

Article

# Reinforcement Learning-Based Lane Change Decision for CAVs in Mixed Traffic Flow under Low Visibility Conditions

Bowen Gong <sup>1</sup>, Zhipeng Xu <sup>1</sup>, Ruixin Wei <sup>1</sup>, Tao Wang <sup>2,\*</sup>, Ciyun Lin <sup>1,3,\*</sup> and Peng Gao <sup>4</sup><sup>1</sup> Department of Traffic Information and Control Engineering, Jilin University, Changchun 130022, China<sup>2</sup> China Academy of Transportation Sciences, Beijing 100029, China<sup>3</sup> Jilin Engineering Research Center for Intelligent Transportation System, Changchun 130022, China<sup>4</sup> Qingdao Transportation Public Service Center, Qingdao Municipal Transport Bureau, Qingdao 266061, China

\* Correspondence: lwwtao@163.com (T.W.); linciyun@jlu.edu.cn (C.L.)

**Abstract:** As an important stage in the development of autonomous driving, mixed traffic conditions, consisting of connected autonomous vehicles (CAVs) and human-driven vehicles (HDVs), have attracted more and more attention. In fact, the randomness of human-driven vehicles (HDV) is the largest challenge for connected autonomous vehicles (CAV) to make reasonable decisions, especially in lane change scenarios. In this paper, we propose the problem of lane change decisions for CAV in low visibility and mixed traffic conditions for the first time. First, we consider the randomness of HDV in this environment and construct a finite state machine (FSM) model. Then, this study develops a partially observed Markov decision process (POMDP) for describing the problem of lane change. In addition, we use the modified deep deterministic policy gradient (DDPG) to solve the problem and get the optimal lane change decision in this environment. The reward designing takes the comfort, safety and efficiency of the vehicle into account, and the introduction of transfer learning accelerates the adaptation of CAV to the randomness of HDV. Finally, numerical experiments are conducted. The results show that, compared with the original DDPG, the modified DDPG has a faster convergence velocity. The strategy learned by the modified DDPG can complete the lane change in most of the scenarios. The comparison between the modified DDPG and the rule-based decisions indicates that the modified DDPG has a stronger adaptability to this special environment and can grasp more lane change opportunities.



**Citation:** Gong, B.; Xu, Z.; Wei, R.; Wang, T.; Lin, C.; Gao, P.

Reinforcement Learning-Based Lane Change Decision for CAVs in Mixed Traffic Flow under Low Visibility Conditions. *Mathematics* **2023**, *11*, 1556. <https://doi.org/10.3390/math11061556>

Received: 5 March 2023

Revised: 19 March 2023

Accepted: 20 March 2023

Published: 22 March 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** reinforcement learning; low visibility and mixed traffic conditions; lane change decision; DDPG

**MSC:** 90B20

## 1. Introduction

In recent years, connected autonomous vehicle (CAV) has been widely discussed in the world. Sensing, decision-making, planning, control and other related autonomous driving technologies have also entered a period of rapid development. Decision technology is one of the most critical modules. It receives information from the perception module and gives guidance to planning and control modules. Therefore, decision-making technology is equivalent to the brain of CAV and plays a key role in the safe driving of the vehicle [1]. Lane change is the main scenario for CAV decision-making, including the determination of lane change intention and execution time. In this paper, we assume that the vehicle has a strong intent to change lanes. The CAV needs to choose the right time to change lanes and perform the lane change. During a lane change, the CAV needs to pay attention to the real-time state of the other vehicles in the area that are affected by the lane change, including velocity, acceleration and relative position. In particular, the success of the lane change is often directly determined by predicting the driving intention of vehicles behind the target lane.

The judgment of lane change timing is the key to a successful lane change. The basis of the judgment includes some hard rules [2], such as the calculation of real-time benefits [3] and the existing experience of the lane change [4,5].

In mixed traffic scenarios, drivers are stochastic, and relying on hard rules may cause the CAV to miss some lane change opportunities. Therefore, lane change decisions based on hard rules are conservative. Benefit calculation-based decision-making usually requires real-time interaction with the driver on the road; however, this approach is difficult to apply in low-visibility weather such as haze, fog, etc. There are two main reasons for this: on the one hand, the sensing sensors (LIDAR, cameras, etc.) may be affected to varying degrees [6], and on the other hand, the low-visibility environment can affect the driver's view. For the lane change experience, it is difficult to collect sufficient marker data from skilled drivers, especially in low-visibility scenarios. Therefore, it is urgent to develop a lane change decision algorithm for autonomous driving under low visibility mixed traffic conditions.

The problem of lane change decisions for CAV under mixed traffic conditions has also been widely studied [7–10]. In the existing research, researchers have used hard rule-based lane change decision methods, such as FSM [11], decision trees (DT) [12] and custom rules [13]. Other researchers have used machine learning methods to guide lane change decisions by learning from successful lane change experiences. These methods are the support vector machines (SVM) [14], XGBoost [4], Gaussian mixture density hidden Markov model (GMM-HMM) [15] and neural networks [16]. There are also many researchers who focus on decision methods for computing benefits, especially game theoretic methods based on real-time interactions [17]. Reinforcement learning, which combines the benefits of computing and machine learning in lane change strategies, has also been extensively studied [18,19]. All the above studies address lane change decisions for CAV in mixed traffic conditions in normal weather. To the best of our knowledge, no studies have been conducted on the lane change decision of CAV in low visibility and mixed conditions.

In this study, we formulated the lane change decision for CAV as a partially observed problem and solved it using modified reinforcement learning methods. First, we considered the effects of trust and visibility on drivers and proposed an FSM to characterize drivers' driving in low visibility and mixed traffic conditions. Second, we used a POMDP to represent the problem of changing lanes for CAV in low visibility and mixed traffic conditions. Third, we performed the solution with a modified reinforcement learning algorithm. Finally, we tested and verified the effectiveness of the modified algorithm.

The main contributions of this research are as follows: First, it introduces a new scenario for lane changing in CAV in mixed traffic and low visibility conditions. As far as we know, this is the first study to address the lane change decision problem in such conditions. Second, the study identifies four car following states that drivers experience and combines the weights of safety and space factors to represent a driver's response to a vehicle ahead. These five states are then integrated into an FSM model. Third, the researchers use POMDP to formulate the entire lane change decision problem for CAV and develop a modified DDPG algorithm to solve this problem.

The remainder of this paper is organized as follows: Section 2 reviews related research on the lane change decision-making in mixed traffic conditions for CAV. Section 3 presents the driver model formulation and problem description. Section 4 describes our proposed solution methodology, the modified DDPG. In Section 5, numerical experiments are conducted to illustrate the effectiveness of our algorithm. This paper is concluded in Section 6 with the summary and possible future research.

## 2. Literature Review

In order to draw a clear distinction between this study and previous research, the literature on rule-based, game-theoretic, machine-learning and reinforcement-learning is reviewed. The influence of low visibility on the perception module of CAV and the

behavior characteristics of drivers under low visibility and mixed traffic conditions are also introduced in this section.

The existing research on rule-based lane change decisions is adequate. In these studies, lane changes are guided by a set of deterministic rules [20]. The main purpose is to ensure the absolute safety of lane change by judging whether vehicles meet lane change conditions. It is based on the idea of absolute safety that Mobileye proposes a liability-sensitive safety policy [21]. The earliest rule-based lane change decision models were mainly concerned with the acceptance of lane change gaps [22,23]. The Gipps model [22] defines the lane change process as a series of rules that use a decision tree and outputs the result as a binary classification result. Of course, the model is no longer satisfied with the complex vehicle composition and complex human-vehicle interaction behavior in the traffic environment. Another typical representative of the rule-based lane change model is the MOBIL model [24], which determines whether to change lanes by setting utility incentives and safety rules that apply to various rules of following and free lane change and forced lane change behavior. Of course, there is also a class of minimum safe distance models model, which determines the safe lane change condition by specifying the minimum safe distance rule between vehicles. At present, the biggest shortcoming of this type of rule-based model is that the driving environment conditions are relatively simple, and the influence of the driver's driving behavior on lane changing is not much considered. In fact, these rules are the basic elements for making lane change decisions, and many new decision methods are further researched based on these rules.

Although these rules guarantee absolute safety for lane changes, they require a large enough clearance, which is unrealistic in the case of traffic congestion or accidents. Therefore, on the basis of accepting gaps, some studies have extended lane change rules, taking into account the interaction between vehicles and the variability of individual drivers [25–30]. Game theory is commonly used to reflect driver differences in inter-vehicle interactions [31]. However, in low visibility, even if the sensing module of the CAV works properly, the driver's restricted field of view causes him/her not to interact properly with the surrounding vehicles, therefore, the result between vehicles cannot be trusted.

With the development of artificial intelligence, many scholars try to improve the safety of lane change decisions by machine learning methods. Chen et al. [32] used a random forest classifier to extract key features from vehicle trajectory data and then used fault analysis and k-mean clustering algorithms to determine the risk level of vehicle lane-changing behavior. Xu et al. [5] developed a convolutional neural network (CNN) model based on the lane-changing image dataset to estimate the driver's intention to change lanes. Jin et al. [16] proposed a Gauss mixture hidden Markov model (GM-HMM) lane change decision model, which is experimentally shown to have a high similarity rate to the actual lane change behavior. Gindele et al. [33] constructed a filter to estimate the behavior of traffic participants and predict their future trajectories using a dynamic Bayesian network (DBN) that takes into account all the information relevant to the driver's decision. Although much progress has been made in applying machine learning to lane change decision problems, lane change decisions under low visibility and mixed traffic conditions have not been considered. In addition, empirical data for learning in such environments are difficult to obtain.

Reinforcement learning not only avoids the use of large amounts of manually marked driving data but also simulates complex lane changes. Therefore, people pay more and more attention to solving lane change decisions by reinforcement learning. Chen et al. [34] designed a hierarchical deep reinforcement learning algorithm to learn lane-changing behavior in dense traffic. Jiang et al. [35] formulated the lane-changing behavior as a POMDP process and proposed a combination of recurrent neural network (RNN) and deep Q network (DQN), which is solved according to the global maximum reward. To overcome the limitations of discrete action space in ordinary reinforcement learning algorithms, Wang et al. [36] established the vehicle lane change behavior under continuous action based on deep deterministic policy gradient (DDPG). Lv et al. [37] developed an

improved deep reinforcement learning algorithm based on DDPG with well-defined reward functions for autonomous driving tasks in highway scenarios where an autonomous vehicle merges into two-lane road traffic flow and realizes the lane-changing maneuvers. Kim et al. [38] proposed a new reinforcement learning-based lane change decision-making technique that uses a safety inspection module and augmented data. Wang et al. [39] studied how to learn a harmonious deep reinforcement learning-based lane-changing strategy for autonomous vehicles without Vehicle-to-Everything (V2X) communication support. Ammourah et al. [40] proposed a reinforcement learning-based framework for mandatory lane changing of automated vehicles in a non-cooperative environment that utilized the double deep Q-learning algorithm structure that takes relevant traffic states as input and outputs the optimal actions (policy) for the automated vehicle. He proposed [41] a novel observation adversarial reinforcement learning approach for robust lane change decision-making of autonomous vehicles. Additionally, a constrained observation-robust Markov decision process is presented to model lane change decision-making behaviors of autonomous vehicles under policy constraints and observation uncertainties. In low visibility and mixed traffic, reinforcement learning algorithms can not only adapt to the randomness of drivers but also enable collaboration between CAVs in the absence of communication. Therefore, a reinforcement learning algorithm is applied in this paper to solve the lane change decision problem in this complex environment.

