Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2022) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

Transport Research Arena (TRA) Conference

Identifying the Automated Vehicle's Driving Policy in the Vicinity of Pedestrians

Foteini Orfanou[a]*, Lars Toettel[b], Eleni I. Vlahogianni[a], George Yannis[a]

[a]*National Technical University of Athens, 5, Iroon Polytechniou, Athens GR-15773 Athens Greece*
[b]*FZI Research Center for Information Technology, Haid-und-Neu-Str. 10-14, 76131 Karlsruhe, Germany*

Abstract

The era of automation has already been launched in the field of transportation, expected to increase road capacity and safety levels by reducing and eliminating crashes while the environmental impacts are also anticipated to be positive. Various studies have tried to analyze the behaviour of automated vehicles and their interaction with the surrounding traffic while concerning pedestrians, research is still limited. The present work aims to enrich existing research by modelling the behaviour of an automated vehicle when it interacts with a pedestrian with the intention to cross the road. For this purpose, vehicle and pedestrian trajectories from a virtual experiment are analyzed and the principles of inverse reinforcement learning are used for developing the model. The results are further discussed along with suggestions for future work.

Peer-review under responsibility of the scientific committee of the Transport Research Arena (TRA) Conference
Keywords: automated vehicles; interaction; pedestrians; inverse reinforcement learning

1. Introduction

One of the challenges in the field of traffic automation is modelling and simulating the behaviour of autonomous vehicles. Till now, researchers were focusing on developing driver behavioural models (DBMs) mimicking the human driver. However, the advent of autonomous vehicles creates the need to observe and model the behaviour of the 'machine-driver', consisting of various sensors exchanging information and acting according to the road context. Various DBMs have been proposed in literature and have formulated the basis for modelling (connected) automated vehicles ((C)AVs) in various simulation software. These models include the (Cooperative) Adaptive Cruise Control (C)ACC models, data driven approaches or are based on existing human behavioural models, whose parameters are

* Corresponding author. Tel.: +30 210 772 1367; fax: +0-000-000-0000 .
E-mail address: forfanou@central.ntua.gr

2352-1465
Peer-review under responsibility of the scientific committee of the Transport Research Arena (TRA) Conference

modified and adapted in order to mimic the automation function (Do et al. 2019). Existing research mainly focuses on modelling the behaviour of autonomous vehicles within the traffic flow and assesses their impact on critical areas such as safety, traffic and network efficiency, public transportation performance and environment. In these cases, the models are tested in simulation environments, where different profiles of the autonomous "driver" (aggressive, normal, conservative) can also be developed. More specifically, autonomous vehicles are studied concerning their interaction with the surrounding traffic (i.e., vehicles) in terms of the distance to their preceding vehicles and the side distances to the right and left

Modelling the interaction of autonomous vehicles with vulnerable road users, such as pedestrians, is still in its infancy, but gradually gains increasing attraction due to its safety implications. Not only must the autonomous vehicle be able to detect the vulnerable road user, it must also react properly and effectively to avoid collisions and allow unobstructed and safe movement of the vulnerable road user. Therefore, this work aims at developing a behavioural model for autonomous vehicles to safely interact with a pedestrian with the intention to cross the road, starting from the sidewalk

2. The Model

Data driven models are flexible and can reveal new variables important for driver behaviour description and modelling that could not be detected through traditional models. Furthermore, traditional models are based on specific formulas, making them more restrictive. In the era of big data, where vehicles can transmit numerous data through V2X communication, data driven models are considered to be the solution towards more evolutionary behavioural models that can be integrated to simulation platforms for training the calibrating and training the model and assess the impact of the corresponding behaviour on critical areas. After the stability and validity of the model is approved, it can be introduced in real test cars. The models are trained using real vehicle data and are validated and calibrated using various machine learning techniques. Data driven models have been used for enhancing existing models describing the car-following (Zhang et al., 2019) as well the lane changing behaviour (Bi et al., 2016, Wang et al., 2017), adaptive cruise control (Lin et al., 2020) and other autonomous driving applications (Di et al., 2021, Bachute et al., 2021, Kiran et al., 2021, Palasinamy, 2020, Talpaert et al., 2019).

