

A Comparative Study of Ego-centric and Cooperative Perception for Lane Change Prediction in Highway Driving Scenarios

Sajjad Mozaffari¹^a, Eduardo Arnold¹^b, Mehrdad Dianati¹^c and Saber Fallah²^d

¹Warwick Manufacturing Group, University of Warwick, Coventry CV4 7AL, U.K.

²Department of Mechanical Engineering Sciences, University of Surrey, Guildford, GU2 7XH, U.K.

Keywords: Lane Change Prediction, Perception Models, Vehicle Behaviour Prediction, Intelligent Vehicles.

Abstract: Prediction of the manoeuvres of other vehicles can significantly improve the safety of automated driving systems. A manoeuvre prediction algorithm estimates the likelihood of a vehicle's next manoeuvre using the motion history of the vehicle and its surrounding traffic. Several existing studies assume full observability of the surrounding traffic by utilising trajectory datasets collected by top-down view infrastructure cameras. However, in practice, automated vehicles observe the driving environment using egocentric perception sensors (i.e., onboard lidar or camera) which have limited sensing range and are subject to occlusions. This study firstly analyses the impact of these limitations on the performance of lane change prediction. To overcome these limitations, automated vehicles can cooperate in observing the environment by sharing their perception data through V2V communication. While it is intuitively expected that cooperation among vehicles can improve environment perception by individual vehicles, the other contribution of this work is to quantify the potential impacts of cooperation. To this end, we propose two perception models used to generate egocentric and cooperative perception dataset variants from a set of uniform scenarios in a benchmark dataset. This study can help system designers weigh the costs and benefits of alternative perception solutions for lane change prediction.

1 INTRODUCTION

Predicting the lane change (LC) manoeuvre of nearby vehicles enables automated vehicles to make proactive decisions and reduce the risk of collisions in highway driving. A vehicle's LC manoeuvre is highly dependent on the behaviour of other vehicles in its vicinity, particularly in congested highway traffic. For example, a slow-moving vehicle motivates its following vehicles to perform an LC manoeuvre, provided there is an available gap in the side lane. Therefore, an Ego Vehicle (EV) needs to observe the states (e.g., location and velocity) of the Target Vehicle (TV), i.e., the vehicle of interest, and its Surrounding Vehicles (SVs) during a time window to predict the next manoeuvre of the TV.

Automated vehicles observe their surrounding using egocentric perception sensors (e.g., camera and LiDAR) which have limited range and are subject to

spatial impairments such as occlusions. The limitations of egocentric perception can prevent tracking some of the SVs which might negatively impacts the LC prediction performance. Figure 1-a illustrates an example of a driving scenario where the EV aims to predict the manoeuvre of a leading TV. The TV is going to perform an LC manoeuvre shortly due to a slow-moving vehicle in front of the TV. However, the EV cannot observe the slow-moving vehicle since the TV is occluding the EV's perception sensors. Consequently, the EV cannot predict the upcoming LC manoeuvre by the TV.

The majority of the existing studies on LC prediction assumes full observability of the driving environment. In such studies, the prediction model is trained using LC scenarios extracted from trajectory datasets such as NGSIM (Colyar and Halkias, 2007; Halkias and Colyar, 2007) and highD (Krajewski et al., 2018). In these datasets, the driving environment is being observed using wide and top-down view cameras installed on infrastructure buildings or drones. Such perception assumption is not realistic for all driving scenarios since the large-scale deployment of infras-

^a <https://orcid.org/0000-0001-8109-6953>

^b <https://orcid.org/0000-0001-7896-7252>

^c <https://orcid.org/0000-0001-5119-4499>

^d <https://orcid.org/0000-0002-1298-1040>

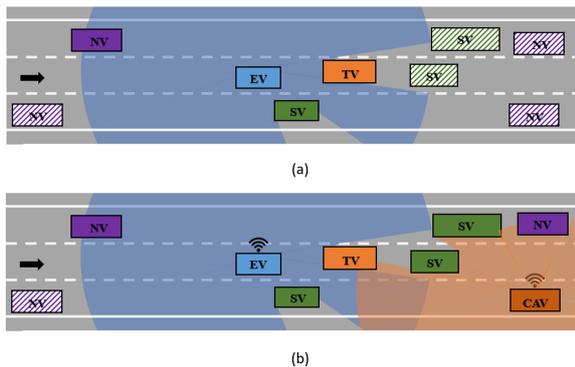


Figure 1: An illustration of a driving environment with (a) egocentric perception, and (b) cooperative perception. Observable areas are depicted with blue and orange colours for egocentric sensor and shared cooperative sensors, respectively. Unobservable vehicles are filled with hashed colours. Following terminology is used in this figure. EV: Ego vehicle, TV: Target vehicle, SVs: Surrounding Vehicles (they have a considerable impact on TV’s behaviour), NVs: Non-effective vehicles (their impact on TV’s behaviour can be neglected). CAVs: Connected Automated Vehicles.

structure sensors to cover all road sections incurs in high cost. Analysing the impact of ego-centric perception limitations on the performance of an LC prediction model has remained a research gap in the literature.