In low-visibility weather, on-board sensors can be disturbed. For example, the functionality of both cameras and LIDAR can be severely degraded in dense fog [42,43]. The sensors can only provide partial information about the observation, which may result in the CAV not obtaining enough information for decision-making. Therefore, it is necessary to consider lane change decisions under partial observation in low-visibility conditions.

In addition, in low visibility, drivers have an increased reaction time and a reduced ability to perceive risks [44,45]. Drivers tend to reduce the headway time distance and shorten the following distance until the view becomes better or they spot the vehicle in front of them [46].

Over the past few years, the rapid arrival of the era of intelligent technology has led to breakthroughs in fuzzy logic technology. Compared to other technologies, fuzzy logic technology has passed through several stages and is more mature, attracting people to research on fuzzy logic technology. Huang et al. [47] distinguished the main differences between interval type 2 (IT2), generalized type 2 (GT2) and interval type 3 (IT3) fuzzy logic systems (FLS) and illustrated the design of IT3-FLS through online identification, offline time series modeling and robot control systems. Jomaa et al. [48] developed a fuzzy logic method to control the actuators that are installed inside the greenhouse for heating, ventilation, humidification and the physical model of the greenhouse used in the Simulink/Matlab environment, which was carefully designed to simulate temperature and indoor humidity. Huang et al. [49] proposed a simple but practical fuzzy logic position tracking controller for a wheeled mobile robot and used particle swarm optimization to turn the parameters of the scaling and membership functions to optimize this controller. Rajagiri et al. [50] designed and developed a fuzzy logic controller in MATLAB Simulink to operate the DC motor. It is experimentally demonstrated that the proposed fuzzy controller results in a better response compared to the normal response of a DC motor. In summary, fuzzy logic technology, as a mature control technique, has broad application prospects and good control performance, especially in the field of intelligent technology where it has been widely applied and researched. With the continuous development and improvement of technology, it is believed that fuzzy logic technology will play its advantages in more fields, such as the research area of mixed traffic flow control of CAV and HDV, and bring more convenience and practical application effects to people.

In summary, the existing research has addressed the lane change decision problem for CAV by various modeling tools and algorithms. However, to our best knowledge, there is no research on the lane change decision for CAV in low-visibility mixed traffic conditions. This paper aims at formulating the lane change decision for CAV as a partially observed

problem and solved it using a reinforcement learning algorithm. Numerical experiments are conducted to test the effectiveness and practicality of our algorithm.

### 3. Problem Formulation

A typical lane change scenario is shown in Figure 1. Lane 1 is the current lane and lane 2 is the target lane. Vehicle A is the CAV preparing to change lanes. Vehicle B and vehicle D are the vehicles ahead of CAV A in the target lane and current lane, respectively, and we do not consider the types of these two vehicles. Vehicle L and vehicle C are the vehicles behind CAV A in the current lane and the target lane, respectively. In fact, the decision timing of CAV A's lane change is mainly influenced by these two vehicles. This paper focuses on the effect of CAV A lane change when vehicle L and vehicle C are HDVs in low visibility.

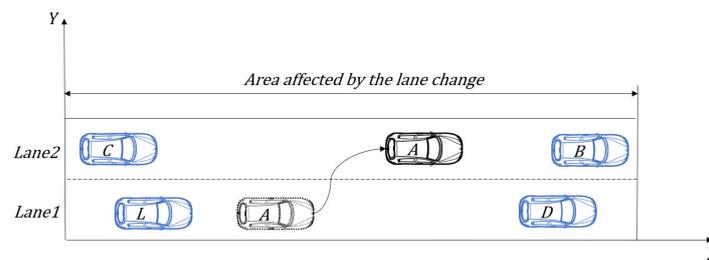


Figure 1. The typical lane change scenario.

In this paper, we assume that the area affected by the lane change is the sensing area of the CAV, i.e., when the CAV changes lanes, the safety judgment will be made for the vehicles in the sensing area. Before describing the problem, the driving behavior of the HDV is modeled: First, we analyzed the influence of low visibility and mixed traffic conditions on drivers. Then, we propose a driver longitudinal control model containing five states under low-visibility mixed traffic conditions. Based on the above, we describe the lane change decision problem to be studied.

#### 3.1. The Longitudinal Control Model for Drivers in Low-Visibility Mixed Traffic Conditions

In the lane change scenario shown in Figure 1, we present the car following states of HDV L and HDV C under low visibility mixing conditions and the reaction states of HDV C when CAV A is changing lanes.

##### 3.1.1. The Influence of Low Visibility and Mixed Traffic Conditions on Drivers

In low visibility, if the driver cannot make a quick and accurate judgment of the surroundings within his or her safe sight distance, the driver will become hesitant about the next driving operation, which will cause a delay in the driver's reaction. Obviously, the reaction delay of the driver is closely related to the safety sight distance and visibility without considering the driver's age, personality and gender. In this paper, we use the stopping sight distance  $dis_0$  to represent the safety sight distance, as shown in Equation (1):

$$dis_0 = vel_0 \times t_0 + \frac{(vel_0/3.6)^2}{2u_{max}} \tag{1}$$

where  $vel_0$  indicates the current velocity,  $t_0$  indicates the normal reaction time and  $u_{max}$  indicates the absolute value of the maximum deceleration.

Quantifying the driver's reaction delay using visibility and safety sight distance in normal weather. To ensure the absolute safety of the vehicle, we express the safety sight distance of the vehicle in low visibility by twice the safety sight distance in normal weather minus the visibility  $dis_v$ , as shown in Equation (2):

$$vel_0 \times (t_0 + t_d) + \frac{(vel_0/3.6)^2}{2u_{max}} \leq 2dis_0 - dis_v \tag{2}$$

where  $t_d$  indicates the driver’s reaction delay due to restricted vision. The left side of the formula indicates the safe driving distance required in the case of a delayed driver reaction. From the above Equation, the smaller the safety sight distance and the greater the visibility, the smaller the driver’s reaction delay.

In mixed traffic conditions, drivers’ trust in CAVs can affect their driving behavior [51]. In low visibility, drivers have fewer targets to refer to, therefore, drivers’ trust in CAVs will be even more important. For example, in low-visibility car-following, drivers tend to follow CAVs at smaller distances when the trust level is high. On the contrary, drivers will treat the behavior of CAVs with caution.

We assume that the trust factor  $p$  obeys a standard normal distribution  $p \sim N(0, 1)$ , and  $\beta(p)$  denotes the cumulative probability density of  $p$ . We let the range of values of  $p$  be  $(-3, 3)$ , which can cover most of the driver aggressiveness. Then we use the distribution function of  $\beta(p)$  to represent the degree of trust of drivers in CAV. When the trust factor  $p$  is large or small, the probability of the distribution of  $\beta(p)$  is small. When  $p$  is near 0, the probability of the distribution of  $\beta(p)$  is large. Additionally, in this paper, we assume that the CAVs under traffic conditions can be clearly recognized by the drivers.

### 3.1.2. The Car following States

Considering the reaction delay and the level of confidence, we classify the car following states of HDVs under low-visibility mixed traffic conditions into four types as shown in Figure 2.

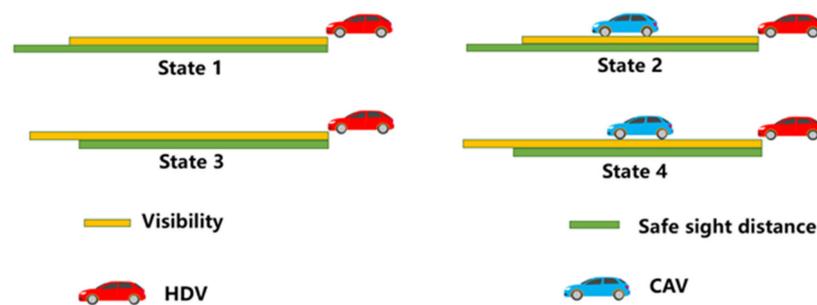


Figure 2. The four car-following states of HDVs under mixed driving conditions in low visibility.

State 1: At the current velocity of the vehicle, the safety sight distance is greater than the visibility and there is no other vehicle in front of the HDV within the visibility range. In this state, the driver has a reaction delay. In addition, since the driver cannot see the situation within the safe time distance, the driver tends to look for the vehicle that he can follow ahead in order to get a sense of security [52], therefore, the driver executes a smaller acceleration  $b_1$ , as shown in Equation (3):

$$a(t) = b_1 (b_1 > 0) \tag{3}$$

where  $a(t)$  is the acceleration taken at moment  $t$ .

State 2: At the current velocity of the vehicle, the safety sight distance of the vehicle is greater than the visibility and there is a CAV in front of the HDV within the visibility range. The HDV will choose to follow the CAV in this state. But when the state of the CAV changes, the HDV will have a reaction delay. In addition, the driver’s trust in the CAV also affects the following state of the HDV. We use the modified full velocity difference following model (MFVD) to represent the behavior of the HDV in this state, as shown in Equations (4)–(7) [53]:

$$a_i(t) = K[V(\Delta x_i(t), vel_i(t), p) - vel_i(t)] + \lambda \Delta v_i(t) \tag{4}$$

$$V(\Delta x_i(t), vel_i(t), p) = \frac{vel_{lim}}{2} [tanh(\Delta x_i(t) - he(vel_i(t), p)) + tanh(he(vel_i(t), p))] \tag{5}$$

$$he(vel_i(t), p) = (1 - \beta(p))(T + \tau_i)vel_i(t) + hc \tag{6}$$

$$\Delta x_i(t) = x_{i-1}(t) - x_i(t) \tag{7}$$

where  $K(K > 0)$  is the sensitivity coefficient of headway spacing,  $V$  is a reasonable value of speed,  $a_i(t)$  is the acceleration taken at moment  $t$  of the  $i$ -th HDV,  $\lambda(\lambda > 0)$  is the response coefficient of velocity difference and  $\Delta x_i(t)$  is the distance between the  $i$ -th HDV and the vehicle in front of it.  $\Delta v_i(t)$  is the velocity difference between the  $i$ -th HDV and the vehicle in front of it.  $he(vel_i(t), p)$  denotes the following distance expected by the  $i$ -th HDV with trust factor  $p$  at a velocity of  $vel_i(t)$ .  $T$  is the driver’s normal reaction time,  $\tau_i$  is the driver’s reaction delay and  $hc$  is the safe following distance with respect to the vehicle type.