A machine learning technique widely used in many applications in transportation engineering is Inverse Reinforcement Learning (IRL) is a modelling framework aiming to learn the reward function based on the states, actions and the optimum policy defined. IRL has been used for modelling interactions between different users in the road sector such as pedestrians and cyclists (Alsaleh and Sayed, 2020) as well as for user behaviour such as pedestrian trajectories (Martinez-Gil et al., 2020), vehicle navigation on a highway (Levine et al., 2010), risk anticipation (Shimosaka et al., 2014) and autonomous vehicle decision making and behaviour (Gao et al., 2018, Sharifzadeh et al., 2017). It has many structures such as Maximum Entropy (Ziebart et al., 2008), Deep Maximum Entropy (Wulfmeier et al., 2016), Adversarial IRL (Fu et al., 2017, Wang et al., 2021) and many more which can be found in the literature.

The algorithm used in the present study for implementing inverse reinforcement learning is the Maximum Entropy (ME) algorithm developed by Ziebart et al. (2008) which assumes optimum behaviour (Alsaleh and Sayed, 2020). In this case, the optimum policy is considered to be extracted by the given n trajectories of an expert which are a sequence of states s and actions a (Eq. 1) . Each trajectory has a temporal horizon of h steps.

$$T = \{(s_1^1, a_1^1), (s_2^1, a_2^1), \dots\dots, (s_h^1, a_h^1), (s_1^2, a_1^2), (s_1^2, a_1^2), \dots\dots, (s_h^2, a_h^2), \dots\dots., (s_1^n, a_1^n), (s_1^n, a_1^n), \dots\dots, (s_h^n, a_h^n)\} \quad (1)$$

For simplifying the model, single agent IRL is implementing for modelling the behavior of an automated car when a pedestrian appears aiming to cross the road and therefore only the trajectories of the vehicle are being taken into consideration assuming the pedestrian is an external object. The existence of the pedestrian is included when defining the states of the agent-vehicle as it is described in the next section.

Generally, it can be assumed that the reward function depends on some features φ_i and is expressed as a linear (as in the Maximum Entropy algorithm) or nonlinear equation (examples in Eq. 2 and Eq. 3):

$$R\_linear(s,a) = \sum_i w_i \varphi_i(s,a) \quad (2) \quad \text{or} \quad R_{non}linear(s,a) = neuralNet(\varphi(s,a); w) \quad (3)$$

3. Application

3.1. The dataset

The data for the model training and evaluation are collected through a virtual reality experiment, led by FZI Research Centre for Information Technology, which took place in Karlsruhe, Germany. In these experiments, a human expert immersed into the scene via a virtual reality (VR) headset as a pedestrian with the aim of crossing the road. At the same time, a simulated vehicle was approaching from the left, and it was either controlled by a human using a steering wheel and pedals or by a highly automated driving function. The digital twin of the real test area used for the virtual simulated experiments is depicted in Fig. 1. The data include the following information:

Fig. 1. Digital twin of the Test Area.

The dataset is visualized using the INTERACTION dataset visualization tool (https://github.com/interaction-dataset/interaction-dataset) as shown in Fig. 2. Rectangles depict vehicles (green circle) and the blue dots are referring to pedestrians (red circle).

Table 1. Data collected from the virtual experiment

Field	Description
Timeframe	data is collected every 100 ms
Agent -type	since the experiment investigates the interaction between passenger cars and pedestrians the “agent-type” parameter takes the values “car” or “pedestrian”
x,y	the x and y position of the agent (m)
vx, vy	speed values of the agent in the x and y dimension (m/s)
psi_rad	the yaw angle of the agent (rad)
length	length of the agent
width	width of the agent
ax, ay	acceleration/deceleration values of the agent in the x and y dimension (m/s2)
time-headway	the temporal distance of the agent from its preceding vehicle (s)
gap	the spatial distance of the agent from its preceding vehicle (m)
lateral_position	the distance of the central axis of the vehicle from the central axis of the lane (m)
side_distance	the distance from the central axis of the agent from the side object (m)
mode	this parameter describes whether automation mode is on or manual vehicle control is applied. Since automation can be activated only in case the agent-type is a car, this parameter takes the values “automated” and “simulated” if automated function is on and off respectively. For pedestrians, the only value for this parameter is “simulated”.

