

# Social Interaction-Aware Dynamical Models and Decision-Making for Autonomous Vehicles

Luca Crosato, Kai Tian, Hubert P. H. Shum, Edmond S. L. Ho, Yafei Wang, and Chongfeng Wei\*

Interaction-aware autonomous driving (IAAD) is a rapidly growing field of research that focuses on the development of autonomous vehicles (AVs) that are capable of interacting safely and efficiently with human road users. This is a challenging task, as it requires the AV to be able to understand and predict the behaviour of human road users. In this literature review, the current state of IAAD research is surveyed. Commencing with an examination of terminology, attention is drawn to challenges and existing models employed for modeling the behaviour of drivers and pedestrians. Next, a comprehensive review is conducted on various techniques proposed for interaction modeling, encompassing cognitive methods, machine-learning approaches, and game-theoretic methods. The conclusion is reached through a discussion of potential advantages and risks associated with IAAD, along with the illumination of pivotal research inquiries necessitating future exploration.

other human road users, potentially leading to traffic dilemmas and safety issues.<sup>[2]</sup> In order to operate in an efficient and safe manner, AVs need to behave in a human-like fashion and generate optimal behaviours that take the interactions with other human road users into account.<sup>[3]</sup> This is critical for the reduction of potential traffic conflicts. For example, cautious but unnecessary stops at intersections might cause rear-end accidents. In order to develop fully automated vehicles, advances in many aspects of AV technology are required, ranging from perception, decision-making, planning, and control.<sup>[4,5]</sup> When it comes to predicting the behaviour of surrounding human road users and taking decisions accordingly for AVs, the interactions with surrounding human road users become increasingly important, as the AV

actions affect their behaviour and vice versa.<sup>[6]</sup>


The purpose of this article is to provide an exhaustive survey of state-of-the-art techniques in interaction-aware motion planning and decision in the context of autonomous driving. In particular, the text first covers human road user behavioural models to highlight what factors influence the decisions that human road users make on the road. Driver and pedestrian behavioural models are relevant to AVs for multiple reasons. Firstly, they can be used to assess and predict what road users that surround the AV will do. Secondly, they can aid in the development of human-like AV behaviour. Therefore, they hold both a predictive value as well as add relevant insights for model/system design.

## 1. Introduction

In the past few years, there has been an increasing interest in the development of technology for autonomous vehicles (AVs), as the recent advancements in robotics and machine learning have enabled autonomous driving (AD) engineers to develop algorithms that could tackle the complexity of the AD task. AVs have the potential to improve traffic quality, reduce traffic accidents, and improve the quality of time spent while travelling.<sup>[1]</sup> Nowadays, more and more AVs are being deployed into the real world, sharing the environment with other human road users. This has raised concerns that AVs could not be able to understand and interact smoothly with

L. Crosato  
Northumbria University  
Newcastle upon Tyne NE1 8ST, UK

L. Crosato, C. Wei  
School of Mechanical and Aerospace  
Queen's University  
Belfast BT7 1NN, UK  
E-mail: c.wei@qub.ac.uk

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.202300575>.

© 2023 The Authors. Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202300575

K. Tian  
Institute for Transport Studies  
University of Leeds  
Leeds LS1 9JT, UK

H. P. H. Shum  
Department of Computer Science  
Durham University  
Durham DH1 3LE, UK

E. S. L. Ho  
School of Computing Science  
University of Glasgow  
Glasgow G12 8QQ, UK

Y. Wang  
School of Mechanical Engineering  
Shanghai Jiao Tong University  
Shanghai 200240, China

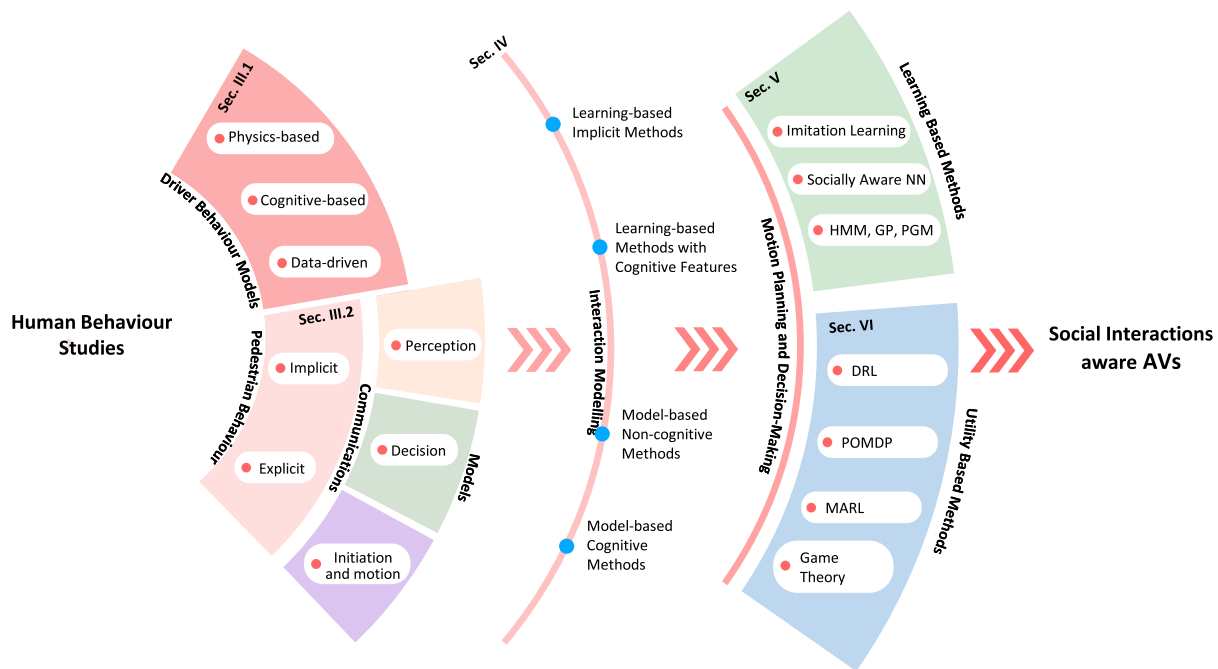


Figure 1. Flow chart: from human behaviour to social interaction aware AVs.

This review consists of five main sections which cover different areas in IAAD. The terminology used in IAAD is introduced in Section 2. Please refer to **Figure 1** for an overview of the article structure.

Section 3 will cover human factors studies on what affects human decision-making while driving, as well as pedestrian behavioural studies. Section 4 gives a broad overview and classification of existing techniques that are used in interaction modeling. Finally, Section 5 and 6 cover state-of-the-art techniques used for motion-planning and decision-making in interactive scenarios.

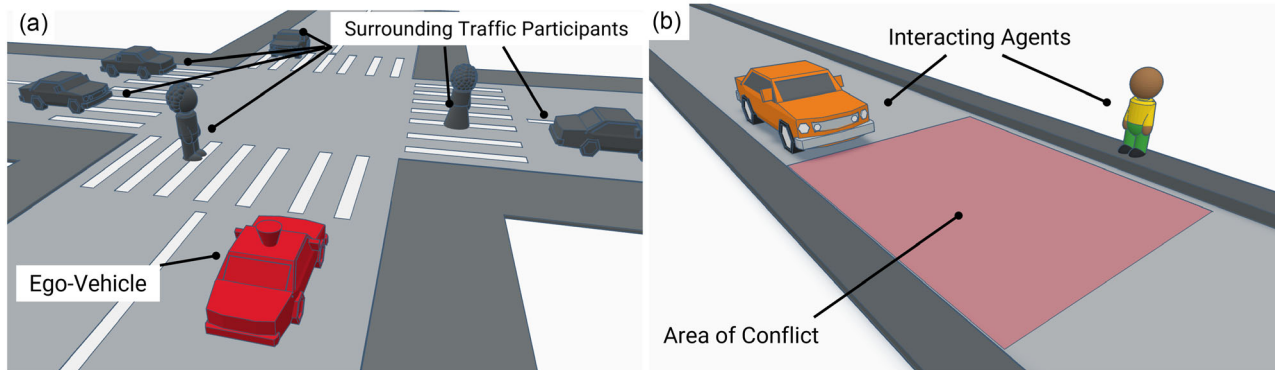
While autonomous driving has been an active research area in recent years, most of it focuses on scenarios involving only vehicles. There is more limited work that addresses heterogeneous scenarios, which include both vehicles and pedestrians. In this article, the focus is on heterogeneous scenarios, but Section 5 and 6 will also cover related work that deals with scenarios without pedestrians. This is because the techniques used in these papers can be easily adapted to mixed traffic scenarios, or they can offer important insights into how to deal with the general problem of mixed traffic scenarios.

## 2. Terminology in Interaction-Aware Autonomous Driving

Before discussing the recent advances in interaction-aware motion-planning and decision-making, the paper first defines some of the terminology used in this field. In the field of autonomous driving, the term ego-vehicle refers to the specific vehicle whose behaviour is to be controlled and studied. All other vehicles, cyclists, pedestrians, etc., that occupy a region of space around the ego-vehicle are treated as interactive obstacles and are

referred to as surrounding traffic participants, see **Figure 2a**. Since road traffic is unlikely to become fully automated in the near future, AVs will inevitably operate in environments mixed with human road users (HRUs), such as human drivers and pedestrians. Therefore, IAAD is a field of research that focuses on developing AVs that can safely and efficiently interact with surrounding HRUs. Traditional autonomous driving approaches often treat surrounding HRUs as dynamic obstacles. However, this is not a realistic approach, as they are constantly changing their behaviour to adapt to the current situation.

Generally, multiple surrounding HRUs can give rise to space-sharing conflicts amongst themselves or with the ego-vehicle: a situation from which it can be reasonably inferred that two or more road users intend to occupy the same region of space at the same time in the near future, see **Figure 2b**. The agents involved in the conflict are said to display an interactive behaviour, which implies that their behaviour would have been different if the space-sharing conflict had not occurred.<sup>[7]</sup> Moreover, interaction does not necessarily involve conflict. It can be explicit or implicit communications that indicate a road user's intentions and affect HRUs. For instance, drivers could plan their driving strategy based on the turning light signals of vehicles in front, even though the ego-vehicle and vehicles in front are not in the same lane, and there will be no conflict in the near future. Hence, interactive behaviour refers to the different courses of action of road users, adapting to the behaviour of others or making requests for reactions and taking actions to achieve their desired goals.<sup>[8]</sup> Since interactions happen all the time when driving, it is crucial that the algorithms developed for AVs be aware of the dynamics of the interactions between agents. Such algorithms are said to be interaction-aware and are often the focus of recent autonomous driving research.<sup>[9]</sup>



**Figure 2.** a) The ego-vehicle is controlled by the autonomous system, whereas surrounding traffic participants act on their own will, b) two agents interacting with each other determine an area of conflict.

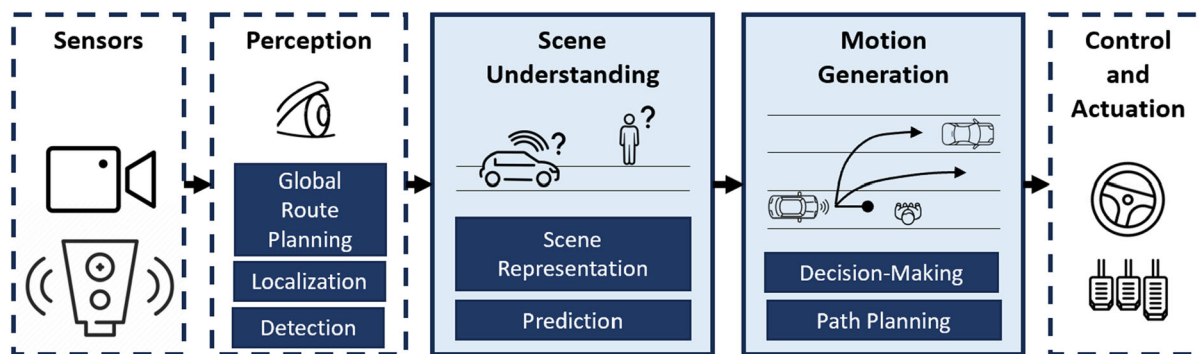
The safe and socially acceptable IAAD systems are currently hampered by a number of challenges.<sup>[10]</sup> One challenge is a lack of innovative theories about how HRUs interact.<sup>[11]</sup> This is a difficult task, as the theories to be developed are not limited to predicting and modeling HRUs' behaviour but also exploring behaviour patterns and their underlying mechanisms. Integrating AVs into road traffic as seamlessly as humans would require more advanced behaviour theories and models. Another challenge is the need to develop algorithms that can safely and efficiently interact with other HRUs and produce an AV behaviour that compels human-like standards. **Figure 3** shows the main parts that make up an AV system. Raw data from sensors is processed by a perception module, which detects the surrounding environment and performs localisation, which allows the generating of a global route plan for the ego-vehicle to reach its target destination. The scene can be further interpreted, and predictions regarding surrounding traffic participants can be made. Interaction-aware models play a major role in prediction tasks, as agents affect each other's trajectories and decisions.

Decision-making and path-planning are two of the most important tasks in autonomous driving. They are responsible for determining how the vehicle will move through its environment. Decision-making is the process of choosing an action from a set of possible options. For example, the vehicle may need to decide whether to change lanes, slow down, or stop. Path-planning is the process of generating a safe and feasible

trajectory for the vehicle to follow. Decision-making and path-planning are closely related. The decision-making process typically outputs a high-level plan, such as "change lanes to the left." The path-planning process then takes this plan and generates a detailed trajectory that the vehicle can follow. Both tasks must take into account the vehicle's current position, the vehicle's capabilities and the surrounding traffic, which is why interaction-aware models are highly relevant to these two tasks. From a control system perspective, the dynamics of the vehicle are represented by its states, i.e., position and orientation, and their time derivatives. The state of the environment is determined by the states of all dynamics and static entities. The strictly physical state-space can also be augmented with additional latent-space variables that capture, for example, the intentions<sup>[12]</sup> or the behavioural preferences of surrounding users,<sup>[13]</sup> which are part of the Scene Understanding system.

### 3. Human Behaviour Studies on Interactions

This section synthesizes empirical and modeling research findings on HRU behaviour, including that of human drivers and pedestrians interacting with AVs or conventional vehicles, especially from a communication perspective. The focus is on research involving road interactions, with the aim of discovering insights that may facilitate the development of interaction-aware



**Figure 3.** Architecture of AV systems. Solid line boxes identify modules that are closely related to interaction-aware models.

AVs. Studies that look at macro-traffic conditions, such as the influences of route choice, weather, or regulation, are beyond this article's scope.

### 3.1. Driver Behaviour Studies

Driver behaviour models are used to predict and understand how drivers will behave in different driving scenarios. These models can be used to improve the safety and efficiency of transportation systems and aid in the process of designing AVs. Many different factors can affect driving behaviour, including individual characteristics (age, gender, personality, and experience), environmental factors, i.e., road and weather conditions, and social factors, which include the driver's interactions with HRUs.<sup>[14]</sup> A comprehensive overview of driver behavioural models (DBM) in vehicle-vehicle interactions can be found in Ref. [15]. Here, the focus will be on DBMs that are relevant to vehicle-pedestrian interactions.

The most common driver behavioural models include:

#### 3.1.1. Driver's Risk Field Models

This model predicts how drivers will perceive risk in different driving situations (Figure 4a). The DRF model is based on the idea that drivers make decisions based on their perception of risk. The results of Ref. [16] suggest that driving behaviour is governed by a cost function that takes into account the effects of noise on human perception and actions. This is similar to how motor-control tasks are governed by cost functions. Risk perception onboard of AVs has also been analysed in Ref. [17] in driving simulator scenarios.

#### 3.1.2. Theory Based Perceptual and Cognitive Models

Models based on perceptual information describe driver's behaviour based on perceptual cues, e.g., distance, vehicle speed, acceleration, expansion angle, reaction times, etc.<sup>[18,19]</sup> (Figure 4b). Cognitive models outline the internal state flow and motive that regulates the driver's behaviour as a psychological human being.<sup>[11,20]</sup>

#### 3.1.3. Data-Driven Models

This set of methods relies on analysing naturalistic driving data with machine learning to analyse driver behaviour (Figure 4c). Data-driven models can learn generative or discriminative<sup>[21–23]</sup>

models of human behaviour to make predictions about the driver's future decisions or preferred driving style. Model validation can be done by comparing predictions with real data and by human-in-the-loop simulations.

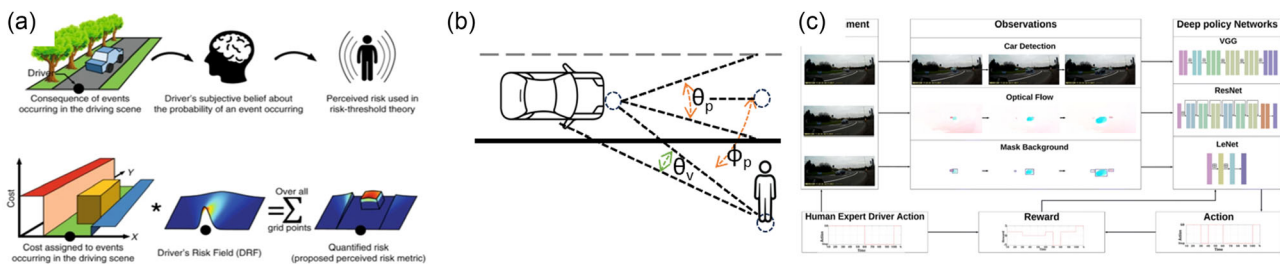
Existing research highlights based on naturalistic driving data analyses how drivers behave in the presence of pedestrians. In Ref. [24], the authors found that drivers tend to maintain smaller minimum lateral clearance and lower overtaking speed when overtaking pedestrians who are walking in the opposite direction, on the lane edge, or when oncoming traffic is present. Minimum lateral clearance and time-to-collision were only weakly correlated with overtaking speed. The results in Ref. [25] show that the vehicle deceleration behaviour is relative to the initial Time To Collision (TTC), subjective judgment of pedestrian crossing intention, vehicle speed, pedestrian position and crossing direction.

