Decision Making for Autonomous Vehicles based on Risk Assessment in a Dynamic Environment

Dany Ghraizi, Reine Talj, Clovis Francis

Decision Making for Autonomous Vehicles based on Risk Assessment in a Dynamic Environment

Dany Ghraizi[1], Reine Talj[1] and Clovis Francis[2]

Abstract— **Recent advancements in autonomous driving technologies have significantly enhanced road safety and collision avoidance. However, ensuring this in dynamic driving environments remains a challenging endeavor. This paper addresses this challenge by proposing a high-level risk-aware decision-making module integrated into the trajectory planner of autonomous vehicles. The module establishes a function for dynamic risk assessment, considering both longitudinal and lateral aspects of the environment. By incorporating various factors such as velocity and relative position, the proposed function enables the vehicle to anticipate and respond to potential hazards proactively. Additionally, the paper incorporates and builds upon previous work of a modular and distinctive AI-based Adaptive Cruise Control (ACC) system with robust generalization capabilities. Results demonstrate the effectiveness of the proposed approach in improving safety, collision avoidance, and adaptability in highly interactive driving environments allowing the vehicle to dynamically re-plan trajectories and speed profiles during lane change maneuvers.**

I. INTRODUCTION

In the last decade, autonomous driving has witnessed tremendous advancements, marking significant strides towards a future of safer transportation systems. However, amid this progress, the assurance of safety in highly interactive driving environments stands out as a formidable challenge [1]. As autonomous vehicles increasingly navigate complex road scenarios alongside human-driven counterparts, the need for robust decision-making frameworks capable of assessing and mitigating risks in real-time becomes paramount [2].

Many vehicles now integrate driver assistance technologies, which are crucial for adapting to traffic changes and providing extensive control over both the vehicle's forward and sideways movement. This significantly improves passenger comfort and safety. Some of these systems give warnings if a crash seems likely, while others actively work to avoid accidents. Together, these advancements will help protect everyone on the road [3].

Lane changes are intricate driving maneuvers influenced by various factors like traffic flow, lane configurations, and driver intentions [4], [5]. They can lead to conflicts with nearby vehicles and prompt drivers to take evasive actions [6]. Additionally, lane changes themselves can sometimes be evasive maneuvers taken to avoid potential risks [7]. Understanding these evasive behaviors during lane changes is crucial due to their complexity and significant impact on safety.

This demand has motivated intensive research into the integration of risk-awareness mechanisms within decision making and trajectory planning algorithms, aiming to empower autonomous systems with the foresight to anticipate and respond to potential hazards proactively. Safety Surrogate Measures (SSM) represent a proactive strategy at capturing near-crash instances. Among the various SSMs, Time Headway (TH) and Time to Collision (TTC) stand out as one of the most commonly employed methods [8]. For instance, TTC is defined as the time needed for two vehicles to collide if they maintain their current speeds along the same path. If the TTC falls below a certain threshold, it indicates an unsafe car-following scenario. Due to its simplicity and practicality, TTC has found widespread use as a safety indicator in numerous studies [9], [10]. However, there are two primary limitations associated with conventional TTC: firstly, it deems a scenario safe even when the following vehicle's speed matches or is slower than the leading vehicle's, despite a potentially very close relative distance; secondly, it assumes that the vehicle pair is traveling in the same lane and only considers longitudinal movements. Therefore, other research efforts have focused on extending the Time to Collision (TTC) concept to the two-dimensional road plane [11], and introducing trajectory-based SSMs. For instance, Anticipated Collision Time (ACT) was introduced as a proactive safety assessment method based on two-dimensional trajectories [12]. However, the process of reconstructing complete vehicle trajectories for all vehicles involved can be complex and costly.

There is also the inclusion of learning-based decision-making which comes with its own set of questions regarding complexity, interpretability, and explainability [13], [14]. For instance in [15], they combine a learning based lateral decision making module with a trajectory planner to safely adjust lane change decisions and dynamically re-plan the trajectory followed. Other methods define functions for measuring risk probabilisitcally. For example, one approach uses a risk distance coefficient model which is a model that considers the distance and speed of other vehicles to understand how they interact with each other. This model, based on a Dynamic Bayesian Network, helps the autonomous vehicle adjust its speed and direction cautiously and comfortably [16]. Another method employs a "loss factor" to show how severe the impact would be if the autonomous vehicle and another vehicle are in the same or different lanes. When they are in different lanes, the loss factor decreases as the distance between lanes increases. Then, the risk of collision is evaluated by looking at the predicted paths of nearby vehicles and the current state of the autonomous vehicle [12]. These methods have proven to be effective at mitigating risks. Therefore, this paper introduces a high-level risk-aware decision-making module to address the safety gap in the trajectory planner [17]. It establishes a function based on various factors to dynamically re-plan the ego vehicle's trajectory through anticipating risks. The contributions of this paper include:

- Establishment of a formula for dynamic risk-assessment that takes into consideration the longitudinal and lateral aspects of the environment.
- Proposing a high-level decision-making module that

*The work was realized within an International Research Project: Approches de Diagnotic et de cONtrôle Intelligent des Systèmes (IRP ADONIS), funded by CNRS Université de technologie de Compiègne (UTC) and Université libanaise (UL).

[1]Sorbonne université, Université de technologie de Compiègne, CNRS, Heudiasyc UMR 7253, CS 60 319, 60 203 Compiègne, France name.surname@hds.utc.fr

[2]Arts et Métiers Paris Tech, 1, Rue Saint Dominique, 51000 Châlons en Champagne, France clovis.francis@ensam.eu

seeks to encompass both lateral decisions, namely "lane keeping" or "lane changing," based on the introduced risk assessment function utilizing both instantaneous risk and predicted risk with a time horizon of 2 seconds.
- Enabling the flexibility to dynamically re-plan both trajectories and speed profiles during lane changes, offering enhanced motion adaptability, and robust collision avoidance compared to the vanilla trajectory planner.
- Building upon previous work of a modular and distinctive AI-based ACC system with robust generalization capabilities that adapts to dynamically changing traffic scenarios for longitudinal decision making.

The rest of the paper is organized as follows: Section II describes the perception, vehicle dynamics, and control systems used. Section III defines the proposed risk assessment function. Section IV presents the risk-aware decision-making and trajectory planning modules, including the lateral risk-aware and longitudinal Deep Reinforcement Learning Adaptive Cruise Control (DRL-ACC) decision frameworks. Section V presents the simulation results, demonstrating the effectiveness of the proposed system in optimizing collision avoidance, dynamic re-planning of trajectories among other behaviors, and speed modulation. Finally, Section VI provides the paper's conclusion and explores potential avenues for future research endeavors.