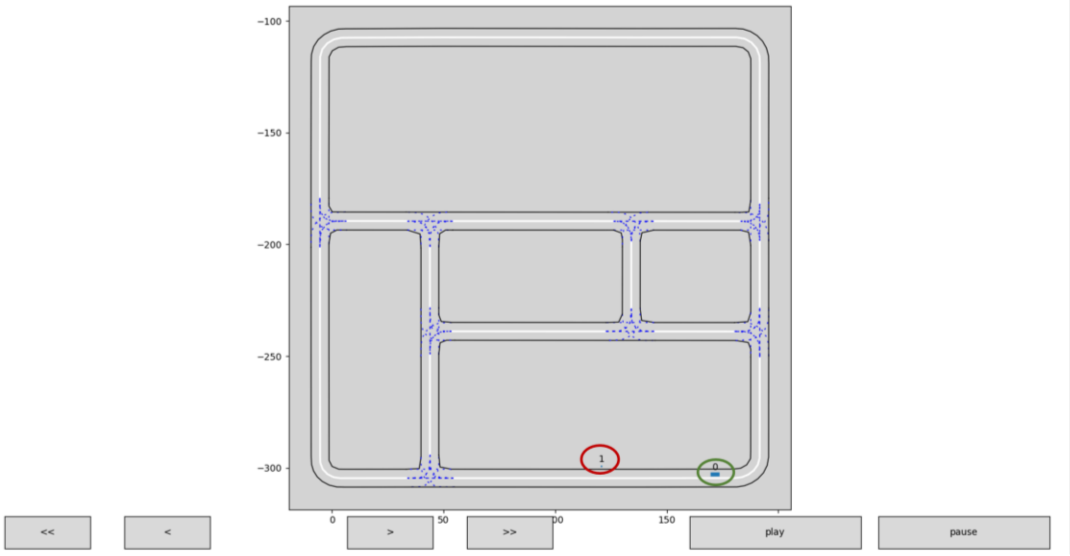

Fig. 2: Visualization of data collected in the VR simulator

3.2. Critical Conditions

Various factors have been used for describing the interaction between vehicles and pedestrians. A literature review was carried out in order to find the most frequent parameters as well as some critical values influencing the decision of a pedestrian to cross the road when a vehicle is approaching. In Petzoldt (2014), the critical gaps were estimated 3.5s and slightly less than 3s for a speed of 30km/h and 50km/h. Clamann et al. (2016) found that a time interval between 4s and 7s is critical for the pedestrian as he may intersect with the vehicle's trajectory. A mean critical gap of around 4.1s - 4.8s and mean critical gap distance of 67m-79m were found in a study conducted by Pawar and Patil (2016) with a vehicle approaching speed of 62km/h. Palmeiro et al. (2017) tried to analyse whether and how the decision to cross the road is influenced based on the vehicle type (traditional or automated). The results showed that the critical time gap in case a conventional vehicle is approaching is around 5.5s, while in case of AV 7s approximately. In the same experiment, the spatial gap was found to range between 20m - 26m for pedestrian interaction with traditional vehicle and 19m – 23m if an AV is approaching. The critical vehicle speed was also recorded and estimated at 16km/h and 12km/h for traditional and automated vehicle, respectively. Recently, a virtual reality experiment was set up by Woodman et al. (2019), which tested the pedestrian behaviour for time gaps of 2s-5s. In their experiment, all participants rejected the gap of 2s, while the highest percentage accepted the gap of 5s. Oxley et al. (2005) carried out an experiment in order to find out how the pedestrian age affects the parameters influencing the crossing decision. For their analysis, they use the spatial gap and the vehicle speed (resulted in the time gap). Apart from the time and spatial gap and the vehicle speed, time to collision has also been used for studying the interaction between pedestrians and vehicles. Schneemann and Gohl (2016) conducted a study for observing the interaction between a driver and a pedestrian under two different TTC values. For their experiment, the authors chose the values of 3s and 4s as the critical ones for assessing pedestrian's gap acceptance. The review of Rasouli et al. (2018) revealed that the gap acceptance in terms of TTC is between 3s and 7s, with the threshold of the 3s meaning that the pedestrian will not decide to cross and if the TTC is more than 7s the pedestrian will not cross.