There is less attention paid to multi-agent settings where multiple vehicles and pedestrians interact with each other. In Ref. [26], the authors develop a multi-agent adversarial Inverse reinforcement learning (IRL) framework based on data collected at a road intersection to simulate driver and pedestrian behaviour at intersections.

Overall, DBMs are a promising area of research with the potential to significantly improve the safety and efficiency of transportation systems. However, there is still much work to be done in developing and validating these models. Future research should focus on developing more comprehensive models that take into account a wider range of factors, such as the driver's internal state, the environment, and the interactions with other HRUs.

### 3.2. Pedestrians Behaviour Studies

Since pedestrians are considered the most vulnerable road users, lacking protective equipment and moving more slowly than other road users,<sup>[27]</sup> investigating pedestrian behaviour is clearly relevant to the safety and acceptance of AVs interacting with pedestrians. Pedestrian behaviour has been the subject of extensive research for decades.<sup>[28]</sup> The emergence of AVs has recently prompted many new research questions about pedestrian behaviour. Given the large body of work in this area and our aims, this Section examines major studies rather than providing an exhaustive survey. The review covers pedestrian behaviour studies regarding interactions with vehicles from three perspectives: communications, theories and models of crossing behaviour, and AV-involved applications. The aim is to identify and summarize their value for developing interaction-aware AVs.



**Figure 4.** Illustration of driver behaviour models. a) Driver risk field. Reproduced with permission.<sup>[16]</sup> Copyright 2020, Springer Nature. b) Joint theory-based model (as in ref. [18]) c) Data-driven model. Reproduced with permission.<sup>[21]</sup> Copyright 2022, Springer Nature.



### 3.2.1. Communications

In dynamic traffic environments, road users intentionally or unintentionally convey signalling information to one another through their movements and spatial cues, giving rise to both explicit and implicit communication. Research findings concur that AVs' kinematics and signalling information significantly impact pedestrian road behaviour due to the absence of a driver role.<sup>[7,29,30]</sup> Therefore, the identification of critical motion cues and signals affecting pedestrian road behaviour holds substantial research significance (see **Figure 5a**).

Implicit communication signals, such as vehicle kinematic cues, involve road user behaviour that affects its own movement but can be interpreted as cues of intention or movement by another road user.<sup>[7]</sup> Distance or TTC between approaching vehicles and pedestrians is the most critical Implicit information influencing pedestrian behaviour.<sup>[31,32]</sup> Evidence showed that pedestrians tend to rely more on distance than TTC.<sup>[33]</sup> That is, for the same TTC, more pedestrians crossed when vehicles approached at higher speeds. A recent study showed pedestrians used multiple sources of information from vehicle kinematics instead of relying on one. The impacts of speed, distance, and TTC on pedestrian behaviour were mutual coupling.<sup>[34]</sup>

Braking manoeuvre is another critical implicit information influencing pedestrian behaviour in interactions. Vehicle movements correlated with pedestrian trust toward vehicles, emotion, and impact on pedestrian decisions.<sup>[35–37]</sup> Pedestrians felt comfortable and started crossing quickly when approaching vehicles slowed down early and braked lightly. Harsh braking led to pedestrian avoidance behaviour.<sup>[35,38,39]</sup> In contrast, early braking manoeuvres

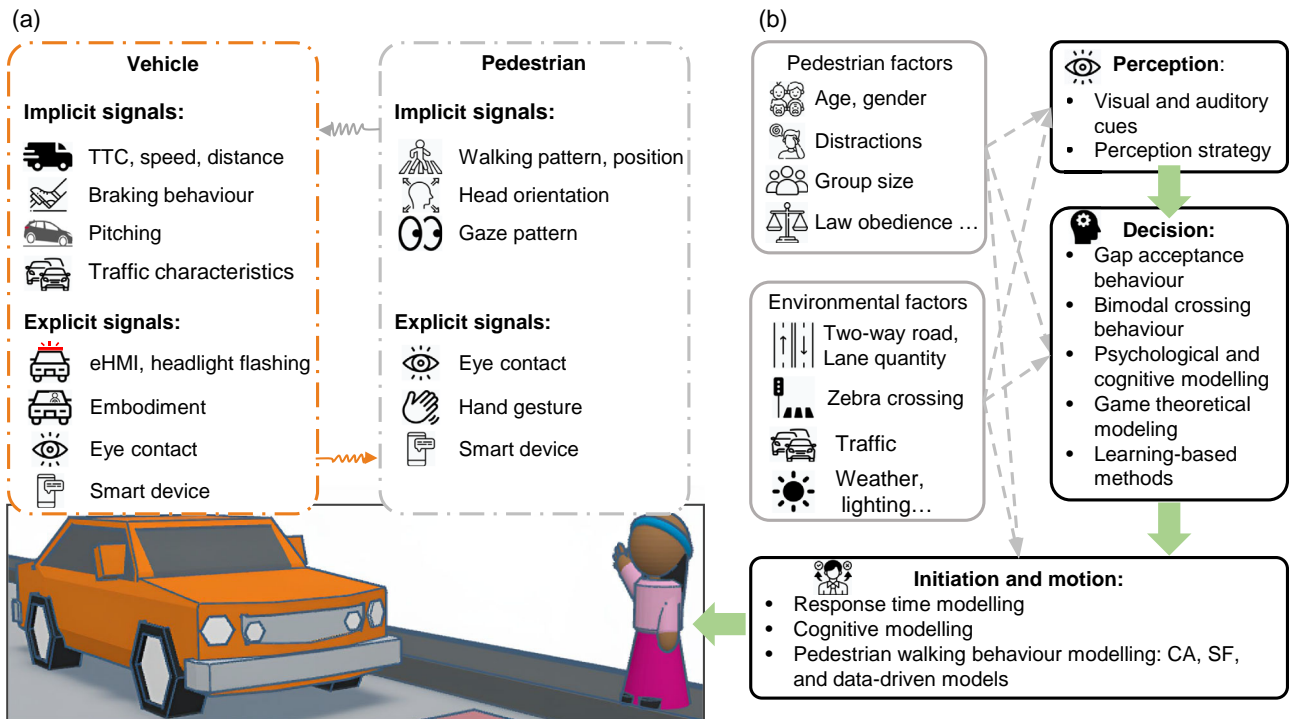
and strong pitching reduced the time required for pedestrians to understand the vehicle's intentions.<sup>[40,41]</sup> Vehicles that approached pedestrians slowly while yielding could hinder understanding.<sup>[35,37]</sup>

Traffic characteristics, such as traffic volume and gap sizes, provide implicit information to pedestrians. High traffic volume forced pedestrians to accept small traffic gaps,<sup>[42]</sup> as the increased time cost increased their propensity for risk-taking.<sup>[43]</sup> However, substantial evidence indicated pedestrians who tended to wait were more cautious and less likely to accept risky gaps.<sup>[44–46]</sup> The relationship between traffic volume and pedestrian crossing behaviour is context-dependent, potentially influenced by the size and order of gaps in traffic.<sup>[47]</sup>

Furthermore, the pedestrian movement toward the road, presence at the curb, and pedestrian head orientation could convey key implicit information to approaching vehicles.<sup>[48,49]</sup> Pedestrians often assert their right of way by stepping onto the road or looking at approaching vehicles.<sup>[30]</sup>

Explicit communication signals involve road user behaviour that conveys signalling information to other road users without affecting one's own movement or perception.<sup>[7]</sup> A common case is vehicles transmitting information to pedestrians through an external human-machine interface (eHMI). In the context of AVs, where there is no human driver, eHMIs have gained importance. Substantial evidence supported eHMI's benefits in pedestrian interactions with AVs.<sup>[35,50,51]</sup> Various types of eHMI prototypes have been proposed, such as headlight,<sup>[52]</sup> light band,<sup>[36]</sup> and anthropomorphic symbols,<sup>[53]</sup> but consensus on the best eHMI form and information to convey remains elusive.

Numerous studies demonstrated that eHMI performance depended on various factors. Pedestrians' familiarity, trust,



**Figure 5.** a) Communication between pedestrians and automated vehicles. b) Theories and models for pedestrian crossing perception, decision, initiation, and motion.

and interpretation of eHMIs could significantly impact eHMIs' effectiveness in communicating information to pedestrians. For instance, pedestrians better understood conventional eHMIs (flashing headlights) as signals for vehicles yielding than novel eHMIs (light bands).<sup>[36]</sup> Pedestrians overtrusting eHMIs might lead them to under-rely on vehicle motion cues, which is dangerous if eHMIs fail.<sup>[54]</sup> Egocentric information transmitted by the eHMI, like "OK TO CROSS", compelled pedestrians more than allocentric information like "STOPPING".<sup>[55]</sup> Moreover, the reliability of eHMIs was questioned as it might be affected by weather,<sup>[56]</sup> light condition,<sup>[29]</sup> and vehicle behaviour.<sup>[35]</sup> For example, in poor weather, pedestrians could not always read vehicle signs.<sup>[57]</sup> Pedestrian willingness to cross was unaffected by eHMIs when vehicles did not yield or decelerated aggressively.<sup>[35,50]</sup> Other concepts, e.g., installing eHMIs on road infrastructures instead of vehicles<sup>[58]</sup> and combining eHMIs with vehicle motion cues,<sup>[59,60]</sup> could outperform pure eHMIs.

Additionally, from the vehicles' perspective, although less frequent, pedestrians also use explicit signals to communicate with AVs. These signals include eye contact and hand gestures, which are used by pedestrians to ensure they are seen by AVs and to request the right of way.<sup>[7,61,62]</sup> To address the absence of a human driver, AVs could utilise a human-like visual embodiment in the driver's seat and wireless communication technology to enhance vehicle-pedestrian communication.<sup>[63–65]</sup>

### 3.2.2. Theories and Models of Crossing Behaviour

Pedestrian crossing behaviour involves various cognitive processes. Previous studies<sup>[47,57,66]</sup> suggested that there might be three levels of processes involved in constructing pedestrian crossing behaviour in interactions, i.e., perception, decision, initiation and motion. In light of this hypothesis, the following sections synthesise the theories and models of pedestrian crossing behaviour regarding the three cognitive processes (Figure 5b) (Table 1).

**Table 1.** Pedestrian models and theories.

Research	Scenario <sup>a)</sup>	Cognitive process	Models	Theories
[90]	1,3,4	Perception	$\tau, \dot{\tau}$	Visual perception
[88]	1,3,4	Perception	$\tau, \dot{\tau}$	Visual perception
[34,47]	1,3	Perception	$\dot{\theta}$	Visual perception
[18]	1,4	Perception	$\tau$ , bearing angle	Visual perception
[84]	1,3,4	Perception	Generalized TTC	Visual perception
[121]	1,4	Perception	TTC, bearing angle	Visual perception
[87]	1,3	Perception	Perceived distance	Visual perception
[78]	1,3	Decision	Critical gap	GA behaviour
[79,224]	1,3	Decision	Critical gap	GA behaviour
[43,47]	1,3,5	Decision	LR	GA behaviour
[82]	1,3	Decision	ANN	GA behaviour Machine learning
[80]	1,2,3, 4,5	Decision	Critical gap	GA behaviour
[87]	1,3	Decision	RL, Bayesian filter	GA behaviour, learning-based
[85]	2,3,4	Decision	Speed-distance	BC behaviour
[86]	1,3,4	Decision	Hybrid perception, LR	Visual perception, BC behaviour
[84,88,90]	1,3,4	Decision	EA	Drift diffusion
[11]	1,3,4	Decision	EA, Bayesian filter	Drift diffusion, Game theory, Theory of Mind, Noisy visual perception
[93]	1,3,4	Decision	DA game	Game theory
[92]	1,4	Decision	SC game	Game theory
[121]	1,4	Decision	Critical gap	GA behaviour, visual perception
[11,84,88,90]	1,3,4	Initiation	EA	Drift diffusion
[47,86]	1,3,4	Initiation	SW distribution	Response time
[87]	1,3	Initiation	RL	Learning-based
[108,114,121]	1,3,4	Motion	SF	Walking behaviour
[115]	2	Motion	ANN	Learning-based
[111,112]	1,2,3,4	Motion	CA	Walking behaviour
[26]	2,3,4	Motion	Adversarial IRL	Learning-based
[117]	2,4	Motion	LSTM	Learning-based

<sup>a)</sup>1. Uncontrolled crossings. 2. Controlled crossings. 3. With non-yielding vehicles. 4. With yielding vehicles. 5. With traffic flow.

**Perception:** The visual perception theory, as established by Gibson,<sup>[67]</sup> explains that as an object approaches an observer, its image on the retina expands, forming the basis for human collision perception. In crossing scenarios, when the rate of image expansion of a vehicle on the retina reaches a certain threshold, pedestrians perceive that the vehicle is approaching, known as the visual looming phenomenon.<sup>[68]</sup> A psychophysical model simplifies this expansion rate as the change in the visual angle subtended by the approaching vehicle at the pedestrian's pupil, denoted as  $\dot{\theta}$  (Figure 6a).<sup>[34,69]</sup> Recent research suggested that pedestrians use  $\dot{\theta}$  as a crucial visual cue to observe approaching vehicles.<sup>[11,34]</sup> However, while  $\dot{\theta}$  provides spatial information, it does not convey when the vehicle arrives at the pedestrian's position.<sup>[70]</sup> In crossing scenarios with yielding vehicles, pedestrians need temporal information to estimate whether the vehicle can stop in time. Lee's mathematical demonstration<sup>[69]</sup> showed that the visual cue  $\tau$ , representing the ratio of  $\theta$  to  $\dot{\theta}$ , may specify the TTC of the approaching vehicle. Moreover, the first temporal derivative of  $\tau$ , denoted as  $\dot{\tau}$ , was relevant for detecting if the current deceleration rate is sufficient to avoid a collision.<sup>[71]</sup> Furthermore, it was found that pedestrians might visually perceive oncoming collision events under a given angle, i.e., bearing angle, which is the angle between the vehicle and the pedestrian's line of regard<sup>[72]</sup> (Figure 6b).

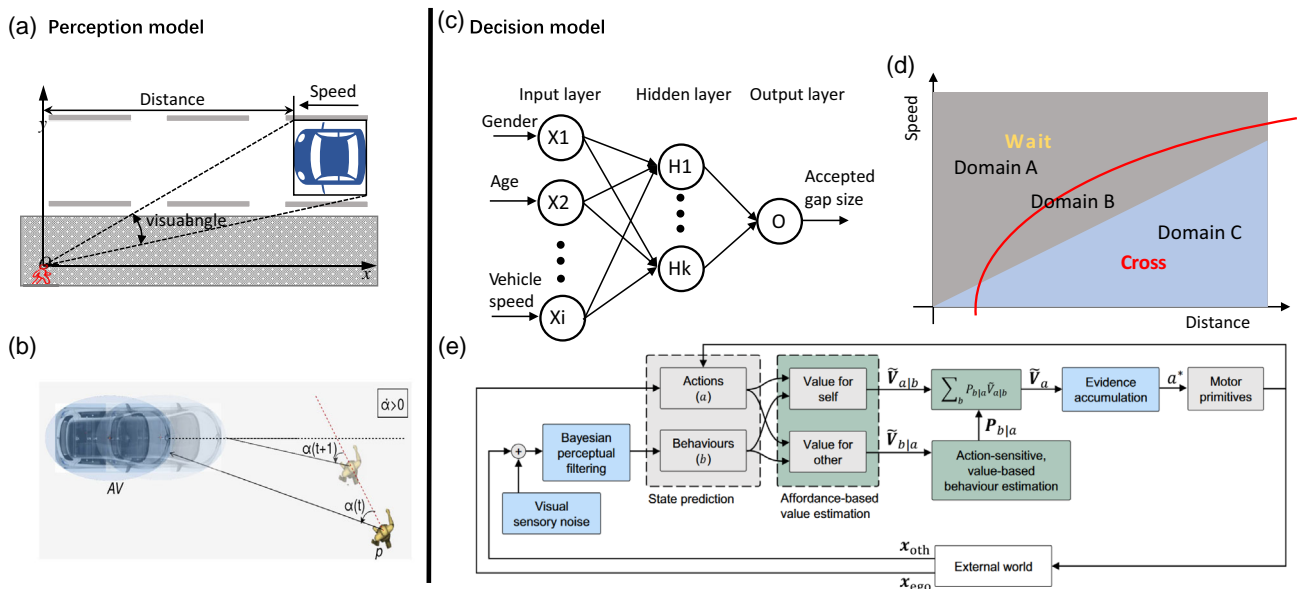
In addition to visual cues, pedestrian perception may depend on perceptual strategies. Research by Tian et al.<sup>[37]</sup> indicated that pedestrian estimation of vehicle behaviour might be a separate process or a sub-process of crossing decision-making. When there is a large traffic gap, pedestrians tend not to rely on vehicle driving behaviour but rather on gap size. Similarly, Delucia<sup>[70]</sup> indicated that when collision events are distant, humans tend to use 'heuristic' visual cues, such as  $\theta$  and  $\dot{\theta}$ . However, as

collisions become imminent, optical invariants like  $\tau$  dominate perception, providing richer spatiotemporal information.

