



HAL
open science

An Overview of Decision-Making in Autonomous Vehicles

Dany Ghraizi, Reine Talj, Clovis Francis

► **To cite this version:**

Dany Ghraizi, Reine Talj, Clovis Francis. An Overview of Decision-Making in Autonomous Vehicles. 22nd International Federation of Automatic Control World Congress (IFAC WC 2023), Jul 2023, Yokohama, Japan. pp.10971-10983, 10.1016/j.ifacol.2023.10.793 . hal-04307145

HAL Id: hal-04307145

<https://hal.science/hal-04307145>

Submitted on 25 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Overview of Decision-Making in Autonomous Vehicles^{*}

Dany Ghraizi^{*} Reine Talj^{*} Clovis Francis^{**}

^{*} Sorbonne université, Université de technologie de Compiègne, CNRS, Heudiasyc UMR 7253, CS 60 319, 60 203 Compiègne, France
(name.surname@hds.utc.fr)

^{**} Arts et Métiers Paris Tech, 1, Rue Saint Dominique, 51000 Châlons en Champagne, France, (clovis.francis@ensam.eu)

Abstract: In this review we have surveyed the literature on the autonomous vehicle focusing on the decision-making architecture. The main contribution of this paper is to provide an up-to-date reference that tries to cover as much as possible this broad topic rapidly. Therefore, after analyzing the body of literature, we were able to present a detailed overview of the usual anatomy of the decision-making system, review the coverage of public reviews and surveys available around the topic, present the environment and vehicle representation, and showcase the literature through: trajectory planning, risk-uncertainty assessment, and mimicking human behaviour mainly using perception-based methods. We also classified the literature according to scenarios covered, action-space used, and whether or not they emphasize the time horizon or frequency or time-step.

Keywords: Intelligent Autonomous Vehicles; Decision Making for Autonomous Driving in Real-Time; Risk Assessment; Trajectory Planning; Deep Learning

1 Introduction

Autonomous vehicles (AV) is a field of study and development of great interest for automotive manufacturers and academics. The rate at which research and development is being done by industry and academia alike is very high for this domain. Several competitions have been organized around the AV, during the last decades, such as the DARPA Challenges in the USA, the Korean competitions, the European GCDC competitions, Bosch Future Mobility Challenge (BFMC) etc.

The vehicles need to build, monitor, and synthesize an internal representation of the environment around them then use this representation to formulate decisions and interact in an appropriate manner that respects the safety of the driver, avoid collisions with vehicles and pedestrians, while also respecting traffic rules. The performance of an autonomous vehicle is assessed based on 3 modes of operation consisting of 5 levels. The first mode is called Assisted mode containing levels 1 and 2, the second mode is called Automated mode containing only level 3, and the third mode is called Autonomous mode containing levels 4 and 5. In these levels the ability of the car to handle different driving conditions, the amount of attention needed from the human driver side, and the safety issues are described.

The work will focus on finding the optimal decision architecture considering the different constraints and challenges inherent in this problem. This allows the AV to transition from the Assisted mode (level 2) to the Automated mode

(levels 3) and the Autonomous mode (level 4). Therefore, it is necessary to assess the methods, with respect to several criterion: Learning-based or Classical-based, the ability to tackle the uncertainty of other vehicles that is affecting both the environment and the autonomous agent's state drastically, their ability to handle the vast number of scenarios that in turn pose additional exploration on the agent to find the safest and best decision (roundabouts and intersections, weather conditions, driving behaviors, etc.)

The goal of this paper is to provide researchers with an up-to-date overview for reference in their research within this broad field. Therefore, this study addresses the literature in Classical- and Learning- based approaches through environment and vehicle representation, and available surveys and reviews. It showcases them across end-to-end and modular designs, trajectory planning, risk-uncertainty assessment, and mimicking Human Behaviour through mainly perception-based methods. It also classifies the literature through scenarios, action space, and lists the time-horizon, timestep, and frequency rate when emphasized.

2 Survey Literature Coverage on Decision Making

Public assessments and surveys on autonomous vehicle decision-making have been made available. The two major types of research techniques are learning-based approaches and traditional methodologies. A complete study (Liu et al. [2021]) is offered, focusing especially on learning-based decision-making strategies that have recently evolved, presenting the different methods within both classical and learning-based approaches alongside

^{*} Sorbonne université, Université de technologie de Compiègne, CNRS, Heudiasyc UMR 7253, CS 60 319, 60 203 Compiègne, France

their pros and cons, as well as an overview of applications in autonomous cars that are presently in operation.

Table 1. Decision Making System in Modular Approach (Badue et al. [2021])

Subsystems	Methods			
Route Planning	Goal Directed	Separator	Hierarchical	Bounded-Hop
Path Planning	Graph Search	Interpolating-Curve		
Behaviour Selector	FSM	Ontology	Markov Decision Process	
Motion Planning	Graph Search	Sampling based	Interpolating Curve based	Numerical Optimization
Obstacle Avoidance	Adjust Speed			
Control	Direct Hardware Actuation	Path Tracking		

Another main aspect is exploring the two main architectures of the system: End-to-End, and Modular. One modular design is offered in (Badue et al. [2021]), where the survey’s authors present a full and in-depth examination of the whole architecture of the pipeline for autonomous cars: The Perception System and the Decision-Making System. Each of which is made up of a number of modules which serve as the modular architecture’s base. As shown in Table .1, the latter is usually separated into a large number of subsystems that are in charge of responsibilities such as route planning, path planning, behavior selection, motion planning, and control. For a more detailed view on the decision-making partition see Fig. 1.

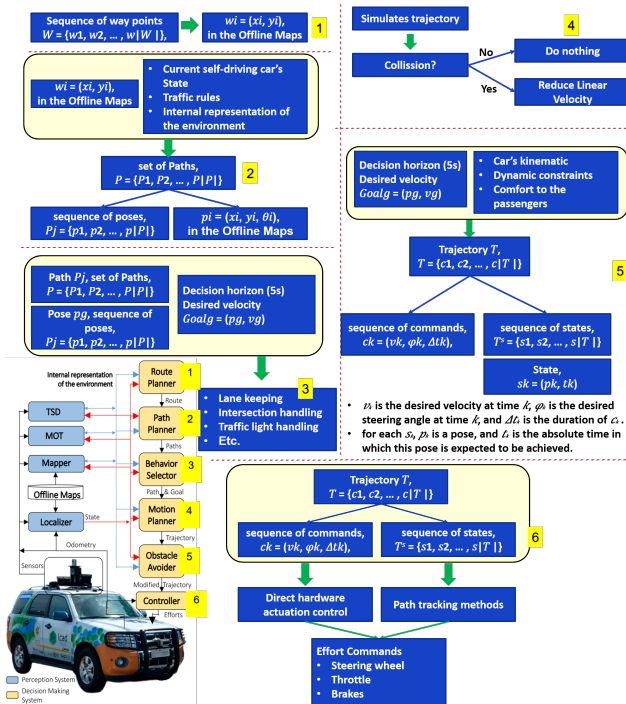


Fig. 1. Illustration of Decision-Making System Architecture Inputs and Outputs of modules presented in (Badue et al. [2021])

In their overview, the Intelligent Autonomous Robotics Automobile (IARA), a self-driving vehicle built by the Universidade Federal do Espirito Santo (UFES), is also

fully discussed. IARA, the country’s (Brazil) first autonomous automobile, traveled 74 kilometers on highways and city streets in the evening of May 12, 2017. They end by outlining important self-driving car research platforms produced by universities and technology firms that have received media attention. Among the most notable are: 1) Baidu, one of China’s largest technological corporations, which is working on the Apollo open source self-driving car project. The project’s source code is accessible on GitHub. 2) Navya and Transdev which are both French companies that are developing self-driving buses. Several of Navya’s buses are already being tested in Australia, Asia, and Europe, and Transdev is holding a public demonstration of its self-driving buses.

Similarly in (Sharma et al. [2021]), the goal is to provide a thorough, exhaustive, and critical analysis but mostly on the available motion and behaviour planning methodologies for AVs in terms of their viability, capacity to deal with dynamic restrictions and impediments, and comfort-maximizing motion. The benefits, capability in handling static and dynamic obstacles, vehicle limits, and limitations in operating situations of the existing behaviour planning systems have also been highlighted.

The End-to-End approach directly maps sensory inputs to steering and acceleration commands as a single machine learning task. A major advantage of such architecture is its simplicity and not having to incorporate expert designers into the system. A major disadvantage is the black-box nature of this method where it is difficult or impossible to figure out why the model misbehaves. On the other hand, the Modular approach, as discussed in the previous surveys, splits this direct mapping into separate modules, each its own entity that contributes to the entire performance of the system Fig. 1.