II. Perception, Vehicle Dynamics, and Control

A. Perception

The discussion of the perception block is beyond the scope of this paper. Here, we consider the output produced by the perception module as an occupancy grid representation, as outlined in the referenced work [17]. To summarize, an initial global occupancy grid is created from a global map. Subsequently, the local occupancy grid is derived from this global representation based on the vehicle's position and orientation. The local grid consists of cells sized at 400 * 400, each representing an area of 25 * 25 cm, in accordance with the perception system's horizon. To improve precision, efficiency, and time optimization in collision checking, the local occupancy grid is converted into a clearance map [18].

B. Vehicle Dynamics Model and Control

Our algorithm integrates a detailed longitudinal and lateral vehicle model developed using the multi-body formalism outlined in [19]. This model takes wheel driving/braking torque (τ_w) and steering angle (δ) as inputs. By employing the Dugoff model to estimate tire forces and conducting model matrix computations, it yields the following outputs: longitudinal ($\ddot{x}$) and lateral ($\ddot{y}$) vehicle accelerations, yaw rate ($\dot{\psi}$), and wheel angular velocities (w_{ij}).

For control purposes, we have selected a second-order sliding mode based on the super-twisting algorithm. This selection ensures robust stability while minimizing chattering, which is a common issue in sliding mode control. The complete model has been validated using the SCANeR studio simulator across various driving conditions. In summary, our algorithm incorporates a comprehensive vehicle model and employs advanced control techniques to enhance performance. Through rigorous validation, we have ensured that the model behaves accurately and reliably across diverse scenarios.

III. Proposed Risk Assessment Function

Speeding increases the likelihood of losing control of a vehicle due to several factors. Higher speeds make it harder to navigate curves, turns, or obstacles on the road, leading to skidding, sliding, or loss of traction. Additionally, excessive speeds reduce the time available for drivers to react and execute evasive maneuvers, increasing the risk of accidents. It also extends braking distances, making it harder to stop in time to avoid collisions. Overall, high speeds jeopardize safety, increase the risk of accidents, and endangers both the driver and other road users [20]. Therefore the risk increases as the ego vehicle speed V_{ego} increases.

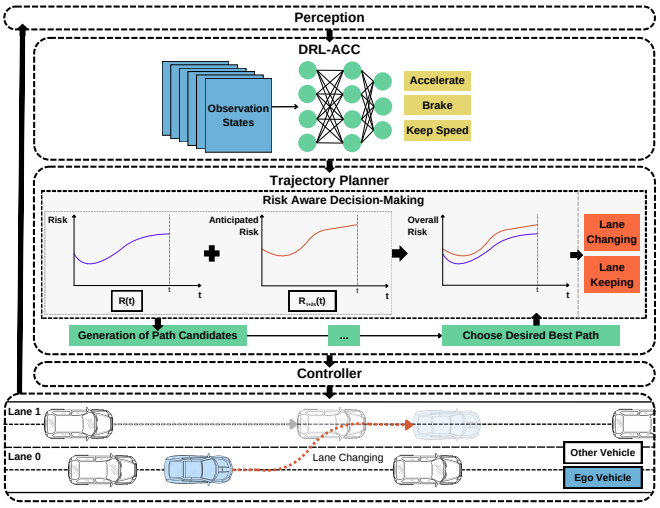

Fig. 1: Overview of the Architecture.

The relative speed between the ego and other vehicle plays a crucial role in road safety. Significant differences in speed can lead to dangerous situations where conflicts and collisions may arise when interacting on the road. For instance, if two vehicles are traveling at different speeds while attempting to navigate the same road space, or if a faster vehicle attempts to overtake a slower vehicle but misjudges the available space or closing speed, it can result in a collision or loss of control [21]. Therefore, maintaining a safe and reasonable relative speed between vehicles is essential to prevent accidents and promote safe driving practices. In summary, the risk is affected by relative speed V_{rel} based on the situation encountered, so it increases when the ego vehicle is faster than a leading vehicle, or when it is slower than a following vehicle and vice-versa.

The proposed risk function (eq. 3) consists of 2 parts, velocity R_v and 2D-position R_d. The first component, R_v (eq. 1), takes into consideration both the velocity of the ego vehicle and the relative velocity of the ego vehicle with respect to the other vehicles. The second component, R_d (eq. 2), takes into consideration both the longitudinal $d_{\text{long-rel}}$ and lateral $d_{\text{lat-rel}}$ relative distances to the ego vehicle. The global risk domain with respect to these variables are depicted in Fig. 3.

During each planning iteration, the local occupancy grid updates to accommodate moving obstacles within its perception zone. However, to take their velocity into consideration and to to ensure a more responsive assessment compared to instantaneous measurements in the rapidly changing dynamics of the environment on the road, a longitudinal widening of their occupancy is carried out relative to the grid. It's equal to the predicted traveled distance during one iteration. The new relative position is then extracted to be used in R_d. A safety zone, presented in Fig.4, with maximum risk is also defined to guarantee that the ego vehicle does not execute extremely dangerous maneuvers within this zone. This is inspired by [22] who proposed the existence of a dynamic space that the driver perceives as an area in which they can navigate safely.

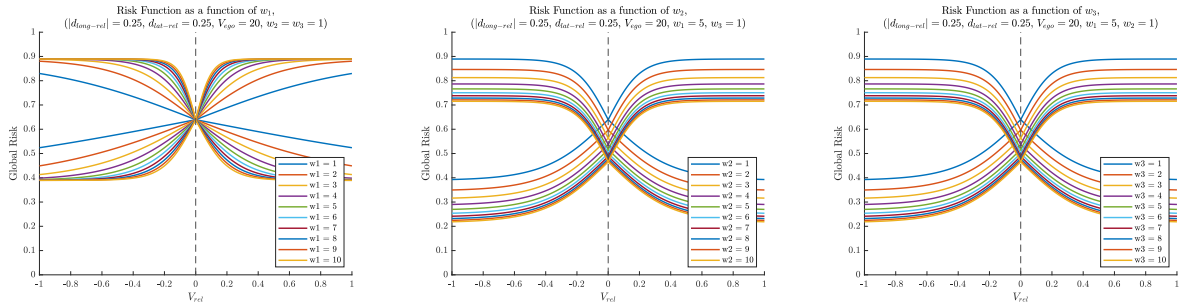

Fig. 2: Global Risk as a function of relative speed as w_1, w_2, and w_3 vary.