Besides perception mechanisms, various factors could influence pedestrians' perceptions. Studies showed that elderly or child pedestrians faced a higher collision risk due to age-related perceptual limitations.<sup>[31,73]</sup> Elderly pedestrians tended to rely more on distance than TTC to judge approaching vehicles, while children struggled to detect vehicles approaching at high speeds.<sup>[33,74]</sup> Distractions, particularly those involving visual and manual components like smartphone usage, diverted significant attention resources and affected pedestrian observation of traffic conditions.<sup>[75]</sup> In comparison, cognitive distractions, such as listening to music, might not significantly impact pedestrian perception.<sup>[44]</sup>

**Decision:** At uncontrolled crossings without signal lights, pedestrians often interact with yielding or non-yielding vehicles.<sup>[31,36,76]</sup> In non-yielding scenarios, pedestrians usually make crossing decisions by evaluating gaps between approaching vehicles, known as gap acceptance behaviour (GA).<sup>[77]</sup> This concept led to the development of critical gap models, including the models by Raff,<sup>[78]</sup> HCM2010,<sup>[79]</sup> and Rasouli.<sup>[80]</sup> Alternatively, binary logit models treat crossing decisions as binary variables, utilising machine-learning algorithms like artificial neural networks (ANN), support vector machines (SVM), and logistic regression (LR).<sup>[81]</sup> For example, Kadali et al.<sup>[82]</sup> used ANN to predict crossing decisions based on various independent variables (Figure 6c), while Sun et al.<sup>[83]</sup> employed LR with variables such as pedestrian age, gender, group size, and vehicle type.

In scenarios involving yielding vehicles, crossing decisions tend to follow a bimodal pattern referred to as bimodal crossing behaviour (BC).<sup>[36,37,84]</sup> Pedestrians prefer to cross when traffic gaps are sufficiently large or when vehicles are about to stop.<sup>[84]</sup> However, making decisions in such scenarios can be challenging



**Figure 6.** Perception and decision models for pedestrians. a) Visual cues, including  $\theta$ ,  $\dot{\theta}$ ,  $\tau$ ,  $\dot{\tau}$ .<sup>[34,86]</sup> b) bearing angle. Reproduced with permission.<sup>[121]</sup> Copyright 2022, AAAI. c) Artificial neural networks.<sup>[82]</sup> d) speed-distance model.<sup>[83]</sup> e) large computational psychological model. Reproduced with permission.<sup>[11]</sup> Copyright 2023, Oxford University Press.

due to the contrasting relationships between decision cues and collision risk, with collision risk negatively correlated to traffic gap and positively correlated to vehicle speed.<sup>[37]</sup> Zhu et al.<sup>[85]</sup> clustered crossing decisions into three groups: crossing, dilemma condition, and waiting, based on vehicle speed and distance (Figure 6 d). Moreover, Tian et al.<sup>[86]</sup> assumed pedestrians applied different decision-making strategies according to BC behaviour and modelled crossing decisions as responses to different visual cues.

While the mentioned approaches model crossing decisions based on observed behaviour patterns, other models delve into the psychological mechanisms that underlie these decisions. Specifically, Tian et al. modelled pedestrian GA behaviour based on pedestrian visual cues<sup>[34]</sup> and extended it to yielding scenarios with a more complex visual perception mechanism.<sup>[86]</sup> Wang et al. utilised a reinforcement learning (RL) model to capture pedestrian crossing behaviour based on limited perception mechanisms.<sup>[87]</sup> Furthermore, a class of models, namely evidence accumulation (EA) models, such as the drift-diffusion model, proposed that crossing decisions result from the accumulation of visual evidence and noise, with the decision determined once a certain threshold was reached.<sup>[84,88–90]</sup> Integrated large-scale psychological theories to explain pedestrian crossing decisions in detail (Figure 6e).<sup>[11]</sup> Additionally, game theory has also been applied to model crossing decisions when pedestrians negotiate the right of way with vehicles. Conventional game theory,<sup>[91]</sup> Sequential Chicken (SC) game,<sup>[92]</sup> and Dual Accumulator (DA) game<sup>[93]</sup> were utilised to characterise the dynamic crossing decisions.

Environmental variability and pedestrian heterogeneity further complicate crossing decision modeling. For example, crossing multiple lanes often involves pedestrians waiting at lane lines and accepting traffic gaps successively, known as rolling gap behaviour.<sup>[94,95]</sup> Pedestrians waiting at lane lines might be more likely to accept smaller traffic gaps than those waiting at curbs.<sup>[43,94]</sup> Another complex scenario is crossing a two-way road, which is physically and cognitively demanding. Pedestrians need to consider vehicles on both sides.<sup>[96]</sup> Similarly, crossing at intersections with dense continuous traffic is also challenging, as pedestrians need to anticipate crossing gaps upstream of traffic and make trade-offs between safety and time efficiency.<sup>[97]</sup> Typically, it was assumed that as waiting time increases, pedestrians tend to accept riskier crossing opportunities.<sup>[43,80]</sup> However, the latest evidence suggested that pedestrians who tended to wait were more cautious and less likely to accept risky gaps.<sup>[44–46]</sup>

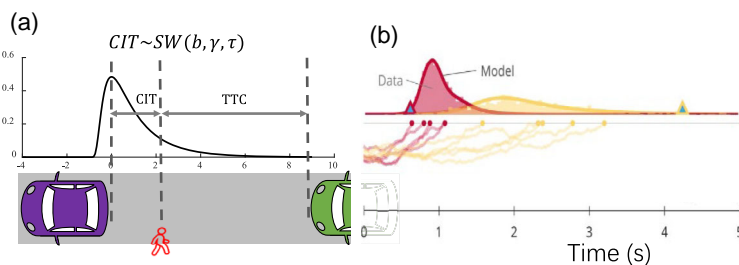
Regarding pedestrian heterogeneity, ANN and LR models were applied to characterise age impact on crossing decisions by Refs. [82,98]. Distractions, such as cellphone usage, could also influence pedestrian crossing decisions.<sup>[82,99,100]</sup> applied ANNs to model the impact of cellphone usage on crossing decisions. Furthermore, pedestrians often cross the road in a group, exhibiting herd behaviour. This behaviour was described by the tendency of group members to maintain a certain distance from the group centre.<sup>[101,102]</sup> used an EA model to characterize information cascades in group decision-making, taking into account the influence of previous agents' decisions.

**Initiation and Motion:** Initiation and motion crossing initiation time (CIT), represents the time it takes for pedestrians to begin crossing the road, reflecting the dynamic nature of their decisions.<sup>[103]</sup> Generally, CIT is the duration between the moment when the crossing opportunity is available and when the pedestrian begins to move.<sup>[31,34]</sup> Drift-diffusion theory posited that CIT is influenced by the accumulation of noisy evidence in the cognitive system, reflecting the efficiency of pedestrian cognitive and locomotor systems.<sup>[104]</sup> Various factors could affect CIT, including vehicle kinematics, age, gender, and distractions. Pedestrians tended to initiate crossing more slowly when faced with higher vehicle speeds.<sup>[34]</sup> Furthermore, female pedestrians tended to initiate crossings more quickly than males,<sup>[105]</sup> and the elderly tended to initiate sooner than young pedestrians.<sup>[45]</sup> The impact of distractions on the CIT depends on their components.<sup>[44]</sup>

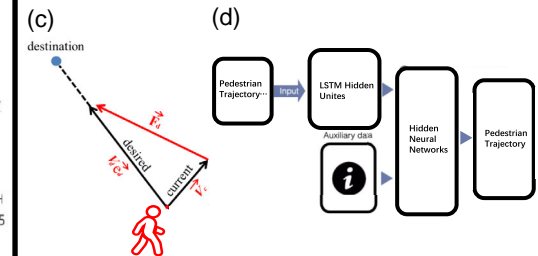
In scenarios where pedestrians face non-yielding vehicles, the risk of collision increases as the distance between the vehicle and the pedestrian decreases. Therefore, pedestrians typically make rapid decisions by assessing “snapshots” of approaching vehicles.<sup>[31,36]</sup> The distribution of CIT in these scenarios is often concentrated, and right-skewed.<sup>[88]</sup> Response time models, such as the Ex-Gaussian and Shifted Wald (SW) distributions, were used to model CITs in these situations.<sup>[106]</sup> For instance,<sup>[107]</sup> modelled CITs as variables following SW distribution (Figure 7a).

In vehicle-yielding scenarios, as discussed in Section 3.2.2, CITs exhibit a bimodal distribution.<sup>[84]</sup> For the early group of CITs, the distribution is similar to that in non-yielding scenarios, as pedestrians employ similar decision-making strategies.<sup>[37]</sup> However, for the late group, the distribution is complex and cannot be described by standard response time distributions.<sup>[84]</sup> EA models with time-varying evidence have been proposed to address this complexity, allowing for the generation of CIT distributions with intricate shapes<sup>[84,88]</sup> (Figure 7b). Moreover,<sup>[86]</sup>

**Initiation model**



**Motion model**



**Figure 7.** Initiation and motion models for pedestrians. a) Response time model.<sup>[47]</sup> b) Evidence accumulation model.<sup>[90]</sup> c) Social force model.<sup>[98]</sup> d) LSTM-ANN.<sup>[117]</sup>



modelled CITs in vehicle-yielding scenarios using the joint distribution of response time models. Additionally,<sup>[87]</sup> applied an RL model to learn the crossing initiation patterns of pedestrians.

After pedestrians initiate their crossings, they need to traverse the road. Walking is a key part of crossing behaviour and is influenced by many factors, such as the presence of approaching vehicles, infrastructures, pedestrian age, and distractions. Pedestrians adjusted their walking trajectories to avoid vehicles.<sup>[108]</sup> In multi-lane crossings, they tended to move to and wait at lane lines, accepting traffic gaps in each lane sequentially.<sup>[94]</sup> Pedestrian walking speeds at crossings were typically faster than normal walking speeds in other scenarios.<sup>[109]</sup> While gender has no significant effect on walking speeds, teenagers and the elderly walk slower than young and middle-aged adults.<sup>[109,110]</sup> Distractions, such as cell-phone use, can reduce pedestrian walking speeds.<sup>[44]</sup>

Walking behaviour can be simulated using microscopic pedestrian motion models, including cellular automata (CA) models, social force (SF) models, and learning-based approaches. CA models are discrete in space, time, and state, making them ideal for simulating complex dynamic systems such as pedestrian-vehicle interactions.<sup>[111,112]</sup> The SF models, based on Newton's second law, were utilised to simulate pedestrian-vehicle interactions and large-scale pedestrian flows<sup>[101,108,113]</sup> (Figure 7c). Ref. [114] used an SF model to simulate the crossing behaviour of pedestrian crowds in complex interaction scenarios involving low-speed vehicles.

In contrast to the above white-box models, there are black-box models based on learning-based approaches, which learn pedestrian walking behaviour from naturalistic datasets or in pre-defined environments. For example, ref. [115] employed ANNs to learn pedestrian walking behaviour by incorporating the relative spatial and motion relationships between pedestrians and other objects extracted from videos. Ref. [116] used the outputs of an SF model as inputs to ANNs to simulate multiple pedestrian walking behaviours. Ref. [117] proposed a Long short-term memory network (LSTM) pedestrian trajectory prediction model (Figure 7d). Additionally, RL and IRL models were also applied to model pedestrian walking behaviour. Ref. [118] applied an RL model to learn multiple pedestrians' walking behaviour in an SF environment. Ref. [26] developed an IRL model to learn pedestrian walking behaviour from video datasets.

### 3.2.3. AV-Involved Applications

In recent years, there has been a growing interest in studying the interactions between autonomous vehicles (AVs) and pedestrians. This interest has led to a multitude of studies that apply pedestrian crossing behaviour theories and models to enhance or assess the performance of AVs in these interactions (Table 2).

One prevalent approach is the use of learning-based methods, which learn pedestrian intention and trajectory from real-world datasets to aid AVs' decision-making. For instance, Ref. [119] proposed a graph convolutional neural network-based pedestrian trajectory prediction model, which considered past pedestrian trajectories to predict both deterministic and probabilistic future trajectories for a range of AV use cases. Other similar models aimed to improve prediction accuracy by considering the social context of interactions. For example, Ref. [117] proposed an LSTM pedestrian trajectory prediction model, which considered past trajectories, pedestrian head orientations, and distance to the approaching vehicle as inputs.<sup>[29]</sup> In addition, there are studies aiming to anticipate pedestrian crossing intentions. Ref. [120] applied SVM, LSTM, and ANN to predict pedestrian crossing intentions separately.

Learning-based approaches have proven effective in predicting pedestrian trajectories and intentions. However, these models demand substantial data for robust performance, limiting their scalability when dealing with interaction cases lacking sufficient data. Moreover, the black box nature of these models can make it challenging to interpret the generated trajectories and intents, which poses a challenge for AV decision-making modeling.<sup>[121]</sup> To address these issues, expert models have been developed. For example, the SF model has been modified to predict pedestrian trajectories for AVs by incorporating more interaction details, such as TTC and the interaction angle between vehicles and pedestrians.<sup>[98,121]</sup> Moreover, SF and CA models have also been embedded in AV decision modules to represent pedestrian crossing behaviour and guide AV decisions in interactions with pedestrians.<sup>[122–124]</sup>

Furthermore, crossing decision models have also been applied in AV research. For example, ref. [123] employed crossing critical gap models to characterise pedestrian crossing decisions in their

**Table 2.** Applications of pedestrian theories and models in AV contexts.

Research	Purpose	Applied theory	Applied model
[119]	Pedestrian trajectory prediction	Learning-based	GCN
[120]	Pedestrian trajectory prediction	Learning-based	SVM, LSTM, Dense NN
[117]	Pedestrian trajectory prediction	Learning-based	LSTM
[98]	Pedestrian trajectory prediction	GA behaviour, walking behaviour	LR,SF
[121]	Pedestrian behaviour modeling	Visual perception, GA behaviour, walking behaviour	Critical gap SF bearing angle
[122]	Pedestrian behaviour modeling	Crossing motivation, walking behaviour	LR,SF
[123]	Pedestrian behaviour modeling	GA behaviour, walking behaviour	Critical gap, CA
[85]	Pedestrian behaviour modeling	BC behaviour	Speed-distance
[92]	Pedestrian behaviour modeling	Game theory	SC game
[18]	Pedestrian behaviour modeling	Visual perception, GA behaviour	Critical gap $\tau$ bearing angle
[124]	Pedestrian behaviour modelling	GA behaviour, Game theory, walking behaviour	LR, critical gap, Stackelberg game, SF

AV decision module.<sup>[85]</sup> applied their speed-distance model to design defensive and competitive interaction behaviour for AVs. Refs. [98,122] used LR models as pedestrian crossing decision models in their proposed AVs decision-making modules. To enhance the dynamic and interactive nature of crossing decisions, game theoretical models were used to model crossing decisions when negotiating the right of way with AVs.<sup>[92,124]</sup>

Researchers also attempted to use pedestrian perception theories or models to design AV decision-making strategies. For instance, ref. [18] simulated AV-pedestrian coupling behaviour using visual cues,  $\tau$  and bearing angle, based on control theory. Ref. [121] modelled the right of way of AVs and pedestrians using a bearing angle.

## 4. Interaction Modeling

Interaction modeling techniques are relevant to a huge variety of autonomous driving tasks, ranging from traffic forecasting to AV planning and decision-making. Understanding and modeling social interactions in autonomous driving is essential for predicting scene dynamics and ensuring safe AV behaviour. Accurate predictions enhance safety, while misunderstood AV behaviour can lead to accidents. Additionally, comprehending the social impact of AV actions can influence surrounding traffic, like encouraging pedestrian crossings through early stopping. As interaction modeling techniques can be utilised across different task domains, the focus will be on dividing existing interaction modeling techniques regardless of the specific driving task they have been designed for.

An introduction to the classification of interaction modeling techniques is now presented for discussion. The first distinction can be made between learning-based methods and model-based methods. There has been extensive research in autonomous driving that makes use of machine-learning and deep-learning-based techniques.<sup>[125]</sup> In a learning-based approach, a model is learned from an extensive dataset. This set of methods does not require any prior knowledge of the system. Data-driven methods are trained on a dataset of examples, and then they are used to make predictions or decisions. On the opposite side of the spectrum, model-based methods start with a theoretical understanding of the system. This a priori knowledge is used to create a mathematical model of the system. Empirical data is then used to validate the model or adjust its parameters to minimize the discrepancy between the model predictions and the data.

A further distinction can be made between methods that explicitly utilise cognitive features of the human mind, which try to explain the rationale that explains human actions, and methods that only implicitly try to model interactions, trying to map environmental inputs to decisions/actions. Human behaviour studies introduced in Section 3 can serve as a guideline for the development of explicit methods. For instance, game theoretic methods (see Section 6.2) take a more explicit approach by considering traffic participants as rational agents who actively consider each other's actions. In contrast, as an example of non-cognitive approaches, social force methods offer a more empirical perspective, capturing the impact of one participant on another without explicitly detailing the reasoning that explains the agent's behaviour during the interaction. We propose to

distinguish existing modeling approaches based on whether they explicitly or implicitly model the interactions.