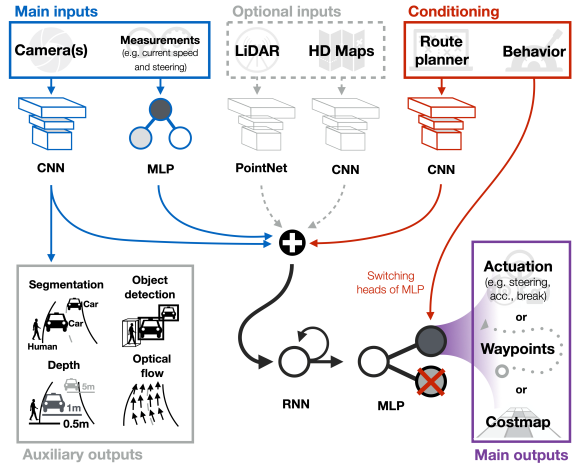


Fig. 2. Candidate End-to-End Architecture (Tampuu et al. [2022])

Thus, End-to-end driving literature on learning strategies, input and output modalities, network architecture, assessment approaches, and evaluation methods are presented in (Tampuu et al. [2022]). Safety and interpretability, two concerns that remain a high priority for modern technology, are also explored. The findings include a detailed examination of the practices already in use and an architecture that incorporates the most promising features

of fully autonomous driving systems. Actually, their approach is a compromise between the modular and end-to-end paradigms Fig. 2.

After outlining the pipeline’s parts, despite being organized around its fundamental components — environment modeling, modeling abstractions, state and perceptual model descriptions, appropriate rewarding, and realization of the underlying neural network — this study discusses DRL theories (Aradi [2022]). On this note, the authors of (Kendall et al. [2019]) showcase the first full-sized autonomous car to use deep reinforcement learning for autonomous driving in order to move away from the need of pre-set logical rules, mapping, and direct supervision. Using solely on-board processing, a continuous deep reinforcement learning system teaches the agent how to operate a real-world autonomous car in a few episodes, where it is able to complete training for lane following in less than 30 minutes, controlling speed and steering angle. The design of the algorithm is a straightforward state machine where the safety driver manages the various duties. The four tasks are defined as train, test, undo, and done. Although this reward function is universal, and have few drawbacks. It excludes conditioning on a predetermined navigational goal. It is quite sparse as well. Interventions will happen much less frequently as the agent gets better, which will lead to weaker training progress.

Furthermore, in (Talpaert. et al. [2019]), they go into further detail on the DRL algorithms, DRL applications for AV, and the components of the AV system. They describe the major issues that must be addressed before DRL may be widely used in AV applications. One study argues that a reliable strategy for navigating un-signalized junctions could be developed by examining the effectiveness of deep reinforcement learning techniques with a focus on responding to inadequate knowledge (Isele et al. [2018]). As a result, they examine several DRL network topologies and present two that they consider successful. The ideal system determines when to start accelerating. As a means of enhancing safety for autonomous vehicles, their second contribution investigates exploratory behaviors, especially creeping propensities. Utilizing the networks Deep Q-Network (DQN), DuelingDQN, and Prioritized Reply DQN (PRDQN), it is possible for them to learn rules that beat frequently-applied heuristic approaches in a number of areas, including task completion time and goal success rate, however they have a narrow range of generalizability.

On the other hand, (Zhu and Zhao [2022]) perform a comprehensive analysis of not only DRL, but also Imitation Learning (IL), or Autonomous Driving Policy Learning. Given that the performance of the decision-making system is greatly influenced by the perception component of the AV, (Yurtsever et al. [2020]) highlights unresolved issues and examines the technological side of autonomous driving. An overview of the coverage of different aspects in the AV with regard to decision making discussed in these surveys is presented in (Table. 2). We can see from this table that the most prevalent topics are Uncertainty-Risk Assessment, Modular and End-to-End designs, Reinforcement Learning and Classical approaches.

Before going into detail about the literature, we present two tables that summarize the action space used (see Table

.3) and the coverage of scenarios (see Table .4). As well as whether or not the time horizon, frequency, and time-step information is provided (see Table .5). We can see that the most used High-level actions are Turn Left-Right, and Lane Change-Keep. While the most used Low-level actions are Acceleration-Deceleration and Steering Angle. In the scenario classification of the literature, we can conclude that the literature consists of Urban (Cross Intersection, Roundabout, Merging) and Lane Change in Highway driving. However, in (Li et al. [2021]), they focus on three common situations that encompass intersectional straight-crossing route crashes as well as rear-end, sideswipe, and angle collisions. Finally, it is clear from the last table that not all papers provide us with information on their planning horizon etc.

3 Environment and Vehicle Representation

The representation of the environment and the ego vehicle impacts the ability to assess the reliability and accuracy of the models and scenarios. Having a simplistic environment does not take into consideration many of the constraints, dynamics, and real-world challenges that autonomous vehicles need to handle. On the other hand, having a complex design is also not beneficial because it is more difficult to scale and interpret, and it is computationally inefficient. Therefore, striking a balance between realism and feasibility is quite important.

3.1 Traffic Participants Representation

One of the simplest and most popular approaches is to represent the environment using cartesian coordinates and velocity. The ego vehicle can see every car within 200 meters of the sensor’s range where the locations, speeds, and orientation of the vehicles traffic scenario relative to the ego vehicle projected onto the center-line of the road define the system’s condition (Speidel et al. [2020], Li et al. [2022], Hoel et al. [2020]).

An occupancy grid provides the agent with richer information about the static and dynamic obstacles in its surroundings through a discretized grid in Cartesian coordinates which is created from a bird’s-eye perspective of space (Isele et al. [2018], Qiao et al. [2018]). Every vehicle in the area is represented by its heading angle, velocity, and an indicator term that is either 1 or 0 depending on whether the vehicle is occupying a spot in the grid. The generated/simulated rays from the ego vehicle are related to the geometric data where a square grid with a fixed size surrounding the vehicle is built at each time step, with the center of the square representing the central front of the ego vehicle. The square is then partitioned into small squares. In conclusion, the agent determines the distance and the velocity at each ray trace’s termination (intersection) point, yielding an input observation vector.

Latent space representation is when data attributes are learned, and representations are simplified, in order to detect patterns. Mostly the information necessary to properly depict the original observation input is included in the latent space representation of the data (Chen et al. [2022]), utilizing past data such as position, velocity, the intents of other road users, drivable zones, and road markers

Table 2. Survey Coverage of Decision Making

Literature	Liu et al. [2021]	Talpaert et al. [2019]	Badue et al. [2021]	Tampuu et al. [2022]	Aradi [2022]	Zhu and Zhao [2022]	Yurtsever et al. [2020]	Sharma et al. [2021]
Year (20 - -)	21	19	21	21	20	21	20	21
Classical Methods	X		X				X	X
Learning Based	X	X		X	X	X		
Reinforcement Learning				X				
Imitation Learning						X		
Components								
Global Planner								
Route and Path Planner								
Behavior Selector								
Motion Planner			X			X	X	X
Obstacle Avoider								
Control								
Local Planner								
End-to-End		X		X		X	X	
Modular			X	X		X	X	X
Modes of Architecture						X	X	
Vehicle Modelling			X			X	X	
Action Space						X	X	
Scenario Classification						X	X	
Uncertainty-Risk Assessment			X	X		X	X	X
Candidate Architecture				X		X	X	

Table 3. Action Space based Classification of Literature

Decision	Literature
High Level	
Turn Left/Right	(Li et al. [2022]), (Mitchell et al. [2020]), (Hoel et al. [2020]), (Duan [2020]), (Chen et al. [2019]), (Chen et al. [2021]), (Hang et al. [2021]), (Naveed et al. [2021])
Lane Change-Keep	(Speidel et al. [2020]), (Li et al. [2022]), (Mitchell et al. [2020]), (Hoel et al. [2020]), (Duan [2020]), (Chen et al. [2019]), (Chen et al. [2021]), (Hang et al. [2021]), (Naveed et al. [2021])
Emergency Braking	(Speidel et al. [2020]), (Hoel et al. [2020])
Car Follow	(Speidel et al. [2020])
Wait	(Isele et al. [2018]), (Naveed et al. [2021])
Low Level	
Accelerate-Deceleration	(Hubmann et al. [2017]), (Speidel et al. [2020]), (Isele et al. [2018]), (Hoel et al. [2020]), (Bernhard et al. [2019]), (Emuna et al. [2020]), (Duan [2020]), (Chen et al. [2021]), (Li et al. [2021]), (Sharma and Sharma [2021]), (Wang et al. [2021]), (Hecker et al. [2020]), (Qiao et al. [2018]), (Wang et al. [2022]), (Duhautbout et al. [2021, 2022]), (Karanam et al. [2022]), (Said et al. [2021]), (Said. et al. [2022])
Maintain Velocity	(Isele et al. [2018]), (Mitchell et al. [2020]), (Duhautbout et al. [2021, 2022]), (Karanam et al. [2022]), (Said et al. [2021]), (Said. et al. [2022])
Steering Angle Change	(Emuna et al. [2020]), (Chen et al. [2021]), (Sharma and Sharma [2021]), (Hecker et al. [2020]), (Wang et al. [2022]), (Duhautbout et al. [2021, 2022]), (Karanam et al. [2022]), (Said et al. [2021]), (Said. et al. [2022])

