance, by creating sufficient space in its front by slowing down for the incoming human-driven vehicle.

To learn the topological relation among involved vehicles in a lane change event, we employ a simple Multi-Layer Perceptron (MLP) model. It works with a 3-vehicle topology that comprises of the CAV, the lane-changing CV, and the vehicle in front of the CAV. The reason not to employ a more sophisticated model is twofold. First, a simpler model will be more advantageous for cheaper implementation and for realtime execution on a typically resource-constrained embedded computing platform like autonomous vehicles. Second, it will be more robust to find the vehicles for the input topology in various traffic situations than larger models that need more vehicles to define the input features. To train the 3-vehicle MLP model, we utilize the publicly available Next Generation Simulation (NGSIM) dataset [11] where 652 lane changes occurs in a span of 15 minutes. The trained model will be shown to predict the lane change made by a human-driven CV with approximately 85% accuracy for the dataset.

NGSIM was published by US Federal Highway Administration and has been the most extensively used dataset in literature. The entire NGSIM dataset contains more than 5,000 vehicle trajectories. Each entry in the original NGSIM dataset records a vehicle's coordinate, lane number, following vehicle's ID, preceding vehicle's ID, vehicle type, speed, and acceleration among others. They are produced from the captured video frames taken at 10 Hz. Consequently, the trajectory of each vehicle captured in the video has the time resolution of 10 Hz. Because the BSMs or CAMs also convey the vehicle dynamics information up to 10 times a second, the NGSIM data is a useful replacement of the future V2X messages that will be exchanged among proximal vehicles within the communication range.

In this paper, we do not attempt to predict the lane changing CV's trajectory itself. Instead, we aim to predict whether there will be a lane changing event attempted by a human driven vehicle to the front of the CAV. This is because our target platform is CAVs. The CAVs will be equipped with an army of sensors (e.g. radars, cameras, and lidars) to monitor the

unfolding lane change event and with the controls to manage the CAV trajectory during the maneuver, once the prediction is made. Therefore, we only task our model to provide a prediction early and precisely enough. The design decision to predict the lane change event itself but not its detailed trajectory lead to the high prediction accuracy even though the model is light-weight. The contributions of this paper are as follows:

- We define a minimal construct that describes a lane change maneuver, composed of only three vehicles directly involved hence a small set of features.
- We provide a light-weight neural network model that produces the accuracy comparable to more complex models, which would facilitate the real-time operation and low-cost deployment on vehicular platforms.

The rest of the paper is organized as follows. Section II introduces related work in the human driver intent prediction in a lane change event. Section III presents the solution approach. In particular, how we process the NGSIM dataset to obtain the training and test data is discussed. Also, the MLP model is described and how it is trained and validated is also discussed. Section IV presents the prediction performance of the MLP model. Finally, Section V concludes the paper.

II. RELATED WORK

Predicting human intents in the context of driving is inherently difficult. Quintanar *et al.* [4] showed that even human drivers have difficulty anticipating lane-change maneuvers, with most detecting them only after they have started. The authors suggested that artificial intelligence may outperform human skills by analyzing hidden cues that go unnoticed, improving the detection time.

Shou *et al.* [12] proposed a prediction model without using any lateral or steering angle information that would allegedly shorten the prediction horizon. A MLP model augmented with prediction stabilization heuristic was shown to capture 75% of real lane change maneuvers with an average prediction horizon of 8.05 seconds. This work also showed that logistic regression and recursive neural network (RNN) do not perform as well. The MLP model used in Shou *et al.* [12], however, can only predict left lane changes. The reason is that the left lane change driving scenario consisted of around 88% of all cases. Mozaffari *et al.* [13] extended the model to predict right lane changes as well. The same MLP model was used and the accuracy of the model decreased to 59% at a prediction horizon of 5 seconds.