State 3: At the current velocity of the vehicle, the visibility is greater than the safety sight distance, and there is no other vehicle in front of the HDV within the visibility range. In this state, since the driver can see the traffic conditions within the safety sight distance, the driver will execute a smaller acceleration  $b_2$  until the safety sight distance is greater than the visibility or the vehicle ahead is detected within the visibility. We assume that  $b_2$  is greater than  $b_1$ , as shown in Equation (8):

$$a(t) = b_2(b_2 > b_1) \tag{8}$$

State 4: At the current velocity of the vehicle, the safety sight distance is less than the visibility, and there is a vehicle in front of the HDV within the visibility range. Similar to state 2, the driver chooses to follow the vehicle ahead, we also use MFVD to represent this; different from state 2 is that the driver has no reaction delay.

### 3.1.3. The Response Model to Lane Change

In fact, drivers can see the driving conditions in the adjacent lane in the visibility range. This section describes the state of drivers’ reactions to the lane change of vehicles in adjacent lanes. As in Figure 1, HDV C will react to the lane change of CAV A. We refer to Yu et al.’s description of the driver’s lane change response: the more aggressive the driver, the greater the concern for spatial benefits. Conversely, there is a greater concern for their own safety interests, as shown in Equations (9)–(18) [10]:

$$\max(f_w(a, a_0) \left( (1 - \beta(q)) \times U_{safety}(a) + \beta(q) \times U_{space}(a) + 1 \right) - 1) \tag{9}$$

$$a_{min} \leq a \leq a_{max} \tag{10}$$

$$f_w(a, a_0) = e^{-\left(T_l^2 \times \frac{(a-a_0)^2}{w_1} + \frac{(v+a \times T_l - v_d)^2}{w_2}\right)} \tag{11}$$

$$U_{safety}(a) = \frac{1}{2}(Sf_{t=T_l} - Sf_{t=0}) \tag{12}$$

$$U_{space}(a) = \frac{1}{2}(SP_{t=T_l} - SP_{t=0}) \tag{13}$$

$$Sf_t = \begin{cases} \frac{2|T_{h1,t}|}{T_b} - 1, & -T_b \leq T_{h1,t} \leq T_b \\ 1, & \text{others} \end{cases} \tag{14}$$

$$SP_t = \begin{cases} -1, & T_{h2} \leq -3 \\ \frac{2T_{h2,t}}{3} + 1, & -3 < T_{h2,t} \leq 0 \\ 1, & T_{h2} > 0 \end{cases} \tag{15}$$

$$T_{h1,t=0} = \frac{P_{A,t=0} - P_{C,t=0}}{v_{C,t=0}} \tag{16}$$

$$T_{h1,t=T_l} = \begin{cases} \frac{P_{A,t=T_l} - P_{C,t=T_l}}{v_{C,t=0} + a T_l}, & P_{A,t=T_l} > P_{C,t=T_l} \\ \frac{P_{C,t=T_l} - P_{A,t=T_l}}{v_{A,t=0} + a_A T_l}, & \text{others} \end{cases} \tag{17}$$

$$T_{h2,t} = \begin{cases} \frac{P_{C,t} - P_{A,t}}{v_{A,t}}, & P_{A,t} \leq P_{C,t} \\ \frac{P_{C,t} - P_{A,t}}{v_{C,t}}, & \text{others} \end{cases} \tag{18}$$

The objective function (9) maximizes security benefits and space benefits. Constraint (10) is an acceleration constraint, where  $a_{max}$  is the value of the maximum acceleration or minimum deceleration that can be taken, and  $a_{min}$  is the value of the minimum acceleration or maximum deceleration that can be taken.  $f_w(a, a_0)$  is the penalty for acceleration and velocity change, as shown in Equation (11). The security benefits  $U_{safety}(a)$  and space benefits  $U_{space}(a)$  are calculated as shown in Equations (12) and (13).  $Sf_t$  is the safety factor, as shown in Equation (14).  $SP_t$  is the space factor, as shown in Equation (15).  $T_l$  is the time for vehicles in adjacent lanes to change lanes.  $a$  and  $a_0$  are the acceleration at the current moment and the next moment, respectively.  $w_1$  and  $w_2$  are the weight values of acceleration change and velocity change, respectively.  $v_d$  is the desired acceleration.  $T_{h1}$  is the time headway between vehicle C and vehicle L, including the time headway at the initial moment  $T_{h1,t=0}$ , the time headway at the moment of lane change completion  $T_{h1,t=T_l}$  and  $a_A$  is the acceleration of the CAV A, as shown in Equations (16) and (17).  $T_{h2,t}$  is the evaluation index of the spatial factor, as shown in Equation (18).  $T_b$  denotes the safe time headway.  $P_{A,t}$  and  $P_{C,t}$  denote the longitudinal position of vehicle A and vehicle C at moment t, respectively.  $v_{A,t}$  and  $v_{C,t}$  denote the longitudinal velocity of vehicle A and vehicle C at moment t, respectively. Please refer to the literature [10] for details of the driver response model.

### 3.1.4. The FSM Model

We compose an FSM model of the driver’s four following states and response states to lane change in low visibility and mixed traffic conditions, as shown in Figure 3. The inputs to the model are whether there is a vehicle ahead in visibility c, whether the safety sight distance is greater than visibility g and whether there is a lane changing vehicle in the adjacent lane in visibility r. When the input is 1, it means certainty, and when the input is 0, it means negation. The output of the model is the five states in the FSM.

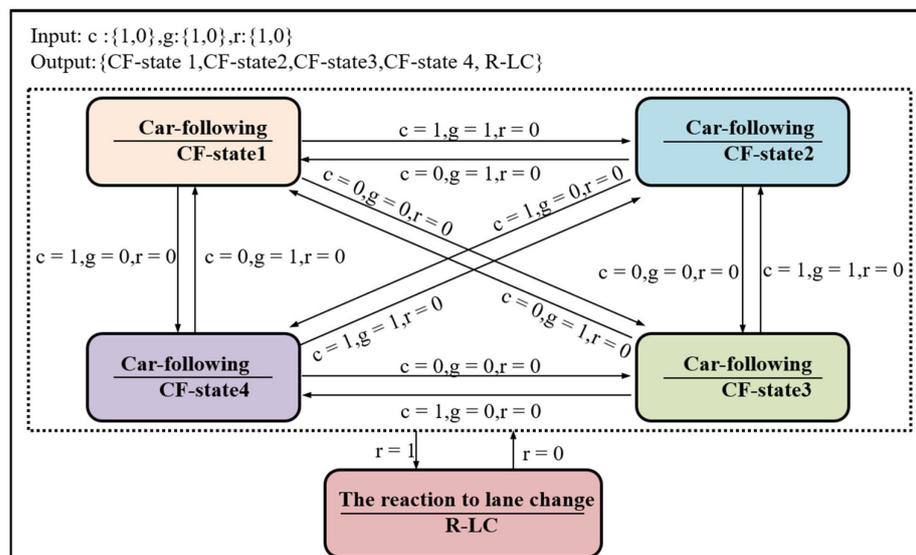


Figure 3. The FSM model of HDV in low visibility and mixed traffic conditions.

### 3.2. The Lane Change Decision Problem

In this section, we first give the rule-based lane change decision method. In a typical lane change scenario as shown in Figure 1, CAV A usually only needs to consider the safe distance  $F_{BA}(t)$  from vehicle B and vehicle D ahead, as shown in Equation (19):

$$F_{BA}(t) = \frac{v_A^2(t)}{2a(t)} - \frac{v_B^2(t)}{2a_{max}} \tag{19}$$

where  $v_A$  and  $v_B$  are the current velocities of vehicle A and vehicle B at moment t, respectively. Considering the driver’s comfort, the deceleration or acceleration value of CAV A should not be too large, as shown in Equation (20):

$$a(t) = a_{min} + \frac{v_A(t)}{v_{lim}}(a_{max} - a_{min}) \tag{20}$$

where  $v_{lim}$  is the limited velocity. The safety distance  $F_{DA}(t)$  between vehicle A and vehicle D is also calculated by Equations (19) and (20).

Within the sensing range of CAV A, if vehicle C is behind the target lane, then CAV A needs to focus on HDV C when changing lanes. There are mainly two cases, one is when the distance between CAV A and HDV C is greater than a specific value  $F_{CA}(t)$ , as shown in Equation (21). In this case, CAV A can directly change lanes without being affected by HDV C.

$$F_{CA}(t) = v_C(t)T + \frac{a_{max}T^2}{2} + \frac{(v_C(t) + a_{max}T)^2}{2a_{min}} - \frac{v_A^2(t)}{2a_{max}} \tag{21}$$

Another case is the distance between CAV A and HDV C is less than  $F_{CA}$ . Under the hard rule, CAV A will drop the lane change. Otherwise, CAV A needs to interact with HDV C and change lanes at the right time according to HDV C’s response. We usually do not need to consider the interference of vehicle L.

In low visibility, although the sensing module will be disturbed, CAV A can still perceive the position of vehicle B, vehicle C and vehicle D most of the time. However, outside of visibility, HDV C cannot find CAV A preparing to change lanes, which will result in no normal interaction between the two vehicles. Therefore, according to Figure 2, we divide the lane change problem of CAV A into four situations based on the relationship between visibility and safety sight distance and the presence or absence of preceding vehicles in visibility, as shown in Figure 4.