Table 4. Scenario based Classification of Literature

Scenario	Literature
Urban Driving	
Cross Intersection	
Signalized	(Ziegler et al. [2014]), (Bansal et al. [2018]), (Chen and Krähenbühl [2022]), (Chen et al. [2019, 2021]), (Chen et al. [2022]), (Sharma and Sharma [2021]), (Hecker et al. [2020]), (Qiao et al. [2018]), (Wang et al. [2022])
Unsignalized	(Hubmann et al. [2017]), (Speidel et al. [2020]), (Isele et al. [2018]), (Bernhard et al. [2019]), (Ziegler et al. [2014]), (Bansal et al. [2018]), (Chen and Krähenbühl [2022]), (Chen et al. [2019, 2021]), (Chen et al. [2022]), (Sharma and Sharma [2021]), (Wang et al. [2021]), (Duhautbout et al. [2021, 2022]), (Karanam et al. [2022])
Roundabout	(Ziegler et al. [2014]), (Bansal et al. [2018]), (Chen and Krähenbühl [2022]), (Chen et al. [2019, 2021]), (Chen et al. [2022]), (Sharma and Sharma [2021]), (Duhautbout et al. [2021, 2022]), (Karanam et al. [2022])
Merging	(Ziegler et al. [2014]), (Bansal et al. [2018]), (Chen and Krähenbühl [2022]), (Chen et al. [2019, 2021]), (Hang et al. [2021]), (Chen et al. [2022]), (Sharma and Sharma [2021]), (Hecker et al. [2020])
Highway Driving	
Lane Change	
One way	(Li et al. [2022]), (Mitchell et al. [2020]), (Hoel et al. [2020]), (Ziegler et al. [2014]), (Emuna et al. [2020]), (Duan [2020]), (Bansal et al. [2018]), (Chen and Krähenbühl [2022]), (Chen et al. [2019, 2021]), (Hang et al. [2021]), (Chen et al. [2022]), (Sharma and Sharma [2021]), (Hecker et al. [2020]), (Naveed et al. [2021]), (Duhautbout et al. [2021, 2022]), (Karanam et al. [2022]), (Said et al. [2021]), (Said. et al. [2022])
Multi way	(Ziegler et al. [2014]), (Bansal et al. [2018]), (Chen and Krähenbühl [2022]), (Chen et al. [2019, 2021]), (Chen et al. [2022]), (Sharma and Sharma [2021]), (Hecker et al. [2020]), (Duhautbout et al. [2021, 2022]), (Karanam et al. [2022]), (Said et al. [2021]), (Said. et al. [2022])
Lane Following	(Kendall et al. [2019])
With Occlusions	(Isele et al. [2018]), (Ziegler et al. [2014]), (Bansal et al. [2018]), (Karanam et al. [2022])
With Other Vehicles	(Hubmann et al. [2017]), (Speidel et al. [2020]), (Li et al. [2022]), (Isele et al. [2018]), (Mitchell et al. [2020]), (Hoel et al. [2020]), (Bernhard et al. [2019]), (Ziegler et al. [2014]), (Duan [2020]), (Bansal et al. [2018]), (Chen and Krähenbühl [2022]), (Chen et al. [2019, 2021]), (Hang et al. [2021]), (Chen et al. [2022]), (Sharma and Sharma [2021]), (Wang et al. [2021]), (Hecker et al. [2020]), (Qiao et al. [2018]), (Wang et al. [2022]), (Duhautbout et al. [2021, 2022]), (Karanam et al. [2022]), (Said et al. [2021]), (Said. et al. [2022])
With Pedestrians	(Ziegler et al. [2014]), (Bansal et al. [2018]), (Chen and Krähenbühl [2022]), (Chen et al. [2019, 2021]), (Sharma and Sharma [2021]), (Hecker et al. [2020]), (Duhautbout et al. [2022])
Real-World Implementation	(Ziegler et al. [2014]), (Bansal et al. [2018]), (Kendall et al. [2019])

etc. The autoencoder, a neural network that serves as an identity function, is one popular deep learning model that modifies the 'closeness' of data in the latent space. After compressing and decompressing the input, the system's speed and performance are enhanced (Iroegbu and Mad-

havi [2021]). Another element is the mask where the main goal is to allow the system to convey comprehensible semantic interpretations of its environment in order to make it understandable. This semantic mask offers a bird's-eye view of the road conditions and objects. Furthermore, low

Table 5. Planning Step-size, Frequency, and Horizon

Literature	Planning Step-Size (s)	Re-planning Frequency (Hz)	Planning Horizon (s)	Notes
Hubmann et al. [2017]	1	1	12 for 1 vehicle and 7.5 for 9 vehicles	Adaptive Belief Tree (ABT) has 11000 particles
Speidel et al. [2020]		0.2		
Isele et al. [2018]	0.2		Max. number of steps per trial/episode = 100 steps = 20s	
Hoel et al. [2020]	1		Max. number of steps per trial/episode = 100 steps	
Bernhard et al. [2019]	0.2			
Duan [2020]		40		Max. Steps = 5000
Kendall et al. [2019]		10 for RL and 100 for controller		
Chen et al. [2021]		H = 5 and in ProcGen H = 30	Where rewards are discretized from t + H - 1 at time t	
Lee et al. [2017]		4	For past trajectory observations t = 2s	
Chen et al. [2022]	0.1		4	Episode Length = 500, action is fixed for every 4 environment steps due to usage of frame skip of 4 (for temporal extension)
Wang et al. [2021]		3, 4, 5, 6 for vehicle passing straight, and 3, 4, 5 for vehicle steering/turning.	80 Trajectories in each.	
(Duhautbout et al. [2021, 2022], Karanam et al. [2022])	0.2			
Said et al. [2021]		The vehicle block (model and controller): 50Hz, and the rest of the system: 5Hz.		
Said. et al. [2022]		The two last modules: 50Hz while the others: 10 Hz.		

dimensional intermediate representations of the environment called "Affordances" (Sharma and Sharma [2021]) have the same effects. This representation is especially used in Deep Reinforcement Learning (DRL) approaches.

There can be many indicators depicting driving conditions (Duan [2020]). The definition and size of the state space can also be dependent sometimes on the performance of the agent (Bernhard et al. [2019]) where the results of the authors' benchmarking of the various observation spaces reveal that the relative state information combined with the ego-vehicle state outperformed the other representations (i.e. cartesian coordinates, distance, orientation, velocity and Time To Collision (TTC)). In (Hubmann et al. [2017]), Frenet-Serret formulae (Erkan and Yüce [2018]) are used to define the locations of the cars along the route of the vehicle at a specific location. In (Emuna et al. [2020]) two observations (apart from the range sensor measurements) constitute the state-space, which has 310 dimensions, in order to infer temporal information.

A privileged AV agent can have direct access to the true state values of the environment through the simulator where it is assumed that the ego vehicle has perfect knowledge in specific areas (Li et al. [2021], Wang et al. [2022, 2021]). Using the driving logs, (Chen et al. [2021]) creates a forward model of the environment that forecasts the next state of the ego-vehicle based on the current ego-vehicle state (compact 2D position, orientation, and speed).

3.2 Traffic Participants Motion Behaviour

In order to define the mobility behavior of traffic participants, Toledo's car-following model (Toledo et al. [2007]) can be used to control the movement of other automobiles. The Intelligent Driver Model (IDM) (Treiber et al. [2000]), used in (Isele et al. [2018]), incorporates randomization

by manipulating the speed distribution, human driving styles, and driver imperfection control settings (based on the Krauss stochastic driving model (Krauss [1998])) used in (Bernhard et al. [2019]). The IDM also serves as the foundation for the longitudinal component if a hard-coded lateral and longitudinal control method for realistic (virtual) background traffic is employed (Mitchell et al. [2020]). The standard SUMO driver model (Lopez et al. [2018]), which includes a lane change model that makes strategic decisions to pass slower moving cars and an adaptive cruise controller for the longitudinal motion, can be used to operate the other vehicles (Hoel et al. [2020]). Simulated automobiles can also be generated using the SUMO simulator (Qiao et al. [2018]), and they are all operated in line with the Krauss Car Following Traffic Model.

