[[1]](#footnote-1)

**Motion Prediction Techniques in Autonomous Vehicles: A Review**

Adeel Ahmad, Zubair Idrees, Ahsan Rahman

Faculty of Computer Sciences, SZABIST(Dubai) UAE

*Abstract*— Autonomous vehicles (AVs) are expected to dramatically redefine the future of transportation. However, there are still significant engineering challenges to be solved before one can fully realize the benefits of self-driving cars. Predicting the Motion of other traffic objects is the most challenging task in autonomous vehicles and self-drive cars, one such challenge is building models that reliably predict the movement of traffic agents around the Autonomous Vehicles, such as cars, cyclists, and pedestrians etc. In this paper, techniques and algorithms used for the motion prediction of other vehicles around autonomous/self-driving vehicles are discussed, later on their review and comparison will be presented along with current challenges and gaps. At the end, the conclusion of this review and direction for future research is presented.

*Index Terms*— Motion Prediction; Autonomous Vehicles; self-driving cars; Motion Planning, Trajectory Planning.

# INTRODUCTION

Self-driving vehicles or Autonomous vehicles are being adopted to reduce road accidents and improve safety for the humans. However, for efficient and safe road operation, a self-driving car should not only interpret the nearby road-users’ current state, but also predict their future actions. The process of designing an autonomous vehicle is not a simple task. Generation of HD maps helped a lot in igniting the process of automation in AVs(Autonomous Vehicles). Firstly, self-driving vehicles should be capable of locating their self in the real world. Secondly, perception process enables the self-driving vehicles in sensing it’s surrounding environment such as traffic signs, road lanes and other traffic objects. Afterwards, self-driving vehicles should be able to predict the future moves or trajectories of other traffic objects i.e., vehicle. This process is also called “prediction”. By knowing the future movements of other traffic objects, self-driving vehicles will be able to plan its future movement and trajectory efficiently. At the end, self-driving vehicles’ trajectory is planned by using control module.

The most challenging and complex step among all the steps mentioned above is motion prediction as it requires real-time environment interaction. In this paper review of the existing studies is being presented, which is categorized in following types i.e., i) physics-based models, ii) deep-learning based models and iii) deep reinforcement-based learning models.

# Study Overview

Lefevre et al. [1] classifies vehicle’s behavior prediction models into physics-based, maneuver-based, and interaction aware models. Physics-based motion models operate at lower level, they state that vehicle’s motion solely depends on physics’ laws. They can efficiently predict the risks, but are only limited to predictions of short-term collisions. Maneuver-based motion models function at a higher-level, they consider the future motion of vehicle is also dependent on the maneuver that the driver is going to perform. They offer reliable estimation of long-term risk and motion, however aren't reliable always as current scene’s dependencies between vehicles are not considered. Interaction-aware motion models add an additional abstraction level by taking vehicles’ maneuvers inter-dependencies in consideration. But the computational complexity of interaction-aware motion models is not compatible for assessment of real-time risks.

Mozaffari et al. [2] suggested state-of-the-art methods and categorized them as (a) input representation; (b) output type; and (c) prediction method. Further deep learning methods for AV’s behavior prediction methods are categories into three classes (i) recurrent neural network, (ii) convolutional neural network, and (iii) other methods (e.g., graph neural networks or combined methods). Long short-term memory (LSTM) model [3] and Gated Recurrent Unit (GRU) [4] are behavior prediction’s main architecture.

Altch'e et al. [5] trained and tested the Long short-term memory (LSTM) model to anticipate the future trajectories (longitudinal and lateral) for vehicles driving on the highways by utilizing the Next Generation Simulation (NGSIM) dataset. Their model gave better results as compared to state-of-the-art methods. The output of their model was a single value for the trajectory prediction. The benefit of their model can be summed up in its speculation considering about 6000 vehicles simultaneously.

