Modelling the Behaviour of Automated Vehicles when Interacting with Pedestrians in Jaywalking

Foteini Orfanou[1], Eleni Vlahogianni[1], George Yannis[1]
[1]School of Civil Engineering, National Technical University of Athens, Greece
E-mail:forfanou;elenivl;geyannis@central.ntua.gr

Abstract

Autonomous vehicles create 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 the literature and have formulated the basis for modelling AVs and CAVs in the various simulation software. The existing research is mainly focusing on modelling the behaviour of autonomous vehicles concerning their interaction with the surrounding traffic (i.e. vehicles) in terms of the keeping distance from their preceding vehicles and the side distances from the vehicles on their right and left side.

The investigation of the interaction of autonomous vehicles with other road users is still in its infancy. Especially, the interaction with vulnerable road users, such as pedestrians, is currently under research and it gains increasing attraction as it is a major pilar in road safety. This work aims to fulfil the need of the development of an additional behavioral model describing the safe interaction between an autonomous vehicle and a pedestrian standing on the curb aiming to cross the road. For this purpose, data were collected through a virtual reality experiment where a human user immersed into the scene as a pedestrian with the aim of crossing the road at an unmarked location. At the same time, a simulated vehicle was approaching, controlled by a highly automated driving function. The behavioural model developed is based on the principles of inverse reinforcement learning and the algorithm implemented is the Maximum Entropy (ME) algorithm which assumes optimum behaviour, and the expert behaviour is modelled as the one with the maximum entropy.

Keywords: *inverse reinforcement learning, maximum entropy, automated vehicle, pedestrian*

Σύντομη Περίληψη

Τα αυτόνομα οχήματα επιβάλλουν την μοντελοποίηση και την συνεχή παρακολούθηση της συμπεριφοράς του «οδηγού-μηχανής» η οποία ενεργεί ανάλογα με τις πληροφορίες που συλλέγονται από το οδικό περιβάλλον και ανταλλάσσονται μεταξύ των ποικίλων αισθητήρων που είναι εγκατεστημένοι στο όχημα. Διάφορα μοντέλα οδηγικής συμπεριφοράς έχουν αναπτυχθεί κατά καιρούς και έχουν αποτελέσει τη βάση για τη μοντελοποίηση της συμπεριφοράς των (συνεργαζόμενων) αυτόνομων οχημάτων και των επιπτώσεων αυτών μέσω των λογισμικών προσομοίωσης. Μέχρι τώρα η έρευνα επικεντρώνεται κυρίως στην μελέτη της αλληλεπίδρασης των οχημάτων αυτών με τα υπόλοιπα οχήματα όσον αφορά την απόστασή τους από το προπορευόμενο όχημα καθώς και τις πλευρικές αποστάσεις από τα οχήματα που βρίσκονται εκατέρωθεν.

Η ανάλυση της αλληλεπίδρασης των αυτόματων και αυτόνομων οχημάτων με τους υπόλοιπους χρήστες της οδού βρίσκεται ακόμα σε πρώιμο στάδιο. Πιο συγκεκριμένα, η αλληλεπίδρασή τους με τους ευάλωτους χρήστες, όπως οι πεζοί, βρίσκεται υπό συνεχή έρευνα και προσελκύει αυξανόμενα το ενδιαφέρον των ερευνητών καθώς αποτελεί σημαντικό πυλώνα της οδικής ασφάλειας. Η παρούσα μελέτη έχει σκοπό να πληρώσει την ανάγκη δημιουργίας ενός επιπλέον μοντέλου οδηγικής συμπεριφοράς που θα περιγράφει την ασφαλή αλληλεπίδραση αυτόματου οχήματος και πεζού, ο οποίος στέκεται στο κράσπεδο με σκοπό να διασχίσει την οδό σε κάποιο σημείο αυτής. Για το σκοπό αυτό, χρησιμοποιήθηκαν δεδομένα τα οποία συλλέχθησαν στο πλαίσιο πειράματος εικονικής πραγματικότητας κατά το οποίο ο συμμετέχων εισέρχεται στο δίκτυο ως πεζός ενώ την ίδια στιγμή ένα όχημα επιπέδου αυτοματισμού 4 κινείται στην οδό και πλησιάζει προς το σημείο που στέκεται ο πεζός. Το μοντέλο που αναπτύχθηκε βασίζεται στις αρχές τις αντίστροφης ενισχυτικής μάθησης και πιο συγκεκριμένα χρησιμοποιήθηκε ο αλγόριθμος της μέγιστης εντροπίας.

Λέξεις κλειδιά: *αντίστροφη ενισχυτική μάθηση, μέγιστη εντροπία, αυτόματο όχημα, πεζός*

1. Introduction

Pedestrians consist a large part of vulnerable road users in transportation networks with unpredictable behaviour and different decision making and motion profiles and patterns. Their interaction with vehicular traffic is a major pillar in the field of transportation as these two type of users can intersect at signalized and non-signalized intersections, mid-block crosswalks and any other unmarked part of the road network (jaywalking). Due to pedestrians' vulnerability it is vital to maintain high safety levels in any type of interaction. Vulnerable road users will be highly affected by the advent of automated vehicles and their interactions can lead to unsafe situations as the behaviour of VRUs is highly unexpected or the behaviour of the AV could be misinterpreted by the VRU. Till now, most studies include field tests (Rothenbuecher et al., 2016, Matthew et al., 2017, Wang et al., 2021) or virtual experiments (Deb et al., 2018, Velasco et al., 2019, Sadraei et al., 2020, Pappas et al., 2021) for observing and analysing these interactions and parameters such as the gap acceptance, Time to Collision (TTC), or AV speed evolution can be recorded and extracted. Additionally, various experiments have been conducted for investigating which types of communication displays (text, sound, etc) are the most suitable and preferable by pedestrians when they interact with AVs and can ensure safe pedestrian crossing (Ackermann et al., 2019a,b, de Clercq et al., 2019, Merat et al., 2018, Langstrom et al., 2016, Clamann et al., 2017, Fridman et al., 2017, Matthews et al., 2017). Various studies have attempted to model the behaviour of pedestrians in the presence of an AVs or the behavior of AVs when interacting with pedestrians aiming to contribute in increasing safety in these types of interactions.

