

Research Article

A Decision-Making Model for Self-Driving Vehicles Based on Overtaking Frequency

Mengyuan Huang , Shiwu Li , Mengzhu Guo , and Lihong Han 

School of Transportation, Jilin University, 5988 Renmin Street, Changchun 130022, China

Correspondence should be addressed to Shiwu Li; shiwu@jlu.edu.cn

Received 30 July 2021; Revised 18 November 2021; Accepted 29 November 2021; Published 14 December 2021

Academic Editor: Ren-Yong Guo

Copyright © 2021 Mengyuan Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The driving state of a self-driving vehicle represents an important component in the self-driving decision system. To ensure the safe and efficient driving state of a self-driving vehicle, the driving state of the self-driving vehicle needs to be evaluated quantitatively. In this paper, a driving state assessment method for the decision system of self-driving vehicles is proposed. First, a self-driving vehicle and surrounding vehicles are compared in terms of the overtaking frequency (OTF), and an OTF-based driving state evaluation algorithm is proposed considering the future driving efficiency. Next, a decision model based on the deep deterministic policy gradient (DDPG) algorithm and the proposed method is designed, and the driving state assessment method is integrated with the existing time-to-collision (TTC) and minimum safe distance. In addition, the reward function and multiple driving scenarios are designed so that the most efficient driving strategy at the current moment can be determined by optimal search under the condition of ensuring safety. Finally, the proposed decision model is verified by simulations in four three-lane highway scenarios. The simulation results show that the proposed decision model that integrates the self-driving vehicle driving state assessment method can help self-driving vehicles to drive safely and to maintain good maneuverability.

1. Introduction

With the significant increase in the feasibility of self-driving technology, decision systems that guarantee the safe and reliable driving state of self-driving vehicles and provide restricted information for their efficiency optimization have become a key affecting factor of the future industry. The driving state refers to the state of a vehicle's lateral velocity, longitudinal velocity, lateral acceleration, and longitudinal acceleration while travelling and the difference between vehicles travelling in adjacent lanes. In traffic flow, vehicles can make different decisions on driving behaviour based on a difference in the driving status compared to the surrounding vehicles. For human drivers, the basis of a pre-decisional judgment is mostly based on personal experience. However, for self-driving vehicles, a number of judgment criteria are required before making decisions. To realize the decision-making of self-driving vehicles in multiple lanes, it is necessary to implement evaluation methods that can assist self-driving vehicles in

judging the difference in the driving status between the ego vehicle and vehicles in other lanes.

There have been many studies on the decision-making of self-driving vehicles, and they have been mainly focused on longitudinal and lateral decision-making.

Considering the longitudinal driving decision-making, Zhu et al. [1] proposed a deep reinforcement learning-based framework for humanoid self-driving vehicles following planning to obtain an optimal policy from the aspects of speed, a relative speed of the vehicles in front and behind, vehicle spacing, and acceleration of the following vehicle in a human-like manner. Wei et al. [2] proposed a decision-making algorithm to assist self-driving vehicles under a single-lane uncertainty, considering the behaviour uncertainty of the in-front vehicles and the uncertainty in the environment perception accuracy, and achieved a significant improvement in system robustness. Ziegler et al. [3] proposed a method for planning the acceleration and deceleration maneuvers of autonomous vehicles during trajectory planning under the condition of deterministic driving

behaviour of surrounding vehicles. However, these decision-making methods consider only the state relationship between the vehicle in front of it and the vehicle and lack comparisons with vehicles in adjacent lanes.

In lateral decision-making, Gao et al. [4] proposed a reinforcement learning-based decision-making method for networked autonomous vehicles, which utilized the advantages of networked information to make autonomous driving decisions more effective. A number of studies have proposed different lane-changing decision-making methods for autonomous vehicles based on the game theory [5, 6]. For instance, Li et al. [7] proposed a game theory-based traffic model for testing and comparing various autonomous vehicle decision-making systems. The lane-changing decisions and lane-changing trajectories have been modeled by learning the driving behaviour of human drivers to build a humanoid lane-changing decision-making system [8].

Comprehensive studies have been conducted on decision-making behaviours, including vertical and horizontal decision-making behaviours. Khattak [9] preinstalled historical road performance data on the navigation map of self-driving vehicles and fused it with the vehicle's multisensor data to help drivers of self-driving vehicles and vehicles with a low automation level to make reasonable driving decisions. Zheng et al. [10] proposed an intelligent vehicle behaviour decision model based on driving risk assessment by analyzing the driver's driving characteristics and selecting safety and high efficiency as the two main factors that drivers pursue when driving. They establish a multiobjective optimal cost function for a decision model based on the least action principle. Hubmann et al. [11] considered the current and future interactions and uncertainties of vehicles and established a multiobjective optimal cost function for a decision-making model that can optimize autonomous driving behaviour under different future scenarios. Bahram et al. [12] proposed a combined optimization prediction-response-based driving strategy selection mechanism, which considered comfort in addition to ensuring the safety of autonomous vehicles. Rauskolb et al. [13] used a hybrid rule-based behavioural modeling approach to model an intelligent vehicle's behaviour decisions. However, this method does not consider differences in the driving state of the surrounding vehicles and its own lane.

Many autonomous driving decisions are based on algorithms designed for the Markov decision process [14, 15]. Zuo et al. [16] proposed a continuous reinforcement learning method that combines deep deterministic policy gradients with live demonstrations. This method accelerates the training process while learning more demonstrator preferences. In several studies, Q-learning and deep learning have been combined to design autonomous driving frameworks [17–19].

Part of the current research on decision-making for autonomous vehicles has been based on an actor-critic model. The DDPG algorithm has good convergence [20]. Based on the DDPG algorithm, Wang et al. [21] built a personalized autonomous driving system and designed driving decision-making methods according to different driving styles. In a multivehicle scenario, based on an actor-

critic learning approach, Xu et al. [22] established an actor-critic model as a decision model for autonomous driving. They used a value network to evaluate the current situation and a strategy network to make the next decision, respectively. By combining the two networks, an intelligent control model, which is in line with the human decision process, was developed.

In the existing autonomous driving following models, the speed of the vehicle in front and the distance between the vehicle and the vehicle in front of it are considered. The lane change model considers whether the driving states of the vehicle in front, the vehicle in front of the target lane, and the vehicle behind allow self-driving vehicles to perform the lane-changing behaviour. However, human drivers make the decision on changing their current state when vehicles in both lanes keep overtaking them or they keep overtaking vehicles in both lanes during the driving process. According to this idea, this paper makes an efficient and safe decision for integrated horizontal and vertical decision-making for autonomous vehicles, based on an OTF approach that considers the driving variability between the ego vehicle (EV) and the vehicles in both lanes.

To overcome the limitations and shortcomings of the existing work, this paper proposes an OTF-based driving state assessment method for autonomous vehicles and, based on this method, designs an autonomous driving decision model. The structure of this paper is shown in Figure 1. The main contribution of this paper is the development of a vehicle state assessment approach based on the OTF parameters to improve the accuracy of self-driving vehicles in judging the driving state, which can quantitatively and objectively measure whether a driving state of a three-lane scenario is appropriate. By using the OTF-based method and the DDPG algorithm, a decision model is established to obtain an optimal action-state by evaluating the combined efficiency of these strategies.

The rest of this paper is organized as follows. Section 2 presents a state evaluation method for self-driving vehicles applicable to three-lane traffic scenarios. Section 3 describes the decision process. Section 4 presents the simulation results of four typical scenarios. Section 5 gives the conclusion.

2. OTF-Based Driving-State Evaluation Approach for Self-Driving Vehicles

In the vehicle driving decision problem in high-speed scenarios, the influencing factors of self-driving vehicles mainly include the speed, position, and driving safety of a vehicle.

Before generating a decision, a self-driving vehicle needs to determine whether the traffic conditions in its current lane are consistent with the traffic conditions in the adjacent lanes and then can decide whether to change the driving status. Therefore, this paper establishes an OTF-based driving state assessment method for self-driving vehicles.