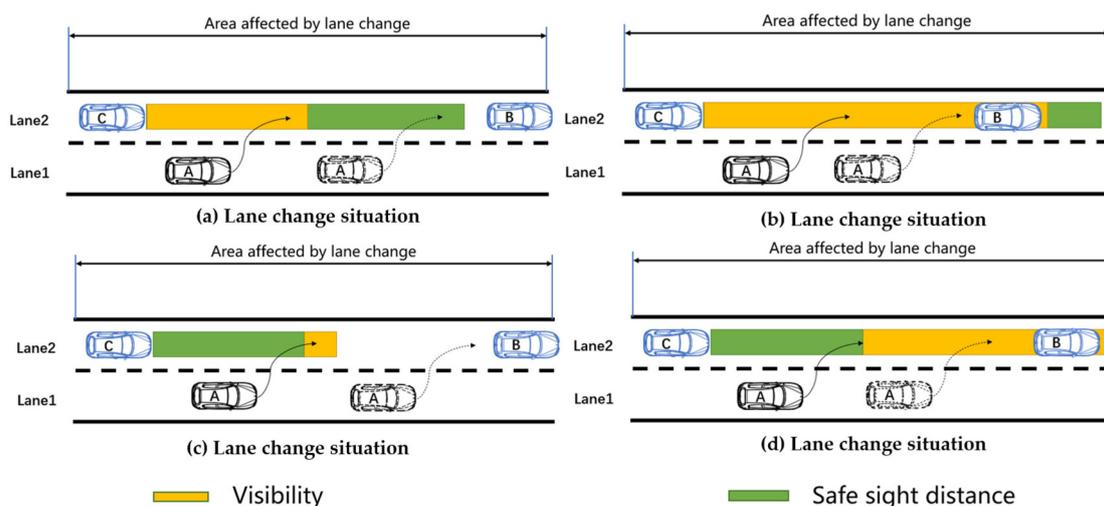


Figure 4. The four lane change scenarios for CAV A in low visibility and mixed traffic conditions.

Figure 4 shows four lane change scenarios for CAV A. Scenario (a) is a lane change situation where the safety sight distance of HDV C is greater than visibility and vehicle B is

not within visibility. Scenario (b) is a lane change situation where the safety sight distance of HDV C is greater than the visibility and the vehicle ahead is visible. Scenario (c) is a lane change situation where the safety sight distance of HDV C is greater than the visibility and vehicle B is not within visibility. Scenario (d) is a lane change situation where the safety sight distance of HDV C is greater than the visibility and vehicle B is not within visibility.

For the lane change in the above special scenarios, the rule-based lane change strategy ensures the absolute safety of the lane change vehicles; therefore, it is still valid. However, it is sometimes too conservative to use a uniform strategy to solve all scenarios; therefore, we try to use reinforcement learning to adopt a corresponding lane change strategy for each scenario.

To show the current problem and the solution method more clearly, we list a schematic diagram, as shown in Figure 5.

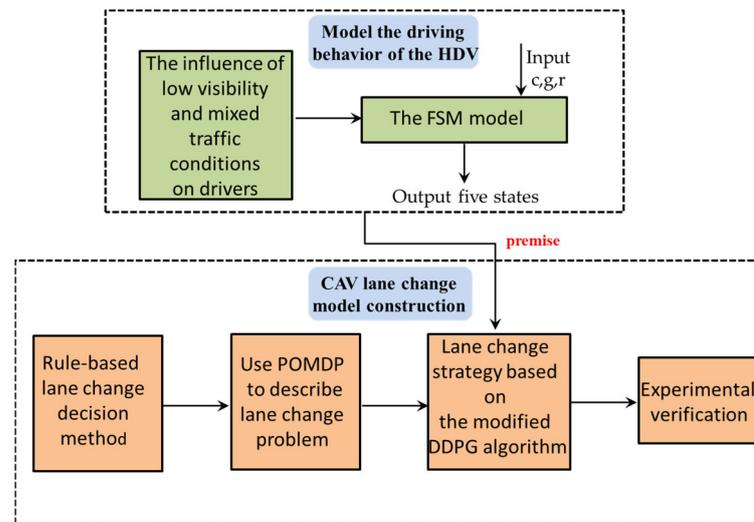


Figure 5. The flow diagram of the problem formulation and methodological logic framework of this paper.

#### 4. Solution Methodology

In this section, first, we use POMDP to describe the CAV A lane change problem in Section 3. Then, we develop a modified DDPG algorithm to solve this problem.

##### 4.1. The POMDP Mathematical Model

Subject to extreme environments, CAV perceives an increased probability of missing or incorrect information; therefore, we use POMDP to describe this process. Typically, POMDP consists of a seven-tuple  $\langle S, A, T, R, Z, O, \gamma \rangle$ , where  $X$  is the set of all environmental states  $s_t$  at moment  $t$ .  $A$  is the set of all possible actions  $a_t$  at moment  $t$ .

$T$  is the transfer function;  $T(s_t, a_{t-1}, s_{t-1}) = P(s_t | a_{t-1}, s_{t-1})$  denotes the probability of taking action  $a_{t-1}$  in state  $s_{t-1}$  to get the state of moment  $t$  as  $s_t$ .  $R$  is the reward function;  $r(s, a)$  denotes the reward received for taking action  $a$  in state  $s$ .  $Z$  is the set of all possible observation states  $z_t$  at moment  $t$ .  $O$  is the observation function;  $O(s_t, a_{t-1}, z_{t-1}) = P(s_t | a_{t-1}, z_{t-1})$  denotes the probability of taking action  $a_{t-1}$  to get the state of moment  $t$  as  $s_t$  when the observation value is  $z_{t-1}$ .  $\gamma$  is the discount factor. Based on the POMDP mathematical model, the lane change problem for CAV is described as follows:

- State  $S$ : State of the entire lane change system, including CAV A and surrounding vehicles. The system coordinate is shown in Figure 1, based on which, the state information can be described by the longitudinal position, lateral position and velocity of all vehicles, as shown in Equation (22):

$$s_t = (x_{A,t}, x_{B,t}, x_{D,t}, x_{C,t}, x_{L,t}, y_{A,t}, y_{B,t}, y_{D,t}, y_{C,t}, y_{L,t}, v_{A,t}, v_{B,t}, v_{D,t}, v_{C,t}, v_{L,t}) \quad (22)$$

where the velocity limit is related to the current visibility of the highway.

- Action A: The longitudinal acceleration and steering angle of the CAV are used as defined action parameters, as shown in Equation (23):

$$a_t = (u_t, \delta_t) \tag{23}$$

- Transfer function T: The dynamic model of the lane change system, which is difficult to describe precisely.
- Observation Z: Due to the influence of the environment or the physical limitations of the sensor itself, CAV A can only obtain the information about other vehicles when sensor noise is present or missing. Z is defined as:

$$z_t = (x_{A,t}', x_{B,t}', x_{D,t}', x_{C,t}', x_{L,t}', y_{A,t}', y_{B,t}', y_{D,t}', y_{C,t}', y_{L,t}', v_{A,t}', v_{B,t}', v_{D,t}', v_{C,t}', v_{L,t}') \tag{24}$$

- Reward function R: The reward function is the key to achieve the goal of lane change. We design the reward function from the perspective of safety, efficiency and comfort:

$$R_{com} = -\alpha \cdot u_{x,t}^2 + \beta \cdot u_{y,t}^2 \tag{25}$$

where  $R_{com}$  denotes the reward function of the comfort level.  $u_{x,t}$  is the longitudinal acceleration rate.  $u_{y,t}$  is the lateral acceleration rate.  $\alpha$  and  $\beta$  are the weighting factors of the longitudinal acceleration rate and the lateral acceleration rate, respectively. The comfort reward is introduced mainly to reduce sudden acceleration and deceleration of the CAV A.

In terms of efficiency, the goal is to get the CAV to the centerline of the target lane as quickly as possible without exceeding the velocity limit, as shown in Equations (26)–(29):

$$R_{tim} = -\Delta t \tag{26}$$

$$R_{pos} = -|y_t - y^*| \tag{27}$$

$$R_{spe} = -|v_t - v^*| \tag{28}$$

$$R_{eff} = \omega_1 \cdot R_{tim} + \omega_2 \cdot R_{pos} + \omega_3 \cdot R_{spe} \tag{29}$$

where  $R_{tim}$  denotes the reward function for the lane change time.  $\Delta t$  denotes the lane change time.  $R_{pos}$  denotes the reward function for position.  $y_t$  is the current lateral position of the vehicle.  $y^*$  is the target lane centerline.  $R_{spe}$  denotes the reward function for lane change velocity.  $v_t$  is the current longitudinal velocity of the vehicle.  $v^*$  is the limiting velocity at the current visibility.  $R_{eff}$  denotes the reward function for efficiency, and  $\omega_1, \omega_2, \omega_3$  are the weights of efficiency, position and velocity, respectively.

Safety is a necessary condition for the vehicle to change lanes. Therefore, when a collision occurs, we give a larger penalty. The safety reward function  $R_{saf}$  is shown in Equation (30):

$$R_{saf} = -m \tag{30}$$

where m denotes a larger penalty.

- Discount factor  $\gamma$ : We use the action value function to evaluate the goodness of the current action, and the value function  $Q_{target}$  is updated to consider future rewards with discounts  $\gamma$ , as shown in Equation (31):

$$Q_{target} = r_t + \gamma \sum_{s_{t+1} \in S} P(s_t, a_t, s_{t+1}) V(s_{t+1}) \tag{31}$$

where  $r_t$  is the reward corresponding to the state  $s_t$  and action  $a_t$  at moment  $t$ .  $P(s_t, a_t, s_{t+1})$  is a dynamic model of the system that represents the probability of the system changing from state  $s_{t+1}$  to state  $s_t$  after taking action  $a_t$ .

#### 4.2. The Modified DDPG Algorithm

DDPG is a deep reinforcement learning algorithm that performs continuous control. To obtain better training stability, the algorithm is designed with four networks, including the value network, the strategy network and their respective target networks, as shown in Figure 6.

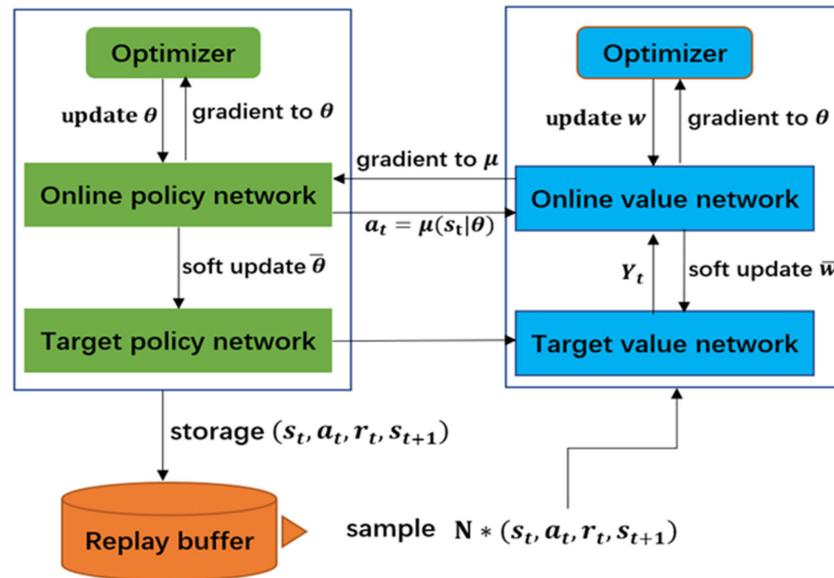


Figure 6. The framework of DDPG algorithm.