Jayaraman et al. (2020) modeled and predicted pedestrian crossing behaviour in the presence of an AV by developing a hybrid automaton model by combining both a discrete and a continuous model. Rad et al. (2020) focused on pedestrian crossing behaviour and modelled the probability for a pedestrian to cross the road based on different parameters using agent - based simulation modelling. An agent – based model was also developed by Predhumeau et al. (2021) for assisting autonomous vehicles in predicting pedestrian trajectories in shared spaces. The authors combine a Social Force Model describing the pedestrian's willingness to reach his destination by avoiding any obstacle (static, pedestrians, vehicles, AVs) and a decision model in case of conflicting interactions with an AV for avoiding a potential collision. Zhu et al. (2022) built an agent – based framework for analyzing encounters between autonomous vehicles and pedestrians at an unsignalized mid-block crosswalk with refuge island and assess the safety levels of this type of conflict point. Cellular automata models consist a popular technique for modelling vehicle and pedestrian behaviour during their encounters. Such a model was developed by Feng et al. (2019) for simulating the interactions between autonomous vehicles and pedestrians at unsignalized mid-block crosswalks combined with a conflict eliminating model for reducing the number of conflicts between the two road user types. Wu et al. (2019) applied in a two-way four-lane cellular automata simulation model for expressing the probability of waiting or crossing of the two "players" based on gap acceptance conditions. Assuming a Poisson pedestrian arrival rate moving upward and downward, the effect of vehicle and pedestrian arrival rates on pedestrian delay was estimated, while the TTC indicator was used for evaluating conflicts and their severity on studied crosswalk. Apart from SFMs and CA models, surrogate safety models (SSM) models have also been used for modelling the conflict between autonomous vehicles and pedestrians (Alghodhaif and Lakshmanan, 2020) in urban environment. The model developed consisted of three sub-models –radar perception, pedestrian classification and decision making model – aiming to adopt the autonomous vehicle behaviour for conflict avoidance with a pedestrian intended to cross.

Pedestrian – autonomous vehicle interaction has also been described and modelled based on the principles of game theory (Michieli and Badia, 2018) using data from virtual reality experiments

(Camara et al., 2018, Fox et al., 2018, Camara et al., 2021), or distributed simulator (Kalantari et al., 2021). A game theory – based model was proposed by Rahmati et al. (2020) for capturing the characteristics and uncertainties of pedestrian movements in shared spaces when interacting with other road users and predict their trajectories for assisting Connected and Automated Vehicles (CAVs) in their navigation through high volume pedestrian areas. Luo et al. (2018) extended the Reciprocal Velocity Obstacle model (Van den Berg et al., 2009) for describing pedestrian movement intentions and his interaction with other users (vehicles or pedestrians). The proposed model, called Optimal Reciprocal Collision Avoidance (ORCA), is integrated within a Partially Observable Markov Process for ensuring safe navigation of the autonomous vehicle among pedestrians. Extensions to the ORCA model have also been introduced by Charlton et al. (2020) to simulate interactions between pedestrians and autonomous vehicles in urban shared spaces. An important shortcoming of these geometric models is that they do not consider the variability and unpredictability of human behaviour and produce optimal collision free trajectories and assume similar reasoning for collision avoidance.

Data driven methods are gaining increasing attention as they are more flexible than the traditional behavioural models and can easily be adapted to any data source and environment as well as reveal patterns ad profiles that other methods and algorithms fail to. Deep reinforcement learning principles have been applied for activating autonomous braking in risky situations (Vasquez and Farroq, 2019) or predicting pedestrian trajectories (Kalatian and Farooq, 2022). Deep survival analysis models and more specifically a deep – neural network CHP model was applied in Kalatian and Farooq (2021) for estimating pedestrian waiting time before starting crossing the road when an autonomous vehicle is approaching at an unsignalized crosswalk under different vehicles' levels of automation and arrival rates, road types, lane widths, time of the day and weather conditions. A deep Q Learning RL model was developed by Elallid et al. (2022) for controlling an autonomous vehicle taking into consideration its interactions with other vehicles and pedestrians. The principles of maximum entropy developed by Ziebart et al. (2008) were used for developing a model for predicting pedestrian behaviour using continuous trajectories recorded in a real – world experiment (Kuderer et al.,2012). The proposed model outperformed the social forces method and was then integrated in a mobile robot, which interacted successfully with humans.

Inverse reinforcement learning (IRL) is also a very valuable tool in modelling vehicle and pedestrian behaviours when interacting with each other. A continuous Gaussian IRL was implemented for describing pedestrian behaviour using trajectories from a signalized intersection in China (Nasernejad et al., 2021). After estimating the pedestrian reward function, the Advantage Actor-Critic DRL method (A2C) was applied for revealing pedestrian policies and then for extracting pedestrian trajectories. The simulate trajectories were compared with the real collected ones and the results revealed the high accuracy and good performance of the proposed model. Zhu et al. (2022) proposed a Convolutional Neural Networks (CNN) Vehicle-Pedestrian detection algorithm aiming to increase the safety levels of pedestrian – autonomous vehicle interactions. Very recently, reinforcement learning techniques have been adopted for assisting autonomous vehicles in planning their trajectories at intersections where they coexist with other road users (vehicles, pedestrians and cyclists) (Zhang et al., 2023). The developed DQN model resulted in safer driving trajectories, increased driving comfort and energy savings while the trajectories planned for a 5s horizon outperformed human driven ones.