Wirthmüller *et al.* [14] proposed a method where autonomous vehicles anticipate the behavior of other drivers by observing the environment as well as other vehicles in traffic. A trajectory planning algorithm was used to model future vehicle positions and maneuvers of up to 5 seconds in advance. A MLP model with one hidden layer was used and the model achieved an accuracy of 75%. Wirthmüller *et al.* [15] also developed a Long Short-Term Memory (LSTM)-based model which estimates the time a lane change happens. The authors argue that this information is more useful than predicting

whether a lane change will occur. The model achieves an accuracy of 79% over a prediction horizon of 5 seconds.

Zhang et al. [16] proposed to exploit both spatial and temporal features in the trajectory changes of vehicles to classify various driving events including lane changes. The authors employed Multi-Scale Convolutional Neural Network (MSCNN) and Bi-directional LSTM (Bi-LSTM) and obtained a prediction performance at around 85% for most of the studied events. The use of Random Oversampling (ROS) boosted the accuracy to at around 90%. Mozaffari et al. [13] also used a CNN model to detect lane change scenarios. The spatial attention used in CNN helped enable the model to obtain an accuracy of 83% over a 5.2 prediction time horizon. Note that CNN-based methods require a visual image of the driving environment, but it will be practically difficult to obtain with line-of-sight (LoS) sensors due to obstructed view. For this reason, our work relies on a model that relies only on V2X communication.

III. SOLUTION APPROACH

To formulate the lane change prediction as a classification problem, we define two classes for a forward human-driven CV in an adjacent lane of an ego CAV: $0 \equiv$ "change lanes (to my front position) within t seconds" and $1 \equiv$ "otherwise." Upon class 0, the ego CAV should slow down if necessary to make a safe room for the CV to cut in. For the classification, the CAV uses the features of only three vehicles: ego, the CAV in the adjacent lane (adj) and the current front vehicle (fro) as depicted in Fig. 1(a).

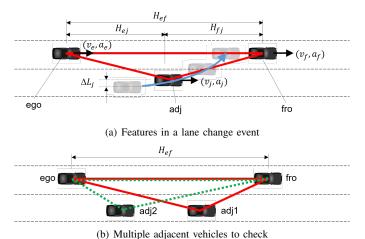


Fig. 1. 3-vehicle topology (the triangle)

If there are multiple adj's within the span H_{ej} , each such adj creates the 3-vehicle relation with ego and fro to be checked for the lane change prediction (Fig. 1(b)). The simple 3-vehicle model is easy to define in most traffic situations and makes the size of the MLP small.

A. NGSIM dataset

The Next Generation Simulation (NGSIM) I-80 vehicle trajectories dataset was captured in Emeryville, CA, on April

13, 2005 [11]. It consists of 45 minutes of trajectory data that was segmented into three 15-minute periods. The 17:15–17:30 period is selected for our study as the traffic was most congested hence with an increased risk of collision when performing a lane changing maneuver [17]. The study area spans approximately 500 meters and consists of six lanes. Seven video cameras were mounted on the top of a 30-story building adjacent to the freeway to record vehicles passing through the study area. A customized software was then used to obtain the vehicle trajectory data from each video frame taken every 100 ms. Each frame can have multiple data entries each of which corresponds to a single vehicle moving in the study area at the time of the recording. Table I shows the structure of an entry for a particular vehicle in a video frame.

Feature	Explanation		
Vehicle ID	ID assigned to each vehicle trajectory		
Frame ID	Frame number where a vehicle appeared		
Total Frame	Number of frames vehicle appears in		
Local X	Lateral coordinate of the vehicle		
Local Y	Longitudinal coordinate of the vehicle		
Length	Length of the vehicle in feet		
Width	Width of the vehicle in feet		
Class	1 = motorcycle, 2 = auto, 3 = truck		
Velocity	Velocity of the vehicle in ft/s		
Acceleration	Acceleration of the vehicle in ft/s ²		
Lane ID	Lane number of the vehicle		
Preceding	Vehicle ID of the leading vehicle		
Following	Vehicle ID of the following vehicle		
Space Headway	Distance to the leading vehicle		
Time Headway	Duration to the leading vehicle		

B. Data preprocessing

1) Labeling: Initially, every entry in the dataset is given the negative (1) label. Then, we find the video Frame ID = F where a lane change occurred for Vehicle ID = V whose Lane ID has changed since F-1. We change the label of V's entry at F to positive (0). In addition, to enable the prediction of the lane change event, we positively label of V from $F-t\cdot 10$ to F-1. Fig. 2 depicts these two steps. Later, we will explore different values for the prediction horizon t, up to 20 seconds.