In Figure 6,  $\theta$  and  $\bar{\theta}$  are the parameters of the policy network and the target policy network, respectively.  $w$  and  $\bar{w}$  are the parameters of the value network and the target value network, respectively.  $\mu$  denotes the strategy network.

The decision-making of DDPG is a deterministic behavior. To ensure that the environment is fully explored during training and thus improves training efficiency, we construct random actions by adding an Ornstein–Uhlenbeck (OU) noise with a mean of 0 and a standard deviation of 0.6. The goal of the value network is to continuously update the parameter  $w$ , so that its evaluation of taking action  $a_t$  under state  $s_t$  is increasingly accurate. The parameter  $w$  is updated by timing differences as shown in Equations (32) and (33):

$$L(w) = \frac{1}{N} \sum_{t=1}^N \left( Y_t - Q(s, a|w)|_{s=s_t, a=a_t} \right)^2 \tag{32}$$

$$Y_t = r_t + \gamma Q(s, \mu(s|\bar{\theta})|\bar{w})|_{s=s_{t+1}, a=a_{t+1}} \tag{33}$$

where  $N$  is the number of experiences used for the update.  $Q$  denotes the value function.  $Y_t$  is the target value, which is the sum of the reward  $r_t$  at the current time step and the future reward with discount  $\gamma$ , as shown in Equation (33).

To ensure the stability of  $Y_t$  during training, we use the target network to calculate the future reward and take a soft update to update the parameters  $\bar{w}$  and  $\bar{\theta}$ , as shown in Equations (34) and (35):

$$\bar{\theta} = \tau\theta + (1 - \tau)\bar{\theta} \tag{34}$$

$$\bar{w} = \tau w + (1 - \tau)\bar{w} \tag{35}$$

where  $\tau$  is the soft update parameter.

The goal of the strategy network is to make the value network score its strategies as high as possible, and the parameters  $\theta$  are updated by gradient descent, as shown in Equation (36):

$$\nabla_{\theta} J = \frac{1}{N} \sum_{i=1}^N \nabla_a Q(s, a|w)|_{s=s_t, a=a_t} \nabla_{\theta} \mu(s|\theta)|_{s=s_t} \tag{36}$$

However, when the randomness of the environment is large, the complexity of reinforcement learning increases, which makes the training process difficult to converge. As shown in Figure 4, the lane change problem is divided into four scenarios based on the relationship between visibility and safe sight distance. In each scenario, the CAV lane-changing strategy needs to learn and complete two tasks, including adapting to different aggressiveness of HDV C and different initial states of CAV A. The combination of these two randomness can complicate the training process.

Therefore, we combine parameter-based transfer learning (TL) with DDPG to split the original complex training task into two subtasks. The first subtask is for CAV A to complete the lane change task for different trust levels of HDV C. Based on the first subtask, the second subtask is for CAV A to learn to complete the lane change at different lane change initial states. In each scenario, the first subtask training converges and then the second subtask learning is performed. Separate training of the tasks will help CAV to complete the lane changing strategy faster. The lane change learning process is shown in Figure 7.

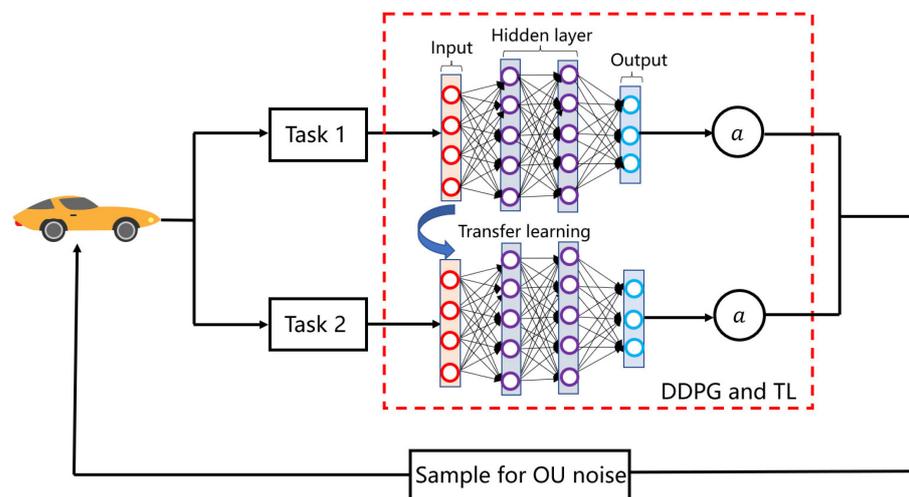


Figure 7. The learning process of lane changing strategy based on DDPG and TL.

## 5. Numerical Experiments

### 5.1. Experiment Design

First, we compared the DDPG and the modified DDPG algorithm so as to verify the effectiveness of the modified algorithm. Second, we tested the modified DDPG algorithm by adjusting the aggressiveness of the HDV driver and the initial lane change state of the CAV A in the four lane change scenarios. Finally, based on some selected metrics, we compared rule-based and DDPG-based lane change algorithms by designing some specific lane change scenarios.

All experiments were performed in MATLAB2021a on a computer equipped with an AMD Ryzen CoreTM16@ 2.90 GHz and 16.0 GB RAM. In addition, the vehicle model used the classical kinematic-based bicycle model. The position of the vehicle is represented by the position of the vehicle’s center of gravity, and the lane width is 3 m. In addition, we assumed that the normal reaction time of the driver is 0.3 s. The goal of the lane change was from the centerline of lane 1 (1.5 m) to the centerline of lane 2 (4.5 m).

Based on the typical lane changing scenario in Figure 1, four lane-changing scenarios were designed. The specific parameter settings for scenario 1 are shown in Table 1. When the visibility is 50 m and the initial velocity of HDV C is 18 m/s, the reaction delay is 0.522 s and the safety sight distance is 68.79 m, obtained by Equations (1) and (2). The safety distance between CAV A and HDV B is 10.5 m, obtained by Equations (19) and (20).

Scenarios 2, 3 and 4 have the same parameters as Table 1 except for the initial velocity of HDV C and the initial longitudinal position of vehicle B, as shown in Tables 2–4. When the initial speed of vehicle C is 14 m/s, its safe sight distance is 37 m, which is less than visibility. Therefore, the driver has no reaction delay.

**Table 1.** Parameter setting for scenario 1.

Parameter	Value	Parameter	Value
Visibility V	50 m	Limit velocity $v_{lim}$	16 m/s
Initial velocity of CAV A $v_a$	14 m/s	Initial velocity of HDV C $v_c$	18 m/s
Initial velocity of HDV B $v_b$	14 m/s	Initial velocity of HDV D $v_d$	14 m/s
Initial longitudinal position of HDV B $x_b$	100 m	Initial longitudinal position of HDV D $x_d$	100 m
Initial longitudinal position of HDV C $x_c$	0 m	Initial lateral position of CAV A $y_a$	1.5 m
Initial lateral position of HDV B $y_b$	4.5 m	Initial lateral position of HDV C $y_c$	4.5 m
Initial lateral position of HDV D $y_d$	1.5 m	Initial longitudinal position of CAV A $x_a$	15~85 m
* Aggression factor of HDV C $q$	−3~3		

\* We assume that the driver’s trust factor  $p$  and aggression factor  $q$  are equal.

**Table 2.** Parameter setting for Scenario 2.

Parameter	Value	Parameter	Value
Initial velocity of HDV C $v_c$	18 m/s	Initial longitudinal position of HDV B $x_b$	40 m

**Table 3.** Parameter setting for Scenario 3.

Parameter	Value	Parameter	Value
Initial velocity of HDV C $v_c$	14 m/s	Initial longitudinal position of HDV B $x_b$	100 m

**Table 4.** Parameter setting for Scenario 4.

Parameter	Value	Parameter	Value
Initial velocity of HDV C $v_c$	14 m/s	Initial longitudinal position of HDV B $x_b$	40 m

### 5.2. An Effectiveness Analysis of the Modified DDPG

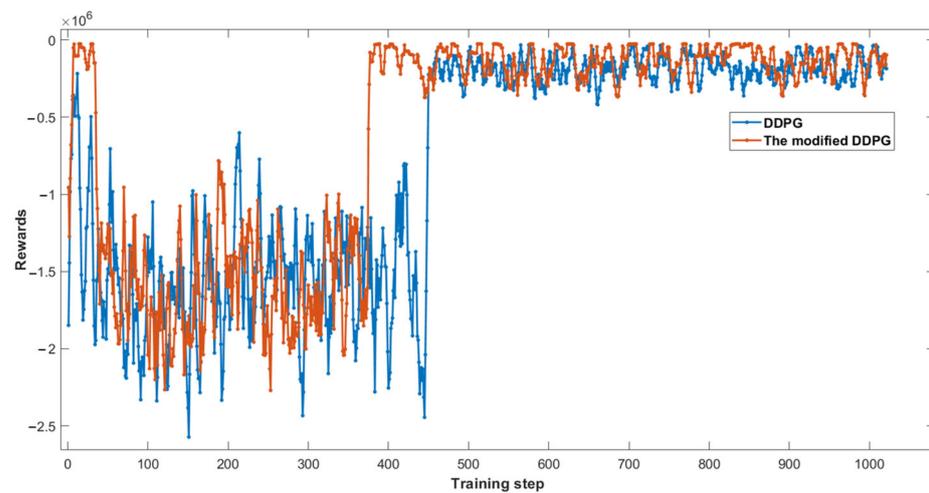
As shown in Figure 8, the blue line represents the reward based on the DDPG algorithm, and the red line represents the reward of the modified DDPG algorithm. Obviously, when trained based on the original DDPG algorithm, the CAV learns poorly and only gets stable reward reporting around 450 episodes, while the modified DDPG can converge to a stable reward around 370 episodes.

### 5.3. Testing of the Modified DDPG Algorithm

In this paper, we mainly focus on whether the modified DDPG can enable CAV A to choose the appropriate lane change opportunity in each specific scenario. Therefore, in this section, we selected several combinations of driver aggression  $q$  (trust factor  $p$ ) and the initial position  $x_a$  of CAV, as shown in the Table 5. In addition, after lane change decision, trajectory optimization and tracking are completed based on polynomial and model predictive control, which is not the focus of this paper.