Pedestrian group behaviour and navigation strategies have also been investigated through reinforcement learning techniques. Focusing on pedestrians, inverse reinforcement algorithm and maximum causal entropy principles have been applied for learning their behaviour from real trajectories when navigating in a virtual 3D environments and predict trustworthy trajectories (Martinez – Gill et al., 2020). The results revealed the model effectiveness and its ability to adequately depict the pedestrian dynamics and its motion features and principles. An autonomous braking system based on deep reinforcement principles was proposed in Chae et al. (2017) for controlling and adjusting vehicle speed in critical situations and especially in urban environments where a vehicle faces a crossing pedestrian. The state space of the vehicle is described by its speed and its distance from the pedestrian, while the action space depends on the value of its deceleration. The system was tested through computer simulations which revealed the good and consistent braking behaviour of the vehicle under different scenarios. LSTM was combined also with a Deep Q-Network for modelling decisions of an autonomous vehicle in urban environments in the presence of pedestrians ay unsignalized crosswalks and jaywalking (Deshpande et al., 2020). The model was trained in CARLA simulator and the comparison with a rule – based model revealed its effectiveness and higher performance. One year later, the same research team developed a multi-objective deep reinforcement learning for assisting autonomous navigation in urban environments (Deshpande et al., 2021). The results from the CARLA simulator revealed the better performance of the new approach compared to the single objective RL proposed in their previous work (Deshpande et al., 2020). Rosati Pappini et al. (2021) used reinforcement learning principles for developing a safe speed function to assist an autonomous driving system when interacting with a distracted pedestrian aiming to ensure that the machine will successfully perform emergency braking maneuvers. The model integrated in a real test vehicle and interacted with emulated pedestrians performed successfully in the scenarios conducted on a real test track as the vehicle reached the minimum speed within a safe temporal margin from the potential conflict area.

Apart from single agent algorithms maximizing the reward and optimize the behaviour of one of the interacting agents (vehicle or pedestrian), multi agent techniques are considered to be more realistic and accurate as they take into consideration the interacting behaviours of the users aiming to optimize both policies and maximize both reward functions. Trumpp et al. (2022) proposed a deep multi agent reinforcement learning algorithm (DMARL) for modelling the behaviour of a pedestrian aiming to cross the road and an approaching autonomous vehicle at an unmarked crosswalk in a simulated driving scenario. Time to collision, pedestrian and vehicle speed, vehicle acceleration, the distance between them are some of the main parameters for defining the state space of the two agents. Vehicle's action space is defined based on its acceleration while pedestrian has two options either walk or wait. A similar study was conducted by Nasernejad et al. (2021) applying Adversarial Inverse Reinforcement Learning for recovering the reward functions of the vehicle and the pedestrian while their policies where further optimized using multiagent Actor-Critic deep-reinforcement-learning. The study of Hu et al. (2023) focused on autonomous and manual vehicles and their interaction with pedestrians on unsignalized intersections using multi-agent deep deterministic policy gradient (MADDPG) algorithm. State of the pedestrian is discrete (wait or walk) based on the TTC values while the vehicles state space is defined based on the TTC values, their own speed and acceleration, the speed of the preceding vehicle (if any) and its type (manual or autonomous), pedestrian speed and the distance between the vehicle and the pedestrian (if there is no leading vehicle) or the leading vehicle. The results showed that the proposed algorithm outperformed single agent DDPG and DRL considering the reward score of all agents and its fluctuation. The addition of LSTM in the MADDPG algorithm resulted in a multi agent recurrent deep deterministic policy gradient algorithm applied for traffic light control at multiple urban intersections where vehicle and crossing pedestrians are interacting with each other (Wu et al., 2020). Apart from vehicles and pedestrians, the authors also include buses giving them higher priority to pass than the regular vehicles. The proposed algorithm succeeded in optimizing the each traffic light control policy to achieve the highest reward and optimal decisions in the whole analyzed network as well as performing well in complicated road conditions and in an unstable external environment. Focusing only on pedestrian behaviour, Martinez et al. (2014) developed a multi – agent reinforcement learning based framework for simulating pedestrian behaviour and navigation strategy in a virtual environment. In the three scenarios tested , i.e. election of the shortest path vs. quickest path, crossing between two groups of pedestrians walking in opposite directions inside a narrow corridor and two agents that move in opposite directions inside a maze, the proposed simulation approach succeeded in learning behaviours and policies that resemble pedestrians in all levels (strategical, tactical and functional) effectively and similarly to the traditional Helbing's social force model. Hsu et al. (2018) tried to express the interaction between a vehicle and an approaching pedestrian aims to cross an unsignalized intersection through a Markov Decision Process model. The authors are focusing on simulating an autonomous vehicle' behaviour resembling human controlled vehicle dynamic that would react appropriately to pedestrian movements and behaviour changes so that the vehicle would avoid any collision. The reward function penalized vehicles for entering the crosswalk area simultaneously with pedestrians.

It is obvious that data driven models have increased and more advanced capabilities succeeding in dealing with huge amount of data and capturing any unrevealed relationship. Instead of just replicating behaviours, they are also able to optimize policies resulting in more reliable and safer behavioural profiles. Additionally, signalized and unsignalized intersections and midblock crosswalks are the major focus point of existing research and the scenarios where pedestrian cross the road at a random location (jaywalking) need to be further investigated as they are riskier and a prompt and safe reaction from the driver is urgent to avid a potential collision. This work aims to further enrich past research by applying maximum entropy inverse reinforcement learning algorithm for describing the encounter between a Level 4 automated vehicle and a pedestrian standing on the curb aiming to cross the road at an unmarked location of the road. The paper is structured as follows: the main principles of maximum entropy inverse reinforcement learning are described in section 2, section 3 includes the description of the data, the model development and results are the focus point of section 4 while section 5 includes the conclusions and further research suggestions.

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.