3.3. State Definition

The data collected from the virtual experiments were used to extract or estimate variables for defining the states space S=(s1, s2, ….sn) where is the number of the given trajectories. Three features are used for defining the states of the interaction between the vehicle and the pedestrian: difference of the vehicle and the pedestrian speeds, spatial gap between the two road users and the vehicle speed. Each feature was divided in various levels based on the k means clustering results. More specifically, the elbow method and the silhouette coefficient were used in order to define the most appropriate number of clusters. Both methods revealed that the features "vehicle speed" and "speed difference" should be discretized in two levels while the feature "gap" in three. After this step, the k means method was applied for defining the thresholds of each level (Table 2). Combining the three features and their levels resulted in 12 states (2x2x3).

Table 2. Features identified for state definition

	Vehicle Speed (m/s)	Speed Difference (m/s)	Spatial Gap (m)
1	(0.005, 10.60)	(0.004, 10.50)	(4.20, 11.80)
2	[10.60, 25.20)	[10.50, 25.40)	[11.80, 19.50)
3			[19.50, 33.10)

3.4. Actions Definition

For the action space, we consider the acceleration as the critical value to define the manner a driver / AV will react to external stimuli (e.g. interaction with pedestrian). Based on the acceleration/deceleration values, two levels are distinguished: (1) smooth and (2) harsh acceleration/deceleration. The thresholds for this classification was found to be around 0.16g – 0.36g (g=9.81m/s2) as it is described in Vlahogianni and Barbounakis (2017). For safety purposes, the value 0.16g (≈1.57m/s2) was chosen as the upper limit for considering that the vehicle accelerates/decelerates smoothly. Besides, there is also the possibility that the driver will not take any action remaining at his current state. Based on the above, 5 actions can be distinguished, as shown in Table 3.

Table 3. Actions identified

Actions	Value
Cruising (no change in speed)	0
Smooth acceleration	(0, 1.57]
Harsh acceleration	(1.57, 4.5]
Smooth deceleration	[-1.57, 0)
Harsh deceleration	[-9, -1.57)

4. Results

4.1. State features and action attributes statistical analysis

Vehicle speed, speed difference and spatial gap were the three features used for describing the vehicle state while the action of the driver was defined by the vehicle acceleration/deceleration. Table 4 shows the descriptive statistics of these five parameters while Fig. 3 presents the distribution of their values and the best fitted distribution the statistical analysis revealed. It should be mentioned that acceleration values over 4.5m/s2 and deceleration values higher than 9m/s2 (emergency braking) were excluded from the analysis.

Table 4. Descriptive Statistics

	Acceleration (m/s^2)	Deceleration (m/s^2)	Vehicle Speed (km/h)	Speed Difference (km/h)	Spatial Gap (m)
Mean	1.101	1.655	7.285	7.375	13.114
Standard deviation	0.989	1.610	8.675	6.774	5.767
Median	0.789	1.226	2.040	4.569	11.736
Min value	0.007	0.040	0.003	0.004	4.272
25%	0.330	0.633	0.130	2.673	8.889
50%	0.789	1.226	2.040	4.569	11.736
75%	1.547	2.050	14.805	11.349	16.026
Mean	1.101	1.655	7.285	7.375	13.114