2.1. Description of the OTF-Based Driving-State Evaluation Approach. The consistency of a vehicle's driving state with the surrounding environment has a significant impact on

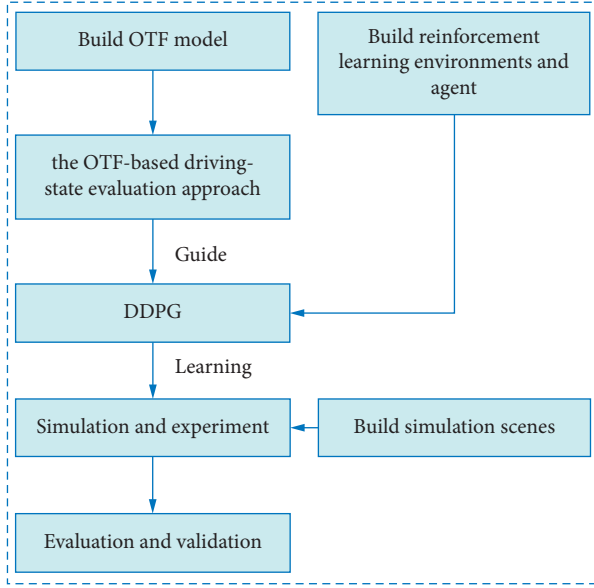


FIGURE 1: The framework of this paper.

whether EV intends to change the vehicle's driving behaviour. In this paper, the term "overtaking frequency" (OTF) is introduced to evaluate the difference between a self-driving vehicle and the surrounding vehicles. The OTF can be used to compare the EV's driving state with the driving states of vehicles in other lanes.

Since the OTF reflects the speed difference between vehicles in other lanes and an EV, it is related to the difference in the number of overtaking and overtaken vehicles on the two sides. The numbers of vehicles overtaking and being overtaken on the left and right sides can be, respectively, calculated by

$$\begin{cases} N_{al}(t) - N_{bl}(t) = N_{nl}(t), \\ N_{ar}(t) - N_{br}(t) = N_{nr}(t), \end{cases} \quad (1)$$

where $N_{al}(t)$ and $N_{ar}(t)$ denote the numbers of vehicles overtaken by an EV in the left and right lanes, respectively; $N_{bl}(t)$ and $N_{br}(t)$ are the numbers of vehicles overtaking the EV in the left and right lanes, respectively; $N_{nl}(t)$ is the difference in the numbers of vehicles overtaking and being overtaken in the left lane, and $N_{nr}(t)$ is the difference in the number of vehicles overtaking and being overtaken in the right lane.

In a three-lane scenario, the number of overtaking vehicles and the number of overtaken vehicles can be calculated as follows:

$$\begin{cases} N_{al}(t) + N_{ar}(t) = N_a(t), \\ N_{bl}(t) + N_{br}(t) = N_b(t), \end{cases} \quad (2)$$

where $N_a(t)$ is the total number of vehicles overtaken by an EV and $N_b(t)$ is the total number of vehicles overtaking the EV.

In this study, the OTF is defined as a difference between the numbers of vehicles overtaking and those being overtaken in the left and right lanes within a unit time interval, which can be expressed as follows:

$$\text{OTF}(t) = \frac{(N_{al}(t) + N_{ar}(t)) - (N_{bl}(t) + N_{br}(t))}{\delta} = \frac{N_a(t) - N_b(t)}{\delta}, \quad (3)$$

where δ is the unit time window.

According to equation (3), the OTF threshold is set to $[-a_o, a_o]$. Therefore, in the OTF evaluation function, when a vehicle's OTF belongs to $[-a_o, a_o]$, the autonomous vehicle's state represents a consistent speed, and when OTF is larger than a_o , the autonomous vehicle's state represents an excessive speed, and when OTF is less than $(-a_o)$, the autonomous vehicle's state represents an insufficient speed. This judgment framework is shown in Figure 2.

2.2. Time-Window Determination. When the time-window length varies, the OTF value in the corresponding time-window also varies. The time-window length is divided into different length ranges to investigate the OTF value ranges under different fixed time-window lengths.

For instance, when $\delta = 5$ s, the OTF is calculated by

$$\text{OTF}(t)|_{\delta=5} = \frac{\sum_{t_i}^{t_i+4} (N_a(t) - N_b(t))}{5}. \quad (4)$$

When $\delta = 10$ s, the OTF is calculated by

$$\text{OTF}(t)|_{\delta=10} = \frac{\sum_{t_i}^{t_i+9} (N_a(t) - N_b(t))}{10}. \quad (5)$$

Therefore, when $\delta = i'$ s, the OTF is calculated by

$$\text{OTF}(t)|_{\delta=i'} = \frac{\sum_{t_i}^{t_i+i'-1} (N_a(t) - N_b(t))}{i'}. \quad (6)$$

In this paper, 1 s is used as a time-window step for each car overtaking an EV. When $\delta_u = 1$ s, the OTF value changes every 1 s, resulting in a dynamic OTF change. When $\delta = i'$ s, the OTF value is calculated for the time-window length of $(N_a(t) - N_b(t))$ in the following time periods, and the division of the time-window under dynamic changes is as follows:

$$\begin{cases} t_1 - t_{i'}, \\ t_2 - t_{i'+1}, \\ t_3 - t_{i'+2}, \\ \vdots \\ t_i - t_{i+i'-1}. \end{cases} \quad (7)$$

2.3. OTF Value Range Analysis. To study the changing trend of the OTF value, a hardware platform consisting of a millimeter-wave radar and GPS device was built to obtain the OTF [23]. The test was conducted using real vehicles with four different speeds, and the obtained data were processed and analyzed under different time-window lengths.

A 1 s time-window step was used to study the OTF values under different time-window lengths. The OTF values for different time-window lengths at four vehicle speeds also

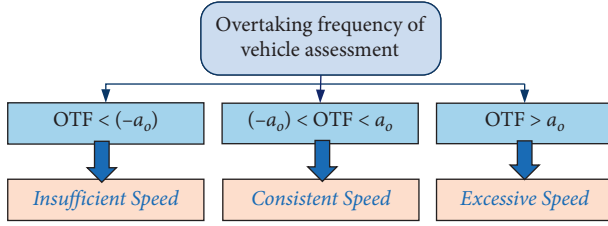


FIGURE 2: The diagram of the OTF-based automatic-vehicle driving status assessment framework.

differed. The time-window lengths were 10 s, 20 s, 30 s, and 60 s. The results are shown in Figure 3.

As shown in Figure 3, at the same vehicle speed, the ranges of OTF values were similar for different time-window lengths. Also, the OTF values showed a large concentration of results for a time-window of 60 s. The ranges of OTF values at different vehicle speeds are given in Table 1.

3. Decision-Making Process

In this study, a learning approach is used to obtain an optimal decision. In this section, a decision-making model based on the policy gradient algorithm is introduced. In this algorithm, the vehicle driving state evaluation function based on the OTF is defined, and reward functions in different scenarios are designed, which can reflect the driving difference between self-driving vehicles and other vehicles. The optimal action-state is searched by the optimal search.

3.1. DDPG. The DDPG [24] represents a realization of the DeepMind research team constructed using the DQN to extend the Q-learning algorithm and a deep neural network to approximate the state-behaviour value function and the deterministic strategy.

The DDPG algorithm separately parameterizes the critic function $Q^\mu(s, a|\theta^Q)$ and the actor function $\mu(s|\theta^\mu)$, where θ^Q and θ^μ are the weight parameters. The critic function is defined by equations (9) and (10), and it is updated by minimization.

$$W(Q) = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i))^2, \quad (8)$$

$$y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'}).$$

The actor function maps the current state to the current best action, and it is updated by

$$\nabla_{\theta^\mu} K \approx \frac{1}{N} \sum_i Q(s, a|\theta^Q)|_{s=S_i, a=\mu(s_i)} \nabla_{\theta^\mu} \mu(s|\theta^\mu)|_{s_i}, \quad (9)$$

where i is the number of training steps, and N is the total number of training steps; Q and μ denote the functions of critics and actors, respectively; Q' and μ' denote the critic and actor functions of the target network, respectively.