Based on these two criteria, four major categories of interaction modeling are identified, and they are reported in **Figure 8**:

### 4.1. Learning-Based Implicit Methods

These types of methods rely on machine-learning or deep-learning techniques. The interactions are implicitly modelled, which means that the agent's behaviour cannot be explained by the model. The model only learns an input-output mapping from the data. Model learning can be facilitated by exploiting interactive model architectures.<sup>[126–129]</sup> In general, deep learning methods that use interaction-specific neural network architectures fall into this category.

In this type of method, the aim is to learn a probabilistic generative model that predicts the agent's future actions  $a$ . The model is a probability distribution conditioned on the environment state  $\mathbf{x}$ , which includes the state of surrounding agents, and a set of learnable parameters  $\theta$ .

$$a \approx p_{\theta}(a|\mathbf{x}) \quad (1)$$

### 4.2. Learning-Based Methods with Cognitive Features

This set of methods relies on explicitly handcrafted interactive features that are used as inputs for a learning-based system. This type of interactive feature can include TTC, relative distance,<sup>[130]</sup> looming and reflecting some cognitive process behind human reasoning. For example, in ref. [131], an LSTM which utilises the inter-vehicle interactions has been developed to classify surrounding vehicles' lane change intentions. The interaction features are composed of risk matrices which account for worst-case TTC with vehicles in surrounding lanes and relative distance. Graph convolutional networks also fall into this category, as interaction features can be explicitly modelled in the adjacency matrix of the graph.<sup>[132,133]</sup>

In this type of method, the aim is to learn a probabilistic generative model that predicts the agent's future actions  $a$ , similarly to 1. In this case, the probability distribution can be conditioned on the environment state  $\mathbf{x}$  and on explicitly handcrafted interactive features  $I(\mathbf{x})$ , which have the purpose of facilitating learning.

$$a \approx p_{\theta}(a|\mathbf{x}, I(\mathbf{x})) \quad (2)$$

### 4.3. Model-Based Non-Cognitive Methods

The modeling is non-cognitive in the sense that the interactions do not actively reason on the cognitive process that is behind the agent's actions. Methods of this group include SF<sup>[114]</sup> and potential fields. The interactions are described by potential functions (or SF), which contain a set of learnable parameters which can be fit from empirical data. Another set of methods includes driver risk fields, which are based on the hypothesis that the driver behaviour emerges from a risk-based field.<sup>[16,134]</sup> The advantage of model-based implicit methods is that they can be easily interpreted and they can embed domain knowledge, such as traffic

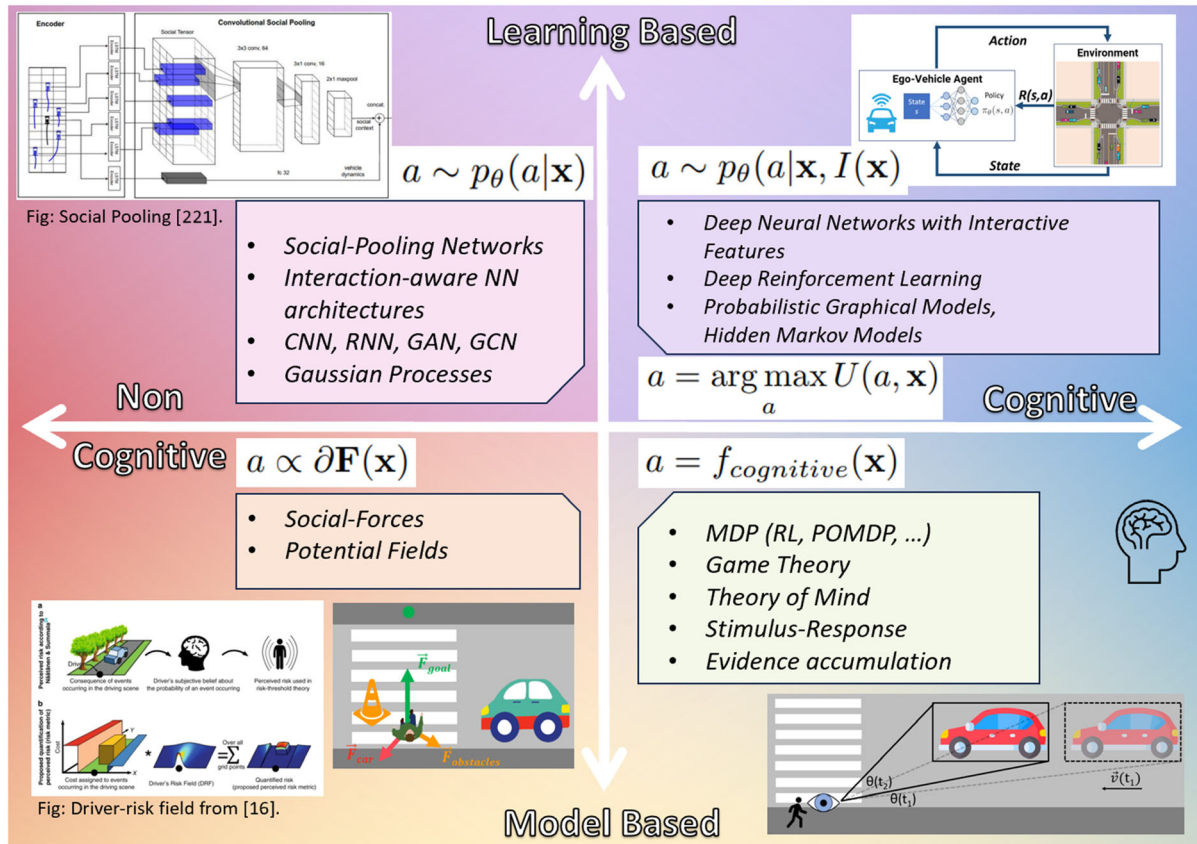


Figure 8. A map of state-of-the-art techniques in IAAD. (Two images reproduced with permission: bottom-left<sup>[16]</sup> Copyright 2020, Springer Nature, and top-left<sup>[221]</sup> Copyright 2018, IEEE).

regulations and scene context. Some models define a potential field and define the agent’s action as proportional to the gradient of such field

$$a \propto \partial F(\mathbf{x}) \tag{3}$$

Otherwise, the forces can be modelled directly so that the gradient operation is not required  $a \propto F(\mathbf{x})$ .

#### 4.4. Model-Based Cognitive Methods

Model-based cognitive methods describe the reasoning behind human decision-making. Two main sets of methods can be distinguished: utility maximisation models and cognitive models.

In utility maximisation methods, humans are modelled as optimizers that select their actions so as to maximise their future utility.

$$a = \arg \max_a U(a, \mathbf{x}) \tag{4}$$

These methods include game theory and Markov Decision Processes (MDPs). In Game theoretic approaches, agents are modelled as players competing or cooperating with each other, thereby taking into account how they react to each other.<sup>[135,136]</sup> The framework of game theory offers a transparent and clear-cut solution for modeling the dynamic interactions among human

drivers, allowing for an understandable explanation of the decision process. However, it still remains hard to satisfy computational tractability as this approach does not scale well with an increasing number of agents. Another possible solution is to model human behaviour as an agent of an MDP, which provides an excellent framework to model decision-making in scenarios where results are influenced by both chance and the decisions made by a decision-maker. Solutions to MDPs can be found with learning-based methods, e.g., DRL algorithms or Monte Carlo Tree Search,<sup>[137]</sup> or with dynamic programming techniques.<sup>[138]</sup>

The second set of methods aims to capture behavioural motivations behind agents’ actions with psychological cognitive processes. This set of methods can include: 1) Stimulus–response models,<sup>[34]</sup> where driver or pedestrian actions are determined, for example, on visual stimuli in the retina; 2) Evidence accumulation,<sup>[84]</sup> where decisions are described as a result of accumulated evidence; 3) Theory of mind, which suggests that humans use their understanding of others’ thoughts and behaviours to make decisions. By predicting others’ actions and inferring their knowledge, humans can drive effectively and safely.<sup>[139,140]</sup>

$$a = f_{cognitive}(\mathbf{x}) \tag{5}$$

In the next sections, each of these classes of interaction modeling will be analyzed in great detail. In particular, Cognitive and Non-Cognitive learning-based methods will be discussed in

Section 5. Model-based cognitive methods have already been thoroughly discussed in Section 3, where Social Force and Potential Fields, Driver Risk Field models, Theory of Mind, Stimulus-Response, and Evidence Accumulation models were included. Section 6 will include Utility-Based methods, which comprise MDPs (Section 6.1) and Game Theory (Section 6.2).

## 5. Learning Based Methods

Machine-learning (ML) methods are widely used in autonomous driving for a variety of tasks, including object detection,<sup>[141]</sup> scene understanding,<sup>[142]</sup> path planning and control.<sup>[143]</sup> By learning from large amounts of data, ML methods can learn to make decisions that are more accurate and efficient than those made by humans.<sup>[144]</sup> This Section will comprise both implicit and explicit learning-based methods identified in the previous Section and give a more detailed view of relevant papers. An overview of some learning-based methods is shown in **Figure 9**.

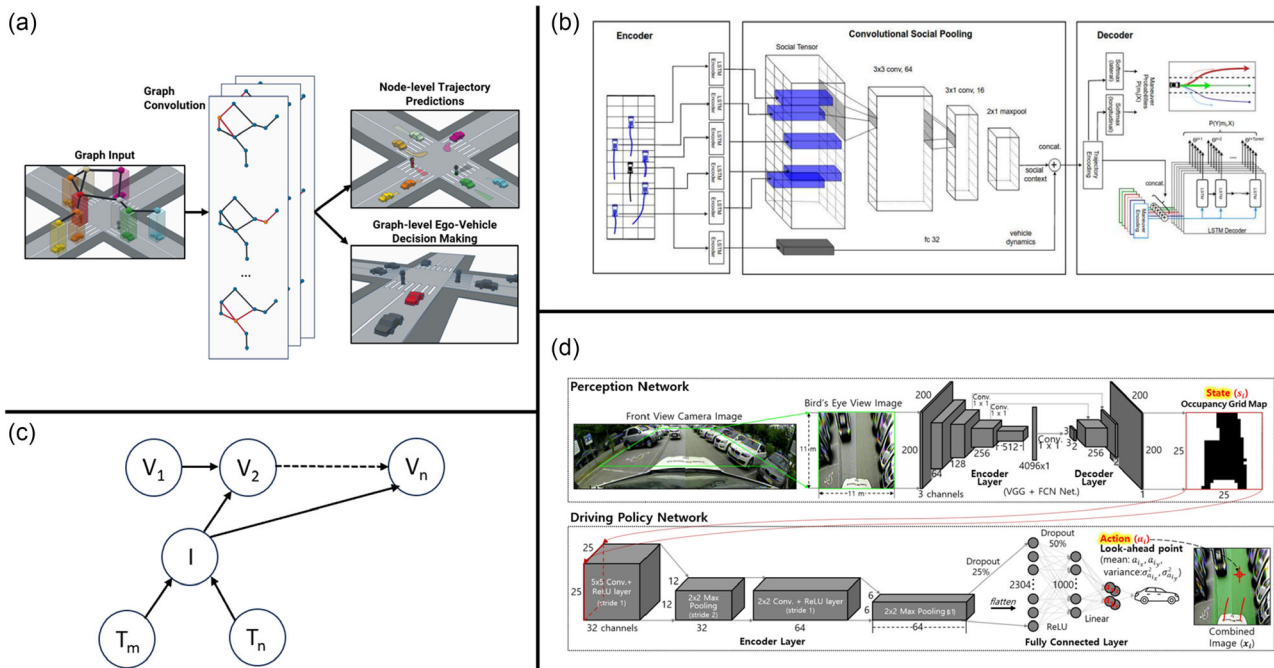
Thanks to recent improvements in neural networks learning representations, it is now possible to use end-to-end driving approaches that take as input the raw sensor readings to output control commands, such as steering and throttle, to solve path-planning and control problems.<sup>[145]</sup> However, it is challenging to learn the entirety of the driving task from high-dimensional raw sensory data (e.g., LiDAR point clouds, camera images) as this involves learning perception and decision-making at the same time. In most of the works, the how-to-act learning process assumes that a scene representation is available to the motion-planning and decision-making module. This actually requires

splitting the end-to-end driving into two main blocks, one in which the AV learns how-to-see and one in which it learns how-to-act.

There are two main approaches to end-to-end self-driving for planning and control tasks (how-to-act): 1) Imitation Learning: in which an agent learns to mimic the behaviour of an expert.<sup>[146–148]</sup> 2) Deep Reinforcement Learning (DRL): in which an agent tries to learn how to act in a trial-and-error process that typically takes place in a simulated environment. DRL methods will be analysed in greater detail in Section 6.

Imitation learning is a machine-learning paradigm in which an agent learns to perform tasks by imitating the behaviour of expert demonstrators, making it a valuable approach for training autonomous systems and robots. In ref. [147], interactive features are learned by means of a graph attention network (GAT). The input to this network consists of surrounding agents' kinematic information as well as a feature vector that encodes scene representation coming from a Bird's Eye View. The model is trained on synthetic data generated by an expert driver in the CARLA simulator. Imitation learning methods tend to work really well in scenarios that are similar to the training scenarios but typically fail when the scenarios diverge from the training distribution. Algorithms like dataset aggregation (Dagger)<sup>[149]</sup> can improve the performances of imitation learning policies by augmenting the initial training dataset with human-labelled data for unseen situations. However, asking an expert to label new training samples can be expensive or unfeasible.

In the context of scene understanding and motion prediction, deep neural networks have been extensively used.<sup>[126]</sup> et al. proposed a social-pooling operation in their neural network



**Figure 9.** Overview figure of deep learning methods in interaction-aware tasks. a) GCNs can be used for both node-level predictions of surrounding agents behaviour as well as ego-vehicle motion generation (graph-level output). b) Social-pooling operation. Reproduced with permission.<sup>[221]</sup> Copyright 2018, IEEE. c) Probabilistic graphical model (as in ref. [222]). d) End-to-end imitation learning network. Reproduced with permission.<sup>[223]</sup> Copyright 2022, Springer Nature.



architecture to account for surrounding neighbours in crowd motion prediction. Similarly, ref. [150] made use of a star-topology network with max-pooling operation to account for interaction features in multi-agent forecasting. CIDNN<sup>[127]</sup> uses LSTM to track the movement of each pedestrian in a crowd and assigns a weight to each pedestrian's motion feature based on their proximity to the target pedestrian for location prediction. The study in ref. [128] created a dataset and proposed a framework called VP-LSTM to predict the trajectories of vehicles and pedestrians together in crowded mixed scenes by exploiting different LSTM architectures for heterogeneous agents. A generative adversarial network (GAN) is applied in ref. [129] to sample plausible predictions for any agent in the scene. The shared feature of these methods is the usage of Recurrent Neural Networks that capture spatiotemporal interaction features in conjunction with pooling operations. The pooling operation allows one to account for surrounding agents by mixing up the hidden states extracted by the LSTMs. During the social-pooling operation, the hidden states of surrounding agents become features that are used to predict the current agent motion. Diffusion models are another set of deep-learning techniques with increasing popularity in modeling spatial-temporal trajectories, which can be used for predicting both pedestrian and car trajectories.<sup>[151]</sup>

Graph convolutional networks (GCNs) have been widely used in trajectory prediction tasks with interacting agents. In these methods, the road structure is represented as a graph, with each node representing a traffic participant. Each node can carry information such as the traffic participant's class (car, truck, pedestrian, etc.), its location, or speed. Explicit interaction can be modelled in the Adjacency Matrix of the graph, whereas the implicit part consists of the graph convolutional layers. GCNs are widely used in traffic forecasting,<sup>[152–155]</sup> and have also been recently used in motion planning,<sup>[156–159]</sup> especially in combination with DRL.

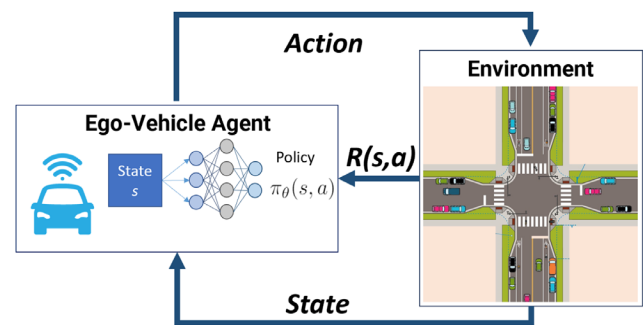
Other machine-learning techniques that can be used to model interactions include Gaussian Processes<sup>[160]</sup> and probabilistic graphical models, including Hidden Markov Models.<sup>[161,162]</sup>

## 6. Utility-Based Methods

Utility-based agents<sup>[163]</sup> employ utility functions to guide decision-making, assigning values to possible world states and selecting actions leading to the highest utility. In contrast to goal-based agents, which evaluate states based on goal satisfaction, utility-based agents can handle multiple goals and factor in probability and action cost. Utility-based methods encompass Markov decision processes (MDPs) and game theoretic models which will be analysed in Section 6.1 and 6.2.