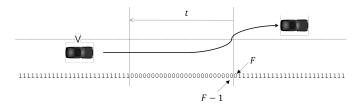


Fig. 2. Labeling method: 0 is positive (imminent lane change) and 1, negative

Naturally, the 0 entries are relatively rare in comparison to 1-labeled entries. For example, at t=5, there are 973,269 entries labeled 1 but only 13,800 labeled 0. We need to balance the 0-labeled and 1-labeled entries used in the model training for higher classification performance. Thus, we randomly sampled only as many 1-labeled entries as there are 0-labeled entries to train our model.

2) Feature selection: Between 17:15–17:30, a total of 1,754 vehicles are recorded in the NGSIM dataset. For each these ego vehicles, their fro is given in the NGSIM dataset ("Preceding" in Table I). However, their adj has to be computed by pre-processing the dataset. Specifically, the adj is identified by the condition $y_e \leq y_j \leq y_e + H_{ef}$ where y_e is the longitude of ego ("Local Y" in Table I) and y_j is the longitude of adj on the adjacent lane in consideration and H_{ef} is the longitude difference of ego to fro (see Fig. 1). For the 3-vehicle topology defined for each ego, we extract or derive features from the NGSIM dataset to associate with label, which are listed in Table II.

TABLE II
SELECTING INPUT FEATURES IN 3-VEHICLE TOPOLOGY

Vehicle	Notation	Feature	Comp. over [Fr.	p-value	Input
ego	v_e	Velocity	-	≈ 0	Y
	a_e^{10}	Acc.	10	≈ 0	Y
	v_f	Velocity	-	≈ 0	Y
fro	a_f^1	Acc.	1	≈ 0	Y
,,,	a_f^{10}	Acc.	10	≈ 0	Y
	ΔH_{fj}^{10}	Dist. change to adj	10	0.001	Y
	v_j	Velocity	-	0.009	Y
adj	ΔL_i^{10}	Latitude change	10	≈ 0	Y
	a_j^{10}	Acc.	10	0.005	Y
ego	a_e^1	Acc.	1	0.162	N
ego	ΔH_{ej}^{10}	Dist. change to adj	10	0.201	N
adj	a_j^1	Acc.	1	0.057	N

Notice that acceleration is computed over either 10 frames or 1 frame and distance/latitude changes are computed over 10 frames. The velocity is directly imported from the NGSIM data. Note that for adj, its latitude change ΔL_j is particularly included as a feature because one of the cues that a human driver uses when judging a neighboring vehicle's intent to make a lane change is how quickly it moves towards her lane. To derive it in the NGSIM dataset, we compute the lateral movement during the last second by $\Delta L_j = L_j[F] - L_j[F-10]$. In our convention, a negative value of ΔL represents a lane change event happening from the right lane, and a positive, from the left. The latitude change of fro and ego are not considered as features since they are driving in the same lane during the lane change event for adj.

In order to apply only relevant features as input to the prediction model, we conduct a logistic regression analysis and remove the features irrelevant to the output label. Logistic regression can estimate the relationship between the label (dependent variable) and one or more input features (independent variables). Null hypothesis H_0 states that a given feature in an entry has no statistical relationship with the label of the entry. The condition $p \leq 0.05$ is used to reject H_0 . We found that ΔH_{ej} , a_e^1 , and a_j^1 have p > 0.05 as seen in Table II, thus we removed them from the input features set.