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

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):

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

3. 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 Figure 1. The data include the following information:

Figure 1: *Digital twin of the Test Area.*

The dataset includes information about the position of the two agents, their kinematic characteristics each time step, their spatial and temporal distance as described in Table 1.

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".

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

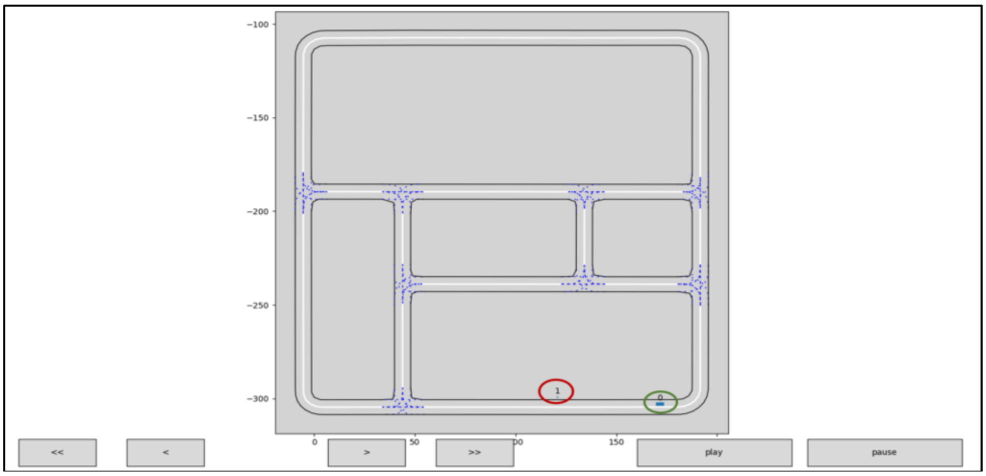

Figure 2: Visualization of data collected in the VR simulator

4. Application

4.1 Parameters for describing encounters between vehicles and pedestrians

Various parameters have been applied for describing and assessing the interactions between vehicles and pedestrians at different type of encounter location such as signalized or unsignalized intersections, crosswalks, and unmarked locations (jaywalking). The Surrogate Safety Measures (SSMs) are tools and indicators are considered to provide a better investigation and interpretation of factors that could lead to a potential collision and are very popular for studying road traffic safety as they can identifying safety critical events without waiting for a crash to occur. The most widely used indicator is Time to Collision (TTC), a time-based and continuous parameter applied for evaluating and classify encounters based on their severity at signalized intersections (El- Basyouny and Sayed, 2013, Sacchi and Sayed, 2016), unsignalized intersections (Kathuria and Vedagiri, 2020) uncontrolled and undesignated mid-block sections in urban environment (Golakiya et al., 2022, Bella and Nobili, 2020) or high speed arterials (Pawar and Patil, 2016). Variations of TTC also used for this purpose are Time to Collision Point (TTCP) (Zhang and Fu, 2022), inverse TTC (1/TTC) introduced by Kiefer et al. (2005), Time Exposed TTC and Time Integrated TTC (Laureshyn et al., 2010), Time-to-Collision with constant acceleration (TTCA) proposed by van der Horst (1990) and Relative Time to Collision (RTTC) (Chen et al., 2017). Post Encroachment Time (PET) has also been extensively used for assessing safety and severity levels of vehicle and pedestrian interactions (Fu et al., 2016, Chen et al., 2017, Marisamynathan and Vedagiri, 2020, Kathuria and Vedagiri, 2020, Chaudhari et al., 2021, Paul and Ghosh ,2020, Golakiya et al., 2022). In multi lane road networks, a full lane is considered as the conflict zone and the Lane based PET (LPET) is applies for classifying conflicts (Almodfer et al., 2016, Zhang et al., 2019). At zebra crossing specifically, the Time-To-Zebra$_{arrive}$ (TTZ$_{arr}$) indicator is applied for describing the behaviour of the driver of the approaching vehicle (Varhelyi, 1998, Bella and Silvestri, 2015) but also for assessing the impact of ADAS systems (Bella and Silvestri, 2017). Time to accident (TA) proposed by Hyden (1987) and Time Advantage (Tadv) suggested by Laureshyn et al. (2010) have also been used for distinguishing conflicts and dangerous events (Yang et al., 2022).

Apart from time- based indicators, speed based indicators are also commonly used in this field of road safety such as Deceleration to Safety Time (DST) introduced by Hupfer et al. (1998) and used by Ismail et al. (2009), Kumar et al. (2019), Olszewski et al. (2020) solely or in combination with other indicators such as speed, acceleration or time based parameters for evaluating conflicts between the two agents. In the same category, conflict speed in combination with TA has been proposed for defining severity levels and establishing thresholds for critical situations (Hyden, 1987, Svensson, 1998, Shbeed, 2000, Laureshyn and Varhelyi, 2020). Few distance based indicators have been used for describing vehicle-pedestrian encounters such as passing distance (Olszewski et al., 2020), relative distance (Pascucci et al., 2020) and minimum distance (Amini et al., 2022).

Finally, risk indicators combining the above-mentioned parameters have been established and used in research studies such as the Pedestrian Risk Index (PRI), a conflict indicator for describing the probability of conflict occurrence as well as the severity of this potential conflict at an unsignalized intersection (Cafiso et al., 2011). The index is defined based on the potential collision vehicle speed and the difference between the time to collision of the vehicle and the time the vehicle needs to come to a complete stop. Additionally, the Risk indicator (RI) was defined for estimating injury risk of a potential collision by dividing the vehicle's approaching speed and PET (Scholl et al., 2019, Govinda et al., 2022).

4.2 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 time 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). Pawar and Patil (2016) estimated temporal and critical gaps accepted by pedestrians when they tend to cross an uncontrolled mid – block crossing on high speed arterials. Their analysis showed that critical accepted values are between 3.6-4.3s and 60-73m and are mainly affected by approaching vehicle speed.