### 6.1. Markov Decision Processes

MDPs are a mathematical framework used to model decision-making problems where the outcomes are partly random and partly under the control of a decision-maker. The modeling framework for MDPs is illustrated in **Figure 10**. Two main methods exist to solve MDPs: dynamic programming and reinforcement-learning.<sup>[138]</sup> Typically the latter set of methods are more used in autonomous driving, as they are more suitable for high-dimensional state spaces (**Table 3**).



**Figure 10.** MDP framework. An agent takes an action that affects the environment state. The updated environment state is used to take the next action, and the cycle repeats. The reward function is used to define the objective of the MDP, which is to maximise the expected cumulative reward over time.

#### 6.1.1. Reinforcement Learning

Reinforcement learning (RL) leverages Markov decision processes (MDPs) to model complex environments and comprises a set of algorithms to learn policies that maximise the expected reward.<sup>[138]</sup>

Traditionally, dynamic programming is a reliable approach for this purpose, iteratively calculating the value of each state, commencing from terminal states and working backwards to the initial state. This method excels in scenarios with modest state spaces. However, it can be computationally burdensome when confronting RL challenges characterized by vast state spaces, such as the domain of autonomous driving. More commonly, RL augmented with deep neural networks (DRL) is used. DRL algorithms can be more sample-efficient and scalable than dynamic programming algorithms, but they can also be more complex and difficult to train. For a more detailed survey on DRL applications for autonomous driving, please refer to ref. [143].

DRL solutions in autonomous driving will be classified based on the scenario used, the state space representation, the action space, and the algorithm used.

Typical state representations used in DRL include see **Figure 11**:

*Vector-Based Representation:* In this type of representation, information regarding surrounding vehicles, such as position and velocity, is included in a vector of fixed length.<sup>[164]</sup>

*Bird's Eye View (BEV) Image:* A 2D image representation of the environment surrounding the ego-vehicle from a top-down perspective.<sup>[165]</sup>

*Occupancy Grid Representation:* Similar to a BEV image, it is a 2D discrete representation of the environment that surrounds the ego-vehicle. It is a 2D or 3D grid of cells, each of which is assigned a probability of being occupied by an obstacle, as well as segmentation information regarding what type of entity is occupying the cell.<sup>[166,167]</sup>

*Graph Representation:* It is a way of representing the state of the environment around an AV as a graph. The nodes in the graph represent objects in the environment, such as vehicles, pedestrians, and traffic lights. The edges in the graph represent the relationships between objects, such as proximity or potential for collision. Graph representations are compact and efficient

**Table 3.** DRL overview table.

Research	Scenario	Observation	Action	Reward	DRL Alg	Network	Simulator
[174]	Pedestrian collision avoidance	Continuous ( $x_p, y_p, v$ )	Discrete (break, keep, change lane)	coll, smooth, succ	DQN	MLP	PreScan
[176]	Pedestrian collision avoidance	Continuous ( $x_p, y_p, v_p, v$ )	Continuous ( $a$ )	coll, speed, succ, SVO	SAC	MLP	Custom
[164]	Highway navigation	Continuous ( $x, y, v$ ) + 8 surr. veh.	Discrete (LLC, RLC, KEEP) or (LLC, RLC, 4 * ACC)	coll, near coll, lc, speed	DQN	CNN	Custom
[172]	Pedestrian collision avoidance	Continuous ( $x_p, y_p, v$ )	Discrete (ACC * 4)	coll, early break penalty	DQN	MLP	PreScan
[175]	Pedestrian collision avoidance	Grid based representation	Discrete (ACC * 4)	speed, coll, near coll	DQN	CNN + LSTM	Custom
[169]	Intersection	Continuous ( $v, path$ ) + 4*surr. veh.	Discrete (5 * ACC)	coll, near coll, acc, succ	D3QN PER	MLP	SUMO
[168]	Dense traffic lane change	Grid	Continuous (jerk, $\omega$ )	coll, speed, jerk, time	PPO	CNN	Custom
[159]	Lane change	Graph ( $p_{rel}, v_{rel}, lane$ index)	Discrete (LLC, RLC, KEEP)	speed, lc, coll	DQN	GCN	SUMO
[158]	Highway lane change	Graph ( $p_{rel}, v_{rel}$ )	Continuous ( $\omega, a$ )	coll, goal, vel, acc	PPO	GCN	BARK
[225]	Intersection, Roundabout	k-nearest, ( $p_{rel}, v_{rel}, \theta$ )	Discrete	distance, coll	H-Ctrl DDQN	MLP	Custom
[181]	Town navigation	Lidar, Camera	Continuous ( $a, \delta$ )	coll, speed acc, time, speed limit, out lane	Proposed	CNN	CARLA
[171]	Crowd-navigation	k-nearest ( $p_{rel}, v_{rel}$ )	Discrete 80 (5 * ACC, 16*angles)	coll, goal, proximity	Deep V- learning	Social- Attentive	Custom
[170]	Crowd-navigation	k-nearest ( $p_{rel}, v_{rel}, d, r$ )	Discrete 11 (angle-speed comb.)	coll, goal, proximity	GA3C- CADRL	LSTM	Custom
[185]	MARL intersection, Roundabout	All within range	-	time, speed, coll, front-car distance	Double DQN	Attention	Custom
[156]	MARL highway navigation	( $p_{rel}, v_{rel}, lane, intention$ )	Discrete 3 (LLC, RLC, keep)	goal, speed, coll., lc	Graph Q (Proposed)	GCN	SUMO
[166]	MARL merge	Cell grid	Discrete (ACC * 5)	goal, coll, flow	Curriculum	MLP	Custom
[165]	MARL connected AD	RGB BEV Image	Discrete nine combinations of (brake, steer, throttle)	goal, speed, coll, lc	IMPALA	CNN	CARLA
[182]	MARL connected AD	Graph ( $s, v, d$ )	Continuous ( $a$ )	speed, action, idle, proximity, coll	TD3	GCN	Highway-env Open Source

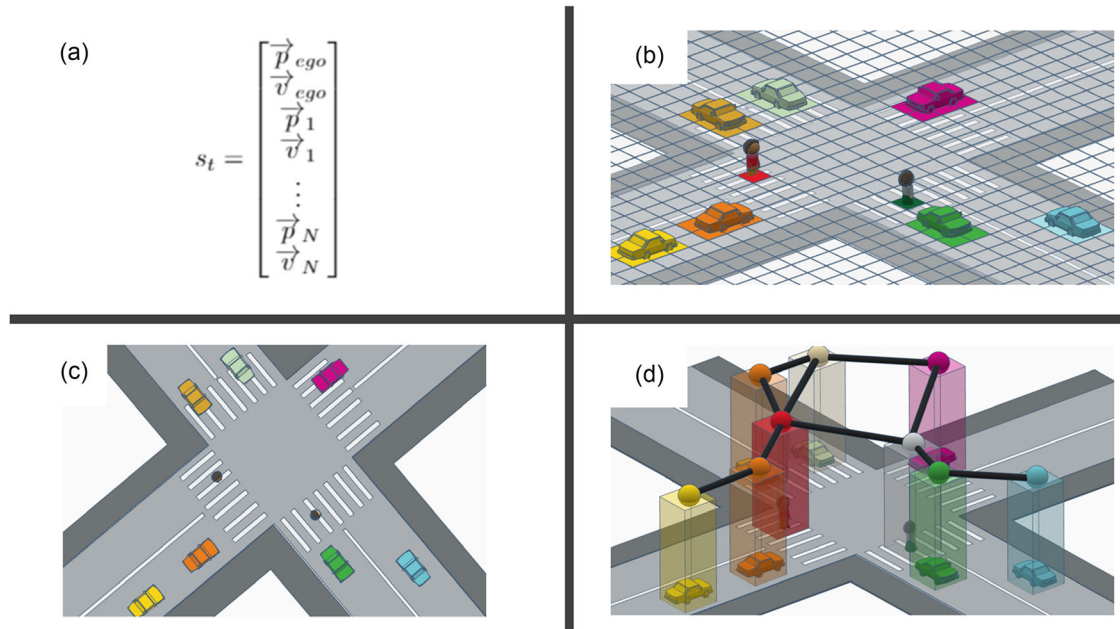
and are a promising approach to representing the state of the environment.<sup>[158,159]</sup>

Vector-based representation offers a compact and efficient representation of objects at the expense of limiting traffic information to a subset of fixed dimensions of surrounding vehicles. BEV images and occupancy grids offer a simple way to represent the environment with fixed and can be easily updated. However, they can be inaccurate in environments with high clutter or uncertainty. Graph representation can easily represent the relationships between agents in a compact way. In contrast, it can be complex and computationally expensive to update the graph as the number of surrounding agents increases.

The action space can be continuous or discrete. Continuous actions usually include the ego vehicle's longitudinal acceleration and steering angle.<sup>[168]</sup> Discrete actions usually depend on the

specific task being solved. For example, in a lane change scenario, discrete actions include left-lane change, keeping the current lane, or right-lane change. The lower-level controller regulates the steering and acceleration of the vehicle to execute the manoeuvre.<sup>[167,169]</sup>

Whilst most DRL papers focus on vehicle-only traffic scenes, the number of papers that deal with mixed traffic scenarios or with vehicle-pedestrian interaction is more limited. Some works exist in the mobile robots' crowd navigation. In ref. [170], DRL is used to navigate a robot in a crowd in a multi-agent setting. In ref. [171], the model in ref. [170] is improved by using attention-based neural networks and social pooling. An autonomous braking system was developed in ref. [172] with a DQN agent. The authors implement a trauma memory, which is used to sample from collision scenarios in a way similar to prioritized experience



**Figure 11.** Illustration of state representations typically used in AD. a) Vector representation, b) grid-based, c) bird's eye view, d) graph.

replay (PER).<sup>[173]</sup> In ref. [174], a DQN agent is trained to avoid collisions with a crossing pedestrian and is further used to develop an ADAS system to aid drivers in pedestrian collision avoidance scenarios. Deshpande et al.<sup>[167,175]</sup> used a grid-state representation with four layers. In a similar scenario, the authors in ref. [176] developed a SAC agent with continuous actions to be used. By integrating SVO component in the reward function, the vehicle can be trained to have different social-compliant behaviours, from pro-social behaviours to more aggressive ones.

Deploying deep reinforcement learning (DRL) in real-world scenarios poses a significant challenge and is an open research field. Some studies, like ref. [177], implement DRL policies directly in real-world applications without additional fine-tuning, showcasing their effectiveness in scenarios like unsignalised intersections. Transfer learning, a sub-field of deep learning, is currently being explored to transfer knowledge from the simulation to the real world. Two main techniques include domain adaptation and domain randomisation.<sup>[178]</sup> With domain randomisation, the approach aims to have a sufficiently large training dataset that encompasses the real world as a specific case.<sup>[179]</sup> With domain adaptation,<sup>[180]</sup> the aim is to learn from a source distribution a model that performs well on a target distribution.

Another issue related to DRL is that the learning-based strategy has high training costs and makes it difficult to achieve semantic interpretation. Recently, some researchers have focused on interpretable learning algorithms and lifelong learning algorithms to solve the above shortcomings.<sup>[181]</sup>

### 6.1.2. Multi-Agent Reinforcement Learning

When multiple RL agents are being deployed into the real world and interact with each other, the problem becomes Multi-agent reinforcement learning (MARL). In order to deal with multi-agent systems, multiple approaches are possible. The first

approach is to have a centralised controller that manages the entire fleet. By increasing the state dimension to include all vehicles and having a joint action vector, the problem can again become a single-agent problem.<sup>[182]</sup> The drawback is the increased dimensionality of the state and action spaces, which can make learning more complex. Recently, graph-based representation has been employed to overcome the curse of dimensionality of the problem.<sup>[182]</sup>

Another approach, which takes inspiration from level-k game theory, is to have a single DRL learner but replace some of the surrounding agents with previous copies of itself.<sup>[169]</sup> This technique is similar to self-play, which is used in competitive DRL scenarios.<sup>[89]</sup> Finally, the last approach is to formulate the problem with a MARL approach, where multiple learners are in parallel. A multi-agent deep deterministic policy gradient (MADDPG) method is proposed in ref. [183], which learns a separate centralized critic for each agent, allowing each agent to have different reward functions. See ref. [184] for an extensive review of MARL. Other applications of MARL in autonomous driving can be found in refs. [156,165,166,185].

### 6.1.3. Partially Observable Markov Decision Processes

Partially observable Markov decision processes (POMDPs) are a generalisation of MDPs. If the process state  $s$  cannot be directly observed by the decision-maker, the MDP is said to be partially observable. POMDPs are computationally expensive but provide a general framework that can model a variety of real-world decision-making processes. POMDP applications for autonomous driving are becoming increasingly popular thanks to hardware improvements. In ref. [186], POMDP has been used to navigate a mobile robot in a crowd. The robot keeps a belief over possible future goals for pedestrians. POMDPs have also been utilised for car decision-making in the presence of pedestrians.<sup>[187]</sup> In a

POMDP, the agents surrounding the ego-vehicle are modelled as part of the environment, and a belief vector is used to model their intentions. In ref. [185], the authors develop a multi-agent interaction-aware decision-making policy with DRL. The problem is modelled as a POMDP, and the interaction is modelled with an attention-based neural network mechanism. POMDPs have also been used to solve decision-making problems under environmental occlusions at intersections.<sup>[188]</sup> Other applications of POMDP to interactive decision-making can be found in refs. [189,190]. Traditional control methods usually deal with sensor uncertainty and planning in a sequential manner, where the sensor noises and uncertainties are dealt with by a state-estimator, and then a deterministic policy is used to determine the action based on the estimated state. POMDPs, in contrast, do not make such separation, and the policy is determined based on a belief state. Surrounding agents can be either explicitly modelled as decision-makers (MARL) or can be treated as part of the environment in which a single agent operates (RL or DRL).<sup>[168,191]</sup>

## 6.2. Game Theoretic Models

Game theory is the study of mathematical models of strategic interactions between rational agents.<sup>[192]</sup> Game theory primarily applies to economics but has also emerged in autonomous driving. In particular, dynamic non-cooperative game theory is of

particular importance for autonomous driving.<sup>[193]</sup> Game theory is dynamic if it involves multiple decisions and the order of such decisions is important, and it is non-cooperative if each person involved pursues their own interest, which is partly conflicting with other people's interest. The dynamic non-cooperative game theory comprises both discrete and continuous time games, and it provides a natural extension of optimal control to multi-agent settings (Table 4).<sup>[192]</sup>

Game theory examines equilibrium solutions under optimal player assumptions, with multiple concepts applicable to trajectory games. Dynamic games classify into open-loop and feedback games based on available information; open-loop assumes that the only information available to each player is the initial state of the game. For feedback games, the information available to each agent is the current state of the game. Although the second type of game describes more accurately the AD setting, open-loop solutions are often preferred for their simplicity. Common equilibria in autonomous driving include open-loop Nash, open-loop Stackelberg, closed-loop Nash, and closed-loop Stackelberg equilibria. For more details on the topic see ref. [192].

When the agents' dynamics have to comply with a set of constraints, such as collision-avoidance constraints, the equilibria are called generalized. Generalized equilibria are studied in ref. [194]. Numerical solutions to the problems of open-loop Nash equilibria can be found in refs. [195–197]. The drawback of the open-loop Nash equilibrium formulation is that the players

**Table 4.** Game theory models for decision making in various scenarios.

Research	Scenario	Game Theory Model	Agents	Action
[226]	Lane change	Nash equilibrium	Up to 4	Both discrete (lane change) and continuous $a$
[214]	Autonomous racing	Nash equilibrium	2	Continuous
[204,205]	Intersection	Nash equilibrium	3	Continuous ( $\omega$ and $a$ )
[198]	Drone racing	Nash equilibrium	2	Continuous ( $\omega$ and $\nu$ )
[213]	Autonomous racing	Nash equilibrium	2	Continuous ( $a$ and $\delta$ )
[206]	Autonomous racing	Nash equilibrium	2	Continuous ( $a$ and $\delta$ )
[13]	Merging	Nash equilibrium, SVO	2	Continuous ( $a$ and $\delta$ )
[208]	Highway, lane change	Nash equilibrium, DRL	2	Discrete ( $a$ and lane change)
[199]	Autonomous racing	Nash Eq. and Stackelberg Eq.	2	Discrete trajectories
[193]	Highway navigation, intersection	Stackelberg Eq.	2	Continuous ( $a$ and $\delta$ )
[200]	Lane change	Stackelberg Eq.	3	Both discrete and continuous
[201]	Lane change	Stackelberg Eq.	2	Discrete (lane change)
[135]	Roundabout	Stackelberg Eq.	2	Continuous
[210]	Truck platooning	Hierarchical Stackelberg Eq.	2	Continuous ( $a$ and $\delta$ )
[227]	Lane change	Stackelberg Equilibrium	2	Discrete (lane change)
[212]	Highway navigation	Hierarchical Feedback Stackelberg Eq.	2+	Continuous
[228]	Highway navigation, merge	Game Tree	2	Discrete
[211]	Merge	Game Tree	2	Discrete
[215]	Highway navigation	Generalised Feedback Nash Eq.	2+	Continuous
[136]	Left turn	Generalized Nash Eq., SVO	2	Continuous ( $a$ and $\delta$ )
[202]	Drone racing	Generalized Nash Eq.	6	Continuous ( $\nu$ )
[195]	Merge, intersection	Generalized Nash Eq.	4	Continuous ( $a$ and $\omega$ )
[209]	Intersection	Level-k	2	Continuous ( $a$ and $\omega$ )
[207]	Roundabout	Level-k	2	Continuous ( $a$ and $\omega$ )



cannot reason directly on how their actions influence the behaviour of surrounding agents. A first simplification of this is the open-loop Stackelberg equilibrium, which has been applied, for instance, in ref. [198] in the context of drone-autonomous racing. In a Stackelberg competition, the leader makes the first move, and subsequent players follow sequentially, allowing those with higher precedence to consider how those with lower precedence will plan their actions. In ref. [199], the authors propose a sequential bimatrix game approach to autonomous racing based on an open-loop Stackelberg game formulation. Other applications of Stackelberg formulation can be found in refs. [200–202]. A formulation for solving generalized feedback Nash equilibria can be found in ref. [203]. Other methods can be found in refs. [204–206]. Sadigh et al.<sup>[193]</sup> model AV-human interaction as a partially observable stochastic game in a Stackelberg competition. The human estimates the AV's plan and acts accordingly, while the AV optimizes its own actions, assuming indirect control over the human's actions.