3.3 Uncertainty of the Environment and Decision

Uncertainty of the surroundings are numerous and the number of possible scenarios that can be formulated is infinite. There is no possible method that allows the ego vehicle to consider the full uncertainties as of yet. Some examples are the intentions of other vehicles and pedestrians, occlusions, and perception uncertainties such as weather conditions. The car is required to make a safe decision and to adapt accordingly to the changes of the environment, this is done by assessing and predicting uncertainties within its environment. Assumptions and simplifications (random speeds and behaviours) and Gaussian noise, used to account for longitudinal uncertainty initialized with unknown probabilities for their trajectories and motions (Hubmann et al. [2017]), can be made; however, this is far from representing the challenges the vehicle will face in real life (Hang et al. [2021], Chen et al. [2022]).

Bayes theory is a popular probabilistic approach when it comes to predicting uncertainties. A stochastic observation

model that accounts for uncertainty based on the Bayes Rule (Hubmann et al. [2017], Speidel et al. [2020], Li et al. [2022]) is used, and takes into account probable future observations of the nearby cars, a 2-dimensional feature vector with a Naive Bayes classifier (velocity and position based). Another probabilistic approach appears in (Emuna et al. [2020]) where the travelled Euclidean distance (arc-length), which serves as the independent variable, is modelled using Gaussian processes (GPs) to create the associated uncertainty distributions.

Game theory can also be used to represent adversarial uncertainty. The learning agent learns a risk-neutral, ideal strategy for the provided Stochastic Bayesian Game (SBG) (Harsanyi [1967, 1968a,b]) by seeing a large number of episodes with various behavior patterns (Bernhard et al. [2019]). Once the state-action distribution has been learned, the distribution shows the inherent uncertainty regarding the precise behavior types that will manifest at a given episode.

One particular type of uncertainty complicating the driving decisions and increasing risk significantly is occlusions. The representation is expanded to include an indication to determine if a certain zone in the environment (such as a cell in the occupancy grid) is occluded and the car's x and y offset to the pixel border to mitigate the effects of discretization (Isele et al. [2018]). In (Karanam et al. [2022]) the occluded zone is represented as a virtual agent. In (Wang et al. [2022]) a threat model of possible danger to neighbouring cars in the occluded region is developed inspired by the potential field technique (Koren and Borenstein [1991]).

4 Trajectory Planning: Path and Speed Planning

4.1 Classical Approaches

After defining the simulation environment representation of both the AV and other participants, the main goal for the AV is to reach its destination. This is achieved through generating a path for the vehicle to follow with a certain speed without any major problems (skidding, collisions etc.). This is done through a trajectory planning method that computes a path based on a collection of predetermined way-points that describe a global map (Said et al. [2021]). The specified way-points serve as the fundamental reference frame for a curvilinear coordinate system, which is used to produce candidate pathways that begin with a transient phase and end with a curve parallel to the road. Each potential path, associated with a desired velocity profile, is weighted against numerous parameters, including passenger comfort, static and dynamic obstacle avoidance, and overall trajectory tracking, using a cost function. Many different multi-way lane scenarios were evaluated, and the findings reveal that the vehicle behaves well in somewhat challenging driving circumstances. The same authors continued this work, with the main goal of studying the algorithm's sensitivity to parameter tuning and determining a generic range of weighting coefficients for the planning algorithm's cost function in order to make the algorithm as reliable as possible under various driving conditions (Said. et al. [2022]). They can then infer that the planning algorithm is resistant to changes

in cost weighting, which is a key benefit for dealing with a variety of driving scenarios and conditions in a dynamic environment without having to re-adjust and tune the planning algorithm.

To investigate the various action spaces that might be employed, which correspond to lateral and longitudinal planning (path and velocity profile planning), dividing the model into high-level and low-level actions provides us with better interpretation and tuning, inspired by how humans are able to think (high-level decisions) unrelated to what they physically do (low-level actions) See Fig. 3. (Duhautbout et al. [2021]) present a fully algorithmic technique, based on a geometrical description of the environment, for computing predictive speed profiles on different pathways. The results reveal that the vehicle reacts well, comfortably, and safely to its static and dynamic environments, with processing speeds that are suitable with real-time control. The same authors went on to improve the algorithmic speed planning approach for an autonomous vehicle coping with moving impediments, proposing to pass before or after while maintaining safety distances (Duhautbout et al. [2022]). The simulation findings suggest that the speed profiles developed are more efficient than the prior technique and less conservative in limited scenarios. The system in (Ziegler et al. [2014]) has two layers. High level: a) Behaviour Generation uses a (Harel) state chart notation to concisely describe constraints that are dependent on the current driving corridor, static barriers, dynamic objects, yield and merging laws, and more. b) Low level: The trajectory planner computes an ideal trajectory that minimizes the integral under the typical type of non-linear inequality constraints. An actual autonomous car by Mercedes named Bertha tested this technique and traveled 104 kilometers.

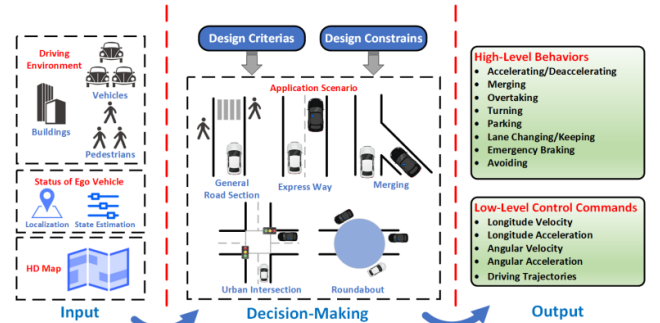


Fig. 3. Design framework of decision-making system (Liu et al. [2021])

4.2 Learning Based Approaches: Hierarchical Reinforcement Learning

Since designing safe routes in constantly changing and unexpected situations is required for autonomous driving, the problem is incredibly difficult. Current sampling-based methods, including Rapidly Exploring Random Trees (RRTs), are not the optimum answer for this problem because of their high processing cost. The universality and safety guarantees of imitation learning and other supervised learning approaches are lacking. To address these challenges and provide a trustworthy framework (Naveed

et al. [2021]) propose a Hierarchical Reinforcement Learning (HRL) structure paired with a Proportional-Integral-Derivative (PID) controller for trajectory planning. They use the PID for tracking waypoints selected by the low-level planner instead of using throttle, steer, and brake as outputs. With the use of HRL, the autonomous driving job may be divided into smaller tasks, and it also encourages the network to learn rules for both high-level and low-level trajectory planner options. The high-fidelity CARLA simulator’s findings show that the suggested strategy speeds up convergence, produces smoother trajectories, and can handle noisy observations and dynamic environments.

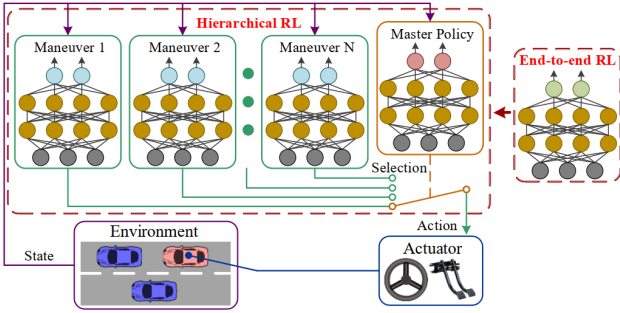


Fig. 4. Hierarchical RL for self-driving decision-making (Emuna et al. [2020])

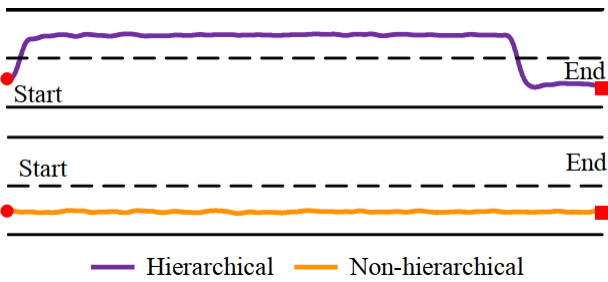


Fig. 5. Driving Trajectory Performance Comparison of RL and HRL (Emuna et al. [2020])

Furthermore, (Emuna et al. [2020]) describes a HRL strategy Fig. 4 that self-driving cars can utilize in place of a vast amount of labeled driving data to make judgments. It considers high-level maneuver selection as well as low-level motion control in both lateral and longitudinal orientations. Each move generates a new state space and reward function. Each maneuver in this study includes a steer policy network (SP-Net and SV-Net) and an accelerate policy network (AP-Net and AV-Net), which govern the lateral and longitudinal motions, respectively. This suggests that self-driving cars can make safe and efficient judgments. Compared to the non-hierarchical technique, each simulation’s driving time is reduced by around 25 percent. As seen in Fig. 5, the HRL’s performance is also less cautious.

5 Risk-Uncertainty Assessment

Trajectory planning can be lacking proper risk and uncertainty representation which highly influences the trajectory of the vehicle. Therefore addressing this aspect, the main points will be predicting other participants’ trajectories, quantifying the level of risk facing the AV, and

handling occlusions safely and efficiently, while also trying to consider different driving styles or behaviours.