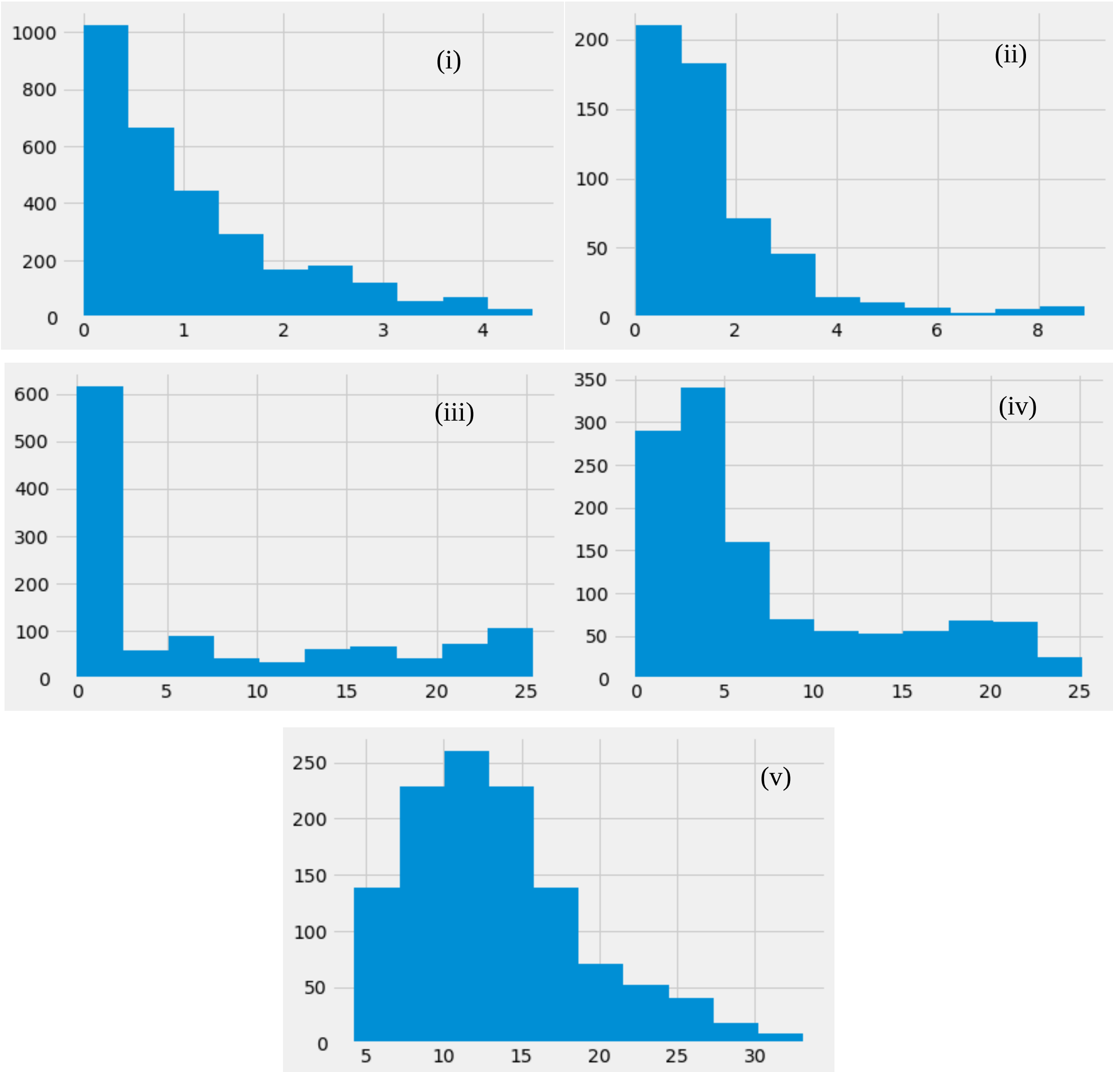

Fig. 3: Distribution of values of (i) acceleration, (ii) deceleration, (iii) vehicle speed, (iv) speed difference and (v) spatial gap

4.2. The Reward function

The implementation of the maximum entropy inverse reinforcement learning algorithm will produce the reward function weights for the features used for the state definition, i.e. vehicle speed, speed difference and spatial gap and for each of the defined levels for each feature. According to the results, shown in Figure 4, states including vehicle speed of level 1 give higher reward than the states with much higher speeds (level 2) indicating that the vehicle should decelerate when the pedestrian appears and starts moving tending to cross the road. Concerning the spatial gap, the highest reward value is observed for level 3, corresponding to higher values of the distance between the pedestrian and the approaching vehicle showing that the pedestrian feel safer to cross when the distance is greater. On the other hand, it seems that the drivers do not prefer to keep very low spatial gaps with the pedestrian as level 1 has the lowest reward function weight. Finally, as far as the speed difference is concerned, drivers seem to prefer not to be very slow compared to the pedestrians and therefore speed differences values belonging the second level give higher reward compared to lower values. This reveals that even though drivers decelerate in the vicinity of a pedestrian, they do not prefer to apply maximum deceleration but to keep a reasonable speed.