**Table 5.** Test combination of  $q$  and  $x_a$ .

q	$x_a$ (m)		
	10	30	70
−3	(−3, 10)	\	\
3	(3, 10)	(3, 30)	(3, 70)



**Figure 8.** Reward changes based on DDPG algorithm and modified DDPG algorithm in scenario 1.

In order to facilitate the discussion during lane change, we subdivide the state of HDV C into the following 13 types.

1. The longitudinal spacing between CAV A and HDV C is greater than the visibility.
  - State 1: visibility is less than the safe sight distance of HDV C and there is no vehicle ahead in visibility.
  - State 2: visibility is less than the safe sight distance of HDV C and there is a vehicle ahead in visibility.
  - State 3: visibility is greater than the safe sight distance of HDV C and there is no vehicle ahead in visibility.
  - State 4: visibility is greater than the safe sight distance of HDV C and there is a vehicle ahead in visibility.
2. The longitudinal spacing between CAV A and HDV C is less than the visibility.
  - State 5: CAV A is in the lane change state.
  - State 6: CAV A has not started lane change, visibility is less than the safe sight distance of HDV C and there is no vehicle ahead in visibility.
  - State 7: CAV A has not started lane change, visibility is less than the safe sight distance of HDV C and there is a vehicle ahead in visibility.
  - State 8: CAV A has not started lane change, visibility is greater than the safe sight distance of HDV C and there is no vehicle ahead in visibility.
  - State 9: CAV A has not started lane change, visibility is greater than the safe sight distance of HDV C and there is a vehicle ahead in visibility.
  - State 10: CAV A completes lane change, visibility is less than the safe sight distance of HDV C and there is no vehicle ahead in visibility.
  - State 11: CAV A completes lane change, visibility is less than the safe sight distance of HDV C and there is a vehicle ahead in visibility.
  - State 12: CAV A completes lane change, visibility is greater than the safe sight distance of HDV C and there is no vehicle ahead in visibility.
  - State 13: CAV A completes lane change, visibility is greater than the safe sight distance of HDV C and there is a vehicle ahead in visibility.

Figure 9 shows the results of the lane change test in scenario 1. In (a), (b) and (c),  $q$  is 3 and  $x_a$  is 10 m. In (d), (e) and (f),  $q$  is  $-3$ , and  $x_a$  is 10 m. In (g), (h) and (i),  $q$  is 3 and  $x_a$  is 30 m. In (j), (k) and (l),  $q$  is 3 and  $x_a$  is 70 m. As shown in (a), when  $q$  is 3, CAV A learns to accelerate during the lane change and also makes itself a sufficiently longitudinal distance from HDV C when the lane change is completed. In contrast, HDV C chooses to yield when  $q$  is  $-3$ . In this case, considering safety and efficiency, CAV A will change the lane decisively, as shown in (d) and (e). Simultaneously, the longitudinal spacing between CAV A and HDV C is less than the visibility at the end of the lane change, as shown in

(c) and (f). When the initial longitudinal position of CAV A is 30 m, CAV A is able to complete the lane change by proper acceleration, as shown in (g) and (h). Due to the acceleration of HDV C during the lane change, the safety sight distance of HDV C is greater than the visibility when CAV A has just completed the lane change, as shown in (i). When the initial longitudinal position of CAV A is 70 m, CAV A has enough space for lane change, even if  $q$  of HDV C is 3, and CAV A can complete the lane change quickly, as shown in (j) and (k). At this moment, the longitudinal spacing between HDV C and CAV A has been greater than the visibility during the whole lane change, as shown in (l).

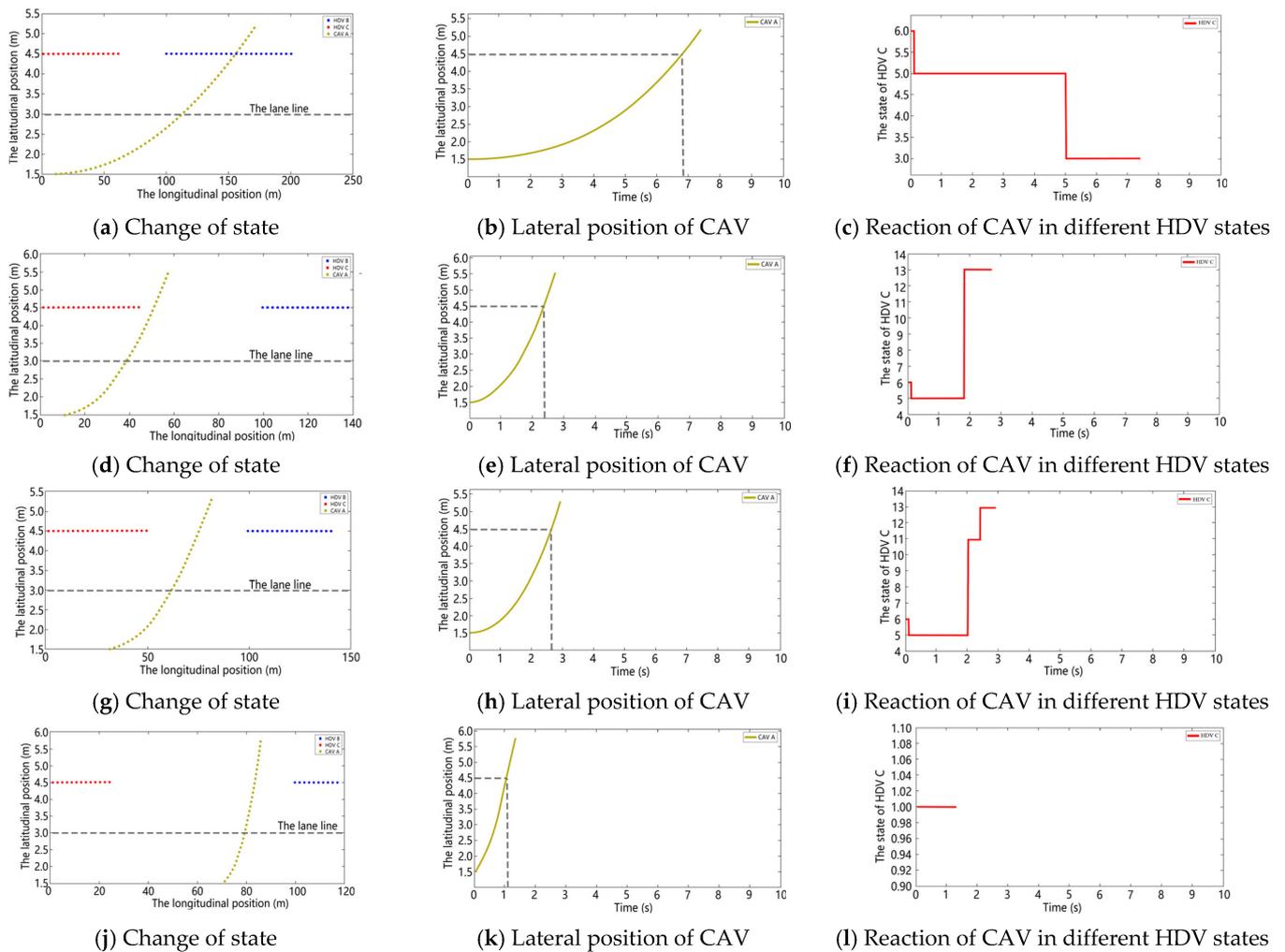


Figure 9. Test results in scenario 1.

Different from scenario 1, the longitudinal position of HDV B in scenario 2 is 40 m. Figure 10 shows the results of the lane change test in scenario 2. In (a), (b) and (c),  $q$  is  $-3$  and  $x_a$  is 10 m. In (d), (e) and (f),  $q$  is 3 and  $x_a$  is 10 m. In (g), (h) and (i),  $q$  is 3 and  $x_a$  is 30 m. When CAV A has just finished the lane change, HDV C is in state 11 due to the acceleration. Comparing (a) and (b), we can find that CAV A chooses to accelerate and complete the lane change after 50 m when HDV C is more aggressive.

After that, HDV C comes to the car following state shown in Figure 3 and comes to state 13, as shown in (f). When CAV A changes lanes at 30 m, CAV A can complete the lane change by accelerating appropriately. However, different from scenario 1, CAV A needs to maintain the distance from vehicle B. The results of lane change are shown in (g) and (h). In scenarios 3 and 4, the initial velocity of HDV C is smaller. Compared with scenarios 1 and 2, only the initial state of HDV C is different, as shown in Figure 11. While CAV A can easily learn the lane change strategy, the specific lane change test results are not given here.

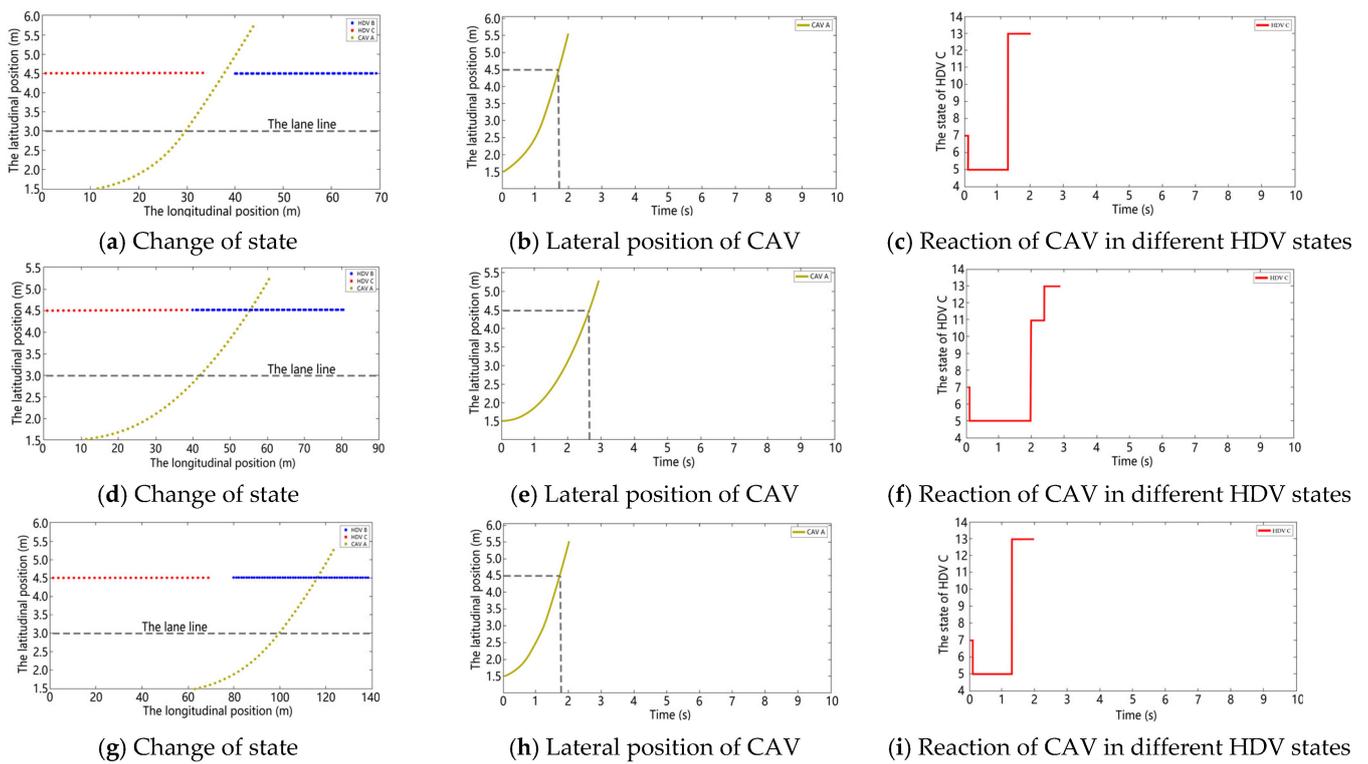


Figure 10. Test results in scenario 2.