Finally, the target network copies the original network's parameters according to the delay factor τ to perform the update by

$$\theta^{Q'} = \tau \theta^Q + (1 - \tau) \theta^{Q'}, \quad (10)$$

$$\theta^{\mu'} = \tau \theta^\mu + (1 - \tau) \theta^{\mu'}. \quad (11)$$

3.2. OTF-Based Self-Driving Vehicle Decision-Making Process in Different Scenarios. When making decisions for autonomous driving, different decisions need to be made according to certain scenarios. According to the OTF-based driving state evaluation function presented in Section 2, this paper establishes four typical scenarios of autonomous vehicle decision-making methods.

This paper discusses the following scenarios in a three-lane highway scenario.

3.2.1. No Cars in front of the EV. When $\text{OTF} \in [-a_o, a_o]$, the driving status of an EV is similar to the driving status of vehicles in the two adjacent lanes.

When $\text{OTF} > a_o$, an EV's driving speed is significantly higher than the driving speeds of vehicles in the adjacent lane, and the EV's state represents the excessive-speed state. Since the EV's status can be assessed as too fast, the EV can make the OTF value meet the OTF threshold range by slowing down and performing the other related actions.

When $\text{OTF} < -a_o$, the EV's state is the insufficient-speed state. Since there is no blocking from the EV in the front, the EV should accelerate to achieve that its OTF is within the threshold range.

The OTF-based self-driving vehicle decision-making process in scenario (1) is shown in Figure 4.

3.2.2. Sudden Insertion of Other Vehicle (OV) from the Adjacent Lane: Consider the Example of the OV Insertion from the Adjacent Left Lane. When $\text{OTF} \in [-a_o, a_o]$, although the EV's state is consistent, the sudden insertion of an OV from the adjacent left lane causes the efficiencies of the left lane and this lane to decrease. If the inserting OV completes the lane-changing action within the unit time-window, the EV first performs the deceleration action and then accelerates until the OTF meets the threshold. If the OV does not complete the lane-changing action within the unit time-window, the EV should consider performing the change to the adjacent right lane to reach the optimal efficiency.

When $\text{OTF} > a_o$, the vehicle can make the OTF meet the threshold range by slowing down and performing the other actions while ensuring the safety distance.

When $\text{OTF} < -a_o$, although the EV's state is the insufficient speed, the OTF threshold cannot be reached by the acceleration maneuver due to the insertion of the OV in the adjacent left lane. Therefore, it is necessary to consider performing a switch to the right lane to achieve optimal efficiency. The same applies to the adjacent right lane with vehicle insertion.

The OTF-based self-driving vehicle decision-making process in scenario (2) is shown in Figure 5.

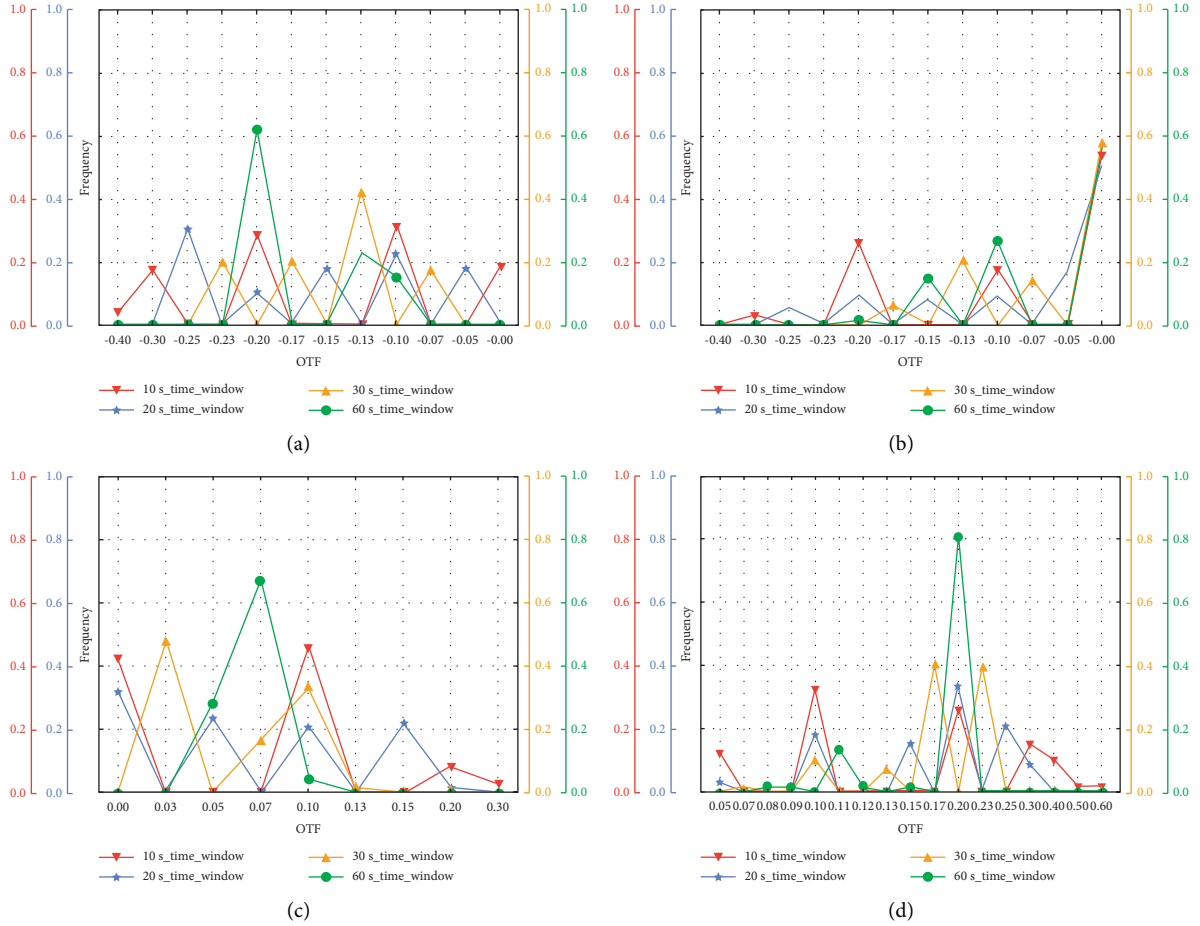


FIGURE 3: The OTF values for different time-window lengths at four vehicle speeds: (a) 60 km/h; (b) 80 km/h; (c) 100 km/h; (d) 120 km/h.

TABLE 1: The OTF value ranges at four vehicle speeds.

Vehicle speed (km/h)	OTF value range
60	$[(-0.40)-0.00]$
80	$[(-0.30)-0.20]$
100	$[0.00-0.30]$
120	$[0.05-0.60]$

3.2.3. Sudden Braking of the Vehicle in Front. When $OTF \in [-a_o, a_o]$, due to the sudden braking of the car in the front, an EV should consider changing lanes and performing the other actions to achieve the best efficiency under these conditions while ensuring a safe distance with the car in front.

When $OTF > a_o$, the EV's state is the excessive speed. Thus, the EV can make the OTF reach the threshold range by deceleration while ensuring the safety distance at the same time.

When $OTF < -a_o$, the EV's state represents the insufficient speed. Therefore, it is impossible to reach the OTF threshold by acceleration action due to the sudden braking of the car in front, so it is necessary to consider executing lane change action.

The OTF-based self-driving vehicle decision-making process in scenario (3) is shown in Figure 6.

3.2.4. The Vehicle in Front Changes Lane: Consider the Example of the Front Car Changing Lanes to the Adjacent Left Lane. When $OTF \in [-a_o, a_o]$, since the vehicle in front executes a lane-changing maneuver, under normal circumstances, the EV will decelerate in its own lane. Therefore, the EV should consider performing a right-lane action to achieve optimal efficiency while ensuring a safe distance from the vehicle in front.

When $OTF > a_o$, the EV can make the OTF reach the threshold range by deceleration action while ensuring the safety distance at the same time.