The 9 features can be defined for every 3-vehicle topology and can be used to consult the MLP model as to a potential lane change in the given time horizon. We believe that the 3-vehicle topology strikes the balance between the model robustness and performance. It is more robustly defined than higher-

complexity models that require the presence of more vehicles to obtain the input feature set. Moreover, we found that adding more vehicles to the topology does not lead to a significant improvement on the performance. By comparing our model with a model using a 4-vehicle topology [12] and an 8-vehicle topology [14], although not shown for space, we found that the 3-vehicle model outperforms them despite having less vehicles in the topology. On the other hand, removing the front vehicle to create a 2-vehicle model with only *ego* and *adj* may be even more easily defined, but the performance degradation is significant. We found that it decreases the prediction accuracy by as much as 6%. Therefore, we use the 3-vehicle topology.

Finally, with each ego who has the 9-feature values computed as in Table II, the label for V=adj as computed in the manner of Fig. 2 is associated. This set of $< features,\ label >$ pairs will be used in the training and test below.

C. Model training

A Multi-Layer Perceptron (MLP) is a fully connected feedforward neural network. For training, MLP uses backpropagation to adjusts connection weights. Our MLP model has 3 hidden layers with 100, 500, and 100 neurons, respectively (Fig. 3). Having fewer hidden layers resulted in poor model convergence, and adding more layers to the network did not improve on the prediction accuracy. The activation function is Rectified Linear Unit (ReLU). The input is the 9-feature vector $\mathbf{X} = [x_0, x_1, \dots, x_8]$ as listed in Table II. The output $\mathbf{Y} = [y_0, y_1]$ from the prediction model is binary; y_0 is 1 if the adj vehicle will change lanes to the front of ego vehicle within the prediction horizon of t seconds; otherwise, y_1 is 1. The parameter t could be set according to the given safety application such as the lane change. We will show below how the prediction accuracy changes as a function of t.

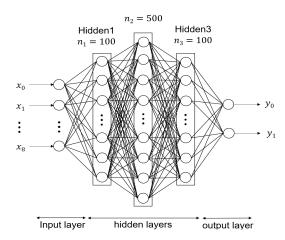


Fig. 3. MLP model with 9 inputs, 3 hidden layers, and 2 outputs

To train the MLP model, we shuffled the < features, label > pair set and split them into training, validation, and test subsets at the ratio of 0.7:0.15:0.15. We used the cross-entropy loss because we tackle a classification problem, and Adam as the optimizer. We also used a batch

size of 30 and stopped training the model at the $40^{\rm th}$ epoch where the model began to overfit as shown in Fig. 4.

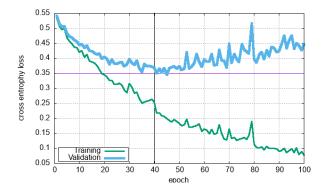


Fig. 4. Training and validation losses over 100 epochs

Finally, each classification operation takes approximately 1 ms on a typical PC environment. Because each ego vehicle can check for an adj vehicle for their potential lane change move at most 10 times a second, it is a minimal overhead even for a less powerful, embedded platform for a few adj vehicles to check in the span of H_{ej} (see Fig. 1(b)).

IV. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the 3-vehicle MLP model in predicting the lane change intent of human drivers as manifested in the NGSIM dataset. The performance metrics are precision (PRE), recall (REC), F1-score (F1), and accuracy (ACC). Fig. 5 presents the performance measures as a function of the length of the prediction horizon (t).