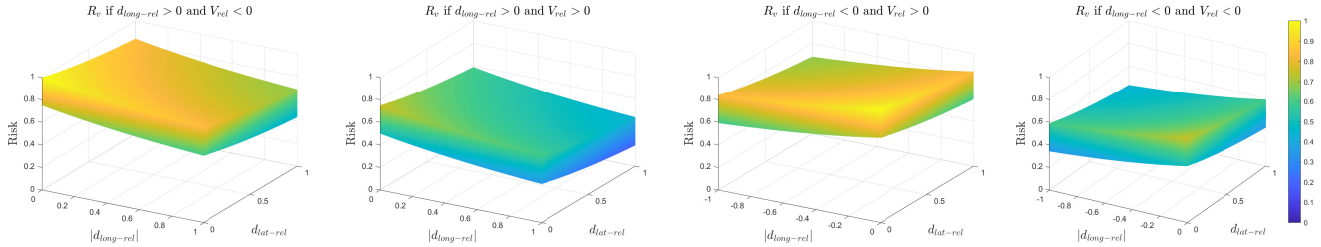

Fig. 3: Domain of Global Risk (z-axis) with respect to relative longitudinal distance $|d_{long-rel}|$ (x-axis), relative lateral distance $d_{lat-rel}$ (y-axis), relative velocity, and ego vehicle velocity.

Termed by the authors, the "field of safe travel" denotes the acceptable trajectories for a vehicle within a defined zone. It depends on the driver's experience, the safety distances they wish to respect and their perception of the size of the car, among other factors. In this case, a time-headway of $0.5s$, based on reaction times [23], is used allowing the safety zone to be adjusted dynamically with respect to the environment.

$$R_v = \begin{cases} \frac{1}{1+e^{w_1 \cdot (1+V_{\text{ego}}) \cdot V_{\text{rel}}}} & \text{if } d_{\text{long-rel}} > 0 \text{ and } V_{\text{rel}} < 0 \\ \frac{1}{1+e^{w_1 \cdot (2-V_{\text{ego}}) \cdot V_{\text{rel}}}} & \text{if } d_{\text{long-rel}} > 0 \text{ and } V_{\text{rel}} > 0 \\ \frac{1}{1+e^{-w_1 \cdot (1+V_{\text{ego}}) \cdot V_{\text{rel}}}} & \text{if } d_{\text{long-rel}} < 0 \text{ and } V_{\text{rel}} > 0 \\ \frac{1}{1+e^{-w_1 \cdot (2-V_{\text{ego}}) \cdot V_{\text{rel}}}} & \text{if } d_{\text{long-rel}} < 0 \text{ and } V_{\text{rel}} < 0 \end{cases} \tag{1}$$

$$R_d = \frac{1}{2}\left(e^{-w_2 \cdot |d_{\text{long-rel}}|} + e^{-w_3 \cdot d_{\text{lat-rel}}}\right) \tag{2}$$

$$Risk = \frac{1}{2}\left(R_v + Rd\right) \tag{3}$$

where w_1, w_2, and w_3 are set to 5, 1, and 1 respectively. In cases where the ego vehicle surpasses others in speed, V_{rel} takes a negative value, and conversely. Likewise, if another vehicle trails the ego vehicle, $d_{\text{long-rel}}$ is negative, while it's positive if it leads.

Figure 2 illustrates the variation of global risk concerning relative speed V_{rel}, with fixed values of $|d_{long-rel}| = 0.25m$, $d_{lat-rel} = 0.25m$, and $V_{ego} = 20m/s$. We explore weights ranging from 1 to 10, demonstrating the tuning flexibility by revealing that higher w_1 values result in a more conservative risk function, while lower w_2 and w_3 tend to a more aggressive driver behavior. Additionally, Figure 3 depicts the domain of the global risk function across all possible scenarios outlined in equation 1. Here, the inter-dependencies among variables are evident, as the risk function is a 3D surface that adjusts in response to variations in ego vehicle velocity and relative velocity across different longitudinal and lateral distances. The differentiation between risk scenarios, such as leading, following, side-swipe, rear-end collision, and front-end collision, enables the risk function to accurately assess each situation. The first two plots (left) depict scenarios where the ego vehicle trails another vehicle, showcasing lower risk domains when car-following at lower relative speeds, and vice versa for the latter two plots (right), which represent scenarios where the ego vehicle leads another vehicle.

IV. Risk-Aware Decision-Making and Trajectory Planning

The ego vehicle makes high-level decisions regarding both lateral and longitudinal actions. Lateral decisions involve maintaining or changing lanes, while longitudinal decisions involve accelerating, decelerating, or maintaining speed. These actions are determined based on risk assessment and passed to the trajectory planner for vehicle guidance alongside the controller.

A. Observation Space

The configuration of the Markov Decision Process (MDP) [24] is defined by a tuple $\langle S_{ego}, S_i \rangle$ within the Frenet Frame [25]. The initial vector S_{ego} encompasses the attributes of the ego vehicle, while S_i comprises the state vector of the surrounding vehicles, denoted as the i^{th} vehicle. $S_{ego} = (v_{ego}, t_{headway}, \psi_{ego}, d_{centerlane}, lane_{ego}, risk_{max})$ includes parameters such as the ego vehicle's speed v_{ego}, the time headway to the leading vehicle $t_{headway}$, the orientation of the ego vehicle ψ_{ego}, the distance to the lane center $d_{centerlane}$, the lane occupancy $lane_{ego}$, and the maximum measured risk $risk_{max}$. Conversely, $S_{others} = (v_i, long_i, lat_i, \psi_i, lane_i, risk_i)$ comprises parameters about other vehicles, including their relative speed to the ego vehicle v_i, relative longitudinal distance $long_i$, relative lateral distance lat_i, relative orientation ψ_i, lane occupancy $lane_i$, and relative risk $risk_i$. In cases where the ego vehicle surpasses others in speed, v_i takes a negative value, and

conversely. Likewise, if another vehicle trails the ego vehicle, $long_i$ is negative, while it's positive if it leads. All normalizable states undergo normalization using the following equation:

$$X_{Normalized} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4}$$

Speed limits in urban areas are established according to traffic regulations in France. A selection of observation states is depicted in Fig. 4.