Primarily concerned with detecting other participants’ trajectories accurately, incorporating probabilistic methods, Gaussian models are used to take safety and efficiency into account when cars intersect (Wang et al. [2021]). In order to precisely measure the probability of a vehicle accident from all angles, they use EPET (Estimating Post-Encroachment Time), which indicates the time difference between automobiles passing through the conflict zone’s center. A multiobjective optimization problem (MOP) is used as a model in order to resolve conflicts between cars interacting at crossings based on conflict resolution theory. Two intersection multi-objective optimization problem solutions are shown, Deep Deterministic Policy Gradient (DDPG) and Non-dominated Sorting Genetic Algorithm (NSGA-II). Results showed DDPG is more dependable and efficient than NSGA-II at solving the MOP model, which offers a theoretical framework for a thorough investigation of decision-making in a challenging intersecting context. Considering driving styles, (Li et al. [2021]) use Conditional Random Field (CRF) in conjunction with Gaussian models to study three common situations encompassing intersectional straight-crossing route crashes as well as rear-end, sideswipe, and angle collisions Fig. 6. The safety metrics used are: Time-to-Collision (TTC), Time-to-Region (TTR), Time-to-Escape (TTE), and Time-to-Stop (TTS). They also address weather circumstances by modifying a factor in TTS.

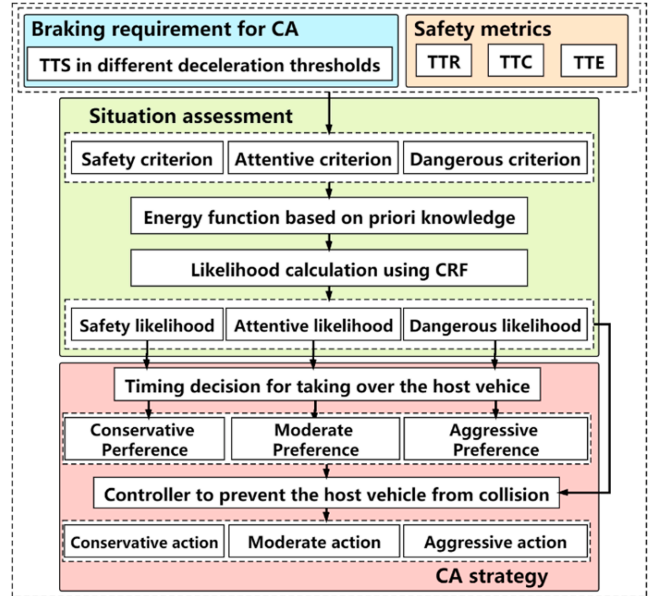


Fig. 6. Risk Assessment and Collision Avoidance (Li et al. [2021])

5.1 Markov Decision Process (MDP)

Interacting with the environment, taking in observations, and formulating judgemental probability distributions for actions to be taken to change the state of the vehicle is what the concept of a Markov Decision Process (van Otterlo and Wiering [2012]) is based on. By establishing the observation set to equal the set of states and specifying the observation conditional probabilities to deterministically

select the observation that fits the real state, an MDP may instead be converted into a Partially Observable Markov Decision Process (POMDP) (Spaan [2012]). Taking the classical approach (Hubmann et al. [2017]), the authors develop a derivation of the most likely future scenario using an interactive, probabilistic motion model for the other cars, which allows for the prediction of uncertainty resulting from poor sensor data and the fact that human drivers' intents cannot be directly detected. For different road configurations and numbers of other vehicles, the POMDP problem can be addressed online with a compact formulation that enables a low-dimensional state-space and uses a Naive-Bayes classifier to simulate trajectory uncertainties of other vehicles. Solutions with interactive behaviour are possible thanks to a basic concept of interaction between the autonomous vehicle and the other users of the road. Their method performs far better than reactive alternatives (Cross Intersection) and almost as well as with complete prior knowledge of the intentions of the other cars.

Incorporating Deep Learning, in (Qiao et al. [2018]), tackles the problem of being computationally inefficient when dealing with the issues of POMDP with RL, for continuous action spaces and the necessity of storing a high number of observations. They use Hierarchical Options for MDP (HOMDP) as a solution rather than modeling the issue as a POMDP. The agent's performance for a four-way junction job with two-way stop signs can be enhanced using the HOMDP technique. With only the observations of the present phase, the HOMDP approach may produce both lower-level continuous actions and higher-level discrete possibilities.

5.2 Bayes Theory

Because the Bayesian interpretation of probability represents a degree of belief or information (knowledge) about an occurrence, its use case in estimating uncertainties or quantifying risks is quite common. (Bernhard et al. [2019]) provide a two-step strategy for accurately taking into account the inherent uncertainties of the traffic environment, such as those resulting from the variability of human driving styles, by risk-sensitive behavior generating algorithms Fig. 7. A method is presented for creating risk-aware behavior that combines online risk assessment with Offline Deep Distributional Reinforcement Learning of different driving styles by modeling the problem as a Stochastic Bayesian Game (SBG). The concept of offline distribution learning implies that actions corresponding to a defined behavior type space/distribution have already been created, and the risk criteria then analyze the actions. In addition to keeping an aggressive driving style, the approach improves safety in unsignalized cross intersection situations when there is a natural ambiguity about how other participants will behave while avoiding overly cautious driving.

(Li et al. [2022]) train a strategic DRL network with the least level of risk to ensure driving safety on one-way highway lane changing scenario. However, it was based on a risk assessment function (not a reward function) that gives realistic likelihood at varying risk levels (Dangerous, Attentive, Safe) See Fig. 8 similar to (Li et al. [2021]).

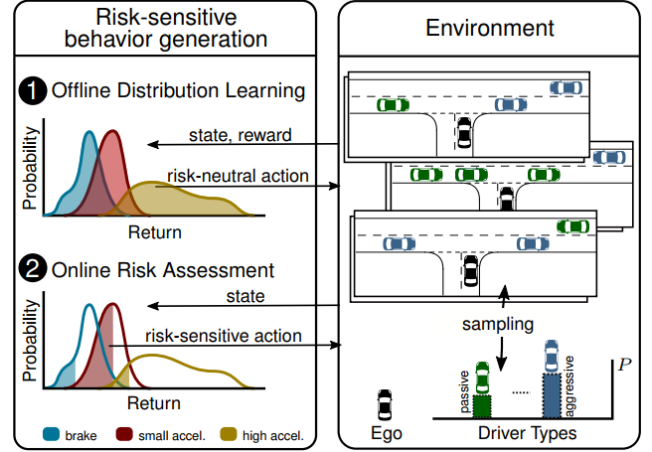


Fig. 7. Driver Type Distribution and Risk Assessment (Bernhard et al. [2019])

To estimate the described driving risk, a quantized, probabilistic model-based technique based on position uncertainty utilizing Bayes theory was presented. The suggested approaches develop robust safe driving strategies and outperform prior methods discussed in the paper. The authors of (Hoel et al. [2020]) describe the agent's level of trust in the suggested activities and how they should be classified as safe or unsafe. An ensemble of neural networks with added randomized prior functions (RPF) was the foundation of the Bayesian RL approach, which was used to calculate the degree of decision uncertainty in autonomous vehicles, and adding a metric (Uncertainty Threshold) to gauge a trained agent's level of confidence in a given course of action. When confronted with circumstances that are outside the training distribution, the trained agent may evaluate the uncertainty of its options and flag one that is undesirable. The identification of scenarios that should be included in the training process may also be done using this knowledge. They show that Bayesian RL is superior to conventional RL for tactical decision-making in autonomous driving in a one-way highway lane changing scenario. In comparison, collisions take place under unknowable circumstances as a result of the regular DQN agent's lack of access to uncertainty information.

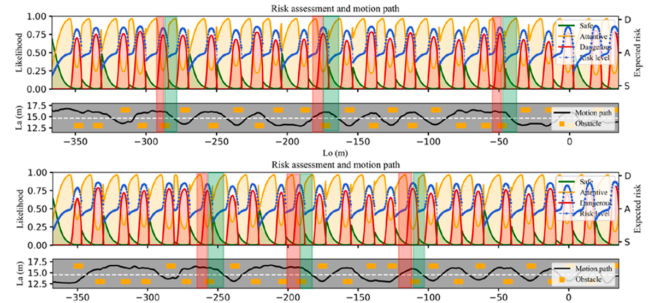


Fig. 8. Risk Assessment and Motion Path Representation (Li et al. [2022])

5.3 Addressing Occlusions

In settings where the probability of fatalities is raised owing to increased traffic frequency and restricted visibility,

path planning algorithms for autonomous cars must take safety and comfort into account. The usage of an extended virtual obstacle within the perception system to account for any beginning location and velocity of a potentially concealed vehicle (Karanam et al. [2022]) results in a safe and pleasant trajectory at the low vision zones, addressing the expectation of the existence of a genuine occlusion. A smoother velocity profile is also observed, which improves comfort by remaining within the prescribed limits. Addressing occlusions here is an extension to the framework presented in (Duhautbout et al. [2021]).