One practical solution for the limitations of egocentric perception is the cooperation of nearby Connected Automated Vehicles (CAVs) in observing the environment through V2V communication. V2V communication has been used to improve the performance of different automated driving applications such as road geometry estimation (Sakr et al., 2017), 3D object detection (Chen et al., 2019), and trajectory planning (Kim et al., 2015). However, sharing perception data through V2V for the task of LC prediction has not been considered in the literature. To investigate the effectiveness of cooperative perception for LC prediction, a large-scale dataset recorded by perception sensors of multiple vehicles driving simultaneously on a road section is needed for training and evaluating the prediction model. To the best of our knowledge, such a cooperative perception dataset does not currently exist. One way to address this problem is to use synthetic data. However, such synthetic datasets with simplified driving behaviour models fail to replicate the naturalistic behaviour of human-driven vehicles. The lack of naturalistic trajectory datasets recorded with cooperative perception is a gap in the literature that prevents further research on this topic.

To address the aforementioned research gaps, this paper carries out a comparative study on the impact

of two perception modes, namely egocentric perception, and cooperative perception on LC prediction performance. In doing so, this paper proposes two perception models used to generate dataset variants with egocentric and cooperative perception from a naturalistic trajectory dataset with full observability (i.e., captured from wide and top-down view camera). These perception models, which are applicable to any top-down view trajectory dataset, enable identifying the impact of partial observability in egocentric perception and preliminary evaluation of cooperative perception for any vehicle behaviour prediction study. In this paper, we use the generated dataset variants to train and evaluate our baseline LC prediction model with different prediction horizons. We are specifically interested in answering the following research questions:

- What is the impact of limited range and occlusion in egocentric perception on the performance of a LC prediction model for different prediction horizons?
- What is the average gain obtained in LC prediction when using cooperative perception with variable penetration rates of automated vehicles?

2 RELATED WORKS

The vehicle behaviour prediction problem has been extensively studied in the literature (Lefèvre et al., 2014; Mozaffari et al., 2020). In this section, we review the related works to LC prediction in highway driving scenarios and highlight their observability assumption under two categories: full observability and ego-centric perception. We then overview the existing studies on the application of offboard V2V data in vehicle behaviour prediction.

2.1 LC Prediction Assuming Full Observability

Several existing studies assume full observability of the TV’s surrounding environment by utilizing naturalistic trajectory datasets collected by wide and top-down view sensors (Yoon and Kum, 2016; Liu et al., 2019; Ding et al., 2019; Hu et al., 2018; Scheel et al., 2019; Rehder et al., 2019a; Gallitz et al., 2020; Mänttari et al., 2018; Deo and Trivedi, 2018). In (Yoon and Kum, 2016) and (Liu et al., 2019), the states of the TV (e.g., lateral position and velocity) are used to predict the TV’s LC manoeuvre. These studies do not consider the interaction between the TV and its surrounding vehicles, which results in a limited

prediction horizon and performance. In (Ding et al., 2019; Hu et al., 2018; Scheel et al., 2019; Rehder et al., 2019a; Gallitz et al., 2020) the interaction between the TV and a fixed number of vehicles around the TV are modelled using fully-connected neural networks (Hu et al., 2018), Bayesian networks (Rehder et al., 2019a), and recurrent neural networks (Ding et al., 2019; Scheel et al., 2019; Gallitz et al., 2020). Instead of considering a constant number of surrounding vehicles, the authors in (Mänttari et al., 2018) and (Deo and Trivedi, 2018) use convolutional neural networks to model the spatial interaction among vehicles within a distance to the TV. In the aforementioned interaction-aware studies, the observability of the TV's nearby vehicles is assumed to be guaranteed, which is not practical given the limited sensor range and potential occlusions of automated vehicles' perception sensors. In this paper, we propose an ego-centric perception model which is used to identify the potential performance loss when moving from a dataset with full observability to an egocentric dataset.