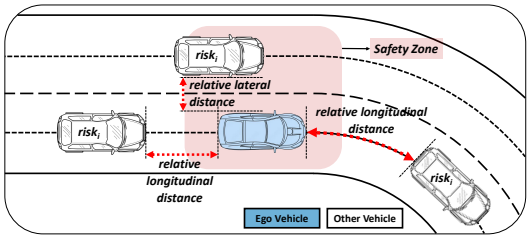

Fig. 4: Some observation states of the system and the safety zone occupancy of the ego vehicle.

B. Lateral Motion

The risk is measured on each vehicle within the perception range of the ego vehicle, then the maximum risk of each lane is assigned based on the maximum risk of the vehicle occupying it. Because of the nature of the proposed risk function, measuring the indirect alongside direct risk, the maximum risk imposed on the ego vehicle could be any of the other vehicles including the ones with no direct trajectory of collision at the time being. This is done through projecting the trajectory of the ego vehicle on the best path chosen, as shown in Fig. 1, and the trajectories of the other vehicles for a time horizon of 2 seconds then calculating the associated predicted risk. Therefore, as outlined in Fig. 5, the ego vehicle hierarchically makes the decision to either change lane or to keep its current lane where it prioritizes the initial risk measured before checking the projected risk in order to finalize its decision to change lanes. Moreover, this process is done every timestep which is equal to $0.125s$ which minimizes the direct impact (on the trajectory of the ego vehicle) of sporadic or irregular decisions during the transition phase towards a steady state when there are sudden changes in the environment prompting the risk to shift relatively quickly, this includes cases with sudden braking, accelerating, or a new obstacle entering the perception zone.

C. Deep Reinforcement Learning Adaptive Cruise Control

The DRL-ACC, extending our prior research [26], utilizes a risk-aware observation space and is coupled with the lateral risk aware decision module. Curriculum learning is adopted for the training methodology consisting of 3 increasingly difficult scenarios lasting betwen $30s$ and $60s$ on a straight 2-lane road that extends up to $1000m$. The initial conditions for the position and speed are randomly generated within the intervals $[1, 100]m$, and $[36, 54]km/h$ respectively and the starting distance headway between the vehicles varies between 15 and 40 meters. It is done on 5 seeds where the seed is randomly chosen. The first scenario contains no vehicles, the second and third scenarios contain low and high density traffic respectively. The behavior of other vehicles adheres to the Intelligent Driver Model (IDM) [27], governed by specific parameters described in Fig. I: a maximum acceleration of $2m/s^2$, a maximum comfort deceleration

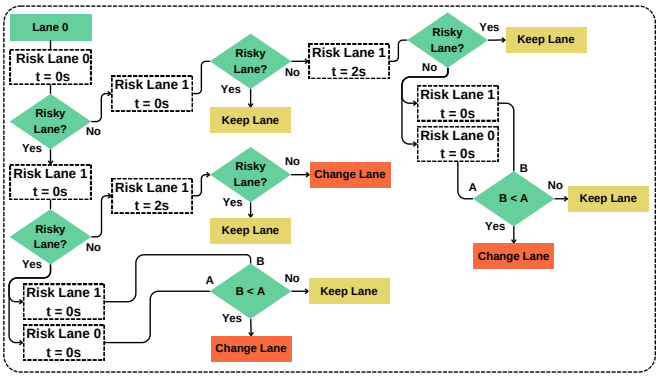

Fig. 5: Overview of the Risk-Aware Lateral Decision-Making.

of $-3m/s^2$, a desired velocity randomly selected from the range $[36, 54]km/h$, an acceleration exponent of 4, and desired distance and time headways of 3 meters and 1.8 seconds respectively. The progression of the ego vehicle can be noticed at the beginning of each scenario where the collision rate spikes. The model discounted long-term reward and the moving average reward $(windowsize = 50)$ converge after around 530 episodes, shown in Fig. 6, with a 96% success rate with no collisions and showcasing the maximization of speed towards convergence.

TABLE I: IDM Parameters [27]

Parameter	Range	Normal	Aggressive
(δ) acceleration exponent	{2, 4}	4	4
(s_{min}) min. desired distance gap	4.0-1.0 m	2.0	1.0
(v^*) desired velocity	54-140 km/h	57.6	64.8
(t_{gap}) desired time gap	1.8-1.0 s	1.5	1.0
(a_{max}) max. acceleration	1.0-2.0 m/s^2	1.4	2.0
(b_{comf}) comfort deceleration	1.0-3.0 m/s^2	2.0	3.0

1) Action Space

The action space, denoted as a_{speed}, comprises three actions: acceleration (a_{acc}), braking (a_{dec}), and maintaining the current speed (a_{hold}). The vehicle follows a predefined velocity profile derived from the desired velocity profile of the base frame, as described in [17]. This profile is computed for each point along the candidate path, taking into account factors like the speed limit ($V_{x\text{limit}}$) [28], velocity limits of the base frame, road curvature, and lateral acceleration. To ensure vehicle stability and passenger comfort, lateral acceleration is restricted to be below a maximum threshold of $|ay_{\max}| = 4m/s^2$, as specified in [29].

2) Reward Function

A comprehensive multi-objective reward function is developed to guide the behavior, emphasizing safety, collision avoidance, and car-following. This function utilizes reward shaping principles and includes safety (R_{safety}) and speed (R_{speed}) rewards, each incorporating weighted sub-rewards to balance agent bias. The reward function is formulated as follows:

$$\begin{cases} R_{safety} & = w_1 \cdot R_{front} + w_2 \cdot R_{back} + w_3 \cdot R_{collision} \\ R & = R_{safety} + w_4 \cdot R_{speed} \end{cases} \tag{5}$$

where w_i = [1, 0.5, 1, 0.35].