When $OTF < -a_o$, if the lane-changing action of the front car is completed within a unit time-window, the strategy of the EV after the lane change of the front car can be referred to case (1). If the front car is unable to complete the lane-changing process within the time-window due to the factor of its target lane and the EV cannot accelerate to within the threshold value, the execution of the lane-changing in the right lane should be considered.

The OTF-based self-driving vehicle decision-making process in scenario (4) is shown in Figure 7.

3.3. Reward Function Design. The measurement of a policy depends on the cumulative reward received by an intelligent body after executing a policy for a long period of time. Since the most important issues to be considered for intelligent

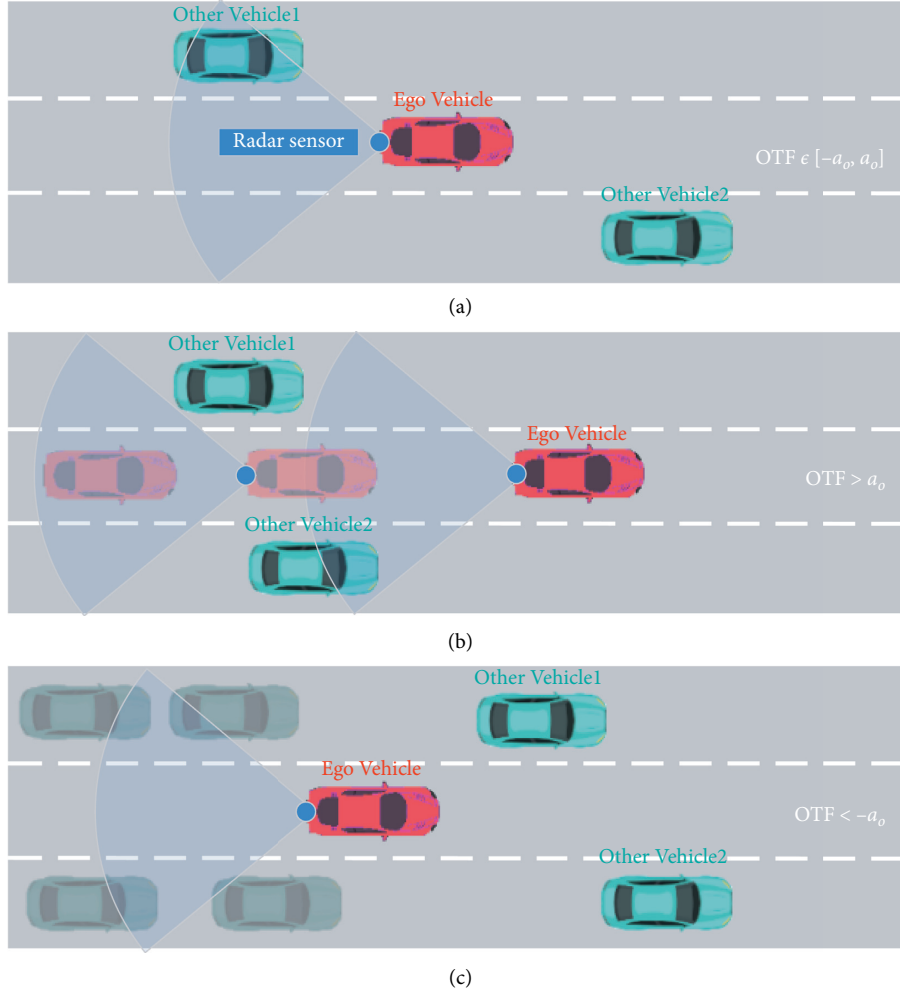


FIGURE 4: The OTF-based self-driving vehicle decision-making process in scenario (1): (a) $OTF \in [-a_o, a_o]$; (b) $OTF > a_o$; (c) $OTF < -a_o$.

vehicles are safety and timeliness, these two aspects should be considered when designing the reward function.

The timeliness is mainly shown by two aspects, OTF function and vehicle speed. Therefore, in this paper, the OTF reward and vehicle speed reward are established separately.

In the OTF-based speed suitability assessment model, the output results in a positive reward value for consistent speed, and the reward value for the rest of the cases is zero. The OTF reward r_o can be defined as follows:

$$r_o = \begin{cases} 1 & OTF(t) = 0, \\ \frac{1}{|OTF(t)|} & -a_o \leq OTF(t) < 0, 0 < OTF(t) \leq a_o, \\ 0 & OTF(t) > a_o, OTF(t) < -a_o. \end{cases} \quad (12)$$

Based on the results in Section 2, a_o was set to 0.05.

Outside the OTF constraint, smart vehicles travel at faster speeds, which is beneficial for timeliness. In their speed constraint, the speed vehicle should be as fast as possible. Therefore, the speed reward can be defined as follows:

$$r_v = \frac{v_E}{v_{\max}}, \quad (13)$$

where v_E is the current vehicle speed and v_{\max} is the maximum vehicle speed in the current lane.

The safety of a smart vehicle is related to the state of a vehicle in front of it, so setting the safety reward is determined by the time-to-collision (TTC) and a relative distance between the two vehicles D [25]. The TTC value is calculated by

$$t_T = \frac{x_F - x_E}{v_E - v_F}, \quad (14)$$

where x_F denotes the longitudinal position of the front vehicle, x_E is the longitudinal position of the EV, v_E is the EV speed, and v_F is the speed of the front vehicle.

The reward r_T for TTC is expressed as

$$r_T = \begin{cases} 0 & t_T < t_{T_{\min}}, \\ 1 & t_T \geq t_{T_{\min}}, \end{cases} \quad (15)$$

where $t_{T_{\min}}$ denotes the minimum threshold of t_T [26]. When the calculated value of t_T is infinite, i.e., when the speeds of the two cars are equal, the reward value is one.

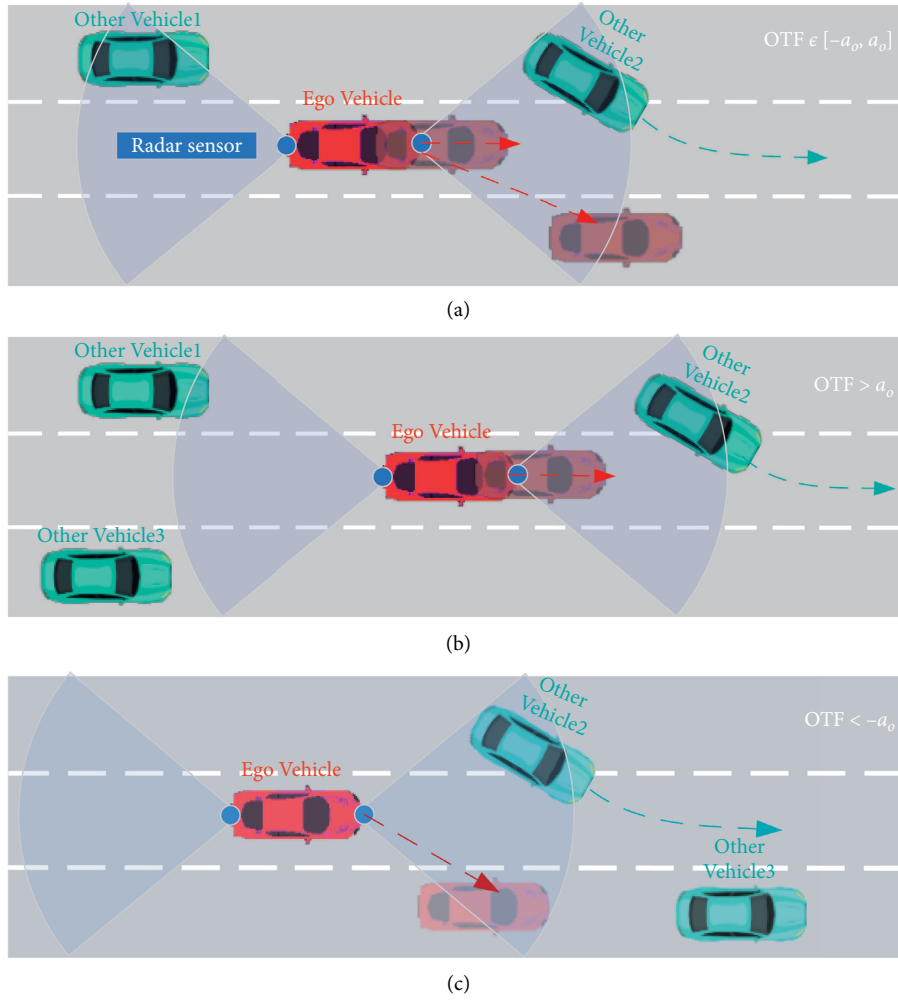


FIGURE 5: The OTF-based self-driving vehicle decision-making process in scenario (1): (a) $OTF \in [-a_o, a_o]$; (b) $OTF > a_o$; (c) $OTF < -a_o$.