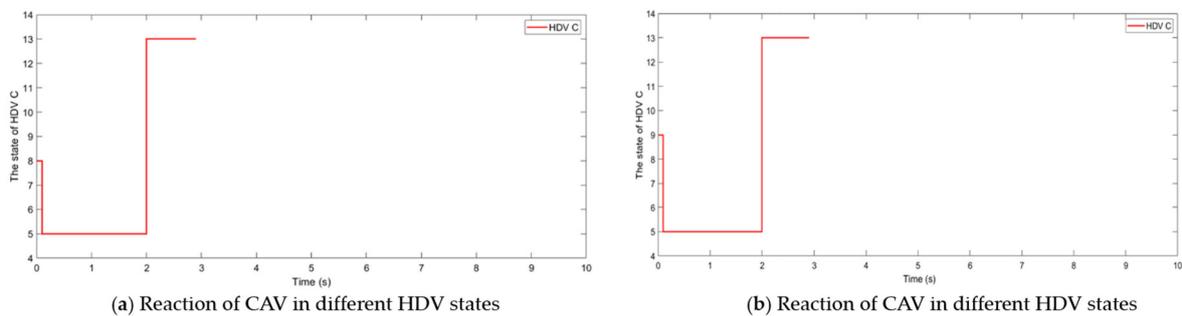


Figure 11. Test results of the state of HDV C in scenario 3 and scenario 4.

5.4. Comparison with the Rule-Based Lane Change Decision Algorithm

To further demonstrate the advantages of the modified DDPG algorithm for lane change in this particular environment, this section compares the rule-based lane change decision with the modified DDPG algorithm. In each of the four scenarios, we take any value in the range of  $(-3, 3)$  as  $q$  of HDV C, and any value in the longitudinal range between HDV C and vehicle B as the initial longitudinal position of CAV A. In each scenario, we conducted 100 test experiments for each decision algorithm and counted the success rate, collision rate and failure rate, as shown in Table 6. In addition, the threshold processing is done for the lane change trajectory learned by the DDPG algorithm so that it can make the wheel steering reduce to 0 after reaching the target lane. The rule-based lane change decision will guarantee the absolute safety of the lane change. When the current conditions cannot achieve the guaranteed safe lane change, CAV A will stay in the current lane.

As shown in Table 6, firstly, the smaller the initial velocity of HDV C, the larger the longitudinal range of the lane change and the higher the success rate. For example, the success rate of scenario 3 can reach 98.5%, while the success rate of Scenario 2 is only 93.1%. Secondly, the success rate of lane change based on the modified DDPG is higher than that of the rule-based in all four scenarios and reaches more than 90%, which indicates that

the modified DDPG is well adapted to the randomness of low visibility and mixed traffic conditions. Thirdly, for the collision cases that occur with the modified DDPG, we analyze that they are caused by the incomplete learning of the DDPG in some cases. Additionally, the incomplete cases are due to the fact that CAV A cannot really complete the lane change under the vehicle kinematic constraints in the current stochastic environment. Especially in scenario 2, the incomplete rate reaches 2.3%. The larger initial velocity of HDV C and the smaller initial lane change range increase the learning difficulty of the CAV lane change strategy. Fourth, we can find that the rule-based lane change decision method requires higher lane change conditions. When CAV A cannot achieve the lane change condition described in Section 3.2, CAV A will continue to drive in the current lane. In particular, the lane change failure rate reaches 59.4% for HDV C with an initial velocity of 18 m/s and an initial lane change range of 40 m.

**Table 6.** Performances of the modified DDPG and the rule-based lane change decision algorithms.

Scenario	The Modified DDPG			The Rule-Based	
	Success	Collision	Unfinished	Success	Unfinished
Scenario 1	97.2%	1.8%	1%	85.4%	14.6%
Scenario 2	93.1%	4.6%	2.3%	40.6%	59.4%
Scenario 3	98.5%	0.4%	1.1%	91.1%	8.9%
Scenario 4	95.8%	2.6%	1.6%	68.9%	31.1%

Two special scenarios are selected to illustrate the two decision algorithms and give the historical trajectory, velocity, steering and yaw variation of the vehicle, respectively. In test scenario 1, the  $q$  of HDV C is  $-3$ , the initial  $v_C$  is 18 m/s,  $v_A$  is 14 m/s, the initial longitudinal spacing between HDV C and CAV A is 20 m, the reaction time  $T$  is 0.3 s and the absolute values of  $a_{max}$  and  $a_{min}$  are  $3 \text{ m/s}^2$ .

According to Equation (21), the rule-based lane change decision is executed when the longitudinal spacing between HDV C and CAV A reaches 32.4 m, which is obviously not satisfied. The CAV A stays in the current lane with no change in speed, steering angle and heading angle throughout the process. The trajectory of CAV A is the red line, and the tracks of other vehicles are blue lines, as shown in Figures 12 and 13. However, the modified DDPG finds the lower aggressiveness of HDV C and completes the lane change just at the beginning of the lane change, as shown in Figures 14 and 15. In test scenario 2, the  $q$  of HDV C is 3, the initial  $v_C$  is 14 m/s and the initial longitudinal spacing between HDV C and CAV A is 10 m. Other conditions are the same as in test scenario 1. Similarly, the rule-based lane change decision will not be executed because the distance condition is not satisfied, as shown in Figures 16 and 17. However, based on the modified DDPG decision algorithm, CAV A accelerates at the beginning. At around 2.7 s,  $v_A$  approaches 18 m/s, and CAV A begins to change lanes, as shown in Figures 18 and 19.

To better illustrate the effect of the modified DDPG, the emergency lane change scenario and the non-emergency lane change scenario are also selected. The success rate and time for lane change are given, respectively, and compared with both rule-based and DQN-based.

The comparison results are shown in Table 7 and Figure 20. The success rate of the proposed method is 90.36% and the time required is 5.2 s in the emergency lane change scenario, both of which are better than the rule-based 45.12% and 7.6 s and the DQN-based 68.23% and 6.2 s. The comparison results are shown in Table 8 and Figure 21. The success rate of the proposed method is 92.15% and the time required is 5.8 s in the non-emergency lane change scenario. The success rate of the proposed method is better than the rule-based 47.98% and the DQN-based 64.58%, but the time required is only better than the rule-based 7.3 s and slightly weaker than the DQN-based 5.7 s. Overall, our proposed algorithm is clearly superior in both emergency and non-emergency lane change scenarios.



Figure 12. The result of the trajectories of test scenario 1 based on the rule-based lane change decision.

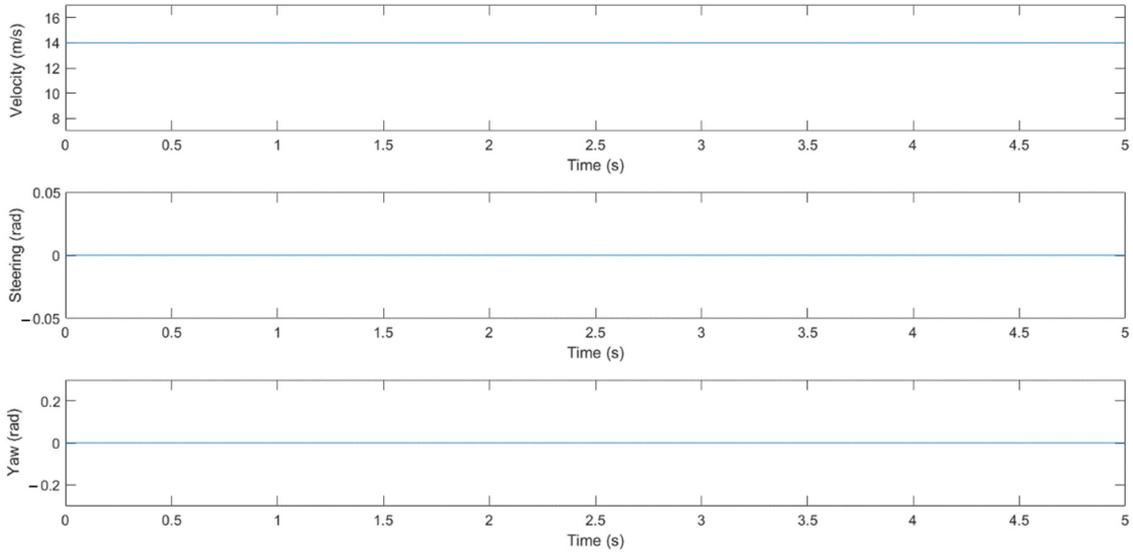


Figure 13. The velocity, steering and yaw of test scenario 1 based on the rule-based lane change decision.

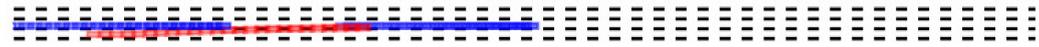


Figure 14. The result of the trajectory of test scenario 1 based on the modified DDPG.