Safety R_{safety}: The safety reward is divided into three components: collision avoidance ($R_{collision}$), maintaining

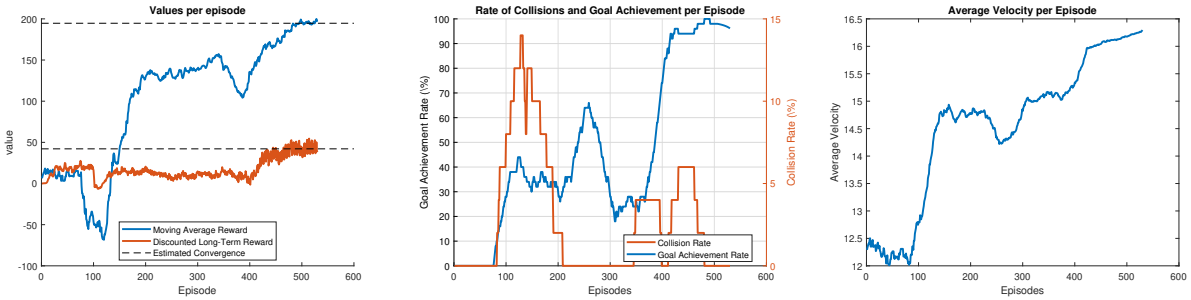

Fig. 6: Training Results for the DRL-ACC.

safe distances in front (R_{front}), and behind (R_{back}) of the ego vehicle in relation to other vehicles. The vehicle's state is categorized into three zones: risk, car-following, and catch-up. In risk and catch-up zones, rewards or penalties are determined based on the time headway and distance error. Calculations are based on the following equations:

$$\begin{cases} TH & = \frac{D}{V_{ego}} \ , \ DTH = 2 \\ DE & = D - D_{s0} + \frac{V_{ego}^2}{2a_{dec-max}} \end{cases} \tag{6}$$

Where $a_{dec-max}$ is the maximum deceleration, D_{s0} is a minimum safety gap and D is the distance between the two vehicles, V_{ego} is the ego vehicle velocity. Together, DE is the error with the safe-stop-distance.

In the risk zone, we apply the penalty P_1, and in the catch-up zone we apply the penalty P_2. The equations are shown below:

$$P_1 = -(k_1 \cdot (TH - DTH)^2 + k_2 \cdot DE^2) \tag{7}$$

$$P_2 = RB - (|D_{front}| \times k) \tag{8}$$

where $k_1 = 0.5$ and $k_2 = 0.5$ are weighting factors, $k = 1$ is the rate of the negative reward, $RB = -0.5$ is the Reward Baseline, D_{front} is the relative distance to the front vehicle.

Based on the relative speed of the ego vehicle, in the car-following zone we apply the reward-penalty SR_1, in the risk zone we apply the reward-penalty SR_2, and in the catch-up zone we apply the reward-penalty SR_3.

$$SR_1 = \begin{cases} \frac{V_i - (-0.05)}{0.05 - (-0.05)} & \text{if } -0.05 \leq V_i \leq 0.05 \\ -abs(V_i) \cdot k & \text{otherwise} \end{cases} \tag{9}$$

$$SR_2 = \begin{cases} \frac{V_i}{0.05} & \text{if } V_i \geq 0.05 \\ -abs(V_i) \cdot k & \text{otherwise} \end{cases} \tag{10}$$

$$SR_3 = \begin{cases} \frac{V_i - (-0.1)}{-0.05 - (-0.1)} & \text{if } -0.1 \leq V_i \leq -0.05 \\ -abs(V_i) \cdot k & \text{otherwise} \end{cases} \tag{11}$$

where V_i is the normalized relative speed, and $k = 2$ is the rate of the negative reward.

For $\alpha = DTH + 0.5$, and the above equations motivating the vehicle to either speed up, slow down, or hold its speed, the front gap reward would be:

$$R_{front} = \begin{cases} SR_1 + 1 & \text{if } a_{hold} \wedge ((DTH < TH \leq \alpha) \\ & \vee (0 \leq DE \leq 3)) \\ SR_2 + P_1 & \text{if } TH \leq DTH \vee DE < 0 \\ SR_3 + P_2 & \text{otherwise} \end{cases} \tag{12}$$

For the vehicles behind the ego vehicle, there is only a penalty for when they are in the risk zone:

$$R_{back} = \begin{cases} SR_3 + P_1 & \text{if } (TH \leq DTH) \vee (DE < 0) \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

If there are collisions, the vehicle receives penalties:

$$R_{collision} = \begin{cases} P_{collision} & \text{if } collision \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

Speed R_{speed}: The speed reward is based on the speed limits of 67 km/h and 36 km/h since we consider a common range in an urban environment [28]. This however does not limit the agent's decision to go beyond these limits.

$$R_{speed} = \begin{cases} \frac{SP - 0.5}{0.8 - 0.5} & \text{if } SP \leq 0.8 \\ -abs(SP) \cdot k & \text{otherwise} \end{cases} \tag{15}$$

where SP is the normalized speed of the ego vehicle.

D. Network Architecture

In our methodology, we utilize a Proximal Policy Optimization (PPO) agent instead of a Double Deep Q-Network (DDQN) as previously employed [26]. This decision stems from several factors. PPO exhibits adeptness in managing environments with high-dimensional state spaces, effectively processing and learning from intricate inputs, and providing inherent generalization capabilities. Additionally, the stability and data efficiency offered by PPO played a significant role in our decision-making process.

The agent's architecture encompasses a 2D convolutional layer configured with a 3x3 kernel and 128 filters, with a stride of 1 and padding set to 'same'. Following the convolution operation, a Rectified Linear Unit (ReLU) activation function is employed to introduce non-linearity, facilitating the extraction of complex features from the provided inputs. Subsequent operations entail a sequence of three fully connected layers. The first layer comprises 128 neurons, followed by a second layer with 64 neurons, and a third layer with 32 neurons, all supplemented by a ReLU activation function. The final layer accommodates a number of neurons equivalent to the size of the action space. The input to the architecture is the tuple $\langle S_{ego}, S_i \rangle$, while the output represents one of the actions defined previously. The hyperparameters specific to the PPO model are outlined in Table II.

E. Trajectory Planner

In this section, the development of the local trajectory planning algorithm is described. Our trajectory planner,

Hyperparameter	Value
Experience Horizon	512
Mini-Batch Size	128
Clip Factor	0.2
Advantage Estimate Method	GAE
GAE Factor	0.95
Optimizer	ADAM
Learning Rate	0.0005
Discount Factor	0.99

TABLE II: Hyperparameters of the PPO Agent

extending our prior research [17], is designed with a multi-step process to ensure comprehensive path generation and evaluation, guiding the vehicle to follow a reference trajectory while avoiding obstacles and prioritizing the safety and comfort of passengers. The navigation strategy encompasses multiple stages, including the generation of a set of candidate paths based on the risk aware decision input, determining the velocity profile, obstacle detection, classification of paths based on navigability, cost calculation, and the selection of the optimal path from this candidate set. The chosen trajectory is then designated as the desired path for execution by the vehicle in each planning iteration and is passed on to the risk aware decision making module to calculate the risk for the next time-step. The entire sequence of the path planning procedure is depicted in Fig. 7.