2.2 LC Prediction with Egocentric Perception

Another category of the existing studies considers more realistic assumption by using datasets collected from egocentric perception sensors. In (Rehder et al., 2019b), the LC behaviour of a TV is predicted using a hybrid Bayesian network applied to interaction-aware features such as time to collision and time headway. The authors collect data using a test vehicle equipped with LiDAR sensors and an object tracking system that provides a list of tracked vehicles within 100 meters of the ego vehicle. However, the prediction horizon of their model is limited to 50 meters, allowing the additional 50 meters to track the TV's surrounding traffic. The authors in (Krüger et al., 2019) apply a neural network to extract relevant patterns from the TV and its SVs' motion features, followed by a Gaussian process to produce a probabilistic LC prediction. D. Lee *et al.* (Lee et al., 2017) apply a six-layer convolutional neural network (CNN) on a binary two-channel representation of the driving environment in front of ego vehicle to predict the left/right cut in manoeuvres of preceding vehicles. This representation is created using front-facing radar and camera sensors installed on the ego vehicle. The binary representation only covers a limited area in front of the EV and, thus, fail to consider other areas that also might influence the next manoeuvres of the preceding vehicles. The authors in (Izquierdo et al., 2019a; Fernández-Llorca et al., 2020) train and evaluate a convolutional

neural network model to predict the LC manoeuvres of TVs using raw egocentric sensor images from Prevention Dataset (Izquierdo et al., 2019b). To reduce the computation cost, the raw sensor images used in these studies are cropped around a close vicinity of the TV, limiting the observation of the surrounding environment. Although the studies presented in this subsection use egocentric perception datasets, they do not quantify the performance drop caused by the partial observability intrinsic to the egocentric perception mode. This paper evaluates the impact of ego-centric perception on the problem of LC prediction. Furthermore, we propose and evaluate a cooperative perception solution to mitigate the limitations of egocentric perception.

2.3 Application of Offboard V2V in Behaviour Prediction

Off-board V2V data (i.e., data from other vehicles) are used to improve the performance of a variety of autonomous driving applications such as road geometry estimation (Sakr et al., 2017), object detection (Chen et al., 2019), and planning (Kim et al., 2015). However, few studies use off-board V2V data for LC detection and prediction. The authors in (Sakr et al., 2018) assume that the TV is sending its states (e.g., position/velocity) to the EV using V2V communication. Therefore, the EV can detect the LC manoeuvres of the TV by observing the transmitted states. N. Williams *et al.* (Williams et al., 2018) develops an LC warning system by extending the observable field of onboard sensors using the assumption that unobservable vehicles can transmit their states to the EV using V2V communication. We extend these works by assuming that CAVs can share processed perception data, i.e., a list of detected objects, with the EV particularly for the task of LC prediction.

3 PROBLEM DEFINITION AND SYSTEM MODEL

The problem of LC prediction consists of estimating the probability of LC manoeuvres of a TV during a prediction window T_{pred} given the available observation of the states of the TV and its SVs during an observation window T_{obs} . The LC manoeuvres can be one of Lane Keeping (LK), Left LC (LLC) or Right LC (RLC). We assume a time delay T_{delay} separates the observation and prediction window. The value of T_{delay} controls the prediction horizon. A low value of T_{delay} corresponds to short-term prediction (i.e., pre-

dicting the LC of the TV in near future) and a large value corresponds to long-term prediction, as illustrated in Figure. 2. The LC prediction problem can be formulated as estimating the following conditional probability mass function:

$$P(m = \bar{m} | \text{Observations}), \quad (1)$$

where

$$\bar{m} \in M = \{LK, LLC, RLC\} \quad (2)$$

We assume that the *observations* in this formulation is obtained using one of the following perception modes:

1. **Full-observability:** baseline mode, where the surrounding vehicles are assumed to be fully observable, as seen from top-down view infrastructure sensors.
2. **Egocentric Perception:** We assume a 360-degree horizontal field of view sensor (e.g., camera or lidar) with R_{sensor} meters effective range installed on the centre of the EV is observing the environment. Note that the range of the sensor is adjustable in our proposed perception modelling method depending on the actual range of the deployed sensor.
3. **Cooperative Perception:** A percentage of vehicles (i.e., P_{CAVs} : CAVs penetration rate) on the road are randomly selected as CAVs which can observe their surrounding environment and share the list of detected vehicles in the environment with the EV.

In all the aforementioned perception modes, we assume that:

- Ideal object detection and tracking modules are estimating the states e.g., (x,y) locations and size of the bounding boxes of observable vehicles.
- The same set of perception sensor models are used in both egocentric and cooperative perception modes.
- Ideal communication channels are being used in the cooperative perception mode.