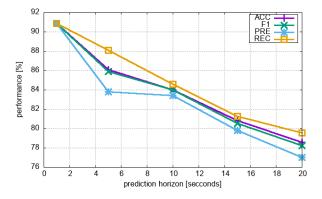


Fig. 5. Prediction performance of the 3-vehicle MLP model

Accuracy is a general measure of prediction performance. For more specific understanding, let us consider the other three measures. Precision is more relevant when the costs of false positives are high. False positives will negatively affect the traffic efficiency by causing *ego* to slow down unnecessarily. On the other hand, recall is more relevant when the costs of false negatives are high. In the safety applications like the lane change, recall is more important than precision ("better safe than sorry"). Fortunately, we observe in Fig. 5 that the recall is higher than precision with our 3-vehicle model. The

recall is over 80% for t < 20, which is in general considered a good performance for a recall. For shorter horizons, the recall is further increased. For instance, it is 88% for t=5 seconds. Finally, a good F1-score means low false positives and low false negatives, correctly identifying real threats and not disturbed by false alarms. The reason that F1-score and accuracy shows similar performance is that the 0-labeled and 1-labeled samples are balanced the classes for higher classification performance (see Section III-B1).

Fig. 5 also shows that the prediction quality naturally degrades as the CAV needs to look further into the future. For instance, the accuracy at t=20 seconds is only approximately 75%. However, it is reported that human drivers spend up to 6.6 seconds on average to prepare for a lane change and 1.5 seconds to execute it [18]. Therefore, assuming t=8 seconds seems sufficient for the prediction horizon for the CAV, the accuracy is 86%. Moreover, just before the actual execution phase that will produce a fast lateral movement by the lane changing vehicle, the prediction accuracy of the proposed model can exceed 90%.

Finally, Table III compares the performance of the proposed method with some existing proposals. We notice that the 3vehicle MLP model obtains the comparable performance as the CNN model [13] despite using a simpler MLP model. Note that [13] does not use the NGSIM dataset, so the comparison is not direct but rather informational. The proposed simple model is slightly outperformed only by a far more sophisticated model [16]. Again, [16] uses a different dataset that the comparison is informational. However, note that the visual road data as required by CNN-based models will not be readily available in practice. We believe that working with the vehicle dynamics information available through V2X communication lends more easily to the realization of the intent prediction mechanisms. Finally, the proposed model achieves a better performance than [12] even when using the same MLP model and the NGSIM dataset at a similar prediction time horizon.

TABLE III
COMPARISON OF THE PROPOSED MODEL WITH EXISTING PROPOSALS

Model	Horizon [s	ACC PRE REC F1
MLP [14]	5	0.75 0.94 0.65 0.77
MLP [12] [13]	5	0.59 0.74 0.52 0.61
LSTM [15]	5	0.79 0.75 0.90 0.82
Bi-LSTM+MSCNN [16]	5	0.85 - 0.85 0.83
Bi-LSTM+MSCNN+ROS [16]	5	0.90 - 0.90 0.88
CNN [13]	5.2	0.83 0.85 0.85 0.85
Proposed	5	0.86 0.84 0.88 0.86
Floposed	8	0.84 0.83 0.86 0.84

V. CONCLUSION

This paper shows that the lane change operation initiated by the human driver can be predicted with high accuracy from the vehicle dynamics information periodically broadcast in vehicle-to-vehicle communication. A simple but robust 3-vehicle topology classified by a MLP model can provide high prediction performance, affordable to run in real time on the embedded platform in connected autonomous vehicles.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2020R1A2C3011888).