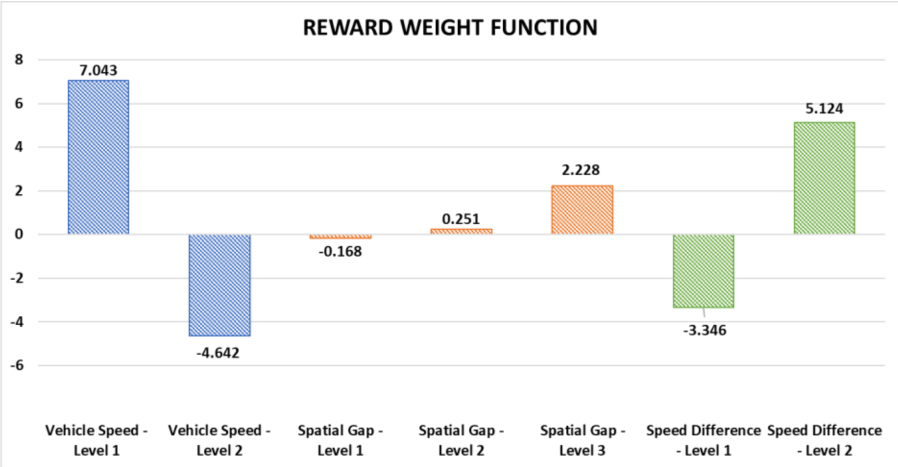

Fig. 4: Reward function weights for the different levels of the three features

5. Conclusion and future work

The present work is a preliminary effort in modelling the behaviour of an automated vehicle when it interacts with a pedestrian standing in the curb aiming to cross the road. Based on trajectories from an automated vehicle collected through a virtual experiment and the principles of inverse reinforcement learning, it was attempted to investigate, study and model the behaviour of an equipped vehicle that will ensure safe interaction with vulnerable road users. The model which is single agent, as it investigates the behavior of the vehicle, can improve safety not only in crossings and shared spaces but also in spots where a pedestrian could unexpectedly try to cross the road and the automated/autonomous vehicles would be required and expected to react immediately and efficiently to avoid a crash and, as a result, increase the safety levels.

The proposed model can be further improved with additional data, collected from the virtual experiment or data from real trajectories collected in real infrastructure pilots. Additionally, it can be integrated in automated vehicles for assisting the machine in case of pedestrian appearance as well as decide the best sequence of actions based on the distance from the pedestrian, his speed as well as the vehicle speed. Additionally, we further intend to implement a multi agent maximum entropy inverse reinforcement learning to introduce a dynamic, instead of static, environment and analyze the behavior of both the pedestrian and the vehicle and what are the principles leading to a safe interaction.

Acknowledgements

The Drive2theFuture project is funded by European Commission under the MG-3.3.2018: "Driver" behaviour and acceptance of connected, cooperative and automated transport; Research and Innovation Action (RIA).

References

Alsaleh, R., & Sayed, T. (2020). Modelling pedestrian-cyclist interactions in shared space using inverse reinforcement learning. Transportation research part F: traffic psychology and behaviour, 70, 37-57.

Bachute, M. R., & Subhedar, J. M. (2021). Autonomous Driving Architectures: Insights of Machine Learning and Deep Learning Algorithms. Machine Learning with Applications, 6, 100164.

Bi, H., Mao, T., Wang, Z., & Deng, Z. (2016). A data-driven model for lane-changing in traffic simulation. In Symposium on Computer Animation (pp. 149-158).

Clamann, M., Aubert, M., & Cummings, M. L. (2017). Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles (No. 17-02119).

Di, X., & Shi, R. (2021). A survey on autonomous vehicle control in the era of mixed-autonomy: From physics-based to AI-guided driving policy learning. Transportation research part C: emerging technologies, 125, 103008.

Do, W., Rouhani, O. M., & Miranda-Moreno, L. (2019). Simulation-based connected and automated vehicle models on highway sections: a literature review. Journal of Advanced Transportation, 2019.

Fu, J., Luo, K., & Levine, S. (2017). Learning robust rewards with adversarial inverse reinforcement learning. arXiv preprint arXiv:1710.11248.

Gao, M., & Shi, G.-Y. (2019). Ship Spatiotemporal Key Feature Point Online Extraction Based on AIS Multi-Sensor Data Using an Improved Sliding Window Algorithm. Sensors, 19. doi:10.3390/s19122706

Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A. A., Yogamani, S., & Pérez, P. (2021). Deep reinforcement learning for autonomous driving: A survey. IEEE Transactions on Intelligent Transportation Systems.

Levine, S., Popovic, Z., & Koltun, V. (2010). Feature Construction for Inverse Reinforcement Learning. In NIPS (Vol. 23, p. 1342).

Lin, Y., McPhee, J., & Azad, N. L. (2020). Comparison of deep reinforcement learning and model predictive control for adaptive cruise control. IEEE Transactions on Intelligent Vehicles, 6(2), 221-231.