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 cross. Hyden (1987), Lord (1996), El- Basyouny and Sayed (2013) and Sacchi and Sayed (2016) used the value of 1.5s as the TTC threshold for defining traffic conflicts at signalized intersections and omitted any value higher than this threshold from their analysis. Golakiya et al. (2022) used Time To Collision (TTC) as a SSM for analysing safety aspects of pedestrians at four different uncontrolled and undesignated mid-block sections at four different cities in India. Their analysis indicated that for TTC values below 3.6 sec the pedestrian is at high risk of collision. Kathuria and Vedagiri (2020) evaluate pedestrian vehicle interactions at unsignalized intersections and classified the severity levels in case either pedestrian or vehicle or both take an evasive action based on their speed profiles. It was found that the interaction is "safe" when TTC>2.5s, "mild" if 1.2s<TTC<2.5s and "critical" for very low TTC values (<1.5s). Under the scenario that neither pedestrian nor vehicle takes an evasive action, it was found that a safe passage is considered when TTC >2.3s and PET >2.6s, the interaction is critical when TTC<1.3s and PET<1s while in any other case we have a mild interaction between the vehicle and the pedestrian. The same approach of different interaction patterns based on the speed changes and profiles of the vehicle and the pedestrian was adopted in Ni et al. (2016) for defining thresholds of conflict indicators for assessing pedestrian safety at unsignalized intersections including interactions between pedestrians and right-turning vehicles. In this study three patterns were defined: (1) either or both users take significant evasive actions, (2) neither user takes evasive reactions and (3) slight speed changes of the two interacting road users. In the first case, TTC is used as conflict indicator with thresholds are 1.5s and 3s while in the second case PET was adopted with thresholds equal to 1s and 3s for characterizing the interaction as "critical", "conflict" and "safe". PET and TTC were considered simultaneously in the third case. If TTC >3.0s and PET > 3.0s, the interaction is safe, if TTC ≤ 1.5s and PET ≤ 1.0s is a conflict, otherwise, it is a critical event.

4.3 State Definition

As described in chapter 2, the implementation of maximum entropy inverse reinforcement learning requires the definition of the agent state space and action space. Based on the results of the literature review, the state of the automated vehicle are described based on the TTC values estimated every time step which includes both the distance between the vehicle and the pedestrian as well their speeds. 7 discrete states are defined as presented in Table 2. The interaction between the two agents is considered to be completed when either the TTC has extremely high values or the vehicle has a very low speed and is considered to be stopped. The threshold for the latter case is 4.32km/h based on previous works. In both cases the agents are not in collision course and no TTC values can be estimated.

Table 2. *Vehicle's state space*

State	TTC (s)	State	TTC (s)
1	<0.5	5	2-3
2	0.5-1	6	3-4
3	1.0-1.5	7	>4
4	1.5-2		

4.4 Action 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 ($g=9.81m/s^2$) as it is described in Vlahogianni and Barbounakis (2017). For safety purposes, the value 0.16g (≈ $1.57m/s^2$) 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. *Vehicle's action space*

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)

5. Results

5.1 State features and action attributes statistical analysis

TTC is the feature chosen to define the different states of the vehicle during its interaction with the pedestrian while the action space was defined by its acceleration/deceleration. Table 4 shows the descriptive statistics of these three parameters and while Fig. 3 presents the distribution of their values. It should be mentioned that acceleration values over 4.5m/s2 and deceleration values higher than 9m/s2 (emergency braking), if any, were excluded from the analysis.

Table 4. *Descriptive Statistics*

	Acceleration (m/s^2)	Deceleration (m/s^2)	Vehicle Speed (km/h)	TTC (s)
Mean	0.174	-1.174	8.260	3.446
Standard Deviation	0.143	1.790	3.292	1.622
Median	0.135	-0.674	7.188	3.316
Min Value	0.000	-10.721	2.936	0.412
25%	0.064	-1.238	6.206	2.219
50%	0.135	-0.674	7.188	3.316
75%	0.262	-0.183	9.360	4.228
Max Value	0.780	-0.002	25.373	10.854