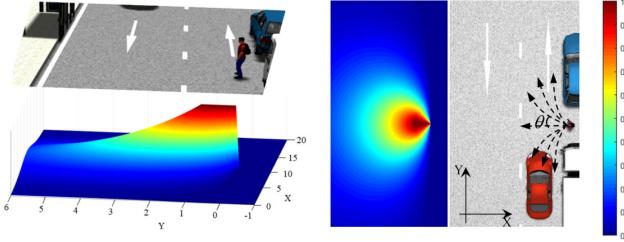


Fig. 9. Distance Coefficient Field Graph (Wang et al. [2022])

In (Wang et al. [2022]) they not only examine how autonomous cars may effectively navigate through hazardous interaction zones but also when scenes are obscured or the view is limited. With the use of a Dynamic Bayesian Network (DBN) based probability model, the possible threats from visually obscured places are put into numerical perspective. A risk distance coefficient model based on distance and velocity is suggested to capture the perceptual interaction function of traffic participants (TPs) See Fig. 9. In real-time, adjusting the speed and heading angle allows the AV to approach the road cautiously and comfortably while gradually expanding the visible area. The model is applicable to junction scenarios without a signal light and scenes with occasional blockage of the visual field. However, working in poor weather conditions like fog or at night when visibility is poor may add difficulties. Additionally, due to limitations on data availability and the influence of indicators on assessment outcomes, it is hard to create an assessment model that incorporates all indications in actual risk assessment. The following summarizes their approach: 1) *Posteriori risk assessment probability P*. Determines the potential risk value (occupancy probability of road cell) of the AV travelling through the interaction area. 2) *The risk distance coefficient*. Abbreviated Kd, is a risk function that incorporates the attention, speed, and perceptual interaction capabilities of TPs and is based on the "conditional risk" concept. 3) *Evaluation of potential risks*. The product of the past 2 metrics represents the possible danger on road cell.

6 Mimicking Human Behaviour

Focusing on achieving human performance by using this human performance as the input, the design of the AV is aimed at mimicking human behaviour but without the downsides of human behaviour. Considering the available routes and maneuver possibilities for other cars while allowing for socially acceptable conduct in crossroads settings, the Intelligent Driver Model (IDM) used as a virtual

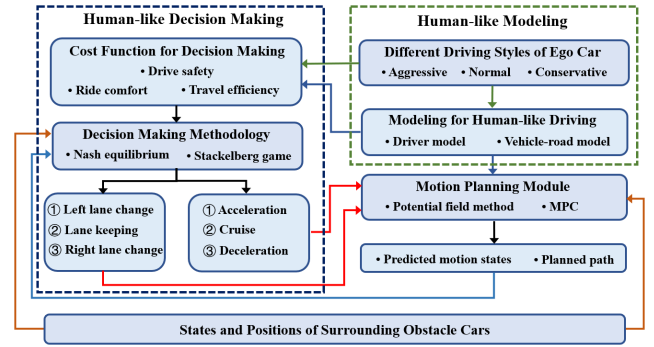


Fig. 10. Human-like decision making framework for AVs (Hang et al. [2021])

leading vehicle (VL) and a local continuous optimization technique are combined in the framework to enable rapid behavior prediction, maneuver creation, and decision making across extended horizons (Speidel et al. [2020]). This enables the ego vehicle to seamlessly merge behind a real leading vehicle (RL) that may arrive from another lane while taking into account its most likely trajectory through the (VL).

Similar to the path-speed split in (Duhautbout et al. [2021]), the framework Fig. 10 suggested in (Hang et al. [2021]) focuses on minimizing the impact of AVs and their mismatch with human drivers while integrating AVs into the traffic ecology of human drivers. The MPC is used for motion prediction and planning, in conjunction with the potential field model. Finally the output is split into high-level decisions and low-level control. Models are created for decision-making using noncooperative game approaches, Nash equilibrium and Stackelberg, that include several driving styles (similar to in (Li et al. [2021]) and (Bernhard et al. [2019])) with associated social interaction features that are collected from real-world driving data. Both game theoretic techniques can provide AVs a decision-making process that is logical and human-like in merging and one-way lane change scenarios. The cost value of decision-making utilizing the Stackelberg game theoretic technique is decreased by about 20 percent when compared to the Nash equilibrium strategy when driving normally.

In the following we will see the use of End-to-End Imitation Learning (IL) and Reinforcement Learning to achieve human-like driving through trajectory/waypoints generation, and directly controlling the speed and steering angle of the vehicle.

6.1 End-to-End Reinforcement Learning: Speed and Steering Angle

To facilitate efficient traffic flow, where AVs use human-like driving policies and negotiation techniques, (Emuna et al. [2020]) incorporates a vehicle management algorithm that receives a stochastic feedback signal from two different sources: a model-driven portion that encodes basic driving laws like lane-keeping and speed control, and a stochastic, data-driven section that adds human expert knowledge from driving data. The machine-driven agent can mimic human driving behaviour, in a one-way lane change scenario, using a reinforcement learning algorithm Proximal Policy Optimization (PPO) in conjunction with

a Mixture Density Network (MDN) and describes both human and automated driving using Gaussian distributions. Thus, creating a framework that is natural for combining data-driven approaches to include expert human driving with rule-based ways to encode fundamental driving norms, driving regulations that are humane and can accurately mimic expert human behavior in a setting with a high dimensional state space, controlling speed and steering angle.

In (Chen et al. [2021]), they greatly facilitate modern reinforcement learning by assuming independence between the agent and the environment, or what they refer to as a (world on rails) by factoring the dynamics into a nonreactive world model and a low-dimensional and compact forward model of the ego-vehicle. They create a world model based on human behaviour that is realistic and expressive enough to allow the agent to investigate its surroundings and the effects of its actions. The policy picks up tips on how to avoid expensive errors or how to correct them. The ultimate driving rule, controlling speed and steering angle, functions effectively in an environment that is both dynamic and reactive. The approach comes in first place on the CARLA scoreboard and uses 40 percent less data while achieving a driving score that is 25 percent higher. On navigational tasks in the benchmark, the approach is also orders of magnitude more sample-efficient than cutting-edge model-free reinforcement learning approaches.

6.2 Perception-based

The perception system is the main input to the decision-making system. Altering its input and output has a high influence on the architecture and performance of decision-making systems. In (Mitchell et al. [2020]), a high-level controller manages velocity modulation and lane-change decisions, while the low-level controller uses this information to monitor the intended lanes at the necessary speeds. Consequently, they develop a sequential choice problem for one way lane change and solve it with actor-critic (A2C) reinforcement learning (1 actor and 2 critics). This study’s primary contribution is its use of mixed reality learning to create intelligent driving strategies that can handle unexpected vehicle encounters and limited navigation. Their findings demonstrate that mixed reality learning may significantly boost performance, which reduces collisions in the learnt policies.

6.2.1 End-to-End Imitation Learning: Waypoints Generation Imitation learning has gained popularity due to advances in computing and sensing. Without the requirement for explicit programming or task-specific reward function design, generic imitation learning techniques may be able to reduce the difficulty of teaching a task to that of delivering demonstrations. In (Bansal et al. [2018]) they train an autonomous driving system that is reliable enough to operate a real car. To maximize the value of the training data, they use mid-level input and output representations that benefit from perception and control elements. Instead of just replicating all data, they enhance the imitation loss with extra losses that discourage poor outcomes and promote advancement. The model predicts the bounding box of the vehicle as a spatial heatmap at each subsequent timestep for each point at time t in the trajectory, where

each point at time t is described by its location, direction, and speed. In simulation, the ChauffeurNet model can manage challenging circumstances, and is then displayed operating a vehicle in the real world.

The study in (Hecker et al. [2020]) investigates ways to make end-to-end driving models more accurate and realistic. It presents the first large-scale dataset with high-precision semantic mappings that is suitable for training end-to-end driving models by adding Here Technologies’ high precision semantic maps to the Drive360 dataset. With the help of an attention mechanism that can support various confidence masks of a semantic segmentation network, they suggest a novel method for incorporating these semantic maps into the end-to-end driving model, allowing for the combination of map information with semantic information from the image. The decision-making problem is handled as an independent and identically distributed (i.i.d) target supervised regression problem (expert actions). Instead of single actions, they suggest matching action chains (which they refer to as drivelets). The drivelet for the following N time steps is represented by $a_t = [a_t^T, \dots, a_{t+N}^T]$, where a_t is calculated using one of the previously suggested driving models, they represent a drivelet for the following N time steps. Extensive testing demonstrates that the driving models outperform traditional techniques in terms of accuracy and behavior in lane change and merging scenarios.