However, when the relative distance between the two cars is less than the minimum safe distance, the reward value is set to negative infinity, which can be expressed as

$$r_d = -\infty \quad \Delta d < d_{D_{\min}}. \quad (16)$$

Therefore, the accumulated total reward value can be calculated by

$$r = r_o + r_v + r_T + r_d. \quad (17)$$

4. Simulation and Validation

In order to verify the effectiveness of the proposed decision algorithm, the reinforcement learning framework provided by MATLAB was used, and four complex high-speed scenes were constructed as an experimental environment.

In the experiment, the reinforcement learning elements, including actions, states, and rewards, were implemented. The OTF-based vehicle DDPG algorithm was used for vehicle driving behaviour decision-making. The control variables (front-wheel angle and acceleration) were output by the neural network in the DDPG. In this section, the three-degree-of-

freedom vehicle dynamics model in Simulink responds to the control variables, and finally outputs the EV's own state variables lateral velocity v_y , longitudinal velocity v_x , and yaw angle ω . The exact flow of the simulation is shown in Figure 8.

The selected high-speed scene was a one-way three-lane scene, and the state space S was the location and motion information of the surrounding 10 vehicles, including the vehicle under test. The vehicles in the scene were free and randomly selected actions. The parameters of the environmental model are shown in Table 2.

The vehicle movement space included: left lane-changing, driving straight ahead, right lane-changing, acceleration, and deceleration. The specific training process refers to the algorithm described in Section 3.1, where the maximum number of epochs in the training phase was set to 10,000.

The tests were divided into four scenarios, and the OTF-based driving behaviour decision model for autonomous vehicles presented in Section 3 was validated in each of the scenarios.

4.1. Scenario (1): No Cars in front of the EV. The results of the self-driving vehicle using the OTF-based driving behaviour decision model in Scenario 1 are displayed in Figure 9. As

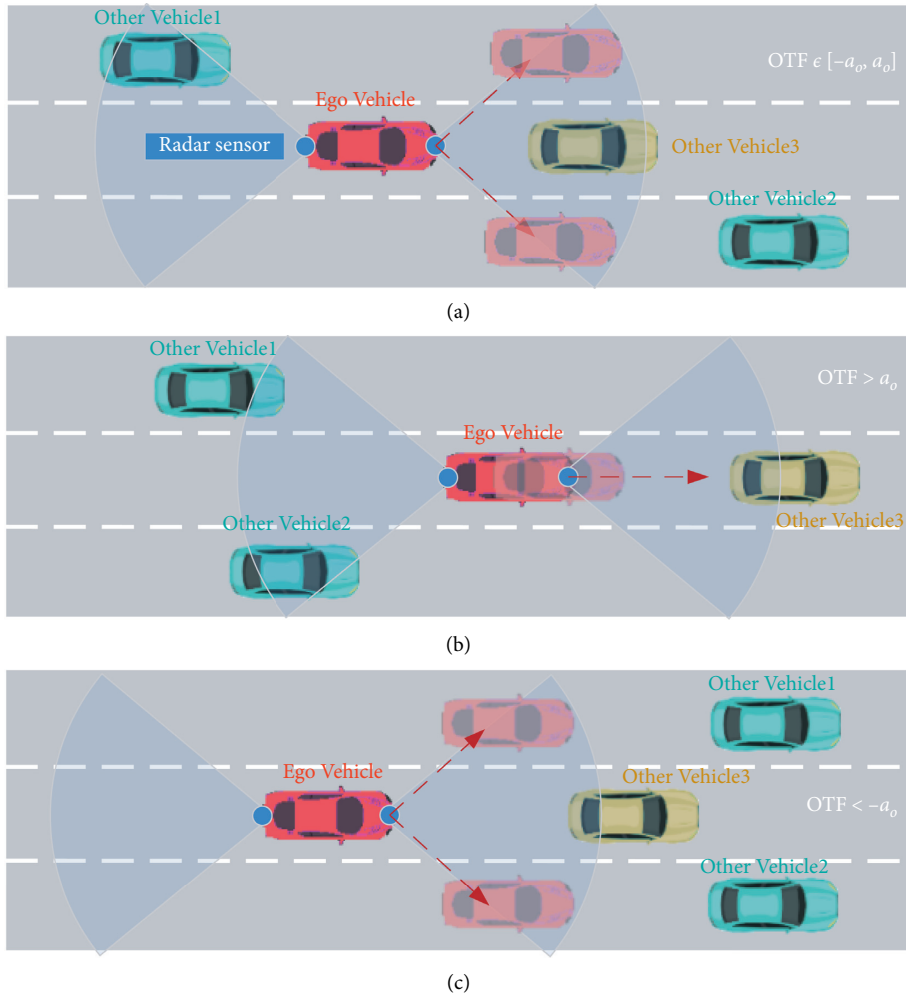


FIGURE 6: The OTF-based self-driving vehicle decision-making process in scenario (1): (a) OTF $\in [-a_o, a_o]$; (b) OTF $> a_o$; (c) OTF $< -a_o$.

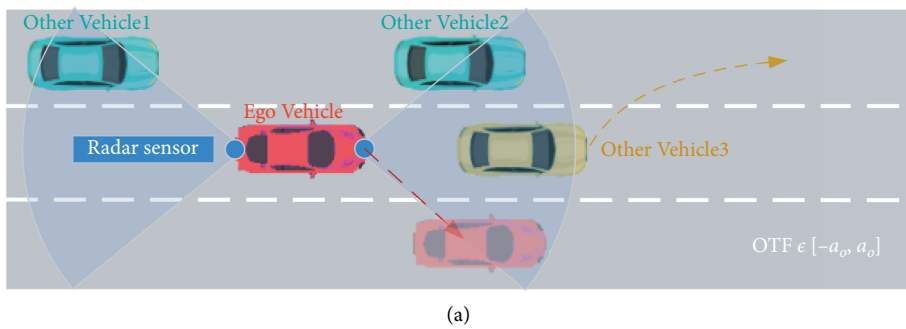


FIGURE 7: Continued.

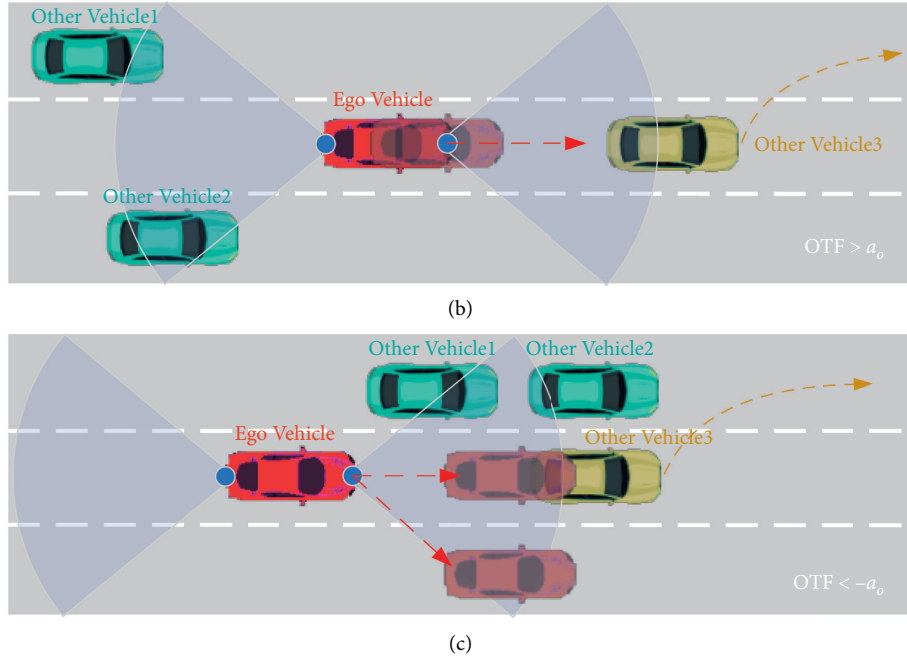


FIGURE 7: The OTF-based self-driving vehicle decision-making process in scenario (1): (a) $OTF \in [-a_o, a_o]$; (b) $OTF > a_o$; (c) $OTF < -a_o$.