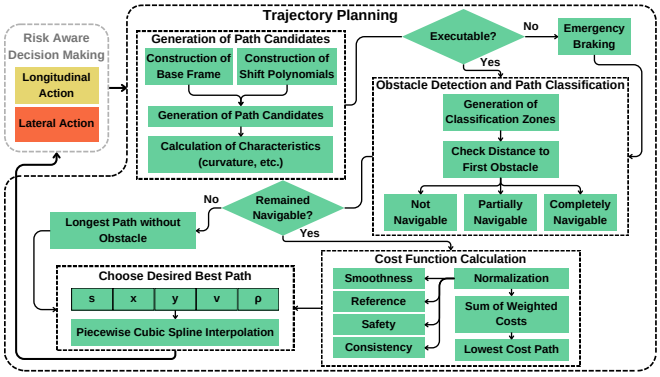

Fig. 7: Overview of the Trajectory Planner.

Initially, paths are created with a lateral offset, including transient and permanent phases. Candidate paths are determined by a fourth-order polynomial lateral shift, considering curvature limitations. Paths exceeding maximum curvature are discarded, selecting the longest viable alternative. Velocity profiles are derived from predefined patterns, considering speed limits, road curvature, and lateral acceleration. Executable candidate paths undergo obstacle detection using large and small circles, determining collision distance (d_{obs}) and obstacle type. Trajectories are classified based on collision distance into no-navigable, partially navigable, and navigable. A low-level decision algorithm selects the optimal trajectory based on the following criteria: smoothness, reference tracking, safety, and consistency. These criteria are combined into a total cost function, minimized to identify the best trajectory which is represented by points facilitating control/model block integration.

V. Simulation Results

This section presents the testing outcomes of the risk-aware system, demonstrating that the designed risk function effectively guided the vehicle toward desired behaviors, particularly in scenarios where it previously lacked consideration or encountered difficulties. It also showcases the ability of the DRL-ACC to maintain an acceptable velocity profile throughout the different situations encountered including car-following, overtaking, lane-changing, and dynamic re-planning of the trajectory. The testing scenarios include: 1.1) Dynamic Environment Lane Change, 2.1) Dynamic Trajectory Re-planning in Anticipated Risk Situations, and 1.2) and 2.2) which introduce curved sections to the road, simulating realistic situations obtained from the SCANeR studio simulator (Global Map in Fig. 8) for which the agent was not trained on.

Scenarios (1.1) and (1.2) involve randomly generated environments with 12 to 24 vehicles controlled by the IDM, allowing for acceleration, deceleration, and car-following behaviors. The ego vehicle is expected to adapt to this dynamic environment by executing various maneuvers such as car following, lane changing, and overtaking multiple vehicles while maintaining relatively high speeds safely and adhering to the speed limits. In contrast, scenarios (2.1) and (2.2) examine the agent's performance when the ego vehicle attempts to overtake a slow-moving vehicle while a speeding vehicle is approaching in the opposite lane either starting to accelerate within or outside the perception zone of the ego vehicle. In other words, the ego vehicle is either aware of the intention of the other vehicle because it can see it before it starts to accelerate, or the ego vehicle is not aware of the other vehicle beforehand as it enters the perception range already accelerating. This situation challenges the agent to assess environmental dynamics and risks from a high-level perspective, such as anticipating the danger posed by the speeding vehicle. The agent demonstrates its ability to make informed decisions, such as choosing not to execute a lane change despite having sufficient time and distance headway. Additionally, it showcases the agent's capability to dynamically adjust its trajectory by aborting a lane change decision if necessary.

These scenarios were tested on 10 different seeds than the ones trained on by the DRL-ACC. The testing conditions mentioned enable us to test several performance indicators considered recalled in Table III. Efficiency: the average velocity of the ego vehicle, and obstacles passed. Safety: the average time headway and distance error (TH and DE) and number of collisions. Comfort: the maximum lateral acceleration, maximum longitudinal deceleration, and lane-change frequency.

In scenario (1.1), both the risk-aware models (TP+RA and TP+RA+DRL-ACC) exhibit comparable efficiency and comfort levels to the model lacking risk-awareness (TP). However, they demonstrate superior safety by navigating the environment without collisions, contrasting with the four collisions experienced by the other model. This is further emphasized by their more conservative behavior, reflected in a lower average velocity and a tendency to follow slower vehicles rather than attempting aggressive lane changes. It's noteworthy that TP+RA+DRL-ACC shows particularly conservative behavior with a significantly low average velocity of $14.3183 m/s$ and a high time headway of $3.1983 s$, alongside fewer lane changes. Conversely, TP+RA displays higher lane change frequency and lateral acceleration, albeit at the cost of comfort compared to TP+RA+DRL-ACC.

In scenario (1.2), the scenario is more challenging with a non-linear road layout. Here, TP+RA+DRL-ACC exhibits even greater conservatism, manifesting in lower speeds, fewer obstacles passed, reduced lane change frequency, and longitudinal deceleration, along with a higher time headway. While TP+RA and TP perform similarly overall, TP+RA again proves safer, avoiding collisions and overtaking more obstacles compared to the other model.

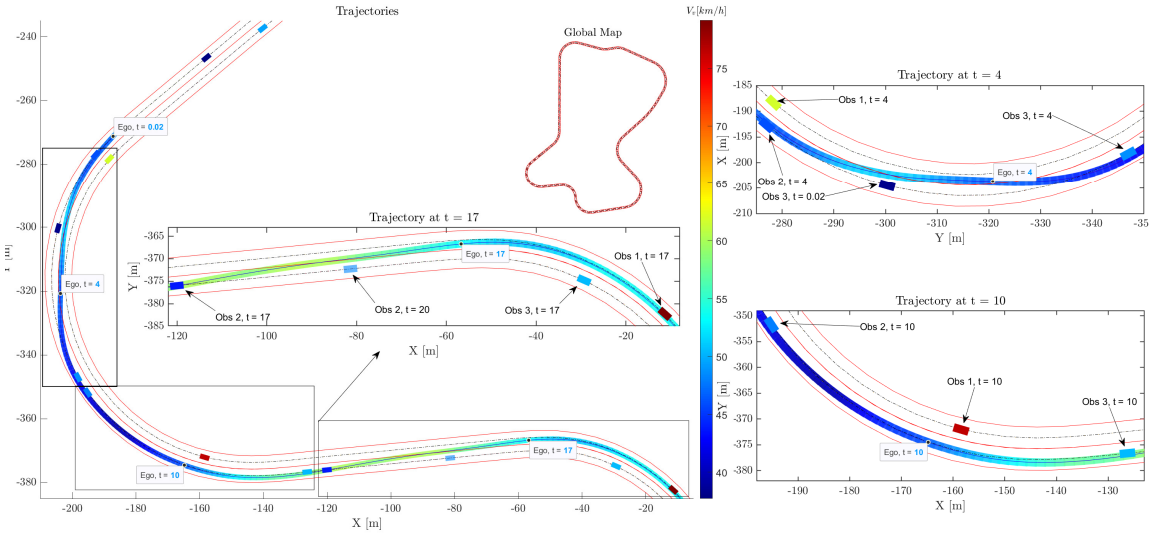

Fig. 8: Simulation Results: Trajectories and Vehicle Dynamics of Risk-Aware Lateral Decision-Making and DRL-ACC System.