4 PROPOSED PERCEPTION MODELS

To evaluate the impact of each perception mode, defined in the previous section, on the performance of LC prediction, it is required to have a dataset captured from the corresponding perception mode. However, a public real-world trajectory datasets collected

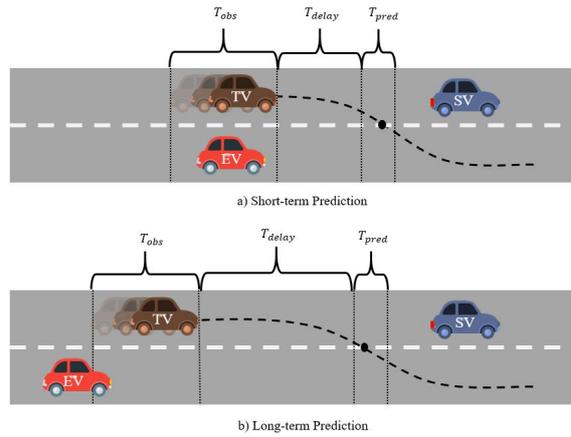


Figure 2: An illustration of the (a) short-term, (b) long-term LC manoeuvre prediction.

from multiple vehicles do not currently exist. Therefore, we propose two perception models to represent egocentric and cooperative perception modes. The proposed models are applied to a base vehicle trajectory dataset captured from top-down view sensors (representing the full-observability mode) to generate variants of the same underlying driving scenarios corresponding to egocentric and cooperative perception, respectively. Figure 4 illustrates the data rendering processes required for generating each dataset variant and Figure. 3 demonstrates an example of the representation in each dataset variant.

4.1 Basic Representation

Using the selected base vehicle trajectory dataset, we populate a binary three-channel image containing a top-down view of the covered road section at each time-step, denoted as the basic representation I_{basic} . The first channel of this representation depicts the 2D bounding boxes of vehicles within the road section. The map data (e.g., lane markings) are indicated in the second channel. The third channel specifies the observability status of each pixel. This channel is initialized with zeros (i.e., all pixels are considered to be unobservable at first) and is populated in later stages according to the considered perception model. The data representation of the first perception mode (i.e., full observability) is identical to the basic representation except that the observability status of all pixels is set to 1 (i.e., observable).

4.2 Egocentric Perception Model

The egocentric perception model estimates the observability status of each area in the driving environment assuming that one of the vehicles, considered

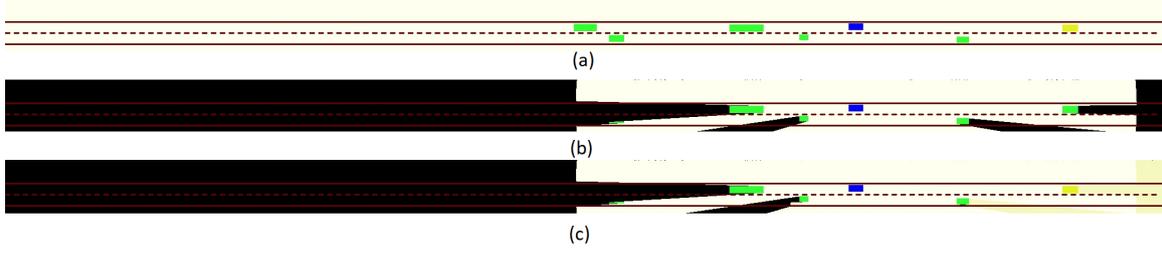


Figure 3: An example of the representation rendered in each perception mode: (a) full observability, (b) egocentric perception, and (c) cooperative perception. The representation is coloured for illustration purposes. Black colour represents unobservable areas. The EV and CAVs are represented with blue and yellow bounding boxes, respectively. Other vehicles are represented with green bounding boxes.

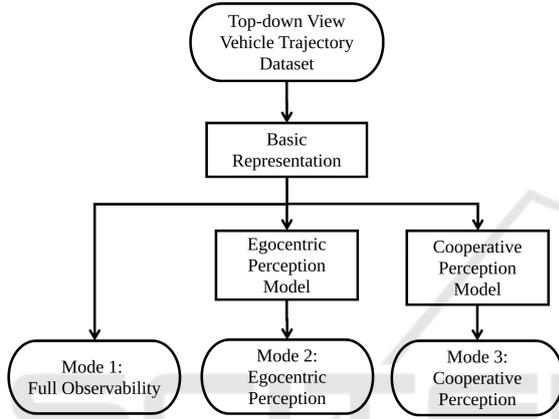


Figure 4: Data generation process for (1) Full observability, (2) Egocentric perception, and (3) Cooperative perception modes.