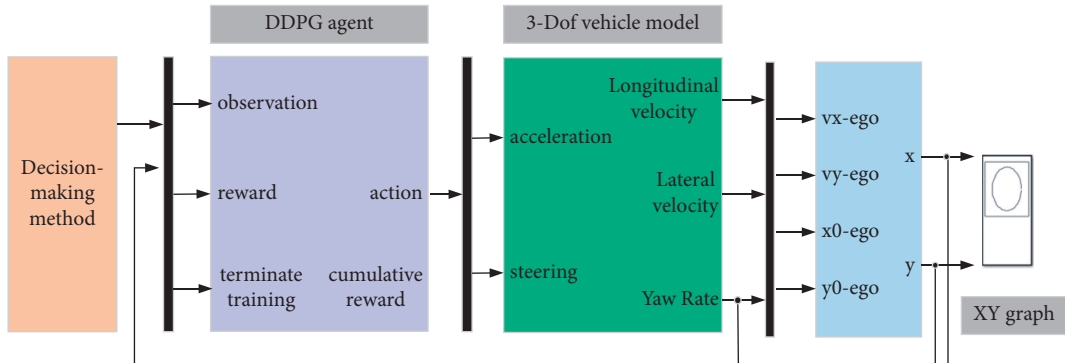


FIGURE 8: The simulation process.

TABLE 2: Simulation parameters.

Variable	Value	Description
L_t	250 m	Total road length
W_s	3.75 m	Single lane width
N_O	10	Total number of other vehicles
W_V	1.8 m	Vehicle width
v_{iE}	22.2/27.7 m/s	Initial speed of the EV

shown in Figure 9(a), the EV was driving in the middle lane, and the EV's driving decision in this scenario was lane-keeping. Since the current driving lane was not blocked by the car in front, the EV performed the acceleration action, as shown in Figures 9(b) and 9(d) in order to improve the driving efficiency. As shown in Figure 9(c), the training result finally converged.

4.2. Scenario (2): Sudden Insertion of the OV from the Adjacent Lane. The results of the autonomous vehicle using the OTF-based driving behaviour decision model in Scenario 2 are

presented in Figure 10. As shown in Figure 10(a), the EV's driving decision in this scenario was lane-keeping. During the simulation process, one of the OVs in the adjacent left lane was changing the lane to the EV's driving lane, and the lane-changing process of the OV was completed in 3 s. Therefore, as shown in Figures 10(b) and 10(d), the EV performed the deceleration action first and then accelerated after the OV's lane change maneuver ended. The training results in Figure 10(c) show that the training results eventually converged.

4.3. Scenario (3): Sudden Braking of the Vehicle in Front.

As shown in Figure 11(a), the decision of the EV was to change lanes to the left. Due to the sudden braking of the vehicle in front, the EV performed the deceleration action first, as shown in Figures 11(b) and 11(d), to ensure driving safety. Then, to drive more efficiently and obtain a larger reward value, the EV performed the lane-changing maneuver. As shown in Figure 11(c), the training results finally converged.

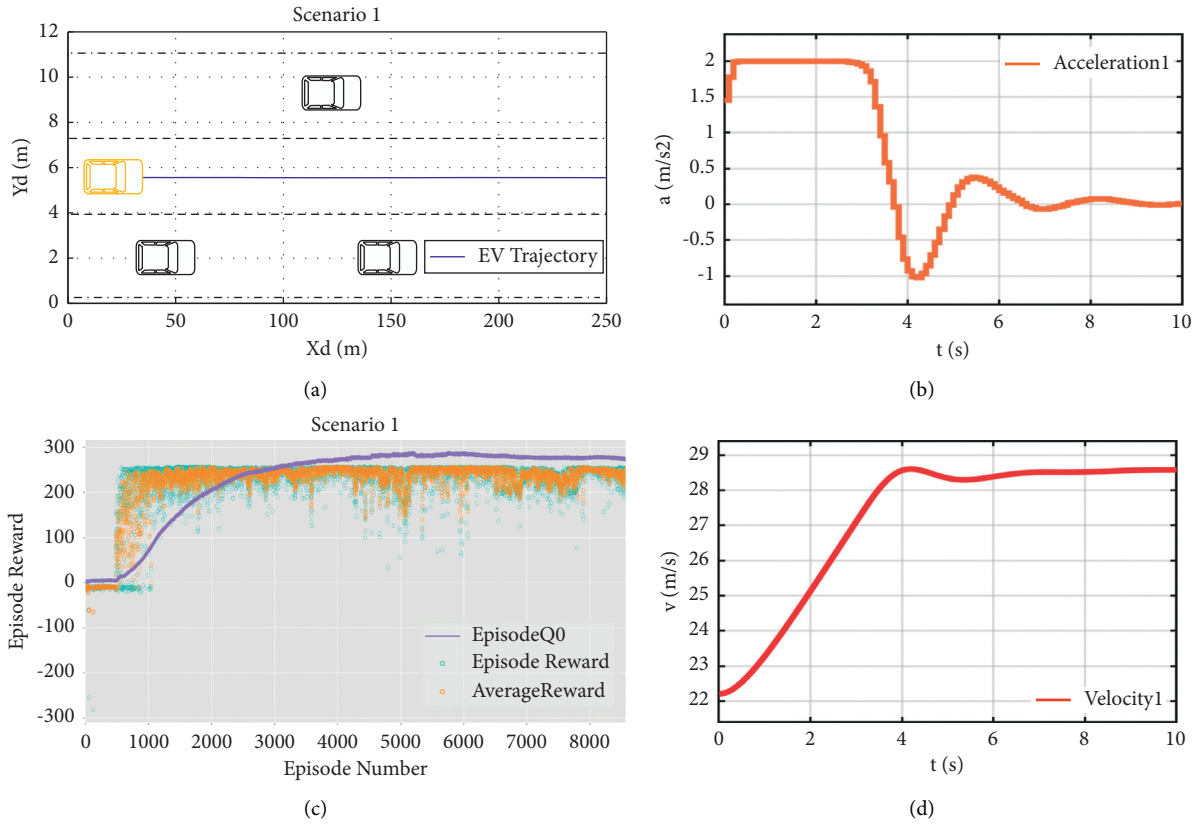


FIGURE 9: Training and simulation results of scenario 1. (a) EV trajectory; (b) EV acceleration; (c) training results; (d) EV velocity.