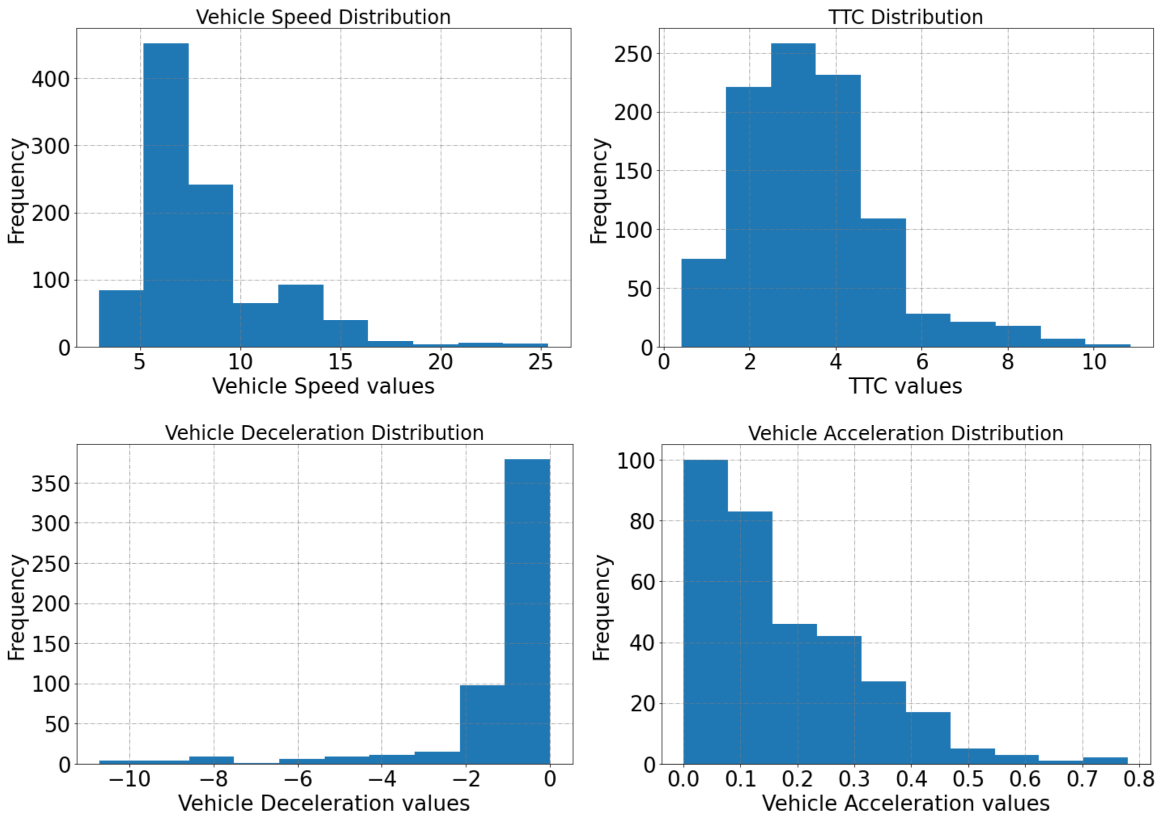

***Figure 3**: Distribution of the (i) vehicle speed, (ii) TTC, (ii) vehicle deceleration and (iv) vehicle acceleration*

5.2 Rewards

The estimated rewards for the different states of the automated vehicle based on the TTC values during its interaction with the pedestrian are presented in Figure 4. Is it noted that the reward is relative to the reward of state 1. The values depict the preferences of the automated vehicle concerning the chosen time based indicator, revealing that states with higher TTC values give higher rewards.

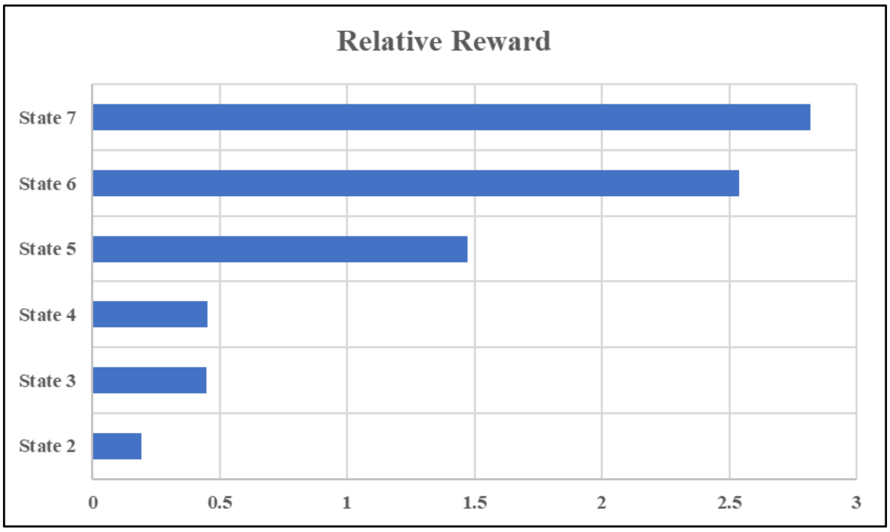

***Figure 4**: Rewards of the automated vehicle states*

6. Conclusions

The present work tries to model the behaviour of a Level 4 automated vehicle during its interaction with a pedestrian standing on the curb at a random location of the network and trying to cross the road. The principles of inverse reinforcement learning were applied for learning the reward function through the analysis of vehicle and pedestrian trajectories collected via a virtual reality experiment. The state space of the automated vehicle were defined based on the TTC values, a surrogate safety measure widely used foe describing and evaluating encounters and conflicts between these two agents. The results showed the preference of the driver/machine for higher TTC values as the reward increases at higher states. The next steps are to feed with these rewards a Dyna Q model for gaining the optimal policy and then produce trajectories that will be compared with the real ones collected from the experiment.

Acknowledgements

The SHOW project (www.show-project.eu) has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 875530. The document reflects only the author's view, the EU is not responsible for any use that may be made of the information it contains.

References

Ackermann, C., Beggiato, M., Bluhm, L. F., L¨ow, A., & Krems, J. F. (2019a). Deceleration parameters and their applicability as informal communication signal between pedestrians and automated vehicles. TransportationResearch Part F: Traffic Psychology and Behaviour, 62, 757–768.

Ackermann, C., Beggiato, M., Schubert, S., & Krems, J. F. (2019b). An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles? Applied Ergonomics, 75, 272–282. https://doi.org/10.1016/j.apergo.2018.11.002

Alghodhaifi, H., & Lakshmanan, S. (2020). Simulation-based model for surrogate safety measures analysis in automated vehicle-pedestrian conflict on an urban environment. In , Vol. 11415. Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure 2020 (p. 1141504). International Society for Optics and Photonics.

Anderson, C., Vasudevan, R., & Johnson-Roberson, M. (2020). Off the beaten sidewalk: Pedestrian prediction in shared spaces for autonomous vehicles. *IEEE Robotics and Automation Letters*, *5*(4), 6892–6899.

Blue, V. and J. Adler, "Cellular automata microsimulation of bidirectional pedestrian flows," Transp. Res. Rec. J. Transp. Res. Board, pp. 135-141, No.1678, Transportation Research Board of the National Academies Washington, DC, 2002

Calvert, S. C., Schakel, W. J., & Van Lint, J. W. C. (2017). Will automated vehicles negatively impact traffic flow?. Journal of Advanced Transportation, 2017. Camara, F., Dickinson, P., & Fox, C. (2021). Evaluating pedestrian interaction preferences with a game theoretic autonomous vehicle in virtual reality. *Transportation Research Part F: Traffic Psychology and Behaviour, 78*, 410–423.