Martinez-Gil, F., Lozano, M., García-Fernández, I., Romero, P., Serra, D., & Sebastián, R. (2020). Using Inverse Reinforcement Learning with Real Trajectories to Get More Trustworthy Pedestrian Simulations. Mathematics, 8(9), 1479.

Oxley, J. A., Ihsen, E., Fildes, B. N., Charlton, J. L., & Day, R. H. (2005). Crossing roads safely: an experimental study of age differences in gap selection by pedestrians. Accident Analysis & Prevention, 37(5), 962-971.

Palmeiro, A. R., van der Kint, S., Vissers, L., Farah, H., de Winter, J. C., & Hagenzieker, M. (2018). Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment. Transportation research part F: traffic psychology and behaviour, 58, 1005-1020.

Palanisamy, P. (2020). Multi-agent connected autonomous driving using deep reinforcement learning. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-7). IEEE.

Panwai, S., & Dia, H. (2007). Neural agent car-following models. IEEE Transactions on Intelligent Transportation Systems, 8(1), 60-70.

Petzoldt, T. (2014). On the relationship between pedestrian gap acceptance and time to arrival estimates, Accident Analysis & Prevention, Volume 72, Pages 127-133, ISSN 0001-4575, https://doi.org/10.1016/j.aap.2014.06.019.

Rasouli, A., Kotseruba, I., & Tsotsos, J. K. (2017). Agreeing to cross: How drivers and pedestrians communicate. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 264-269). IEEE.

Schneemann, F., & Gohl, I. (2016). Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments. In 2016 IEEE intelligent vehicles symposium (IV) (pp. 38-43). IEEE.

Sharifzadeh, S., Chiotellis, I., Triebel, R., & Cremers, D. (2016). Learning to drive using inverse reinforcement learning and deep q-networks. arXiv preprint arXiv:1612.03653.

Shimosaka, M., Kaneko, T., & Nishi, K. (2014). Modeling risk anticipation and defensive driving on residential roads with inverse reinforcement learning. In 17th International IEEE Conference on Intelligent Transportation Systems (ITSC) (pp. 1694-1700). IEEE.

Talpaert, V., Sobh, I., Kiran, B. R., Mannion, P., Yogamani, S., El-Sallab, A., & Perez, P. (2019). Exploring applications of deep reinforcement learning for real-world autonomous driving systems. arXiv preprint arXiv:1901.01536.

Vlahogianni, E. & Barmpounakis, E. (2017). Driving analytics using smartphones: Algorithms, comparisons and challenges, Transportation Research Part C: Emerging Technologies, Volume 79, Pages 196-206.

Wang, Z., G. Wu, P. Hao, K. Boriboonsomsin and M. Barth, "Developing a platoon-wide Eco-Cooperative Adaptive Cruise Control (CACC) system," 2017 IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, 2017, pp. 1256-1261, doi: 10.1109/IVS.2017.7995884.

Wang, P., Liu, D., Chen, J., Li, H., & Chan, C. Y. (2021, May). Decision making for autonomous driving via augmented adversarial inverse reinforcement learning. In 2021 IEEE International Conference on Robotics and Automation (ICRA) (pp. 1036-1042). IEEE.

Woodman, R., Lu, K., Higgins, M. D., Brewerton, S., Jennings, P. A., & Birrell, S. (2019). Gap acceptance study of pedestrians crossing between platooning autonomous vehicles in a virtual environment. Transportation research part F: traffic psychology and behaviour, 67, 1-14.

Wulfmeier, M., Ondruska, P., & Posner, I. (2015). Maximum entropy deep inverse reinforcement learning. arXiv preprint arXiv:1507.04888.

Zhang, X., Sun, J., Qi, X., & Sun, J. (2019). Simultaneous modeling of car-following and lane-changing behaviors using deep learning. Transportation research part C: emerging technologies, 104, 287-304.

Zheng H, Son Y-J, Chiu Y-C, Head L, Feng Y, Xi H, Kim S, Hickman M. (2013). A primer for agent-based simulation and modelling in transportation applications. U.S. Department of Transportation - Federal Highway Administration. Report FHWA-HRT13-054.

Ziebart, B. D., Maas, A. L., Bagnell, J. A., & Dey, A. K. (2008). Maximum entropy inverse reinforcement learning. In Aaai (Vol. 8, pp. 1433-1438).