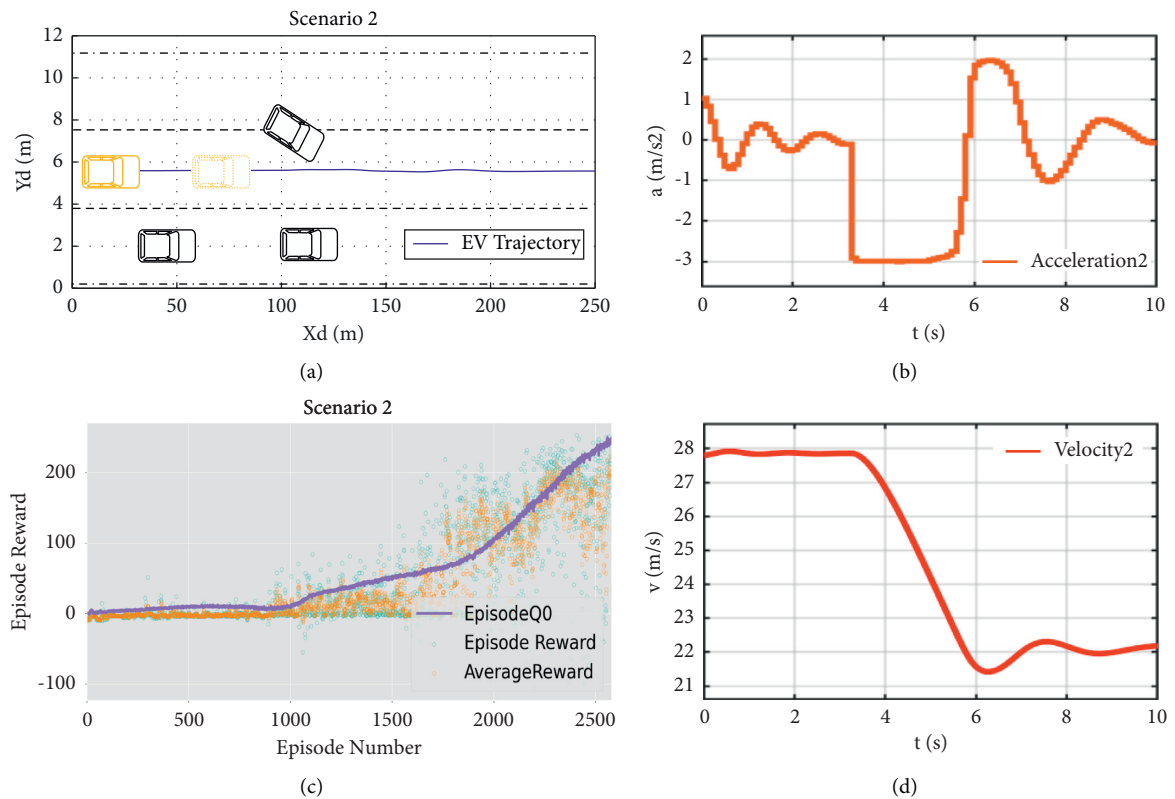


FIGURE 10: Training and simulation results of scenario 2. (a) EV trajectory; (b) EV acceleration; (c) training results; (d) EV velocity.

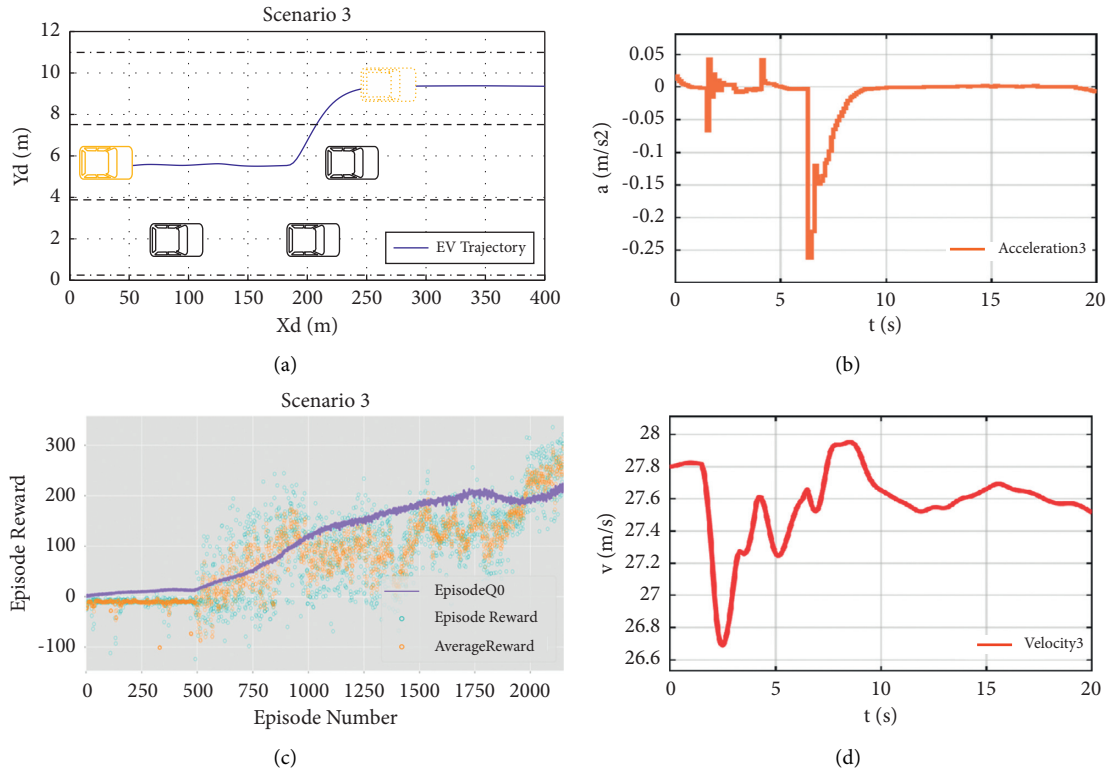


FIGURE 11: Training and simulation results of scenario 3. (a) EV trajectory; (b) EV acceleration; (c) training results; (d) EV velocity.

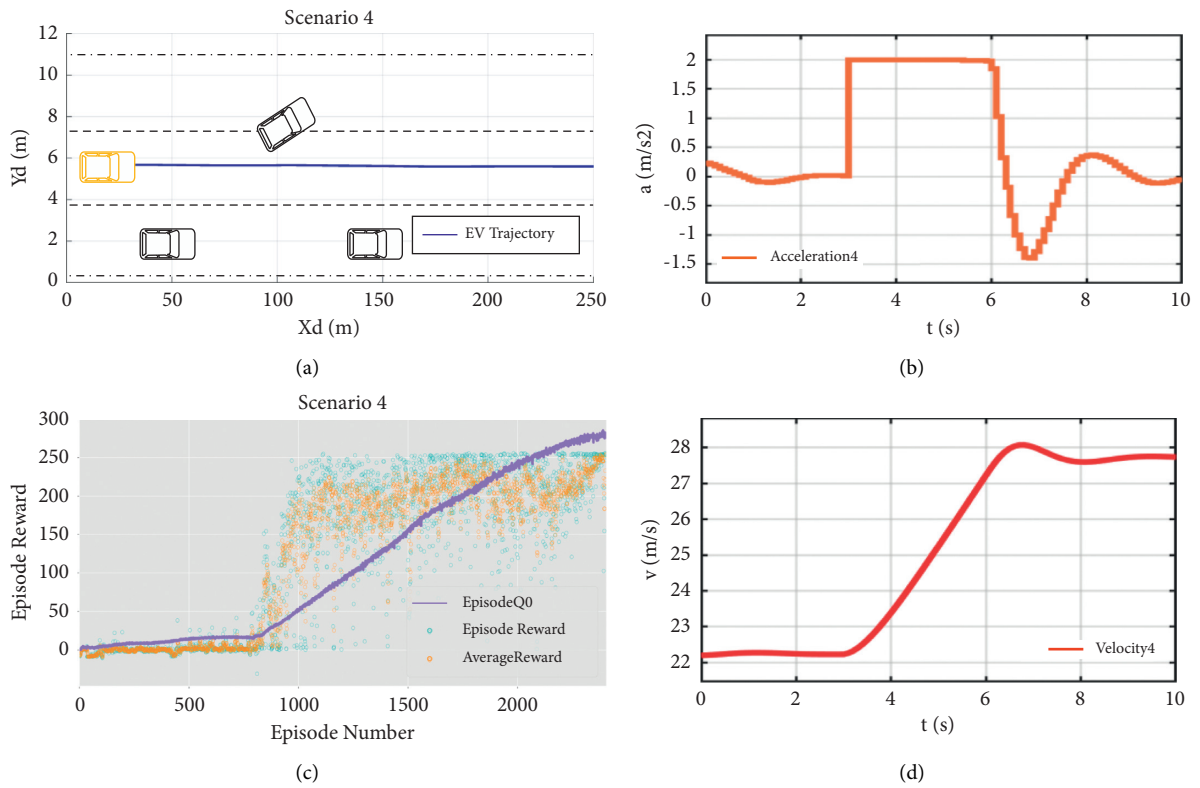


FIGURE 12: Training and simulation results of scenario 4. (a) EV trajectory; (b) EV acceleration; (c) training results; (d) EV velocity.