Typically, game-theoretic approaches suffer from the following problems: 1) the computational complexity grows exponentially in the number of agents and with the increasing temporal horizon, 2) they assume that the utility function that justifies other agents' actions is known to the ego-vehicle and that the agents act rationally with respect to those reward functions—however it is a known fact in game-theoretic finance problems that humans often do not act rationally;<sup>[19]</sup> 3) the behaviour of the agents might be stochastic and solving for mixed or behavioural strategies makes the computations even more intractable. Naturally, game theory also has the great advantage of capturing the interdependence of actions and some exact solutions exist for a restricted number of problems. Many papers in the field of game-theoretic autonomous driving try to alleviate those issues by further simplifying the problem or by finding approximate solutions. Now, a look will be taken at some papers in the field to analyze their simplifying hypotheses.

Level-k theories break down the Nash-equilibrium rational expectations logic by assuming people see others as being less sophisticated than themselves. This is level-k reasoning, in which the iteration process is stopped after  $k$  steps. Other agents are modelled as level  $k-1$  actors. A level- $k$  agent assumes that all the other agents are level- $(k-1)$ , makes predictions based on this assumption, and responds accordingly. In ref. [207], level- $k$  reasoning is applied to a roundabout scenario. This approach has also been incorporated in an RL framework in ref. [208]: the authors restrict the problem to two interacting agents and solve the Markov game with two vehicles with a DQN RL-based approach. In ref. [209], level- $k$  reasoning is adopted to resolve conflicts at intersections. The authors showed that conflicts can be resolved easily in the case when the ego-vehicle is a level- $k$  agent and all surrounding vehicles are level- $k-1$  or inferior. However, the number of collisions increases when both agents are of the same level, which indicates that further improvements need to be added to tackle scenarios with agents of the same kind, which would be crucial in the case of multiple AVs.

In order to keep computational complexity tractable, the number of agents can be reduced by identifying a subset of all agents that interact with the ego-vehicle.<sup>[210,211]</sup> The time horizon can also be limited by considering a receding horizon controller

or by implying hierarchical game-theoretic planning. The latter consists of having a short-horizon tactical planner in combination with a long-horizon strategic planner. The first one is responsible for accurate dynamics modeling of the problem, and the second one is responsible for deciding the strategy with approximate dynamics. Examples of this approach can be found in refs. [210,212].

Iterative linear-quadratic (LQ) methods are increasingly common in robotics and control communities. The authors of ref. [204] formulate the problem as a general-sum differential game characterized by nonlinear system dynamics. In ref. [205] extend their methods to systems with feedback-linearizable dynamics.

Another method to solve game theoretic problems is to use iterative best response to calculate pure Nash equilibria, i.e., Nash equilibria in pure strategies. The authors of ref. [202] propose a “sensitivity enhanced” iterative best response solver. In ref. [213] an online game-theoretic trajectory planner based on the IBR is presented. The planner is suitable for online planning and exhibits complex behaviours in competitive racing scenarios. Williams et al.<sup>[214]</sup> propose an IBR algorithm together with an information-theoretic planner for controlling two ground vehicles in close proximity.

In ref. [13], Schwarting et al. propose an alternative method to the iterative best response to solve the Nash equilibrium problem based on a reformulation of the optimisation problem as a local single-level optimisation using the Karush–Kuhn–Tucker conditions. In ref. [136], game theory is used to model other vehicles' decision-making. They propose a parallel-game interaction model (PGIM) to serve active and socially compliant driving interactions. To address uncertainty in the environment,<sup>[206]</sup> extend the game theoretic Nash equilibrium concept to POMDPs. In ref. [215] the authors take uncertainty about the intention of other agents by constructing multiple hypotheses about the objectives and constraints of other agents in the scene.

## 7. Discussion and Future Challenges

In this comprehensive review, two critical sections central to the advancement of autonomous driving were introduced: Human Behaviour Studies and Interaction Modeling. These sections form the base for understanding and optimizing the intricate dynamics of interactions within autonomous driving scenarios. In this Section, challenges and research directions are highlighted for future AV research in interaction scenarios.

### 7.1. Human Behaviour Studies

Driven by society's strong desire for autonomous driving, human behaviour research has once again become a hot topic in recent years, especially research in AV contexts. In order to better understand pedestrian behaviour during AV interactions, many challenges still need to be overcome.

In general, the exploration of driver behavioural models holds promise as a research area with the potential to bring about substantial enhancements in the safety and efficiency of transportation systems. Nevertheless, there is a substantial amount of work yet to be undertaken in the development and validation of these models. Future research should prioritise the creation of more

holistic models that encompass a broader spectrum of factors, including the driver's psychological state, the surrounding environment, and interactions with other human participants on the road.

For pedestrian behaviour studies, one critical challenge is communication. Firstly, although most researchers agree on the effectiveness of eHMIs, a consensus is still lacking regarding their contents, forms, and perspectives.<sup>[216]</sup> An open question persists regarding whether eHMIs should be anthropomorphic or non-anthropomorphic. A similar question arises for textual and non-textual eHMIs. Furthermore, given the presence of multiple pedestrians on the road, the current eHMIs are primarily designed for one-to-one encounters, which may mislead other pedestrians.<sup>[217]</sup> Numerous comparable problems persist, which impede the standardisation of the eHMI. In contrast, since implicit signals, such as vehicle kinematics, are widely accepted, pervasive, common, and reliable communication methods, their critical roles cannot be ignored.<sup>[218]</sup> Although researchers have made initial attempts to influence pedestrians by manipulating implicit signals, such as vehicle deceleration rate, lateral distance, and pitch, these efforts are insufficient to ensure safe and efficient communication.<sup>[35,219,220]</sup> These communication methods lack relevant theoretical support to demonstrate the accurate and effective delivery of communication information. Additionally, there is a dearth of reliable research paradigms to guide the research methodology, including vehicle driving behaviour design, subjective and objective experimental design, and more. Additionally, how to combine eHMI and implicit signals efficiently and smoothly to take advantage of both parties is also an interesting research direction.

Another challenge is pedestrian behaviour studies. Pedestrian decision-making and behavioural patterns are influenced by the diversity of interaction situations, traffic environments, and participants. However, these aspects currently lack sufficient research attention. Existing studies often focus on specific and simple interaction scenarios to control variables or simplify research complexity. However, real-life situations involve a plethora of complex scenarios, including crossings at multi-lane, two-way, or unstructured roads, crossings facing dense continuous traffic flow, multi-pedestrian crossing scenarios, and more. Moreover, pedestrian heterogeneity, such as gender, age, distraction, and group effects, also plays a significant role in interactions. It is noteworthy that many influential factors, such as waiting time and distraction, still lack a consensus. Accordingly, due to the dearth of sufficient and reliable results, research conclusions mostly rely on assumptions, highlighting the inadequacy of understanding the underlying mechanisms of pedestrian road behaviour.

Regarding pedestrian behaviour modeling, learning-based methods have been appealing in recent years. The end-to-end deep neural networks can effectively capture complex behavioural mechanisms and have made considerable progress in the field of pedestrian intention prediction and trajectory prediction. However, its black-box nature cannot be ignored. These approaches require a significant amount of data to achieve robust performance, which limits their scalability to sporadic cases with insufficient data. Additionally, the black-box models have difficulties in explaining their decision-making and behaviour logic, which brings new problems to modelling. In contrast, expert

models, such as the Social Force models, evidence accumulation models, or game theoretical models, have solid psychological and behavioural foundations, and their behavioural decision-making logic is clear and interpretable. However, most of these models have only been validated on limited datasets or are still in the stage of laboratory validation, lacking extensive engineering practice. Hence, the theories of expert models need to be further refined and extensively verified on a large number of real datasets in the future. In addition, expert models and data-driven models have advantages in different aspects. A possible future trend is to find a balance point where the two models are used together.

Finally, considering that only a small fraction of the overall literature on autonomous driving explicitly considers the behaviour of pedestrians, there is a need to increase the application of pedestrian behaviour models, potentially including but not limited to pedestrian behaviour prediction, AV behaviour design, and virtual AV validation.

## 7.2. Interaction Modeling

As autonomous driving technology continues to evolve, research in interaction modeling will play a critical role in addressing challenges and driving the development of safer and more reliable AVs.

One prominent approach that has garnered attention in autonomous driving research is the use of learning-based methods. These methods offer the appeal of end-to-end solutions,<sup>[5]</sup> directly mapping sensory inputs and destination knowledge into the ego vehicle's actions. However, such systems can behave like a black-box, leading to issues with interpretability in case of faults and the validation of their models. Furthermore, the enormity of the task, i.e., learning the entire driving process, poses a significant challenge. Hence, current research endeavours are breaking this down into sub-tasks, including route planning, perception, motion planning, and control and utilising learning-based methods to address these partial challenges.

The advantage of learning interactive behaviours through imitation learning or simulation in Deep reinforcement learning (DRL) approaches is also gaining momentum. Nevertheless, challenges persist. Most decision-making with Deep learning assumes ideal road scenarios and a perfect perception of the surrounding environment. Yet, real-world conditions often involve occlusions, sensor noise, and environmental anomalies. Maintaining system performance in these sporadic events and handling partial or noisy information is an ongoing research challenge. Uncertainty seeps in from unpredictable behaviours in surrounding traffic participants, as well as sensor noise and vehicle models. Furthermore, models trained in simulated environments (like DRL models) raise the question of how to bridge the gap between simulation and reality. Several strategies have been proposed, including making simulations more realistic, domain randomisation, and domain adaptation.<sup>[179]</sup> These approaches aim to prepare models for the unpredictability and complexity of the real world, ensuring that what they've learned in simulation can be applied effectively on the road.

An alternative approach to learning-based methods is model-based methods. This set of methods includes game theoretic models, behavioural models (which have been discussed in the previous Section), social forces and potential fields.

Game theory offers versatility and adaptability and can effectively handle various scenarios without relying on specific data distributions. One of its key advantages is the ability to address planning and prediction for agents in a given situation. However, there's a trade-off in terms of computation. As the number of agents and the time horizon grow, the computational burden increases. Researchers have put forth several strategies to enhance game-theoretic solutions, including hierarchical game-theoretic formulations,<sup>[212]</sup> limiting the optimisation problem for surrounding agents to approximate solutions,<sup>[193]</sup> level-k game theory,<sup>[207]</sup> or improving nonlinear optimisation solvers' performances.<sup>[195,204,205]</sup>

In contrast, Social Force or Potential Field methods offer a fast-computing solution. They can be used to predict surrounding agents' behaviour but also for ego-vehicle control. SFMs rely on simplified assumptions about human behaviour. They often treat pedestrians as particles or agents with fixed characteristics, neglecting the cognitive aspects of human decision-making, which can lead to unrealistic representations of complex and dynamic human behaviours. Future research directions for these methods include incorporating cognitive elements or contextual information, such as road rules and traffic signals. Exploring the integration of machine-learning techniques to improve the adaptability and predictive power of SFMs can also be a possible future research direction.

The research predominantly concentrates on vehicle-vehicle interactions, which undoubtedly play a crucial role in autonomous driving. However, there is a pressing need to develop methods that address interactions with human road users, particularly pedestrians. As the realm of autonomous driving continues to evolve, the elucidation of theories and models that govern communication and interaction with diverse road users takes on an increasingly technical relevance, promising to propel safety and efficiency within autonomous driving scenarios.

## Acknowledgements

Luca Crosato and Kai Tian contributed equally to this work as co-first authors. This project is supported in part by the EPSRC NorthFutures project (ref: EP/X031012/1) and the European Regional Development Fund.

## Conflict of Interest

The authors declare no conflict of interest.

## Keywords

behavioral models, interaction-aware autonomous driving, multi-agent interactions, pedestrians, socially-aware decision making

Received: September 17, 2023

Revised: October 27, 2023

Published online:

[1] E. Yurtsever, J. Lambert, A. Carballo, K. Takeda, *IEEE Access* **2020**, *8*, 58443.

[2] A. Millard-Ball, *J. Plan. Educ. Res.* **2018**, *38*, 6.

- [3] M. Maurer, J. C. Gerdes, B. Lenz, H. Winner, *Autonomous Driving: Technical, Legal and Social Aspects*, Springer Nature, London **2016**.
- [4] S. D. Pendleton, H. Andersen, X. Du, X. Shen, M. Meghiani, Y. H. Eng, D. Rus, M. H. Ang Jr., *Machines* **2017**, *5*, 6.
- [5] W. Schwarting, J. Alonso-Mora, D. Rus, *Annu. Rev. Control Robot* **2018**, *1*, 187.
- [6] C. Hubmann, M. Becker, D. Althoff, D. Lenz, C. Stiller, in *2017 IEEE Int. Veh. Symp.*, IEEE, Piscataway, NJ **2017**, pp. 1671–1678.
- [7] G. Markkula, R. Madigan, D. Nathanael, E. Portouli, Y. M. Lee, A. Dietrich, J. Billington, A. Schieben, N. Merat, *Theor. Issues Ergon. Sci.* **2020**, *21*, 728.
- [8] L. A. Swan, M. B. Owens, *Mid-Am. Rev. Sociol.* **1988**, p. 59.
- [9] A. Turnwald, D. Wollherr, *Int. J. Soc. Robot.* **2019**, *11*, 151.
- [10] L. Liu, S. Lu, R. Zhong, B. Wu, Y. Yao, Q. Zhang, W. Shi, *IEEE Internet Things J.* **2020**, *8*, 6469.
- [11] G. Markkula, Y.-S. Lin, A. R. Srinivasan, J. Billington, M. Leonetti, A. H. Kalantari, Y. Yang, Y. M. Lee, R. Madigan, N. Merat, *PNAS Nexus* **2023**, *2*, pgad163.
- [12] T. Bandyopadhyay, K. S. Won, E. Frazzoli, D. Hsu, W. S. Lee, D. Rus, in *Algorithmic Found. Robot. X*, Springer, New York **2013**, pp. 475–491.
- [13] W. Schwarting, A. Pierson, J. Alonso-Mora, S. Karaman, D. Rus, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 24972.
- [14] N. AbuAli, H. Abou-Zeid, *Int. J. Veh. Tech.* **2016**, *2016*, 12.
- [15] N. M. Negash, J. Yang, *IEEE Access* **2023**, *11*, 22788.
- [16] S. Kolekar, J. de Winter, D. Abbink, *Nat. Commun.* **2020**, *11*, 4850.
- [17] J. Petit, C. Charron, F. Mars, *Front. Neuroergonomics* **2021**, *2*, 682119.
- [18] J. E. Domeyer, J. D. Lee, H. Toyoda, B. Mehler, B. Reimer, *IEEE Trans. Hum.-Mach. Syst.* **2022**, *52*, 557.
- [19] L. Sun, W. Zhan, Y. Hu, M. Tomizuka, in *2019 IEEE Int. Intell. Transp. Syst. Conf.* IEEE, Piscataway, NJ **2019**, pp. 4329–4335.
- [20] A. Varhelyi, *Accid. Anal. Prev.* **1998**, *30*, 731.
- [21] W. Wang, L. Jiang, S. Lin, H. Fang, Q. Meng, *Multimed. Tools Appl.* **2022**, *81*, 39873.
- [22] D. Zhou, H. Liu, H. Ma, X. Wang, X. Zhang, Y. Dong, *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 2669.
- [23] M. Shahverdy, M. Fathy, R. Berangi, M. Sabokrou, *Expert Syst. Appl.* **2020**, *149*, 113240.
- [24] A. Rasch, G. Panero, C.-N. Boda, M. Dozza, *Accid. Anal. Prev.* **2020**, *139*, 105494.
- [25] S. Sun, Z. Zhang, Z. Zhang, P. Deng, K. Tian, C. Wei, *Sensors* **2022**, *22*, 7860.
- [26] P. Nasernejad, T. Sayed, R. Alsaleh, *Transportmetrica A: Transp. Sci.* **2022**, *19*, 3.
- [27] S. El Hamdani, N. Benamar, M. Younis, *Transp. Res. C: Emerg. Technol.* **2020**, *121*, 102856.
- [28] R. L. Moore, *J. Oper. Res. Soc.* **1953**, *4*, 3.
- [29] A. Rasouli, J. K. Tsotsos, *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 900.
- [30] D. Dey, J. Terken, in *9th AutomotiveUI*, **2017**, pp. 109–113.
- [31] R. Lobjois, V. Cavallo, *Accid. Anal. Prev.* **2007**, *39*, 934.
- [32] J. A. Oxley, E. Ihsen, B. N. Fildes, J. L. Charlton, R. H. Day, *Accid. Anal. Prev.* **2005**, *37*, 962.
- [33] T. Petzoldt, *Accid. Anal. Prev.* **2014**, *72*, 127.
- [34] K. Tian, G. Markkula, C. Wei, Y. M. Lee, R. Madigan, N. Merat, R. Romano, *Safety Sci.* **2022**, *154*, 105837.
- [35] D. Dey, A. Matviienko, M. Berger, B. Pflöging, M. Martens, J. Terken, *IT-Info. Tech.* **2021**, *63*, 123.
- [36] Y. M. Lee, R. Madigan, C. Uzundu, J. Garcia, R. Romano, G. Markkula, N. Merat, *J. Safety Res.* **2022**, *80*, 270.
- [37] K. Tian, A. Tzigieras, C. Wei, Y. M. Lee, C. Holmes, M. Leonetti, N. Merat, R. Romano, G. Markkula, *Accid. Anal. Prev.* **2023**, *190*, 107173.
- [38] C. Creech, D. Tilbury, J. Yang, A. Pradhan, K. Tsui, L. Robert, et al., *Front. Robot. AI* **2019**.