TABLE III: Simulation Results.

		Efficiency		**Safety**					**Comfort**		
Model	**Scenario**	V_{avg} (km/h)	Obs_{passed} (s)	TH_{front}	DE_{front} (s)	TH_{back} (m)	DE_{back} (s)	Collisions (m)	$acc_{lat-max}$	$dec_{max-longi}$ (m/s^2)	LC_{avg} (m/s^2)
TP+RA	**(1.1)**	16.5957	**13**	2.1539	0.6435	1.7195	-0.6112	**0**	0.9501	-2.5347	4
TP	**(1.1)**	**17.9183**	11	2.1584	0.7173	1.8890	-0.7024	4	1.0070	-0.7568	2.7
TP+RA+DRL-ACC	**(1.1)**	14.3183	**13**	3.1982	0.7735	2.6138	-0.7339	**0**	**0.0396**	**-1.0118**	**1.2**
TP+RA	**(2.1)**	**17.0939**	10	1.9320	0.5540	2.2079	-0.6999	**0**	0.4426	**-0.0008**	1.9
TP	**(2.1)**	16.0949	9	1.8432	0.5176	1.7824	-0.6109	9	0.4328	-0.0522	0.9
TP+RA+DRL-ACC	**(2.1)**	15.2471	**12**	2.4056	0.6070	2.2773	-0.6644	**0**	**0.0114**	-1.2192	2
TP+RA	**(1.2)**	15.1265	**14**	2.2715	0.6603	2.2781	-0.6586	**0**	3.8998	-2.8706	2.3
TP	**(1.2)**	**15.3550**	10	1.9854	0.6325	2.3223	-0.6836	6	**3.8650**	-2.9601	1.3
TP+RA+DRL-ACC	**(1.2)**	13.8327	7	2.9487	0.7766	2.1820	-0.5928	**0**	4.1585	**-2.0551**	**1**
TP+RA	**(2.2)**	14.2937	**10**	2.1977	0.5812	2.6154	-0.7027	**0**	**2.1752**	-2.4086	2.2
TP	**(2.2)**	14.5980	8	1.9021	0.5356	2.6189	-0.7380	6	3.9503	-0.1952	0.9
TP+RA+DRL-ACC	**(2.2)**	**14.8645**	9	2.7622	0.8048	3.0694	-0.8711	**0**	4.8745	**-0.3863**	**1.1**

Moving to scenario (2.1), the TP model's lack of dynamic re-planning leads to a notably high failure rate due to limited environmental awareness. Introducing risk-aware decision making in TP+RA (and TP+RA+DRL-ACC) addresses this deficiency, excelling in all metrics and avoiding collisions altogether, rendering the TP model obsolete. Although TP+RA+DRL-ACC maintains a conservative stance with lower speeds and higher time headway, it surpasses TP+RA in the number of obstacles passed.

Similarly, in scenario (2.2), both TP+RA and TP+RA+DRL-ACC navigate without collisions, contrasting with the non-risk aware model. While their performances are comparable across most metrics, TP+RA+DRL-ACC exhibits superior comfort in terms of longitudinal deceleration and lane change frequency, whereas TP+RA exhibits in lateral acceleration.

For a detailed analysis, let's examine the behavior of the vehicle in the dynamic re-planning scenario depicted in Figures 8 and 9. Initially, the ego vehicle trails another vehicle (Obs 3) while leading another (Obs 2) behind it. Meanwhile, a vehicle outside its perception zone (Obs 1) approaches in the adjacent lane. At $t = 4s$, the ego vehicle is in the process of overtaking Obs 3, but suddenly, Obs 1 enters its perception zone at high speed, still accelerating. Recognizing the risk (above the threshold), the ego vehicle promptly aborts the maneuver, dynamically re-planning its trajectory to return to the safe lane by adjusting speed and following distance. By $t = 10s$, the ego vehicle adopts a car-following behavior, waiting to resume the overtaking maneuver while Obs 1 passes in the safety zone. Finally, at $t = 17s$, the ego vehicle successfully executes the lane change, accelerating afterward with Obs 1 positioned further ahead. Additionally, the agent maintains acceptable longitudinal speed and lateral acceleration. A video illustrating this scenario is available here.

These findings highlight the efficacy of the proposed risk assessment function and its seamless integration into the decision-making module across various velocity profile methods. Furthermore, they underscore the adaptability of the DRL-ACC system in responding to dynamic environ-

mental changes.

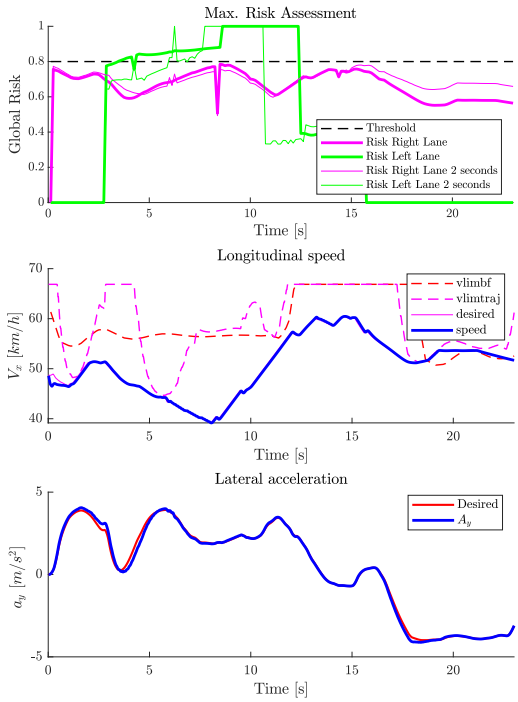

Fig. 9: Simulation Results: Risk and Vehicle Dynamics of Risk-Aware Lateral Decision-Making and DRL-ACC System.