Camara, F., Romano, R., Markkula, G., Madigan, R., Merat, N., & Fox, C. (2018). In *June). Empirical game theory of pedestrian interaction for autonomous vehicles* (pp. 238–244). Manchester Metropolitan University.

Camara, F., Dickinson, P., & Fox, C. (2021). Evaluating pedestrian interaction preferences with a game theoretic autonomous vehicle in virtual reality. Transportation Research Part F: Traffic Psychology and Behaviour, 78, 410–423.

Chae, H.; Kang, C.M.; Kim, B.; Kim, J.; Chung, C.C.; Choi, J.W. Autonomous braking system via deep reinforcement learning. In Proceedings of the IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017.

Charlton, J., Gonzalez, L. R. M., Maddock, S., & Richmond, P. (2020). Simulating crowds and autonomous vehicles. In Transactions on Computational Science XXXVII: Special Issue on Computer Graphics (pp. 129-143). Berlin, Heidelberg: Springer Berlin Heidelberg.

De Clercq, K., Dietrich, A., Nú˜nez Velasco, J. P., De Winter, J., & Happee, R. (2019). External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. Human factors, 61(8), 1353–1370.

Deb, S., Strawderman, L., & Carruth, D. (2018). Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment. Transportation Research Part F Traffic Psychology and Behaviour., 59, 135–149.

Deshpande, N.; Vaufreydaz, D.; Spalanzani, A. Behavioral decision-making for urban autonomous driving in the presence of pedestrians using deep recurrent Q-network. In Proceedings of the 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), Shenzhen, China, 13–15 December 2020; pp. 428–433.

Deshpande, N.; Vaufreydaz, D.; Spalanzani, A. Navigation in urban environments amongst pedestrians using multi-objective deep reinforcement learning. In Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–22 September 2021; pp. 923–928

Elallid, B. B., Benamar, N., Mrani, N., & Rachidi, T. (2022, November). DQN-based Reinforcement Learning for Vehicle Control of Autonomous Vehicles Interacting With Pedestrians. In 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT) (pp. 489-493). IEEE.

El-Basyouny, K., & Sayed, T. (2013). Safety performance functions using traffic conflicts. Safety science, 51(1), 160-164.

Feng, C., Cunbao, Z., & Bin, Z. (2019). In *July). Method of pedestrian-vehicle conflict eliminating at unsignalized mid-block crosswalks for autonomous vehicles* (pp. 511–519). IEEE.

Fox, C., Camara, F., Markkula, G., et al. (2018). When should the chicken cross the road?: Game theory for autonomous vehicle-human interactions. *In: Proc. 4th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS)*.

Fridman, L., Mehler, B., Xia, L., Yang, Y., Facusse, L. Y., & Reimer, B. (2017). To walk or not to walk: Crowdsourced assessment of external vehicle-to-pedestrian displays. arXiv preprint arXiv:1707.02698

Golakiya, H., Chauhan, R., Bari, C. S., & Dhamaniya, A. (2022). Pedestrian safety analysis at urban midblock section under mixed traffic conditions using time to collision as surrogate safety measure. Current Science (00113891), 123(9).

Hsu, Y. C., Gopalswamy, S., Saripalli, S., & Shell, D. A. (2018, August). An MDP model of vehicle-pedestrian interaction at an unsignalized intersection. In 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall) (pp. 1-6). IEEE.

Hu W, Mu H, Chen Y, Liu Y, Li X. Modeling Interactions of Autonomous/Manual Vehicles and Pedestrians with a Multi-Agent Deep Deterministic Policy Gradient. Sustainability. 2023; 15(7):6156. https://doi.org/10.3390/su15076156

Hydén, C. (1987). The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique. Bulletin Lund Institute of Technology, Department, (70).

Jayaraman, K., Tilbury, D. M., Yang, X. J., Pradhan, A. K., & Robert, L. P. (2020). In *May). Analysis and prediction of pedestrian crosswalk behavior during automated vehicle interactions* (pp. 6426–6432). IEEE.

Kalantari, Amir Hossein & Markkula, Gustav & Uzondu, Chinebuli & Lv, Wei & Pedro, Jorge & Madigan, Ruth & Lee, Yee Mun & Holmes, Christopher & Merat, Natasha. (2021). Vehicle-Pedestrian Interactions at Uncontrolled Locations: Leveraging Distributed Simulation to Support Game-Theoretic Modeling. 10.13140/ RG.2.2.15469.72165.

Kalatian, A., & Farooq, B. (2021). Decoding pedestrian and automated vehicle interactions using immersive virtual reality and interpretable deep learning. *Transportation research part C: emerging technologies, 124*, Article 102962.

Kalatian, A., & Farooq, B. (2022). A context-aware pedestrian trajectory prediction framework for automated vehicles. *Transportation Research Part C: Emerging Technologies, 134*, Article 103453.

Kathuria, A., & Vedagiri, P. (2020). Evaluating pedestrian vehicle interaction dynamics at un-signalized intersections: A proactive approach for safety analysis. Accident Analysis & Prevention, 134, 105316.

Kuderer, M., Kretzschmar, H., Sprunk, C., & Burgard, W. (2012). *Feature-Based Prediction of Trajectories for Socially Compliant Navigation*. In Robotics: science and systems.

Luo, Y., Cai, P., Bera, A., Hsu, D., Lee, W. S., & Manocha, D. (2018). Porca: Modeling and planning for autonomous driving among many pedestrians. *IEEE Robotics and Automation Letters*, *3*(4), 3418–3425.

Martinez-Gil, F., Lozano, M., & Fernández, F. (2014). MARL-Ped: A multi-agent reinforcement learning based framework to simulate pedestrian groups. Simulation Modelling Practice and Theory, 47, 259-275.