- [39] M. Risto, C. Emmenegger, E. Vinkhuyzen, M. Cefkin, J. Hollan, in *Driv. Asse. Conf.*, Vol. 9, University of Iowa, Iowa City **2017**.
- [40] A. Dietrich, P. Maruhn, L. Schwarze, K. Bengler, in *Int. Conf. Human Sys. Eng. Des.*, Springer, New York **2019**, pp. 176–181.
- [41] C. Ackermann, M. Beggiato, L.-F. Bluhm, A. Low, J. F. Krems, *Transp. Res. F: Traffic Psychol.* **2019**, *62*, 757.
- [42] S. Chandra, R. Rastogi, V. R. Das, *KSCE J. Civ. Eng.* **2014**, *18*, 284.
- [43] J. Zhao, J. O. Malenje, Y. Tang, Y. Han, *Accid. Anal. Prev.* **2019**, *129*, 76.
- [44] K. Tian, G. Markkula, C. Wei, E. Sadraei, T. Hirose, N. Merat, R. Romano, *Accid. Anal. Prev.* **2022**, *174*, 106770.
- [45] R. Lobjois, N. Benguigui, V. Cavallo, *Accid. Anal. Prev.* **2013**, *53*, 166.
- [46] G. Yannis, E. Papadimitriou, A. Theofilatos, *Transp. Plan. Technol.* **2013**, *36*, 450.
- [47] K. Tian, G. Markkula, C. Wei, Y. M. Lee, R. Madigan, T. Hirose, N. Merat, R. Romano, arXiv preprint arXiv:2301.10419, **2023**.
- [48] A. Rasouli, I. Kotseruba, J. K. Tsotsos, *IEEE Trans. Intell. Veh.* **2017**, *3*, 61.
- [49] B. Völz, H. Mielenz, G. Agamennoni, R. Siegart, in *2015 IEEE Int. Conf. Intell. Transp. Syst.*, IEEE, Piscataway, NJ **2015**, pp. 854–860.
- [50] K. De Clercq, A. Dietrich, J. P. Núñez Velasco, J. De Winter, R. Happee, *Hum. Factors* **2019**, *61*, 1353.
- [51] J. Carmona, C. Guindel, F. Garcia, A. de la Escalera, *Sensors* **2021**, *21*, 2912.
- [52] T. Petzoldt, K. Schleinitz, R. Banse, *IET Intell. Transp. Syst.* **2018**, *12*, 449.
- [53] Semcon, <https://www.semcon.com/article/the-smiling-car>, **2016**.
- [54] A. A. Kaleefathullah, N. Merat, Y. M. Lee, Y. B. Eisma, R. Madigan, J. Garcia, J. d. Winter, *Hum. Factors* **2022**, *64*, 1070.
- [55] P. Bazilinskyy, D. Dodou, J. C. de Winter, in *Advances in Transportation, AHFE International Conference 2022*, Vol. 60, pp. 141–155.
- [56] L. Kooijman, R. Happee, J. C. de Winter, *Information* **2019**, *10*, 386.
- [57] A. R. Palmeiro, S. van der Kint, L. Vissers, H. Farah, J. C. de Winter, M. Hagenzieker, *Transp. Res. F: Traffic Psychol. Behav.* **2018**, *58*, 1005.
- [58] H. Farah, I. Postigo, N. Reddy, Y. Dong, C. Rydergren, N. Raju, J. Olstam, *IEEE Trans. Intell. Transp. Syst.* **2022**, *24*, 6.
- [59] M. Rettenmaier, K. Bengler, *IEEE Open J. Intell. Transp. Syst.* **2021**, *2*, 282.
- [60] A. Dietrich, M. Tondera, K. Bengler, in *Road Safety Sim. Conf.*, **2019**.
- [61] V. Onkhar, P. Bazilinskyy, D. Dodou, J. De Winter, *Transp. Res. F: Traffic Psychol. Behav.* **2022**, *84*, 194.
- [62] M. R. Epke, Master Thesis, Delft University of Technology **2019**.
- [63] A. Rouchitsas, H. Alm, *Information* **2022**, *13*, 420.
- [64] H. Furuya, K. Kim, G. Bruder, P. J. Wisniewski, G. F. Welch, in *Extended Abstracts of the 2021 CHI Conf. on Hum. Factors in Computing Systems*, **2021**, pp. 1–7.
- [65] P. Sewalkar, J. Seitz, *Sensors* **2019**, *19*, 358.
- [66] S. Cœugnet, B. Cahour, S. Kraiem, *Transp. Res. F: Traffic Psychol. Behav.* **2019**, *65*, 141.
- [67] J. J. Gibson, *The Ecological Approach to Visual Perception: Classic Edition*, Psychology Press, London, England **2014**.
- [68] E. R. Hoffmann, R. G. Mortimer, *Accid. Anal. Prev.* **1994**, *26*, 511.
- [69] D. N. Lee, *Perception* **1976**, *5*, 437.
- [70] P. R. DeLucia, *Hum. Factors* **2008**, *50*, 811.
- [71] B. G. Bardy, W. H. Warren Jr., *J. Sports Sci.* **1997**, *15*, 607.
- [72] P. R. DeLucia, in *The Cambridge Handbook of Applied Perception Research*, Vol. 1, Cambridge University Press **2015**, pp. 568–591.
- [73] A. Dommès, G. Merlhlot, R. Lobjois, N.-T. Dang, F. Vienne, J. Boulo, A.-H. Oliver, A. Cretual, V. Cavallo, *Accid. Anal. Prev.* **2021**, *159*, 106256.
- [74] J. P. Wann, D. R. Poulter, C. Purcell, *Psychol. Sci.* **2011**, *22*, 429.
- [75] K. Jiang, F. Ling, Z. Feng, C. Ma, W. Kumfer, C. Shao, K. Wang, *Accid. Anal. Prev.* **2018**, *115*, 170.
- [76] D. S. Pawar, A. K. Yadav, *J. Safety Res.* **2022**, *80*, 87.
- [77] A. Theofilatos, A. Ziakopoulos, O. Oviedo-Trespalacios, A. Timmis, *J. Transp. Health* **2021**, *22*, 101108.
- [78] M. S. Raff, J. W. Hart, *A Volume Warrant For Urban Stop Signs*, The ENO Foundation for Highway Traffic Control, Saugatuck, CT, USA **1950**.
- [79] H. C. Manual, *Transp. Res. Board* **2010**, 1207.
- [80] A. Rasouli, I. Kotseruba, arXiv preprint arXiv:2203.07324, **2022**.
- [81] V. Himanen, R. Kulmala, *Accid. Anal. Prev.* **1988**, *20*, 187.
- [82] B. R. Kadali, N. Rathi, V. Perumal, *J. Traffic Transp. Eng.* **2014**, *1*, 111.
- [83] D. Sun, S. Ukkusuri, R. F. Benekohal, S. T. Waller, in *Transportation Research Record*, TRB Annual Meeting CD-ROM, Washington, DC **2003**.
- [84] J. Pekkanen, O. T. Giles, Y. M. Lee, R. Madigan, T. Daimon, N. Merat, G. Markkula, *Comput. Brain Behav.* **2022**, *5*, 60.
- [85] H. Zhu, W. Alhajyaseen, M. Iryo-Asano, H. Nakamura, C. Dias, *Physica A: Stat. Mech. Appl.* **2022**, *606*, 128083.
- [86] K. Tian, *Ph.D. Thesis*, University of Leeds, **2023**.
- [87] Y. Wang, A. R. Srinivasan, J. P. Jokinen, A. Oulasvirta, G. Markkula, arXiv preprint arXiv:2301.11737, **2023**.
- [88] O. Giles, G. Markkula, J. Pekkanen, N. Yokota, N. Matsunaga, N. Merat, T. Daimon, in *41st Annu. Meeting Cogn. Sci. Society*, Cognitive Science Society, **2019**, pp. 366–372.
- [89] B. Baker, I. Kanitscheider, T. Markov, Y. Wu, G. Powell, B. McGrew, I. Mordatch, arXiv preprint arXiv:1909.07528, **2019**.
- [90] G. Markkula, R. Romano, R. Madigan, C. W. Fox, O. T. Giles, N. Merat, *Transp. Res. Rec.* **2018**, *2672*, 153.
- [91] W. Wu, R. Chen, H. Jia, Y. Li, Z. Liang, *Int. J. Mod. Phys. C* **2019**, *30*, 1950025.
- [92] F. Camara, S. Cosar, N. Bellotto, N. Merat, C. W. Fox, in *Human Factors in Intelligent Vehicles*, River Publishers, Denmark **2022**, pp. 1–20.
- [93] A. H. Kalantari, Y. Yang, N. Merat, G. Markkula, *IEEE Access* **2023**, *11*, 110707.
- [94] C. Zhang, B. Zhou, T. Z. Qiu, S. Liu, *J. Safety Res.* **2018**, *64*, 145.
- [95] M. A. Brewer, K. Fitzpatrick, J. A. Whitacre, D. Lord, *Transp. Res. Rec.* **2006**, *1982*, 132.
- [96] A. Dommès, *Accid. Anal. Prev.* **2019**, *128*, 175.
- [97] M. Sucha, D. Dostal, R. Risser, *Accid. Anal. Prev.* **2017**, *102*, 41.
- [98] X. Zhang, H. Chen, W. Yang, W. Jin, W. Zhu, *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 3025.
- [99] Y. Liu, R. Alsaleh, T. Sayed, *Transp. Res. F: Traffic Psychol. Behav.* **2021**, *81*, 115.
- [100] L. L. Thompson, F. P. Rivara, R. C. Ayyagari, B. E. Ebel, *Inj. Prev.* **2013**, *19*, 232.
- [101] M. Moussaid, N. Perozo, S. Garnier, D. Helbing, G. Theraulaz, *PLoS One* **2010**, *5*, e10047.
- [102] A. N. Tump, T. J. Pleskac, R. H. Kurvers, *Sci. Adv.* **2020**, *6*, eabb0266.
- [103] J. Domeyer, A. Dinparastdjadid, J. D. Lee, G. Douglas, A. Alsaïd, M. Price, *Transp. Res. Rec.* **2019**, *2673*, 70.
- [104] G. Markkula, Z. Uludag, R. M. Wilkie, J. Billington, *PLoS Comput. Biol.* **2021**, *17*, e1009096.
- [105] R. Lobjois, V. Cavallo, *Accid. Anal. Prev.* **2009**, *41*, 259.
- [106] R. Anders, F. Alario, L. Van Maanen, et al., *Psych. Methods* **2016**, *21*, 309.
- [107] K. Tian, G. Markkula, C. Wei, R. Romano, in *2022 IEEE 25th Int. Conf. Intell. Transp. Syst.*, IEEE, Piscataway, NJ **2022**, pp. 183–188.
- [108] W. Zeng, P. Chen, H. Nakamura, M. Iryo-Asano, *Transp. Res. C: Emerg. Technol.* **2014**, *40*, 143.
- [109] J. Montufar, J. Arango, M. Porter, S. Nakagawa, *Transp. Res. Rec.* **2007**, *2002*, 90.
- [110] A. Forde, J. Daniel, *J. Traffic Transp. Eng.* **2021**, *8*, 57.



- [111] M. Layegh, B. Mirbaha, A. A. Rassafi, *Physica A: Stat. Mech. Appl.* **2020**, 556, 124843.
- [112] L. Lu, G. Ren, W. Wang, C.-Y. Chan, J. Wang, *Accid. Anal. Prev.* **2016**, 95, 425.
- [113] D. Helbing, P. Molnar, *Phys. Rev. E* **1995**, 51, 4282.
- [114] D. Yang, Ü. Özgüner, K. Redmill, *ACM Trans. Spat. Algorithms Syst.* **2020**, 6, 1.
- [115] Y. Ma, E. W. M. Lee, R. K. K. Yuen, *IEEE Trans. Intell. Transp. Syst.* **2016**, 17, 3159.
- [116] X. Song, D. Han, J. Sun, Z. Zhang, *Physica A: Stat. Mech. Appl.* **2018**, 509, 827.
- [117] A. Kalatian, B. Farooq, *Transp. Res. C: Emerg. Technol.* **2022**, 134, 103453.
- [118] F. Martinez-Gil, M. Lozano, F. Fernández, *Simul. Model Pract. Theory* **2014**, 47, 259.
- [119] K. Li, S. Eiffert, M. Shan, F. Gomez-Donoso, S. Worrall, E. Nebot, in *2021 IEEE Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2021**, pp. 14241–14247.
- [120] B. Völz, K. Behrendt, H. Mielenz, I. Gilitschenski, R. Siegwart, J. Nieto, in *2016 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2016**, pp. 2607–2612.
- [121] M. Prédhumeau, L. Mancheva, J. Dugdale, A. Spalanzani, *J. Artif. Intell. Res.* **2022**, 73, 1385.
- [122] L. Crosato, H. P. Shum, E. S. Ho, C. Wei, *IEEE Trans. Intell. Veh.* **2022**, 8, 1339.
- [123] Y. Wang, H. Huang, B. Zhang, J. Wang, *IET Intell. Transp. Syst.* **2023**, 17, 454.
- [124] Y. Chen, S. Li, X. Tang, K. Yang, D. Cao, X. Lin, *IEEE Trans. Transp. Electr.* **2023**, 9, 4704.
- [125] M. R. Bachute, J. M. Subhedar, *Mach. Learn. App.* **2021**, 6, 100164.
- [126] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, S. Savarese, in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, **2016**, pp. 961–971.
- [127] Y. Xu, Z. Piao, S. Gao, in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, **2018**, pp. 5275–5284.
- [128] H. Bi, Z. Fang, T. Mao, Z. Wang, Z. Deng, in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, **2019**, pp. 10383–10392.
- [129] J. Amirian, J.-B. Hayet, J. Pettré, in *IEEE/CVF Conf. Comput. Vis. Pat. Recog. Workshops*, Long Beach, CA, USA **2019**.
- [130] T.-Y. Chen, H.-J. Lin, C.-S. Shih, K.-T. Kuo, Q. Liu, R. H. Chan, in *2021 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2021**, pp. 64–69.
- [131] Y. Yoon, K. Yi, in *2021 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2021**, pp. 196–203.
- [132] B. Rainbow, Q. Men, H. P. H. Shum, in *2021 IEEE Int. Conf. Syst. Man Cyber., SMC '21*, IEEE, Piscataway, NJ **2021**, pp. 2959–2966.
- [133] R. Li, S. Katsigiannis, H. P. H. Shum, in *2022 IEEE Int. Conf. Image Proc., ICIP '22*, IEEE, Piscataway, NJ **2022**, pp. 2346–2350.
- [134] L. Li, J. Gan, X. Ji, X. Qu, B. Ran, *IEEE trans. Intell. Transp. Syst.* **2020**, 23, 122.
- [135] L. Wang, L. Sun, M. Tomizuka, W. Zhan, *IEEE Robot. Autom. Lett.* **2021**, 6, 3421.
- [136] X. Zhao, Y. Tian, J. Sun, in *2021 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2021**, pp. 453–459.
- [137] G. M. J.-B. Chaslot, *Monte-Carlo Tree Search*, University Maastricht, Netherlands **2010**.
- [138] R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA **2018**.
- [139] L. Sun, W. Zhan, M. Tomizuka, in *2018 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2018**, pp. 2111–2117.
- [140] R. Chandra, A. Bera, D. Manocha, *IEEE Trans. Intell. Transp. Syst.* **2021**, 23, 2572.
- [141] D. Feng, A. Harakeh, S. L. Waslander, K. Dietmayer, *IEEE Trans. Intell. Transp. Syst.* **2021**, 23, 9961.
- [142] Z. Guo, Y. Huang, X. Hu, H. Wei, B. Zhao, *Electronics* **2021**, 10, 471.
- [143] S. Aradi, *IEEE Trans. Intell. Transp. Syst.* **2020**, 23, 740.
- [144] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, D. Hassabis, *Nature* **2015**, 518, 529.
- [145] Y. Hu, H. P. H. Shum, E. S. L. Ho, *IET Intell. Transp. Syst.* **2020**, 14, 1845.
- [146] L. Sun, C. Peng, W. Zhan, M. Tomizuka, in *Dynamic Systems and Control Conf. (DSCC)*, Vol. 3, American Society of Mechanical Engineers, New York, NY **2018**, p. V003T37A012.
- [147] P. Cai, H. Wang, Y. Sun, M. Liu, in *2021 IEEE/RSJ Int. Conf. Intell. Robots Syst.* IEEE, Piscataway, NJ **2020**, pp. 8979–8984.
- [148] E. Bronstein, M. Palatucci, D. Notz, B. White, A. Kuefler, Y. Lu, S. Paul, P. Nikdel, P. Mouglin, H. Chen, J. Fu, A. Abrams, P. Shah, E. Racah, B. Frenkel, S. Whiteson, D. Anguelov, H. Chen, in *2022 IEEE/RSJ Int. Conf. Intell. Robots Syst.* IEEE, Piscataway, NJ **2022**, pp. 8652–8659.
- [149] S. Ross, G. Gordon, D. Bagnell, in *Proc. Int. Conf. Artif. Intell. Stat. JMLR Workshop and Conf. Proc.*, **2011**, pp. 627–635.
- [150] Y. Zhu, D. Ren, D. Qian, M. Fan, X. Li, H. Xia, in *2021 IEEE Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2021**, pp. 3255–3261.
- [151] Z. Chang, G. A. Koulteris, H. P. H. Shum, arXiv preprint arXiv:2306.04542, **2023**.
- [152] Y. Lu, W. Wang, X. Hu, P. Xu, S. Zhou, M. Cai, *IEEE Trans. Intell. Transp. Syst.* **2022**, 24, 8452.
- [153] X. Li, X. Ying, M. C. Chuah, in *2019 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2019**, pp. 3960–3966.
- [154] S. Casas, C. Gulino, R. Liao, R. Urtasun, in *2020 IEEE Int. Conf. Robot. Autom.* IEEE, Piscataway, NJ **2020**, pp. 9491–9497.
- [155] X. Mo, Y. Xing, C. Lv, in *2021 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2021**, pp. 1934–1939.
- [156] S. Chen, J. Dong, P. Ha, Y. Li, S. Labi, *Comput.-Aided Civ. Infrastruct. Eng.* **2021**, 36, 838.
- [157] K. Jin, X. Han, arXiv preprint arXiv:2203.09952, **2022**.
- [158] P. Hart, A. Knoll, in *2020 IEEE Int. Veh. Symp.*, IEEE, Piscataway, NJ **2020**, pp. 1589–1594.
- [159] M. Hügle, G. Kalweit, M. Werling, J. Boedecker, in *2020 IEEE Int. Conf. Robot. Autom.* IEEE, Piscataway, NJ **2020**, pp. 4329–4335.
- [160] H. Su, J. Zhu, Y. Dong, B. Zhang, in *IJCAI*, Melbourne, Australia, Vol. 1, **2017**, p. 2.
- [161] S. Liu, K. Zheng, L. Zhao, P. Fan, *Comput. Comm.* **2020**, 157, 143.
- [162] H. Jin, C. Duan, Y. Liu, P. Lu, *IET Intell. Transp. Syst.* **2020**, 14, 401.
- [163] F. Garrido, P. Resende, *IEEE Access* **2022**, 10, 100348.
- [164] C.-J. Hoel, K. Wolff, L. Laine, in *2018 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2018**, pp. 2148–2155.
- [165] P. Palanisamy, in *2020 Int. Joint Conf. on Neural Networks (IJCNN)*, IEEE, Piscataway, NJ **2020**, pp. 1–7.
- [166] Y. Hu, A. Nakhaei, M. Tomizuka, K. Fujimura, in *2019 IEEE/RSJ Int. Conf. Intell. Robots Syst.*, IEEE, Piscataway, NJ **2019**, pp. 151–158.
- [167] N. Deshpande, A. Spalanzani, in *2019 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2019**, pp. 2081–2086.
- [168] D. M. Saxena, S. Bae, A. Nakhaei, K. Fujimura, M. Likhachev, in *2020 IEEE Int. Conf. Robot. Autom.* IEEE, Piscataway, NJ **2020**, pp. 5385–5392.
- [169] M. Yuan, J. Shan, K. Mi, *IEEE Robot. Autom. Lett.* **2021**, 7, 818.
- [170] M. Everett, Y. F. Chen, J. P. How, in *2018 IEEE/RSJ Int. Conf. Intell. Robots Syst.*, IEEE, Piscataway, NJ **2018**, pp. 3052–3059.