Xin et al. [6] predicted the interaction of EV with SVs by predicting the long-horizon trajectory for surrounding vehicles (SVs) using dual long short-term memory (LSTM). As an indicator driver’s intention is identified by the very first LSTM block, further future trajectory is predicted by second LSTM block. Lateral positions were predicted in small bounds by their work which led their model to be adaptable for different road geometries. The recommended future work by them was, generalization of model for various road scenarios, like unstructured roads and intersections and adding noise to model’s output.

Deo and Trivedi [7] provided an interaction-aware prediction model by presenting multi-layer long short-term (LSTM) model. A multi-modal distribution is provided by the model using I-80 and NGSIM US-101 dataset. They correlated six highway maneuvers by using six LSTM decoders. An encoder LSTM is applied to the past direction of vehicles. First, on vehicle’s past trajectory another LSTM encoder is applied. Each decoder’s LSTM is initialized by concatenating encoder LSTM’s last hidden state with one-hot vector comprising specific maneuvers to each decoder. The parameters of maneuvers-conditioned bivariate Gaussian-distribution are predicted by decoder LSTM for TVs future locations. The probability of each one of six maneuvers is predicted by another encoder LSTM.

Zhao et al. [8] used combination of CNN and LSTM networks by concatenating scene status with agent’s movement vector encoded by CNN and LSTM and fed to network like U-net. Further it’s I/O is fed to LSTM encoder for predicting agent’s future trajectories.

Lee et al. [9] proposed six CNNs for the prediction of lane change intention and control. Simplified BEV was used to reduce the computational cost. But serial data cannot be modeled by these types of CNNs.

Hu et al. [10] used fully connected neural network for the prediction of various scenarios possible by introducing the Semantic-based Intention and Motion Prediction (SIMP) model. They concluded competitive results can be achieved while utilizing semantic approach by combining several tasks in single framework as compared to traditional approaches.

Casas et al. [11] suggested IntentNet as a trainable end to end model by combining semantic high-resolution maps and LiDAR’s 3D point clouds. They experimented and proved that complex algorithms can model statistical dependencies among continuous and discrete intentions. Their model can be extended for bicycle and pedestrian’s intentions.

Kiran et al. [12] reviewed the different reinforcement learning approaches for autonomous vehicle. In Reinforcement learning methods, an agent learns by experiencing different scenarios and interacting with real environment. Agents performance is improved by maximizing the received accumulative reward during its lifetime. It improves knowledge gradually for long term by exporting previous knowledge and exploring new knowledge. The main challenge faced in reinforcement learning is to balance tradeoff among exploitation and exploration. By using this approach, AVs predict future trajectories by experiencing various scenarios and its reaction w.r.t its environment is improved a lot. Prediction methods for the trajectories can be classified into following categories (a) Value based methods (b) Policy based methods (c) Actor critic methods (d) Model based (vs. Model free) and On/Off Policy methods (e) Deep reinforcement learning.

Reinforcement learning can be applied in following autonomous driving areas: path planning, controller optimization, trajectory optimization, dynamic path planning, motion planning, high level driving policies development to predict complex navigation scenarios, highway’s scenario based policy learning, intersections split and merge, reward learning with inverse reinforcement learning using expert data for traffic actors intent prediction like other vehicles , pedestrian and learning policies which assure safety and perform risk estimation.

Leurent [13] presented a comprehensive review on various state and action representations that being used in autonomous cars. In AVs, most common state space features used are heading, position, ego-vehicle’s velocity and obstacles in the sensor view of ego-vehicle.

Keselman et al. [14] suggested a model-based reinforcement learning method that for dynamic planning of the trajectories and for providing vehicle’s smooth controlling behavior. This learning model can be combined with learnable heuristic of A\* algorithm and trees by using a Deep Q-Network over obstacle map of image-based input.

Yung and Ngai [15] provided a solution for vehicles collision and overtaking problem by using a multiple-goal reinforcement learning (MGRL). Q-learning (QL) or double-action QL was used for determining the action-based decisions and to figure out whether surrounding vehicles and AV could reach to same point.

Isele et al. [16] proposed an effective technique to solve the problems related to intersections for AVs specially in jammed intersections and highway merging related problems. An optimal driving policy can be found by using LSTM and applying deep reinforcement learning to produce an internal state having historic driving data and Deep Q-Network for approximation of Q-function.