VI. Conclusion

In conclusion, the integration of risk assessment into decision-making frameworks for autonomous vehicles represents a significant advancement in enhancing safety and adaptability in dynamic driving environments. By systematically evaluating risks, autonomous vehicles can make informed decisions to mitigate potential hazards and navigate complex scenarios more effectively. The formulation presented in this paper provides a transparent and systematic approach to risk-aware decision making, offering greater confidence in the behavior of autonomous systems. The function was able to properly assess the risk across 2 different speed profiles generated by the DRL-ACC [26] and the traditional trajectory planner [17]. Moving forward, continued research and development in this area will be crucial for realizing the full potential of risk assessment through extending the proposed function either by reshaping it to include other parameters within the environment, developing a more complex decision-making module, or deriving a probabilistic assessment.

References

[1] O.-R. A. D. O. Committee, *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, apr 2021.

[2] D. Ghraizi, R. Talj, and C. Francis, "An overview of decision-making in autonomous vehicles," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 10 971–10 983, 2023, 22nd IFAC World Congress.

[3] J. Nidamanuri, C. Nibhanupudi, R. Assfalg, and H. Venkataraman, "A progressive review: Emerging technologies for adas driven solutions," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 2, pp. 326–341, 2022.

[4] C. Zhang, J. Zhu, W. Wang, and J. Xi, "Spatiotemporal learning of multivehicle interaction patterns in lane-change scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6446–6459, 2022.

[5] Y. Ali, M. C. Bliemer, Z. Zheng, and M. M. Haque, "Cooperate or not? exploring drivers' interactions and response times to a lane-changing request in a connected environment," *Transportation Research Part C: Emerging Technologies*, vol. 120, p. 102816, 2020.

[6] Q. Chen, H. Huang, Y. Li, J. Lee, K. Long, R. Gu, and X. Zhai, "Modeling accident risks in different lane-changing behavioral patterns," *Analytic Methods in Accident Research*, vol. 30, p. 100159, 2021.

[7] G. Xu, L. Liu, Y. Ou, and Z. Song, "Dynamic modeling of driver control strategy of lane-change behavior and trajectory planning for collision prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1138–1155, 2012.

[8] W. M. D. Chia, S. L. Keoh, C. Goh, and C. Johnson, "Risk assessment methodologies for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 16 923–16 939, 2022.

[9] D. Deveaux, T. Higuchi, S. Uçar, C.-H. Wang, J. Härri, and O. Altintas, "Extraction of risk knowledge from time to collision variation in roundabouts," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021, pp. 3665–3672.

[10] L. Zeng, W. Hu, B. Zhang, Y. Wu, and D. Cao, "Risk-aware deep reinforcement learning for decision-making and planning of autonomous vehicles," in *2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 2022, pp. 1–6.

[11] H. Guo, K. Xie, and M. Keyvan-Ekbatani, "Modeling driver's evasive behavior during safety–critical lane changes: Two-dimensional time-to-collision and deep reinforcement learning," *Accident Analysis Prevention*, vol. 186, p. 107063, 2023.

[12] S. P. Venthuruthiyil and M. Chunchu, "Anticipated collision time (act): A two-dimensional surrogate safety indicator for trajectory-based proactive safety assessment," *Transportation Research Part C: Emerging Technologies*, vol. 139, p. 103655, 2022.

[13] H. Wang, B. Lu, J. Li, T. Liu, Y. Xing, C. Lv, D. Cao, J. Li, J. Zhang, and E. Hashemi, "Risk assessment and mitigation in local path planning for autonomous vehicles with lstm based predictive model," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 4, pp. 2738–2749, 2022.

[14] X. He, H. Yang, Z. Hu, and C. Lv, "Robust lane change decision making for autonomous vehicles: An observation adversarial reinforcement learning approach," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 184–193, 2023.

[15] S. Li, C. Wei, and Y. Wang, "Combining decision making and trajectory planning for lane changing using deep reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 16 110–16 136, 2022.

[16] D. Wang, W. Fu, Q. Song, and J. Zhou, "Potential risk assessment for safe driving of autonomous vehicles under occluded vision," *Scientific Reports*, vol. 12, no. 1, p. 4981, Mar 2022.

[17] A. Said, R. Talj, C. Francis, and H. Shraim, "Local trajectory planning for autonomous vehicle with static and dynamic obstacles avoidance," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021, pp. 410–416.

[18] P. F. Felzenszwalb and D. P. Huttenlocher, "Distance transforms of sampled functions," *Theory Comput.*, vol. 8, pp. 415–428, 2012.

[19] A. Chebly, R. Talj, and A. Charara, "Coupled longitudinal/lateral controllers for autonomous vehicles navigation, with experimental validation," *Control Engineering Practice*, vol. 88, pp. 79–96, 07 2019.

[20] E. Commission, "Road safety thematic report – speeding," European Road Safety Observatory, Brussels, Tech. Rep., 2021. [Online]. Available: https://road-safety.transport.ec.europa.eu/system/files/2021-07/road_safety_thematic_report_speeding.pdf

[21] X. Wu, X. Xing, J. Chen, Y. Shen, and L. Xiong, "Risk assessment method for driving scenarios of autonomous vehicles based on drivable area," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022, pp. 2206–2213.

[22] J. J. Gibson and L. E. Crooks, "A theoretical field-analysis of automobile-driving," *The American Journal of Psychology*, vol. 51, no. 3, pp. 453–471, 1938.

[23] V. V. Dixit, S. Chand, and D. J. Nair, "Autonomous vehicles: Disengagements, accidents and reaction times," *PLoS One*, vol. 11, no. 12, p. e0168054, 2016.

[24] M. van Otterlo and M. Wiering, *Reinforcement Learning and Markov Decision Processes*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 3–42.

[25] M. Werling, J. Ziegler, S. Kammel, and S. Thrun, "Optimal trajectory generation for dynamic street scenarios in a frenet frame," 06 2010, pp. 987 – 993.

[26] D. Ghraizi, R. Talj, and C. Francis, "A deep reinforcement learning decision-making approach for adaptive cruise control in autonomous vehicles," in *2023 21st International Conference on Advanced Robotics (ICAR)*, 2023, pp. 71–78.

[27] A. Kesting, M. Treiber, and D. Helbing, "Agents for traffic simulation," *Multi-Agent Systems: Simulation and Applications*, 05 2009.

[28] "Article r413-3 - code de la route - légifrance," accessed on: 19th July 2023, 17h43.

[29] R. Rajamani, *Vehicle Dynamics and Control*, ser. Mechanical Engineering Series. Springer US, 2014.