Addressing the drawbacks of the integration of complementary sensor representations for autonomous driving, geometry-based sensor fusion by itself might not be adequate to successfully fuse representations in end-to-end driving models (Prakash et al. [2021]). In order to include the whole context of the 3D scene into the feature extraction layers of various modalities, they present TransFuser, a unique Multi-Modal Fusion Transformer, as a means of fusing attention with image and LiDAR representations. The two primary parts of the architecture they suggest for end-to-end driving are an auto-regressive waypoint prediction network and a MultiModal Fusion Transformer. The method offers cutting-edge driving performance with 76 percent less collisions as compared to geometry-based fusion.

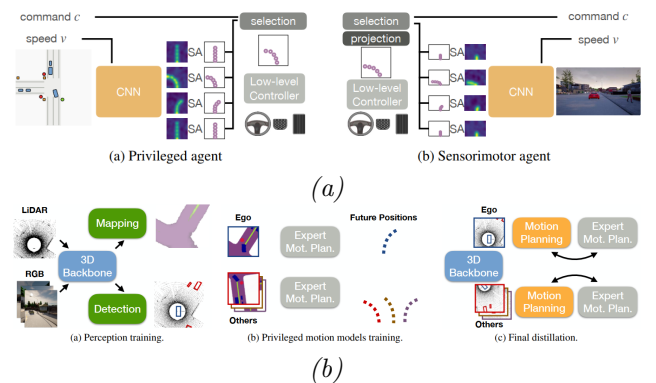


Fig. 11. (a) Decomposition of Imitation Learning (Chen et al. [2019]) (b) Overview of the training pipeline (Chen and Krähenbühl [2022])

The authors of (Chen et al. [2019]) divided the learning process into two steps, in the first step they train an

agent who has access to privileged information such as the environment’s layout and the positions of other traffic participants. This special agent has been taught to mimic the expert motions. In the second step, a sensorimotor agent without any access to any privileged information is trained to mimic the privileged agent (waypoints generation) See Fig. 11. On the CARLA test and the NoCrash benchmark, this imitation learning process for vision based urban driving (cross intersections, roundabouts, merging lane change,...) significantly outperforms the state of the art.

A sensing module, a motion planner, and a low-level controller comprise an end-to-end differentiable three-stage modular pipeline that handles the partial observability of both perception and motion (Chen and Krähenbühl [2022]). They describe a system for training driving policies (waypoints generation) using data gathered from all the other vehicles the ego-vehicle sees, not just the its own data See Fig. 11. The approach significantly exceeds all previous techniques on the open CARLA Leaderboard, raising driving score by 25 points and route completion rate by 24 points. Limitations and possibly detrimental societal effects: Continue to commit driving violations. It would probably result in traffic accidents if used in the real world (negative social impacts). Technically speaking, the conditional motion planner’s current behavior predictor does not take multimodality into account beyond the high-level directives.

6.2.2 End-to-End Reinforcement Learning: Latent Space Representation All the crucial information required to represent the original input observation is included in the latent space representation of data. This is the fundamental idea behind the term "representation learning" which refers to a collection of methods that let a machine extract the representations required for feature identification or classification from unstructured input. In this use case, more complicated forms of raw data (such as photos and videos) are converted into simpler representations that are easier to handle and evaluate using latent space representations See Fig. 12. One main advantage is shown in (Iroegbu and Madhavi [2021]) where the use of latent space representation significantly shortens training time and improves the quality of deep reinforcement learning. In (Sharma and Sharma [2021]), they take into account numerous dynamic aspects, including multi-agent interactions, various scene perceptions, complicated road geometry, and other infrequently occurring real-world occurrences, by the decision-making system. They tackle the issue where complexity of urban driving situations does not generalize well with current DRL approaches through step-by-step learning of various driving tasks, harsh episode termination policy, and dense reward system. A unique strategy for object segmentation by pre-detecting and labeling crucial objects/states in the environment is presented. Their agents achieve a success rate of 100 percent on all driving tasks in the original CARLA benchmark and set a new record of 82 percent on further complex NoCrash benchmark, outperforming the state-of-the-art model by more than 30 percent on NoCrash benchmark. Making long-range predictions about how various interacting agents would behave in dynamic situations is presented in (Lee et al. [2017]). In order to implement DESIRE, a broad framework that may be used for

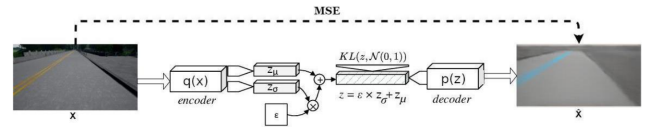


Fig. 12. Latent Space Representation Iroegbu and Madhavi [2021]

any future prediction challenge, they suggest the following innovative mechanisms: Scene Context Fusion, Inverse Optimal Control (IOC)-based Ranking and Refinement, and Diverse Sample Generation. The suggested approach, when compared to existing standard techniques, greatly enhances prediction accuracy, according to experiments.

The authors of (Chen et al. [2022]) offer an interpretable, complex-urban-scenario-capable deep reinforcement learning approach for end-to-end autonomous driving. By adding an entropy regularization to the reward, they get Maximum Entropy Reinforcement Learning (MaxEnt RL) which conducts better exploration, typically outperforms traditional RL algorithms, and is more reliable. It can also be modeled as Probabilistic Graphical Model (PGM), which enables joint learning of the environment model and policy, and Latent Space Representation, which enables a comprehensible explanation of how the policy reasons about the environment. This in turn is used to address issues with autonomous driving, such as the learned model’s lack of interpretability and its inability to accept multiple sources of sensor inputs, which are crucial for autonomous driving systems. The performance of their system in most urban settings with congested surrounding cars outperforms numerous baselines, using a simulated autonomous automobile in CARLA. The learnt policy is also able to give a better description of how the automobile thinks about the driving environment by using masked outputs. Although their framework can offer understandable justifications for how the model perceives its surroundings, it cannot offer any insight into how it decides because the driving policy is acquired in a model-free manner.

Instead of using a learning-based approach to extract the important features of a latent space, the authors of (Xu et al. [2020]) try to introduce explanations to the perception of the environment around the car and how they relate to the decision-making capabilities. The new paradigm, which draws its inspiration from how people solve problems, sits between end-to-end and pipelined techniques. It also depends on scene knowledge, the latter solely takes into account potential hazard-causing things. Additionally, they specify a collection of justifications for these actions, which have to be created in tandem with the latter. These are referred to as action inducing since modifications to their condition ought to cause the vehicle to act. The necessity of explanations enhances the detection of items that prompt actions, which in turn results in improved action predictions. With a greater understanding of the object-scene linkages, the system is better able to identify the objects that cause actions and, as a result, make action predictions simpler. For example, the system can determine what to localize in the global feature map and relate local to global features.

7 Conclusion

An overview of the research community direction has been provided where learning based techniques have dominated the current stage of development. The main outcome of this work is the classification of the coverage literature surveys, action space and scenarios, and a breakdown of both the modular and end-to-end approaches which will hopefully aid in the researcher’s work in finding a solution for the decision-making problem in AVs.

Acknowledgements

The work was realized as under an International Research Project: Approches de Diagnostique et de cONtrôle Intelligent des Systèmes (IRP ADONIS)