Wang et al. [17] suggested a Deep Reinforcement Learning to provide the solutions for problems related to lane changing in AVs by using Q-learning. These methods are based on Q-function approximation having action space assumption and continuous state space. Their results show that the Deep Reinforcement Learning model provides an efficient and smooth driving policies for lane-changing maneuvers in an interactive driving environment. For future work, it is suggested that model can be improved further by training it with traffic rules and various road geometries. In addition, it is suggested to combine Reinforcement Learning (RL) and Model predictive control (MPC) to get best of both techniques.

Sallab et al. [18] proposed a Deep Reinforcement Learning (DRL) framework by the combining RNN/LSTM, Q-network and CNN/DNN using spatial and temporal aggregation. Eventually, they tested their model in Torcs an Open-source Racing Car Simulator.

# Problem Statement

Motion of vehicles is well structured as it’s governed by driving rules and surrounding conditions. Desired trajectories are not possible in vehicles in a real-time manner. However, task of vehicles’ behavior prediction is not simple because of certain challenges. Firstly, behavior of one vehicle is dependent on behavior of other vehicles around i.e., behavior of one vehicle affects behavior of other vehicles and vice versa. Therefore, it’s very important to observe the behavior of surrounding vehicles to predict their future moves. Secondly, traffic rules and road geometries could reshape the behavior of vehicles. For instance, if a give-way sign is placed at an intersection, behavior of vehicles approaching it will be changed completely. Therefore, it’s very important to consider road geometries and traffic rules while training a model, as model trained in one specific driving conditions would not be able to produce same results in other driving conditions. Thirdly, vehicles’ future behavior is multimodal, i.e., based on history of motion of vehicles there can be more than one possible future behaviors. For instance, a slowing down vehicle at an intersection which hasn’t changed its heading direction, can turn right or left. For a smooth and safe travel, an autonomous vehicle should be capable to predict all possible moves or trajectories of other traffic objects i.e., cars. In addition to challenges mentioned above, there are several limitations in implementing vehicle behavior prediction. For instance, in autonomous vehicles computational resources are limited, further, an autonomous vehicle cannot fully observe it’s driving environment due to on-board sensor limitations (i.e., object occlusion, sensor range limitations and noise).

# Mathematical Problem Formulation

To handle uncertainties in the nature of problem we will be using probabilistic formulation approach for other vehicles’ maneuver prediction. The words “maneuver” and “behavior” are used interchangeably in literature. The future behavior of TVs can be represented using following equation.

(1)

Where x is defining the states (i.e., positions) of vehicle i at a particular time step t, N represents the number of the Target Vehicles, and length of the prediction window is represented by m.

Our target is to compute the conditional distribution P(XTV|OEV) where OEV represents all the available observations for the EV. it’s a mutual distribution over series of states of several interdependent vehicles. To estimate the conditional distribution P(XTV|OEV) with less computational cost, most of the researchers drop the interdependence among vehicles’ future behavior. Such as, each TVs behavior can be predicted individually with affordable hardware. One vehicle is selected on each step as TV and its conditional distribution P(XTV|OEV) is calculated, such as:

(2)

Where T represents the TV.

# Objectives

The core objectives of this research work are to present a comprehensive review of existing work related to behavior or maneuver prediction of traffic objects more specifically other vehicles for autonomous vehicles. This prediction will further help the decision module to plan future trajectories or movement of the autonomous vehicle for the efficient and safe autonomous driving.

# Scope and Limitations of the Study

Scope of this research is to present a review of existing work related to behavior prediction of traffic objects like other vehicles in autonomous vehicles based on most recent literature or research papers available openly on different research Journals. For now, this research is only limited to behaviors prediction of other vehicles, further it can be extended to include other traffic objects also for example pedestrians, animals and pets etc.