REFERENCES

- H. Wang, S. Avedisov, T. Molnar, A. H. Sakr, O. Altintas and G. Orosz, "Conflict Analysis for Cooperative Maneuvering with Status and Intent Sharing via V2X Communication," *IEEE Transactions on Intelligent* Vehicles, doi: 10.1109/TIV.2022.3149796.
- [2] J. I. Ge, S. S. Avedisov, C. R. He, W. B. Qin, M. Sadeghpour, and G. Orosz, "Experimental validation of connected automated vehicle design among human-driven vehicles," *Transportation Research Part C*, vol. 91, pp. 335–352, 2018.
- [3] S. S. Avedisov, G. Bansal, and G. Orosz, "Impacts of connected automated vehicles on freeway traffic patterns at different penetration levels," *IEEE Transactions on Intelligent Transportation Systems*, 2021. [Online]. Available: https://doi.org/10.1109/TITS.2020.3043323
- [4] A. Quintanar, R. Izquierdo, I. Parra, D. Fernandez-Llorca, and M. A. Sotelo, "The PREVENTION Challenge: How Good Are Humans Predicting Lane Changes?" in Proceedings of IEEE Intelligent Vehicles Symposium (IV), October 20-23, 2020. Las Vegas, USA.
- [5] F. You, R. Zhang, G. Lie, H. Wang, H. Wen, and J. Xu, "Trajectory planning and tracking control for autonomous lane change maneuver based on the cooperative vehicle infrastructure system," *Expert Syst. Appl.*, vol. 42, no. 14, pp. 5932–5946, 2015.
- [6] C. Rodemerk, S. Habenicht, A. Weitzel, H. Winner, and T. Schmitt, "Development of a general criticality criterion for the risk estimation of driving situations and its application to a maneuver-based lane change assistance system," in Proceedings of IEEE Intelligent Vehicles Symposium (IV), pp. 264–269, 2012.
- [7] SAE, Taxonomy and Definitions for Terms Related to Cooperative Driving Automation for On-Road Motor Vehicles, J3216, Apr. 2021.
- [8] ETSI TR 103 578, Intelligent Transport Systems (ITS); Vehicular Communication; Informative Report for the Maneuver Coordination Service, May 2020, draft V0.0.5.
- [9] SAE, V2X Communications Message Set Dictionary, J2735, July 2020.
- [10] ETSI, Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service, EN 302 637-2, Apr. 2019.
- [11] US DoT, Next Generation Simulation (NGSIM). Available at: https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm.
- [12] Z. Shou, Z. Wang, K. Han, Y. Liu, P. Tiwari, and X. Di, "Long-Term Prediction of Lane Change Maneuver Through a Multi-layer perceptron," in Proceedings of IEEE Intelligent Vehicles Symposium (IV), October 20-23, 2020. Las Vegas, USA.
- [13] S. Mozaffari, E. Arnold, M. Dianati and S. Fallah, "Early Lane Change Prediction for Automated Driving Systems Using Multi-Task Attentionbased Convolutional Neural Networks," *IEEE Transactions on Intelligent* Vehicles, doi: 10.1109/TIV.2022.3161785.
- [14] F. Wirthmüller, J. Schlechtriemen, J. Hipp and M. Reichert, "Teaching Vehicles to Anticipate: A Systematic Study on Probabilistic Behavior Prediction Using Large Data Sets," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 11, pp. 7129-7144, Nov. 2021, doi: 10.1109/TITS.2020.3002070.
- [15] F. Wirthmüller, M. Klimke, J. Schlechtriemen, J. Hipp and M. Reichert, "Predicting the Time Until a Vehicle Changes the Lane Using LSTM-Based Recurrent Neural Networks," in IEEE Robotics and Automation Letters, vol. 6, no. 2, pp. 2357-2364, April 2021, doi: 10.1109/LRA.2021.3058930.
- [16] H. Zhang, Z. Nan, T. Yang, Y. Liu and N. Zheng, "A Driving Behavior Recognition Model with Bi-LSTM and Multi-Scale CNN," in Proceedings of IEEE Intelligent Vehicles Symposium (IV), October 20-23, 2020. Las Vegas, USA.
- [17] L. Li, D. Zhang, Z.-G. Xu, P. Wang, G.-P. Wang, "The Roles of Car Following and Lane Changing Drivers' Anticipations during Vehicle Inserting Process: A Structural Equation Model Approach", Journal of Advanced Transportation, vol. 2018, Article ID 6372861, 19 pages, 2018. https://doi.org/10.1155/2018/6372861
- [18] P. Finnegan and P. Green, "The Time to Change Lanes: A Literature Review," *Intelligent Vehicle-Highway Systems Technical Report 90-13*, U. of Michigan, Sept. 1990.