References

- Aradi, S. (2022). Survey of deep reinforcement learning for motion planning of autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 23(2), 740–759.
- Badue, C., Guidolini, R., Carneiro, R.V., Azevedo, P., Cardoso, V.B., Forechi, A., Jesus, L., Berriel, R., Paixão, T.M., Mutz, F., de Paula Veronese, L., Oliveira-Santos, T., and De Souza, A.F. (2021). Self-driving cars: A survey. *Expert Systems with Applications*, 165, 113816.
- Bansal, M., Krizhevsky, A., and Ogale, A. (2018). Chauffeurnet: Learning to drive by imitating the best and synthesizing the worst.
- Bernhard, J., Pollok, S., and Knoll, A. (2019). Addressing inherent uncertainty: Risk-sensitive behavior generation for automated driving using distributional reinforcement learning. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, 2148–2155.
- Chen, D., Koltun, V., and Krähenbühl, P. (2021). Learning to drive from a world on rails. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 15570–15579.
- Chen, D. and Krähenbühl, P. (2022). Learning from all vehicles. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 17201–17210.
- Chen, D., Zhou, B., Koltun, V., and Krähenbühl, P. (2019). Learning by cheating. In L.P. Kaelbling, D. Kragic, and K. Sugiura (eds.), *3rd Annual Conference on Robot Learning, CoRL 2019, Osaka, Japan, October 30 - November 1, 2019, Proceedings*, volume 100 of *Proceedings of Machine Learning Research*, 66–75. PMLR.
- Chen, J., Li, S.E., and Tomizuka, M. (2022). Interpretable end-to-end urban autonomous driving with latent deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems*, 23(6), 5068–5078.
- Duan, J., E.L.S.G.Y.S.Q..C.B. (2020). Hierarchical reinforcement learning for self-driving decision-making without reliance on labelled driving data.
- Duhautbout, T., Talj, R., Cherfaoui, V., Aioun, F., and Guillemard, F. (2021). Generic trajectory planning algorithm for urban autonomous driving. In *2021 20th International Conference on Advanced Robotics (ICAR)*, 607–612.
- Duhautbout, T., Talj, R., Cherfaoui, V., Aioun, F., and Guillemard, F. (2022). Efficient speed planning in the path-time space for urban autonomous driving. intelligent transportation systems. Unpublished.
- Emuna, R., Borowsky, A., and Biess, A. (2020). Deep reinforcement learning for human-like driving policies in collision avoidance tasks of self-driving cars.
- Erkan, E. and Yüce, S. (2018). Serret-frenet frame and curvatures of bézier curves. *Mathematics*, 6(12).
- Hang, P., Lv, C., Xing, Y., Huang, C., and Hu, Z. (2021). Human-like decision making for autonomous driving: A noncooperative game theoretic approach. *IEEE Transactions on Intelligent Transportation Systems*, 22(4), 2076–2087.
- Harsanyi, J. (1967). Games with incomplete information played by "bayesian" players, i-iii part i. the basic model. *Management Science*, 14(3), 159–182.
- Harsanyi, J.C. (1968a). Games with incomplete information played by "bayesian" players, i-iii. part ii. bayesian equilibrium points. *Management Science*, 14(5), 320–334.
- Harsanyi, J.C. (1968b). Games with incomplete information played by "bayesian" players, i-iii. part iii. the basic probability distribution of the game. *Management Science*, 14(7), 486–502.
- Hecker, S., Dai, D., Liniger, A., Hahner, M., and Van Gool, L. (2020). Learning accurate and human-like driving using semantic maps and attention. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2346–2353.
- Hoel, C.J., Wolff, K., and Laine, L. (2020). Tactical decision-making in autonomous driving by reinforcement learning with uncertainty estimation. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, 1563–1569.
- Hubmann, C., Becker, M., Althoff, D., Lenz, D., and Stiller, C. (2017). Decision making for autonomous driving considering interaction and uncertain prediction of surrounding vehicles. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, 1671–1678.
- Iroegbu, E. and Madhavi, D. (2021). Accelerating the training of deep reinforcement learning in autonomous driving. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 10, 649.
- Isele, D., Rahimi, R., Cosgun, A., Subramanian, K., and Fujimura, K. (2018). Navigating occluded intersections with autonomous vehicles using deep reinforcement learning. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2034–2039.
- Karanam, S.K., Duhautbout, T., Talj, R., Cherfaoui, V., Aioun, F., and Guillemard, F. (2022). Efficient speed planning in the path-time space for urban autonomous driving. intelligent transportation systems. Unpublished.
- Kendall, A., Hawke, J., Janz, D., Mazur, P., Reda, D., Allen, J.M., Lam, V.D., Bewley, A., and Shah, A. (2019). Learning to drive in a day. In *2019 International Conference on Robotics and Automation (ICRA)*, 8248–8254.
- Koren, Y. and Borenstein, J. (1991). Potential field methods and their inherent limitations for mobile robot navigation. In *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, 1398–1404 vol.2.
- Krauss, S. (1998). Microscopic modeling of traffic flow: investigation of collision free vehicle dynamics.

- Lee, N., Choi, W., Vernaza, P., Choy, C.B., Torr, P.H.S., and Chandraker, M. (2017). Desire: Distant future prediction in dynamic scenes with interacting agents. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2165–2174.
- Li, G., Yang, Y., Li, S., Qu, X., Lyu, N., and Li, S.E. (2022). Decision making of autonomous vehicles in lane change scenarios: Deep reinforcement learning approaches with risk awareness. *Transportation Research Part C: Emerging Technologies*, 134, 103452.
- Li, G., Yang, Y., Zhang, T., Qu, X., Cao, D., Cheng, B., and Li, K. (2021). Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios. *Transportation Research Part C: Emerging Technologies*, 122, 102820.
- Liu, Q., Li, X., Yuan, S., and Li, Z. (2021). Decision-making technology for autonomous vehicles: Learning-based methods, applications and future outlook. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 30–37.
- Lopez, P.A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., and Wiessner, E. (2018). Microscopic traffic simulation using sumo. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2575–2582.
- Mitchell, R., Fletcher, J., Panerati, J., and Prorok, A. (2020). Multi-vehicle mixed reality reinforcement learning for autonomous multi-lane driving. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '20*, 1928–1930. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC.
- Naveed, K.B., Qiao, Z., and Dolan, J.M. (2021). Trajectory planning for autonomous vehicles using hierarchical reinforcement learning. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 601–606.
- Prakash, A., Chitta, K., and Geiger, A. (2021). Multi-modal fusion transformer for end-to-end autonomous driving. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 7073–7083.
- Qiao, Z., Muelling, K., Dolan, J., Palanisamy, P., and Mudalige, P. (2018). Pomdp and hierarchical options mdp with continuous actions for autonomous driving at intersections. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2377–2382.
- Said, A., Talj, R., Francis, C., and Shraim, H. (2021). Local trajectory planning for autonomous vehicle with static and dynamic obstacles avoidance. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 410–416.
- Said, A., Talj, R., Francis, C., and Shraim, H. (2022). Tuning and costs analysis for a trajectory planning algorithm for autonomous vehicles. In *Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems - VEHITS*, 88–95. INSTICC, SciTePress.
- Sharma, A. and Sharma, S. (2021). Wad: A deep reinforcement learning agent for urban autonomous driving.
- Sharma, O., Sahoo, N., and Puhan, N. (2021). Recent advances in motion and behavior planning techniques for software architecture of autonomous vehicles: A state-of-the-art survey. *Engineering Applications of Artificial Intelligence*, 101, 104211.
- Spaan, M.T.J. (2012). *Partially Observable Markov Decision Processes*, 387–414. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Speidel, O., Graf, M., Kaushik, A., Phan-Huu, T., Wedel, A., and Dietmayer, K. (2020). Trajectory planning for automated driving in intersection scenarios using driver models. *CoRR*, abs/2010.03345.
- Talpaert, V., Sobh, I., Kiran, B., Mannion, P., Yogamani, S., El-Sallab, A., and Perez, P. (2019). Exploring applications of deep reinforcement learning for real-world autonomous driving systems. In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 5: VISAPP*, 564–572. INSTICC, SciTePress.
- Tampuu, A., Matiisen, T., Semikin, M., Fishman, D., and Muhammad, N. (2022). A survey of end-to-end driving: Architectures and training methods. *IEEE Transactions on Neural Networks and Learning Systems*, 33(4), 1364–1384.
- Toledo, T., Koutsopoulos, H.N., and Ben-Akiva, M. (2007). Integrated driving behavior modeling. *Transportation Research Part C: Emerging Technologies*, 15(2), 96–112.
- Treiber, M., Hennecke, A., and Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. E*, 62, 1805–1824.
- van Otterlo, M. and Wiering, M. (2012). *Reinforcement Learning and Markov Decision Processes*, 3–42. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Wang, D., Fu, W., Song, Q., and Zhou, J. (2022). Potential risk assessment for safe driving of autonomous vehicles under occluded vision. *Scientific Reports*, 12.
- Wang, Z.j., Chen, X.m., Wang, P., Li, M.x., Ou, Y.j.x., and Zhang, H. (2021). A decision-making model for autonomous vehicles at urban intersections based on conflict resolution. *Journal of Advanced Transportation*, 2021, 1–12.
- Xu, Y., Yang, X., Gong, L., Lin, H.C., Wu, T.Y., Li, Y., and Vasconcelos, N. (2020). Explainable object-induced action decision for autonomous vehicles. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9520–9529.
- Yurtsever, E., Lambert, J., Carballo, A., and Takeda, K. (2020). A survey of autonomous driving: Common practices and emerging technologies. *IEEE Access*, 8, 58443–58469.
- Zhu, Z. and Zhao, H. (2022). A survey of deep rl and il for autonomous driving policy learning. *IEEE Transactions on Intelligent Transportation Systems*, 23(9), 14043–14065.
- Ziegler, J., Bender, P., Schreiber, M., Lategahn, H., Strauss, T., Stiller, C., Dang, T., Franke, U., Appenrodt, N., Keller, C.G., Kaus, E., Herrtwich, R.G., Rabe, C., Pfeiffer, D., Lindner, F., Stein, F., Erbs, F., Enzweiler, M., Knöppel, C., Hipp, J., Haueis, M., Treppe, M., Brenk, C., Tamke, A., Ghanaat, M., Braun, M., Joos, A., Fritz, H., Mock, H., Hein, M., and Zeeb, E. (2014). Making bertha drive—an autonomous journey on a historic route. *IEEE Intelligent Transportation Systems Magazine*, 6(2), 8–20.