as the EV, is observing the environment with onboard perception sensors. The egocentric perception model uses I_{basic} : the top-down view basic representation, X_{EV} : the location of the selected EV, and R_{sensor} : the range of the perception sensor, to generate I_{ego} : a top-down view representation with the modelled perception of the EV. The I_{ego} is an updated version of I_{basic} , where the observability status of each pixel is estimated according to the perception model described in Algorithm 1. This algorithm considers the sensor range and occlusion effect in estimating the observability status of all pixels. To model the range of the observing vehicle's sensors, we first consider all the lines connecting the vehicle centre point to the points on the boundary of the sensor range. Then, we use Bresenham's line algorithm (Earnshaw, 1985) to select the pixels on each line in the input representation. To model occlusion, we then select the segments of each line from the centre of the EV until the first occupied pixel by other vehicles. Finally, the pixels on the final line segments are marked as observable by setting their value to 1 in the observation channel of the representation.

Algorithm 1: Egocentric Perception Model.

Input: $I_{basic}, X_{EV}, R_{sensor}$
Output: I_{ego}

- 1 Initialisation:
- 2 $I_{ego} \leftarrow I_{basic}$
- 3 $\{B_i\}_{i=1}^P \leftarrow$ List of pixels in I_{basic} on the border of the EV's sensor range.
- 4 **for** $i \leftarrow 1$ to P **do**
- 5 $L_i \leftarrow$ List of pixels in I_{basic} on the line connecting X_{EV} to B_i obtained using Bresenham's line algorithm (Earnshaw, 1985).
- 6 $\hat{L}_i \leftarrow$ Pixels on L_i line from X_{EV} until reached the first occupied pixel in I_{basic} by other vehicles.
- 7 Update I_{ego} by setting \hat{L}_i pixels in third channel of I_{ego} to 1.

4.3 Cooperative Perception Model

The cooperative perception model estimates the observability status of each pixel in the driving environment assuming a group of CAVs, including the EV, are observing the environment with their onboard perception sensors. The cooperative perception model uses the estimated I_{ego} and the location of N vehicles, selected uniformly at random as CAVs, to generate the cooperative perception mode representation denoted as I_{coop} . The percentage of CAVs w.r.t. all vehicles is denoted as P_{CAVs} or CAVs penetration rate. This perception model, described in Algorithm 2, iteratively uses the Algorithm 1 to update the observability status of the I_{coop} based on the location of each CAVs. As we assumed ideal object detection and tracking, there is no misalignment error among the detected vehicles from different CAVs and the EV. Therefore, the fusion of the representations is done using pixel-wise *logical OR* operation among the representations.

Algorithm 2: Cooperative Perception Model.**Input:** $I_{ego}, \{X_{CAV\#1}, X_{CAV\#2}, \dots, X_{CAV\#N}\}, R_{sensor}$ **Output:** I_{coop}

- 1 Initialisation:
- 2 $I_{coop} \leftarrow I_{ego}$
- 3 **for** $i \leftarrow 1$ **to** N **do**
- 4 $I_{coop} \leftarrow I_{coop}$ *logical OR*
 Algorithm 1($I_{coop}, X_{CAV\#i}, R_{sensor}$)

5 DATASET AND PREDICTION MODEL

This section describes the base vehicle trajectory dataset used to generate the three dataset variants corresponding to each perception mode. Next, the prediction model used for performing the comparative study on perception modes is introduced.

5.1 Vehicle Trajectory Dataset Description

The highD dataset (Krajewski et al., 2018), a publicly available naturalistic vehicle trajectory dataset is used as the base dataset from which the three dataset variants are obtained. This dataset is recorded on German highways using a top-down view camera installed on a drone. The highD Dataset, compared to similar existing datasets such as NGSIM (Colyar and Halkias, 2007; Halkias and Colyar, 2007), contains more data samples and more accurate annotations (typical positioning error is less than 10 cm). The data is reported in 60 spreadsheets and includes several features of vehicles (e.g., x-y position, velocity, lane ID, etc) for each time frame. In this study, we select the first 40 spreadsheets of the highD dataset as training data, the next 10 spreadsheets as validation data, and the final 10 spreadsheets as test data. After applying the data preparation process on the highD dataset, we have 14K training samples, 6K validation samples and 5K test samples in each dataset variant, where a sample is a group of T_{obs} frames.