# Evaluation

Conventionally vehicles’ motion predication methods are categorized in three types, i.e., physics based, maneuver based and interaction aware. Physics based motion prediction model function at very low level, and consider vehicles’ motion is mainly dependent on physics laws. They can compute the risks efficiently, but can predict only short-term collisions. Maneuver based motion prediction models function at relatively higher level, and consider that vehicle’s future motion is also dependent on maneuver that driver is going to perform. More reliable estimates for long term movement and risk can be obtained from them, but those are not much reliable as dependencies among different vehicles in scene are not considered. Lastly, interaction aware motion prediction models take a step ahead by considering dependencies among vehicles’ maneuvers. However, as they are computationally complex therefore not compatible always with risk assessment in real time manner.

Deep learning-based behavior/maneuver prediction methods give better performances as compared to solutions discussed earlier, more specifically they gives better results in complex driving situations, it is achieved by applying Intellectually appealing input and output type, still there are different challenges which are required to be dealt for allowing their adoption in AVs. Deep learning-based approaches like i) Single, RNN ii) Multiple RNNs, iii) CNN, iv) Combination of RNNs and CNNs, v) Graph Neural Networks.

Mozaffari et al. [2] further classified the interaction-aware models in four types i.e., i) Track history of the TV, ii) Track history of the TV and SVs, iii) Simplified Bird’s Eye View, and iv) Raw sensor data.

Further advantages and disadvantages of all the techniques will be discussed.

|  |  |  |
| --- | --- | --- |
| **Class** | **Advantages** | **Disadvantages** |
| Physics based | - Easy to compute as they depend on the low-level dynamic and kinematic properties of motion | - Limited to short-term motion prediction (usually less than a second).  - Unable to predict changes in vehicle’s motion induced due to execution of a particular manoeuvre (e.g., slowing down, maa turn) |
| Manoeuvre  Intention | - It has low computational cost. | - Provides only high-level understanding of the vehicle behaviour.  - Only covers single driving scenarios manoeuvres. |
| **Interaction aware** | | |
| **Track History of the TV** | - Complies with limited observability of the EV. | - Environment’s impact and interaction between vehicles is not considered for TVs behaviour.  - EVs perception module limitations are inherited. |
| Track History of the TV and SVs | - Impact of interaction between vehicles is considered for TVs behaviour. | - Environment’s impact is not considered for TVs behaviour.  - EVs perception module limitations are inherited. |
| Simplified Bird’s Eye View | - Environment’s impact and interaction between vehicles is considered for TVs behaviour.  - Enables blending the data gathered from different sensors  on the EV.  - Simple in terms of representation.  Complies with limited observability of the EV. | - EVs perception module limitations are inherited. |
| Raw Sensor  Data | - Complies with limited observability of the EV.  - Minimal information loss. | - Computational cost is high. |
| **Neural Networks** | | |
| Recurrent  Neural  Networks | - Temporal dependencies are processed in better way. | - Requires additional mechanism to model  interaction and contextual features. |
| Convolutional Neural Networks | - Spatial dependencies are processed in better way. | - CNNs lacks mechanism for modelling data series. |
| Combination of RNNs  and CNNs | - Benefits from advantages of both CNNs and RNNs |  |
| Graph Neural Networks | - Complies to graph structure of traffic. | - Usually neglects static scene context. |

# CONCLUSION

We discussed different techniques for motion/trajectory prediction like physics-based models, maneuver-based models, interaction aware models and different deep learning-based models (Single RNN, Multiple RNNs, CNN, Combination of RNNs and CNNs, Graph Neural Networks). Deep learning-based behavior/maneuver prediction methods give better performances as compared to other techniques, more specifically they give better results in complex driving situations. In particular most of the available solutions consider interaction between vehicles, but other factors like traffic laws and environmental conditions are not inputted to prediction model. Additionally, real world limitations like sensor damages and limitations related to computational resources are not fully considered.

# Future work

Several algorithms with different approaches were introduced in area of self-driving/autonomous cars. Different models from basic physics-based to advanced level i.e., deep reinforcement learning based models are presented. Currently benchmarks are not available for the evaluation of the performance of deep learning and reinforcement learning approaches. There is a need for defining benchmarks for behavior/mauver prediction to evaluate different algorithms’ performance. Sensor impairment can cause problems in real world AV operations, there should be some mechanism which takes care of such impairments. It can be through some reports sent to owner/authorities of the AVs about the sensor from which AVs are not getting any data for some time.