4.4. Scenario (4): The Vehicle in Front Changed Lanes to the Adjacent Lane. As shown in Figure 12(a), the EV's decision in this scenario was lane-keeping. During the simulation, since the lane-changing maneuver of the car in front was completed in 3 s, there was no obstruction in front of the EV after the front car left the lane. Therefore, the EV performed the acceleration action after the front car changed the lane, as shown in Figures 12(b) and 12(d). As shown in Figure 12(c), the training results converged.

5. Conclusions

In this paper, a method based on overtaking frequency is proposed for solving the autonomous decision-making problem of self-driving vehicles in highway scenarios. The degree of difference in the driving state between the self-driving vehicle and surrounding vehicles is evaluated. This difference is quantified by the evaluation method of the driving state of autonomous vehicles based on the overtaking frequency. With the assistance of this evaluation method, a decision-making model based on the DDPG is established. A driving decision-making method of self-driving vehicles based on the overtaking frequency in different typical scenarios is designed to make self-driving decisions more efficient and reasonable. The proposed model is verified by simulations, and simulation results prove the applicability and effectiveness of the decision-making model in four typical driving scenarios. The method can provide a theoretical basis for further research in uncertainty decision-making. However, whether the application scenarios of the algorithm have broad applicability remains to be studied in the future research. In the future research, the training amount of the model will be further increased, and the application of the decision model will be expanded.

Data Availability

The data presented in this study are available on request from the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

This work was supported by National Natural Science Foundation Item (Grant nos. 52172385 and 52002143).

References

- [1] M. Zhu, X. Wang, and Y. Wang, "Human-like autonomous car-following model with deep reinforcement learning," *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 348–368, 2018.
- [2] J. Wei, J. M. Dolan, J. M. Snider, and B. Litkouhi, "A point-based MDP for robust single-lane autonomous driving behavior under uncertainties," in *Proceedings of the IEEE International Conference on Robotics & Automation*, pp. 8–20, IEEE, Shanghai, China, 9 May 2011.
- [3] J. Ziegler, P. Bender, M. Schreiber et al., "Making bertha drive an autonomous journey on a historic route," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 2, pp. 8–20, 2014.
- [4] H. Gao, Z. Wu, and K. Qiu, "Driving polices of V2X autonomous vehicles based on reinforcement learning methods," *IET Intelligent Transport Systems*, vol. 14, no. 7540, pp. 8–20, 2020.
- [5] T. Alireza, H. S. Mahmassani, and S. H. Hamdar, "Modeling lane-changing behavior in a connected environment: a game theory approach," *Transportation Research Procedia*, vol. 59, pp. 216–232, 2015.
- [6] W. Daamen, M. Wang, S. P. Hoogendoorn, B. van Arem, and R. Happee, "Game theoretic approach for predictive lane-changing and car-following control," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 73–92, 2015.
- [7] N. Li, D. W. Oyler, M. Zhang, Y. Yildiz, I. Kolmanovsky, and A. R. Girard, "Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 5, pp. 1782–1797, 2018.
- [8] L. Li, K. Ota, and M. Dong, "Humanlike driving: empirical decision-making system for autonomous vehicles," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 6814–6823, 2018.
- [9] J. L. A. Khattak, "Informed decision-making by integrating historical on-road driving performance data in high-resolution maps for connected and automated vehicles," *Journal of Intelligent Transportation Systems*, vol. 24, pp. 11–23, 2020.
- [10] X. Zheng, H. Huang, J. Wang, X. Zhao, and Q. Xu, "Behavioral decision-making model of the intelligent vehicle based on driving risk assessment," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 3, pp. 820–837, 2021.
- [11] C. Hubmann, M. Becker, D. Althoff, D. Lenz, and C. Stiller, "Decision making for autonomous driving considering interaction and uncertain prediction of surrounding vehicles," in *Proceedings of the Intelligent Vehicles Symposium*, pp. 1671–1678, IEEE, Los Angeles, CA, USA, 11 June 2017.
- [12] M. Bahram, A. Wolf, M. Aeberhard, and D. Wollherr, "A prediction-based reactive driving strategy for highly automated driving function on freeways," in *Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings*, pp. 400–406, IEEE, Dearborn, MI, USA, 8 June 2014.
- [13] F. W. Rauskolb, K. Berger, C. Lipski et al., "Caroline: an autonomously driving vehicle for urban environments," *Journal of Field Robotics*, vol. 25, no. 9, pp. 674–724, 2008.
- [14] S. Ulbrich and M. Maurer, "Probabilistic online POMDP decision making for lane changes in fully automated driving," in *Proceedings of the 2013 16th International IEEE Conference on Intelligent Transportation Systems*, pp. 2063–2067, IEEE, 6 October. 2013.
- [15] C. Hubmann, J. Schulz, M. Becker, D. Althoff, and C. Stiller, "Automated driving in uncertain environments: planning with interaction and uncertain maneuver prediction," *IEEE transactions on intelligent vehicles*, vol. 3, no. 1, pp. 5–17, 2018.
- [16] S. Zuo, Z. Wang, X. Zhu, and Y. Ou, "Continuous reinforcement learning from human demonstrations with integrated experience replay for autonomous driving," in *Proceedings of the 2017 IEEE International Conference on Robotics And Biomimetics*, pp. 2450–2455, IEEE, Macao, Macao, 5 December 2017.
- [17] K. Min, H. Kim, and K. Huh, "Deep Q learning based high level driving policy determination," in *Proceedings of the IEEE*

- Intelligent Vehicles Symposium*, pp. 226–231, IEEE, Changshu, China, 26 June 2018.
- [18] A. Sallab, M. Abdou, E. Perot, and S. Yogamani, “Deep reinforcement learning framework for autonomous driving,” *Electronic Imaging*, vol. 2017, no. 19, pp. 70–76, 2017.
- [19] V. Mnih, K. Kavukcuoglu, D. Silver et al., “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [20] P. G. Tai L and M. Liu, “Virtual-to-real deep reinforcement learning: continuous control of mobile robots for mapless navigation,” in *Proceedings of the 2017 IEEE International Conference on Intelligent Robots And Systems*, pp. 31–36, IEEE, Vancouver, BC, Canada, 24 September 2017.
- [21] X. Wang, C. Wu, J. Xue, and Z. Chen, “A method of personalized driving decision for smart car based on deep reinforcement learning,” *Information*, vol. 11, no. 6, p. 295, 2020.
- [22] C. Xu, W. Zhao, Q. Chen, and C. Wang, “An actor-critic based learning method for decision-making and planning of autonomous vehicles,” *Science China Technological Sciences*, vol. 64, no. 5, pp. 984–994, 2021.
- [23] S. Li, M. Huang, M. Guo, and M. Yu, “Evaluation model of autonomous vehicles’ speed suitability based on overtaking frequency,” *Sensors*, vol. 21, no. 2, pp. 1–16, 2021.
- [24] T. P. Lillicrap, J. J. Hunt, A. Pritzel et al., “Continuous control with deep reinforcement learning,” in *Proceedings of the 3th Annual Meeting of the International Conference on Learning Representations*, Neptuneblog, San Juan, California, 14 October 2021.
- [25] J. Kim and D. Kum, “Collision risk assessment algorithm via lane-based probabilistic motion prediction of surrounding vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 9, pp. 2965–2976, 2018.
- [26] X. Shi, Y. D. Wong, M. Z.-F. Li, C. Palanisamy, and C. Chai, “A feature learning approach based on XGBoost for driving assessment and risk prediction,” *Accident Analysis & Prevention*, vol. 129, pp. 170–179, 2019.