5.2 Prediction Model

A convolutional neural network (CNN) is used to estimate the likelihood of the LC manoeuvres using the representations generated in each dataset variant. The CNN model consists of 3 layers of 2D convolution, each with 16 filters with a kernel size of 3×3 and followed by a 2×2 pooling layer and a ReLU activation function. These layers are followed by two fully-

connected layers with 512 and 3 output neurons, corresponding to three classes of LC manoeuvres (i.e., RLC, LLC, and LK). To feed the representations to the CNN prediction model, first, the TV is selected from one of the observed vehicles exactly next to the EV (i.e., the EV's preceding/following vehicle in its lane or adjacent lanes or the right/left alongside vehicles). Then, A rectangle $\hat{L} \times \hat{W}$ crop of the representation is selected which is centred on the TV and moves with it. A stack of cropped images for the time-steps in $[t_0 - T_{obs}, t_0]$ forms a data sample in the corresponding dataset variant, which is fed as input to the CNN for prediction query at t_0 . The sample is labelled as an RLC or LLC if the TV keeps the lane during T_{obs} and T_{delay} and its centre crosses the right or left lane marking during prediction window T_{pred} , respectively. The data sample in which the TV does not cross any lane marking during T_{obs}, T_{delay} and T_{pred} , is labelled as lane-keeping. We perform random under-sampling on the LK class to balance the dataset.

6 COMPARATIVE EVALUATION

This section describes the implementation details, evaluation metrics and discusses the experiments and results.

6.1 Implementation Details

We empirically set the length of the observations time (T_{obs}) and prediction window (T_{pred}) to 1 second. In the analysis, we train and test the baseline model separately with different time delays (T_{delay}) from 0 to 4 seconds. The width and height of cropped images are selected as 100 and 90, respectively. The selected height assures all the driving lanes are covered in the cropped image, regardless of the TV's current lane.

We adopt Adam optimizer (Kingma and Ba, 2017) with the initial learning rate of 0.001 and set the maximum number of epochs to 10. Using the early-stopping technique on validation data, we stop training before over-fitting occurs. All the models are trained and tested on an NVIDIA RTX 2080 TI using PyTorch platform (Paszke et al., 2017).

6.2 Evaluation Metrics

Given a balanced dataset (i.e., equal number of samples in each class), accuracy can be a trustworthy metric to identify the utility of each perception mode for the problem of vehicle behaviour prediction. The accuracy (ACC) is defined as the percentage of correctly

Table 1: LC prediction accuracy and OBS metrics for different perception modes.

Perception Mode	OBS	Time Delay (T_{delay})				
		0s	1s	2s	3s	4s
Mode 1. Full Observability	100%	98.3%	90.77%	80.33%	77.11%	73.49%
Mode 2. Egocentric Perception	63%	97.03%	88.69%	78.18%	73.35%	69.2%
Mode 3. Cooperative Perception ($P_{CAVs} = 20\%$)	85%	97.59%	90.34%	80.53%	72.69%	72.28%

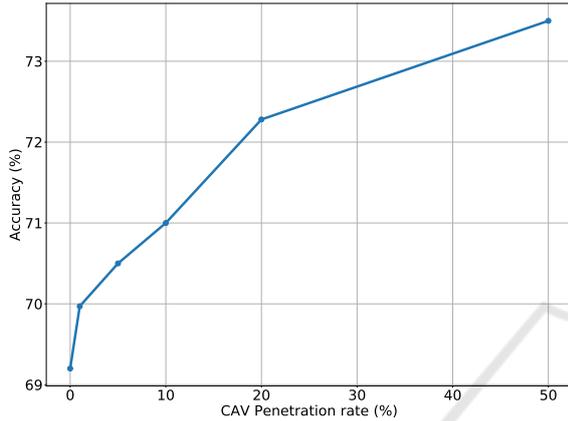


Figure 5: Long-term prediction accuracy for different penetration rate of cooperative perception.

predicted LC and LK manoeuvres of the TV to the total number of LC and LK manoeuvres in a dataset variant corresponding to a perception mode. We also report the average percentage of observable pixels in cropped input image for each dataset variant as the observability (OBS) metric. This metric shows the utility of perception mode in observing the TV's surrounding.

6.3 Model Performance for Each Perception Mode

This experiment aims to quantify the impact of using each of the defined perception modes (i.e., full observability, egocentric perception, and cooperative perception modes) on the baseline prediction model with different prediction horizons. To this end, we train and test the baseline model separately on each dataset variants from the highD dataset and for time delays T_{delay} of 0,1,2,3, and 4 seconds. The results are reported in Table 1.