- [171] C. Chen, Y. Liu, S. Kreiss, A. Alahi, in *2019 Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2019**, pp. 6015–6022.
- [172] H. Chae, C. M. Kang, B. Kim, J. Kim, C. C. Chung, J. W. Choi, in *2017 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2017** pp. 1–6.
- [173] T. Schaul, J. Quan, I. Antonoglou, D. Silver, arXiv preprint arXiv:1511.05952, **2015**.
- [174] J. Li, L. Yao, X. Xu, B. Cheng, J. Ren, *Info. Sci.* **2020**, 532, 110.
- [175] N. Deshpande, D. Vaufreydaz, A. Spalanzani, in *2020 Int. Conf. Control Autom. Robot. Vis.*, IEEE, Piscataway, NJ **2020**, pp. 428–433.
- [176] L. Crosato, C. Wei, E. S. Ho, H. P. Shum, in *2021 IEEE Int. Conf. Human-Machine Syst.*, IEEE, Piscataway, NJ **2021** pp. 1–6.
- [177] H. Seong, C. Jung, S. Lee, D. H. Shim, in *2021 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2021**, pp. 559–566.
- [178] K. Zhou, Z. Liu, Y. Qiao, T. Xiang, C. C. Loy, *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, 24, 8452.
- [179] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, P. Abbeel, in *2017 IEEE/RSJ Int. Conf. Intell. Robots Syst.*, IEEE, Piscataway, NJ **2017**, pp. 23–30.
- [180] Y. Shan, W. F. Lu, C. M. Chew, *Neurocomputing* **2019**, 367, 31.
- [181] J. Chen, S. E. Li, M. Tomizuka, *IEEE Trans. Intell. Transp. Syst.* **2021**, 23, 5068.
- [182] M. Klimke, B. Völz, M. Buchholz, in *2022 IEEE Int. Veh. Symp.*, IEEE, Piscataway, NJ **2022**, pp. 167–174.
- [183] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, I. Mordatch, in *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA **2017**.
- [184] K. Zhang, Z. Yang, T. Basar, in *Handbook of Reinforcement Learning and Control* (Eds: K. G. Vamvoudakis, Y. Wan, F. L. Lewis, D. Cansever), Springer, Cham **2021**, pp. 321–384.
- [185] X. Chen, P. Chaudhari, in *2021 IEEE Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2021**, pp. 7980–7986.
- [186] Y. Luo, P. Cai, A. Bera, D. Hsu, W. S. Lee, D. Manocha, *IEEE Robot. Autom. Lett.* **2018**, 3, 3418.
- [187] Y.-C. Hsu, S. Gopalwamy, S. Saripalli, D. A. Shell, in *2020 IEEE/RSJ Int. Conf. Intell. Robots Syst.*, IEEE, Piscataway, NJ **2020** pp. 1984–1991.
- [188] C. Hubmann, N. Quetschlich, J. Schulz, J. Bernhard, D. Althoff, C. Stiller, in *2019 IEEE Int. Veh. Symp.*, IEEE, Piscataway, NJ **2019**, pp. 2172–2179.
- [189] M. Bouton, A. Cosgun, M. J. Kochenderfer, in *2017 IEEE Int. Veh. Symp.*, IEEE, Piscataway, NJ **2017**, pp. 825–830.
- [190] F. Pusse, M. Klusch, in *2019 IEEE Int. Veh. Symp.*, IEEE, Piscataway, NJ **2019** pp. 1013–1020.
- [191] K. B. Naveed, Z. Qiao, J. M. Dolan, in *2021 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2021**, pp. 601–606.
- [192] T. Basar, G. J. Olsder, *Dynamic Noncooperative Game Theory*, SIAM, **1998**.
- [193] D. Sadigh, N. Landolfi, S. S. Sastry, S. A. Seshia, A. D. Dragan, *Auton. Robots* **2018**, 42, 1405.
- [194] F. Facchinei, C. Kanzow, *Ann. Oper. Res.* **2010**, 175, 177.
- [195] S. L. Cleac'h, M. Schwager, Z. Manchester, arXiv preprint arXiv:1910.09713, **2019**.
- [196] B. Di, A. Lamperski, in *2019 IEEE Conf. Decis. Control.*, IEEE, Piscataway, NJ **2019**, pp. 4073–4078.
- [197] F. Facchinei, A. Fischer, V. Piccialli, *Math. Program.* **2009**, 117, 163.
- [198] R. Spica, E. Cristofalo, Z. Wang, E. Montijano, M. Schwager, *IEEE Trans. Robot.* **2020**, 36, 1389.
- [199] A. Liniger, J. Lygeros, *IEEE Trans. Control Syst. Technol.* **2019**, 28, 884.
- [200] P. Hang, C. Lv, C. Huang, J. Cai, Z. Hu, Y. Xing, *IEEE Trans. Veh. Technol.* **2020**, 69, 14458.
- [201] J. Yoo, R. Langari, arXiv preprint arXiv:2003.09786, **2020**.
- [202] Z. Wang, T. Taubner, M. Schwager, *Robot. Auton. Syst.* **2020**, 125, 103410.
- [203] F. Laine, D. Fridovich-Keil, C.-Y. Chiu, C. Tomlin, arXiv preprint arXiv:2101.02900, **2021**.
- [204] D. Fridovich-Keil, E. Ratner, L. Peters, A. D. Dragan, C. J. Tomlin, in *2020 IEEE Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2020**, pp. 1475–1481.
- [205] D. Fridovich-Keil, V. Rubies-Royo, C. J. Tomlin, in *2020 IEEE Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2020**, pp. 2216–2222.
- [206] W. Schwarting, A. Pierson, S. Karaman, D. Rus, *IEEE Trans. Robot.* **2021**, 37, 2157.
- [207] R. Tian, S. Li, N. Li, I. Kolmanovsky, A. Girard, Y. Yildiz, in *2018 IEEE Conf. Decis. Control.*, IEEE, Piscataway, NJ **2018**, pp. 321–326.
- [208] G. Ding, S. Aghli, C. Heckman, L. Chen, in *2018 IEEE/RSJ Int. Conf. Intell. Robots Syst.*, IEEE, Piscataway, NJ **2018**, pp. 3640–3647.
- [209] N. Li, I. Kolmanovsky, A. Girard, Y. Yildiz, in *2018 Annu. Amer. Control Conf.*, IEEE, Piscataway, NJ **2018**, pp. 3215–3220.
- [210] E. Stefansson, J. F. Fisac, D. Sadigh, S. S. Sastry, K. H. Johansson, in *2019 Euro. Control Conf.*, IEEE, Piscataway, NJ **2019**, pp. 3165–3172.
- [211] D. Isele, in *2019 IEEE Int. Intell. Transp. Syst. Conf.*, IEEE, Piscataway, NJ **2019**, pp. 3981–3986.
- [212] J. F. Fisac, E. Bronstein, E. Stefansson, D. Sadigh, S. S. Sastry, A. D. Dragan, in *2019 Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2019**, pp. 9590–9596.
- [213] M. Wang, Z. Wang, J. Talbot, J. C. Gerdes, M. Schwager, *IEEE Trans. Robot.* **2021**, 37, 1313.
- [214] G. Williams, B. Goldfain, P. Drews, J. M. Rehg, E. A. Theodorou, in *2018 IEEE Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2018**, pp. 2403–2410.
- [215] F. Laine, D. Fridovich-Keil, C.-Y. Chiu, C. Tomlin, in *2021 IEEE Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2021**, pp. 8016–8023.
- [216] P. Bazilinskyy, D. Dodou, J. De Winter, *Transp. Res. F: Traffic Psychol. Behav.* **2019**, 67, 175.
- [217] D. Dey, A. van Vastenoven, R. H. Cuijpers, M. Martens, B. Pflöging, in *13th AutomotiveUI*, **2021**, pp. 274–286.
- [218] W. Tabone, J. De Winter, C. Ackermann, J. Bärgrman, M. Baumann, S. Deb, C. Emmenegger, A. Habibovic, M. Hagenzieker, P. A. Hancock, R. Happee, J. Krems, J. D. Lee, M. Martens, N. Merat, D. Norman, T. B. Sheridan, N. A. Stanton, *Transp. Res. Interdiscip. Persp.* **2021**, 9, 100293.
- [219] A. Sripada, P. Bazilinskyy, J. de Winter, *Ergonomics* **2021**, 64, 1416.
- [220] A. Dietrich, P. Maruhn, L. Schwarze, K. Bengler, in *Human Systems Engineering and Design II: Proc. of the 2nd Int. Conf. on Human Systems Engineering and Design (IHSED2019): Future Trends and Applications, September 16-18, 2019, Universität der Bundeswehr München, Munich, Germany*. Springer, New York, **2020** pp. 176–181.
- [221] N. Deo, M. M. Trivedi, in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, IEEE, Piscataway, NJ **2018**, pp. 1468–1476.
- [222] C. Dong, J. M. Dolan, B. Litkouhi, in *2017 IEEE Int. Veh. Symp.*, IEEE, Piscataway, NJ **2017**, pp. 1584–1589.
- [223] J. Ahn, M. Kim, J. Park, *Sci. Rep.* **2022**, 12, 21285.
- [224] D. S. Pawar, G. R. Patil, *Safety Sci.* **2016**, 86, 295.
- [225] J. Li, L. Sun, J. Chen, M. Tomizuka, W. Zhan, in *2021 IEEE Int. Conf. Robot. Autom.*, IEEE, Piscataway, NJ **2021**, pp. 2660–2666.
- [226] M. Wang, S. P. Hoogendoorn, W. Daamen, B. van Arem, R. Happee, *Transp. Res. C: Emerg. Technol.* **2015**, 58, 73.
- [227] H. Yu, H. E. Tseng, R. Langari, *Transp. Res. C: Emerg. Technol.* **2018**, 88, 140.
- [228] M. Bahram, A. Lawitzky, J. Friedrichs, M. Aeberhard, D. Wollherr, *IEEE Trans. Veh. Technol.* **2015**, 65, 3981.



**Luca Crosato** received a bachelor's degree in mechanical engineering and a master's degree in robotics and automation engineering from the University of Pisa, Pisa, Italy, in 2018. He is currently working toward a Ph.D. degree with Northumbria University, Newcastle upon Tyne, UK. He is currently a Research Assistant at Queen's University Belfast, UK. His research interests include decision-making and control of AVs, pedestrian motion simulation, and machine learning.



**Kai Tian** received an M.Sc. degree in automotive engineering from Chongqing University, China, and a Ph.D. degree in transport Studies from the University of Leeds, UK. His main research interests include pedestrian-automated vehicle interaction, human factors and safety, and decision modeling.



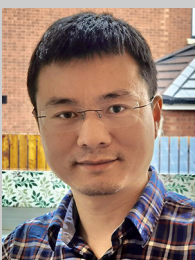
**Hubert P. H. Shum** (Senior Member, IEEE) is an Associate Professor in Computer Science at Durham University, UK. Before this, he worked as the Director of Research/Associate Professor/Senior Lecturer at Northumbria University, UK, and a Postdoctoral Researcher at RIKEN, Japan. He received his Ph.D. degree from the University of Edinburgh, UK. His research interests include computer vision, computer graphics, motion analysis, and machine learning.



**Edmond Shu-lim Ho** is currently a Senior Lecturer (Associate Professor) in the School of Computing Science (IDA-Section) at the University of Glasgow, Scotland, UK. Prior to joining the University of Glasgow in 2022, he was an Associate Professor in the Department of Computer and Information Sciences at Northumbria University, Newcastle upon Tyne, UK (2016–2022) and a Research Assistant Professor in the Department of Computer Science at Hong Kong Baptist University (2011–2016). He has been an Associate Editor of Computer Graphics Forum (CGF) since 2023. He received a BSc degree in Computer Science from the Hong Kong Baptist University, an MPhil degree from the City University of Hong Kong, and a Ph.D. degree from the University of Edinburgh.



**Yafei Wang** (Member, IEEE) received a B.S. degree in internal combustion engine from Jilin University, Changchun, China, in 2005, an M.S. degree in vehicle engineering from Shanghai Jiao Tong University, Shanghai, China, in 2008, and the Ph.D. degree in electrical engineering from The University of Tokyo, Tokyo, Japan, in 2013. From 2013 to 2016, he was a Post-Doctoral Researcher at The University of Tokyo. He is currently an Associate Professor at the School of Mechanical Engineering, Shanghai Jiao Tong University. His research interests include state estimation and control for connected and automated vehicles.



**Chongfeng Wei** received a B.S. degree in computational and applied mathematics and an M.S. degree in vehicle engineering from Southwest Jiaotong University, Chengdu, China, in 2009 and 2011, respectively, and a Ph.D. degree in mechanical engineering from the University of Birmingham, Birmingham, UK, in 2015. He is now a lecturer (assistant professor) at Queen's University Belfast, UK. His current research interests include decision-making and control of intelligent vehicles, human-centric autonomous driving, cooperative automation, and dynamics and control of mechanical systems. He is also serving as an Associate Editor of IEEE TITS, IEEE TIV, IEEE TVT, and Frontier on Robotics and AI.