References

1. S. Lef`evre, D. Vasquez, and C. Laugier, “A survey on motion prediction and risk assessment for intelligent vehicles,” ROBOMECH Journal, vol. 1, no. 1, p. 1, Jul 2014. [Online]. Available: <https://doi:org/10:1186/s40648-014-0001-z>
2. S. Mozaffari, O. Y. Al-Jarrah, M. Dianati, P. Jennings, and A. Mouzakitis, “Deep learning-based vehicle behavior prediction for autonomous driving applications: A review,” IEEE Transactions on Intelligent Transportation Systems, 2020.
3. S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.
4. K. Cho, B. Van Merri¨enboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” arXiv preprint arXiv:1406.1078, 2014.
5. F. Altch´e and A. de La Fortelle, “An lstm network for highway trajectory prediction,” in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2017, pp. 353–359.
6. L. Xin, P. Wang, C.-Y. Chan, J. Chen, S. E. Li, and B. Cheng, “Intention-aware long horizon trajectory prediction of surrounding vehicles using dual lstm networks, ” in 2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2018, pp. 1441–1446.
7. N. Deo and M. M. Trivedi, “Multi-modal trajectory prediction of surrounding vehicles with maneuver based lstms,” CoRR, vol. abs/1805.05499, 2018. [Online]. Available: http://arxiv.org/abs/1805.05499
8. T. Zhao, Y. Xu, M. Monfort, W. Choi, C. Baker, Y. Zhao, Y. Wang, and Y. N. Wu, “Multi-agent tensor fusion for contextual trajectory prediction,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 12 126–12 134.
9. D. Lee, Y. P. Kwon, S. McMains, and J. K. Hedrick, “Convolution neural network-based lane change intention prediction of surrounding vehicles for acc,” in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2017, pp. 1–6.
10. Y. Hu, W. Zhan, and M. Tomizuka, “Probabilistic prediction of vehicle semantic intention and motion,” CoRR, vol. abs/1804.03629, 2018. [Online]. Available: <http://arxiv.org/abs/1804.03629>
11. S. Casas, W. Luo, and R. Urtasun, “Intentnet: Learning to predict intention from raw sensor data,” in Conference on Robot Learning, 2018, pp. 947–956.
12. B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogamani, and P. P´erez, “Deep reinforcement learning for autonomous driving: A survey,” arXiv preprint arXiv:2002.00444, 2020.
13. E. Leurent, “A survey of state-action representations for autonomous driving,” 2018.
14. A. Keselman, S. Ten, A. Ghazali, and M. Jubeh, “Reinforcement learning with a\* and a deep heuristic,” CoRR, vol. abs/1811.07745, 2018. [Online]. Available: <http://arxiv.org/abs/1811.07745>
15. D. C. K. Ngai and N. H. C. Yung, “A multiple-goal reinforcement learning method for complex vehicle overtaking maneuvers,” IEEE Transactions on Intelligent Transportation Systems, vol. 12, no. 2, pp. 509–522, 2011.
16. D. Isele, A. Cosgun, K. Subramanian, and K. Fujimura, “Navigating intersections with autonomous vehicles using deep reinforcement learning,” CoRR, vol. abs/1705.01196, 2017. [Online]. Available: <http://arxiv.org/abs/1705.01196>
17. P. Wang, C. Chan, and A. de La Fortelle, “A reinforcement learning based approach for automated lane change maneuvers,” CoRR, vol. abs/1804.07871, 2018. [Online]. Available: http://arxiv.org/abs/1804.07871
18. A. Sallab, M. Abdou, E. Perot, and S. Yogamani, “Deep reinforcement learning framework for autonomous driving,” Electronic Imaging, vol. 2017, no. 19, p. 70?76, Jan 2017. [Online]. Available: <http://dx.doi.org/10.2352/ISSN.2470-1173.2017.19.AVM-023>

1. [↑](#footnote-ref-1)