The accuracy of short-term LC prediction for a time delay of 0 seconds in all perception modes are above 97% and the difference between modes is negligible. Nonetheless, the gap between the accuracy of the prediction model applied to Mode.1 and Mode.2 increases from 2% to 4% when the time delay in-

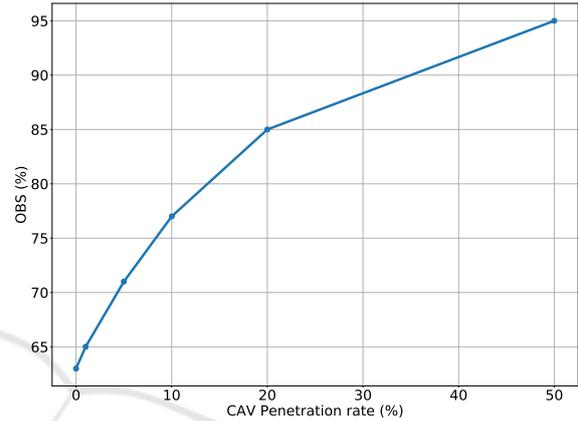


Figure 6: observability of TV's surrounding (OBS metric) for different penetration rate of cooperative perception.

creases from 1 to 4 seconds, indicating the importance of observing the TV's surrounding for longer prediction horizons. Fusing the CAVs perception with onboard perception can compensate the performance gap completely for time delays of 1s and 2s and decrease it to 1% for time delay of 4s; however, no improvement is observed for time delay of 3s. In addition, augmenting onboard perception with offboard CAVs perception can increase the observability metric (OBS) from 63% to 85%. These results achieved by assuming the penetration rate of CAVs on the road to be $P_{CAVs} = 20\%$.

Comparing the LC prediction performance of the CNN model for different perception modes suggests that comprehensive perception of the driving environment is especially important in longer prediction horizons. This is mainly because long-term predictions require a deep understanding of the driving environment and of the emerging traffic status. This is where cooperative perception has the most impact due to the increased observability of the driving environment

6.4 Model Performance with Different CAVs Penetration Rate

In this experiment, we measure the performance of the long-term LC prediction model ($T_{delay} = 4$) and

the OBS metric for different penetration rates of CAVs. The model's performance is only reported for the long-term prediction since such samples have the highest performance drop in the egocentric perception model based on the aforementioned experiment. To prepare the data, we model the cooperative perception mode by assuming 1%, 5%, 10%, 20%, and 50% percentages of CAVs. Figure 6 and Figure 5 show the relation between percentage of CAVs and OBS metric, and percentage of CAVs and accuracy, respectively. Both figures demonstrate a logarithmic increase in performance and observability with increasing the penetration rate. With 20% penetration of CAVs on average, 85% of TV's surrounding become observable and the drop in performance compared to full observability mode is decreased to 1%.

7 CONCLUSION

In this paper, we proposed two perception models, applicable to any vehicle trajectory dataset recorded by a top-down view sensor, to model the egocentric and cooperative perception. Then, a comparative study has been performed to quantify the impact of each perception mode on the problem of lane change prediction. The results showed a 4% performance drop in our long-term LC prediction model (i.e., $T_{delay} = 4s$) when using ego-centric perception instead of the full-observable original dataset. Also, the results indicated that cooperative perception with 20% penetration of CAVs can almost compensate for the performance drop of our prediction model related to ego-centric perception limitation.

Future work should consider extending the data generation method by considering the errors in object detection and tracking modules. The binary representation used in this work can be extended to a probabilistic representation which enables encoding the error and uncertainty in vehicles states estimation. Furthermore, the 2D occlusion model used in this study can be extended to a 3D occlusion model to decrease the modelling error.

ACKNOWLEDGEMENTS

This work was supported by Jaguar Land Rover and the U.K.-EPSRC as part of the jointly funded Towards Autonomy: Smart and Connected Control (TASCC) Programme under Grant EP/N01300X/1. We would like to thank Omar Al-Jarrah for reviewing previous version of this paper.