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.

Merat, N., Lee, Y. M., Markkula, G., Uttley, J., Camara, F., Fox, C., et al. (2018). In How do we study pedestrian interaction with automated vehicles? Preliminary findings from the European interACT project (pp. 21–33). Cham: Springer.

Michieli, U., & Badia, L. (2018, September). Game theoretic analysis of road user safety scenarios involving autonomous vehicles. In 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) (pp. 1377-1381). IEEE.

Nagel, K. and M. Schreckenberg, "A cellular automaton model for freeway traffic," J. Physique I, vol. 2, pp. 2221-2229, 1992.

Nasernejad, P., Sayed, T., & Alsaleh, R. (2021). Modeling pedestrian behavior in pedestrian-vehicle near misses: A continuous Gaussian Process Inverse Reinforcement Learning (GP-IRL) approach. *Accident Analysis & Prevention*, *161*, 106355.

Nasernejad, P., Sayed, T., & Alsaleh, R. (2022). Multiagent modeling of pedestrian-vehicle conflicts using Adversarial Inverse Reinforcement Learning. Transportmetrica A: Transport Science, 1-35.

Ni, Y., Wang, M., Sun, J., & Li, K. (2016). Evaluation of pedestrian safety at intersections: A theoretical framework based on pedestrian-vehicle interaction patterns. Accident Analysis & Prevention, 96, 118-129.

Orfanou F., Vlahogianni E., Yannis G. and Mitsakis E. (2022). Humanizing Autonomous Vehicle Driving: Understanding, Modeling and Impact Assessment, Transportation Research Part F: Traffic Psychology and Behaviour.

Papini, G.P.R.; Plebe, A.; Da Lio, M.; Dona, R. A Reinforcement Learning Approach for Enacting Cautious Behaviours in Autonomous Driving System: Safe Speed Choice in the Interaction with Distracted Pedestrians. IEEE Trans. Intell. Transp. Syst. 2021, 23, 8805–8822. [Google Scholar] [CrossRef]

Pawar, D. S., & Patil, G. R. (2016). Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials. Safety science, 86, 295-303.

Predhumeau, M., Mancheva, L., Dugdale, J., & Spalanzani, A. (2021, May). An Agent-Based Model to Predict Pedestrians Trajectories with an Autonomous Vehicle in Shared Spaces. In *20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021)*.

Rad, S. R., de Almeida Correia, G. H., & Hagenzieker, M. (2020). Pedestrians' road crossing behaviour in front of automated vehicles: Results from a pedestrian simulation experiment using agent-based modelling. *TransportationResearch Part F: Traffic Psychology and Behaviour*, *69*, 101–119.

Rahmati, Y., Talebpour, A., Mittal, A., & Fishelson, J. (2020). Game theory-based framework for modeling human–vehicle interactions on the road. Transportation research record, 2674(9), 701-713.

Rothenbücher, D., J. Li, D. Sirkin, B. Mok and W. Ju, "Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles," 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), New York, NY, 2016, pp. 795-802, doi: 10.1109/ROMAN.2016.7745210.

Sacchi, E., & Sayed, T. (2016). Conflict-based safety performance functions for predicting traffic collisions by type. Transportation Research Record, 2583(1), 50-55.

Trumpp, R.; Harald, B.; David, S. Modeling interactions of autonomous vehicles and pedestrians with deep multi-agent reinforcement learning for collision avoidance. In Proceedings of the 2022 IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany, 4–9 June 2022

Van Den Berg, J., Guy, S. J., Lin, M., & Manocha, D. (2011). Reciprocal n-body collision avoidance. In Robotics Research: The 14th International Symposium ISRR (pp. 3-19). Springer Berlin Heidelberg.

Vasquez, R., & Farooq, B. (2019). In *Multi-objective autonomous braking system using naturalistic dataset* (pp. 4348–4353). IEEE.

Velasco, J. P. N., Farah, H., van Arem, B., & Hagenzieker, M. P. (2019). Studying pedestrians' crossing behavior when interacting with automated vehicles using virtual reality. Transportation Research Part F: Traffic Psychology and Behaviour, 66, 1–14.

Wang, P., Motamedi, S., Qi, S., Zhou, X., Zhang, T., & Chan, C. Y. (2021). Pedestrian interaction with automated vehicles at uncontrolled intersections. Transportation Research Part F: Traffic Psychology and Behaviour, 77, 10–25.

Wu, W., Chen, R., Jia, H., Li, Y., & Liang, Z. (2019). Game theory modeling for vehicle–pedestrian interactions and simulation based on cellular automata. International Journal of Modern Physics C, 30, 1950025.

Wu, T., Jiang, M., & Zhang, L. (2020). Cooperative multiagent deep deterministic policy gradient (CoMADDPG) for intelligent connected transportation with unsignalized intersection. Mathematical Problems in Engineering, 2020.

Yang, D., ¨Ozgüner, Ü., & Redmill, K. (2020). A social force based pedestrian motion model considering multi-pedestrian interaction with a vehicle. *ACM Transactions on Spatial Algorithms and Systems (TSAS)*, *6*(2), 1–27.

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

Zhang, E., Zhang, R., & Masoud, N. (2023). Predictive trajectory planning for autonomous vehicles at intersections using reinforcement learning. Transportation Research Part C: Emerging Technologies, 149, 104063.

Zhu, H., Alhajyaseen, W., Iryo-Asano, M., Nakamura, H., & Dias, C. (2022). Defensive or competitive Autonomous Vehicles: Which one interacts safely and efficiently with pedestrians?. Physica A: Statistical Mechanics and its Applications, 606, 128083.

Zhu, Z., Hu, Z., Dai, W., Chen, H., & Lv, Z. (2022). Deep learning for autonomous vehicle and pedestrian interaction safety. Safety science, 145, 105479.