REFERENCES

- Chen, Q., Tang, S., Yang, Q., and Fu, S. (2019). Cooper: Cooperative perception for connected autonomous vehicles based on 3d point clouds. *arXiv preprint arXiv:1905.05265*.
- Colyar, J. and Halkias, J. (2007). Us highway 101 dataset. Technical report, U.S. Department of Transportation, Federal Highway Administration (FHWA).
- Deo, N. and Trivedi, M. M. (2018). Convolutional social pooling for vehicle trajectory prediction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.
- Ding, W., Chen, J., and Shen, S. (2019). Predicting vehicle behaviors over an extended horizon using behavior interaction network. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 8634–8640.
- Earnshaw, R. A. (1985). *Fundamental Algorithms for Computer Graphics: NATO Advanced Study Institute Directed by JE Bresenham, RA Earnshaw, MLV Pitte-way*. Springer.
- Fernández-Llorca, D., Biparva, M., Izquierdo-Gonzalo, R., and Tsotsos, J. K. (2020). Two-stream networks for lane-change prediction of surrounding vehicles. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–6.
- Gallitz, O., De Candido, O., Botsch, M., Melz, R., and Utschick, W. (2020). Interpretable machine learning structure for an early prediction of lane changes. In Farkaš, I., Masulli, P., and Wermter, S., editors, *Artificial Neural Networks and Machine Learning – ICANN 2020*, pages 337–349, Cham. Springer International Publishing.
- Halkias, J. and Colyar, J. (2007). Us highway 80 dataset. Technical report, U.S. Department of Transportation, Federal Highway Administration (FHWA).
- Hu, Y., Zhan, W., and Tomizuka, M. (2018). Probabilistic prediction of vehicle semantic intention and motion. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 307–313.
- Izquierdo, R., Quintanar, A., Parra, I., Fernández-Llorca, D., and Sotelo, M. A. (2019a). Experimental validation of lane-change intention prediction methodologies based on cnn and lstm. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 3657–3662.
- Izquierdo, R., Quintanar, A., Parra, I., Fernández-Llorca, D., and Sotelo, M. A. (2019b). The prevention dataset: a novel benchmark for prediction of vehicles intentions. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 3114–3121.
- Kim, S., Liu, W., Ang, M. H., Frazzoli, E., and Rus, D. (2015). The impact of cooperative perception on decision making and planning of autonomous vehicles. *IEEE Intelligent Transportation Systems Magazine*, 7(3):39–50.
- Kingma, D. P. and Ba, J. (2017). Adam: A method for stochastic optimization.

- Krajewski, R., Bock, J., Kloeker, L., and Eckstein, L. (2018). The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2118–2125.
- Krüger, M., Novo, A. S., Nattermann, T., and Bertram, T. (2019). Probabilistic lane change prediction using gaussian process neural networks. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 3651–3656.
- Lee, D., Kwon, Y. P., McMains, S., and Hedrick, J. K. (2017). Convolution neural network-based lane change intention prediction of surrounding vehicles for acc. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–6.
- Lefèvre, S., Vasquez, D., and Laugier, C. (2014). A survey on motion prediction and risk assessment for intelligent vehicles. *ROBOMECH Journal*, 1(1):1.
- Liu, J., Luo, Y., Xiong, H., Wang, T., Huang, H., and Zhong, Z. (2019). An integrated approach to probabilistic vehicle trajectory prediction via driver characteristic and intention estimation. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 3526–3532.
- Mänttari, J., Folkesson, J., and Ward, E. (2018). Learning to predict lane changes in highway scenarios using dynamic filters on a generic traffic representation. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1385–1392.
- 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*, pages 1–15.
- Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A. (2017). Automatic differentiation in pytorch.
- Rehder, T., Koenig, A., Goehl, M., Louis, L., and Schramm, D. (2019a). Lane change intention awareness for assisted and automated driving on highways. *IEEE Transactions on Intelligent Vehicles*, 4(2):265–276.
- Rehder, T., Koenig, A., Goehl, M., Louis, L., and Schramm, D. (2019b). Lane change intention awareness for assisted and automated driving on highways. *IEEE Transactions on Intelligent Vehicles*, 4(2):265–276.
- Sakr, A. H., Bansal, G., Vladimerou, V., and Johnson, M. (2018). Lane change detection using v2v safety messages. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3967–3973.
- Sakr, A. H., Bansal, G., Vladimerou, V., Kusano, K., and Johnson, M. (2017). V2v and on-board sensor fusion for road geometry estimation. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–8.
- Scheel, O., Nagaraja, N. S., Schwarz, L., Navab, N., and Tombari, F. (2019). Attention-based lane change prediction. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 8655–8661.
- Williams, N., Wu, G., Boriboonsomsin, K., Barth, M., Rajab, S., and Bai, S. (2018). Anticipatory lane change warning using vehicle-to-vehicle communications. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 117–122.
- Yoon, S. and Kum, D. (2016). The multilayer perceptron approach to lateral motion prediction of surrounding vehicles for autonomous vehicles. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pages 1307–1312.