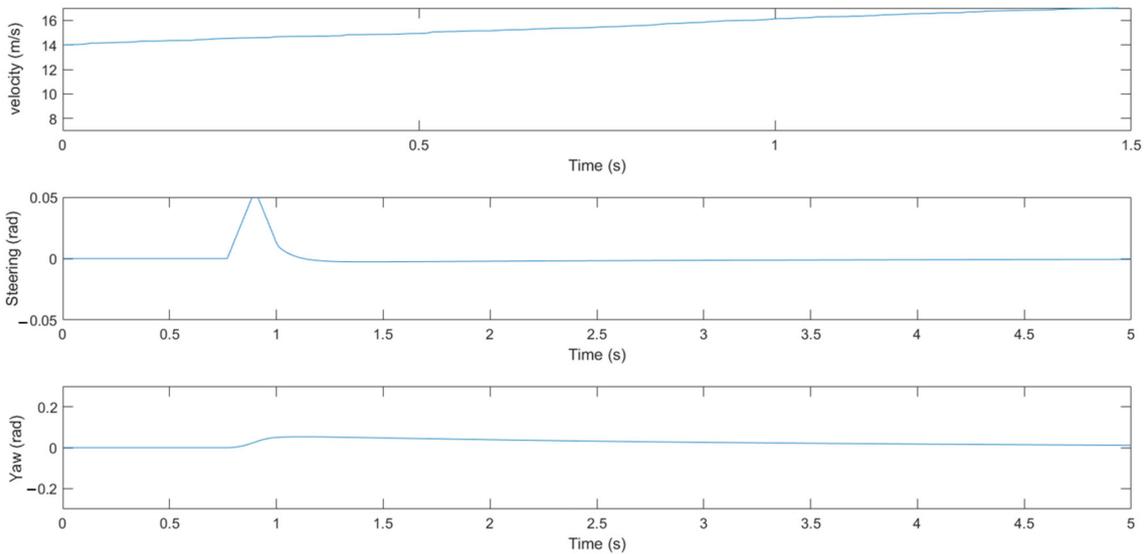


Figure 15. The velocity, steering and yaw of test scenario 1 based on the modified DDPG.



Figure 16. The result of the trajectory of test scenario 2 based on the rule-based lane change decision.

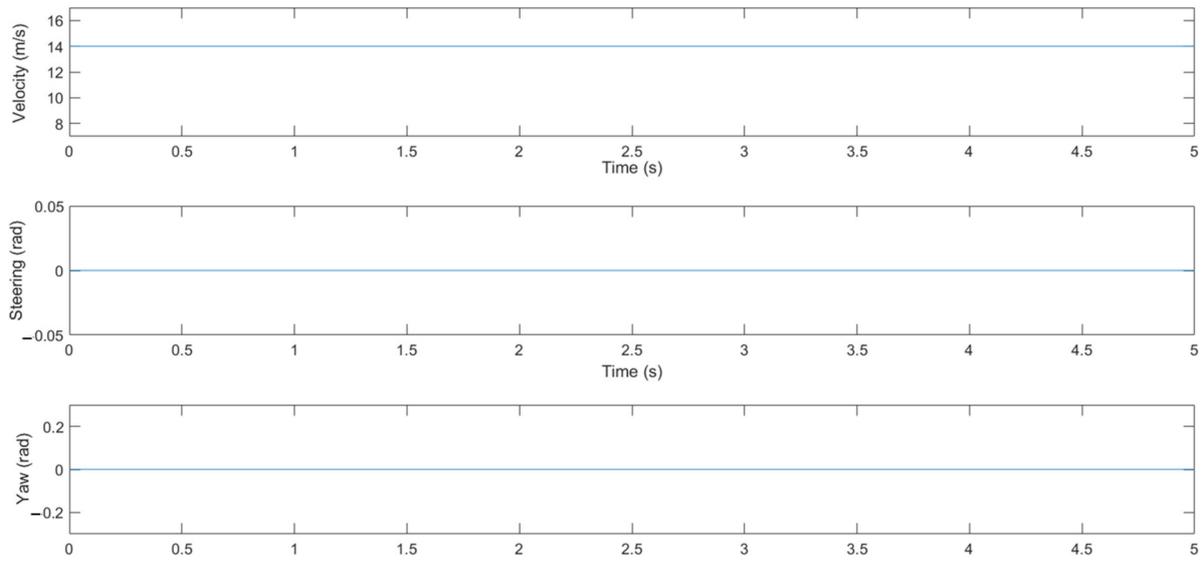


Figure 17. The velocity, steering and yaw of test scenario 2 based on the rule-based lane change decision.

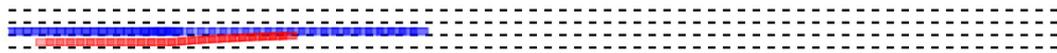


Figure 18. The result of the trajectory of test scenario 2 based on the modified DDPG.

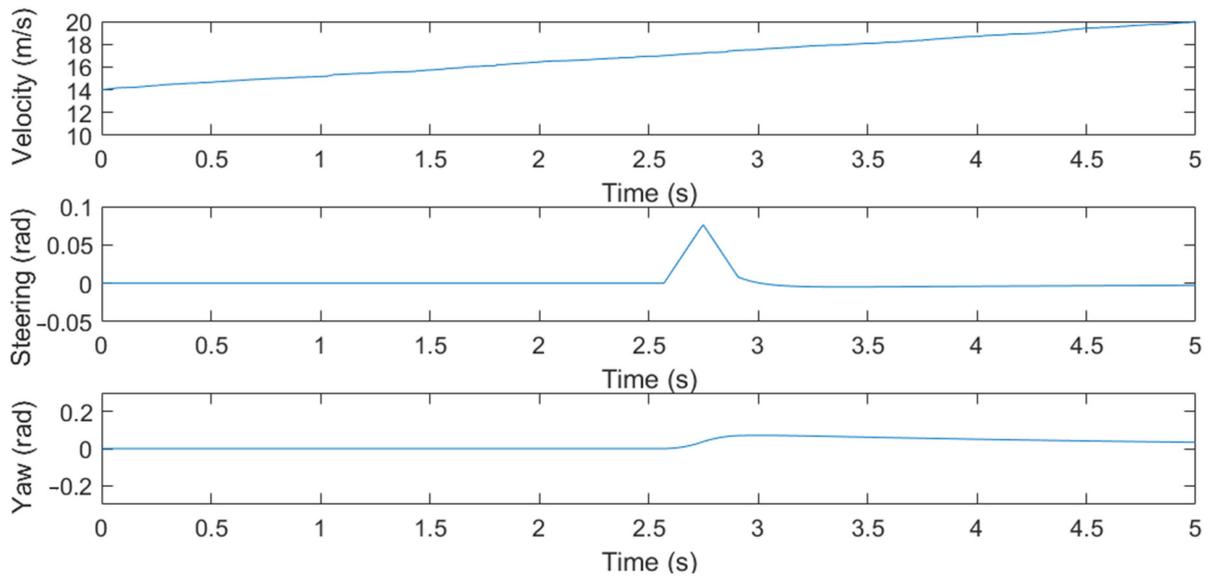
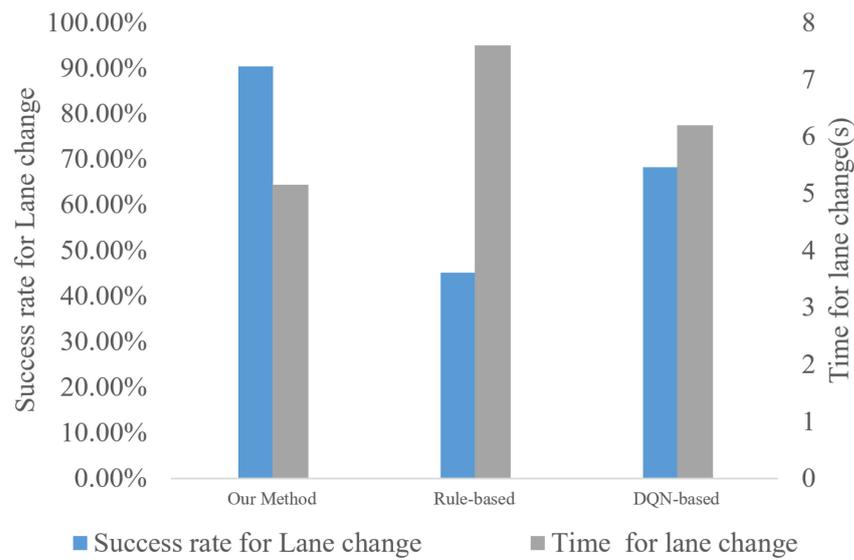


Figure 19. The velocity, steering and yaw of test scenario 2 based on the modified DDPG.

Table 7. Comparison of modified DDPG and rule-based and DQN-based in the emergency lane change scenario.

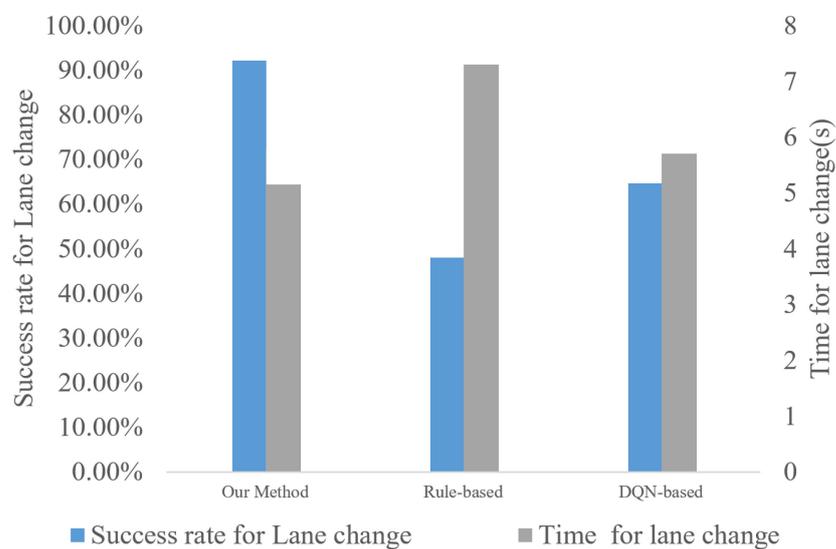
	Success Rate for Lane Change	Time for Lane Change
Our Method	90.36%	5.2 s
Rule-based	45.12%	7.6 s
DQN-based	68.23%	6.2 s



**Figure 20.** The success rate and time for lane change of the emergency lane change scenario based on three methods.

**Table 8.** Comparison of modified DDPG and rule-based and DQN-based in the non-emergency lane change scenario.

	Success Rate for Lane Change	Time for Lane Change
Our Method	92.15%	5.8 s
Rule-based	47.98%	7.3 s
DQN-based	64.58%	5.7 s



**Figure 21.** The success rate and time for lane change of the non-emergency lane change scenario based on three methods.

### 6. Conclusions

In this paper, we present a lane change decision problem for CAVs in low visibility and mixed traffic conditions.

We first consider the effects of low visibility and mixed traffic conditions on the CAV, and then propose an FSM model to describe the driving states of the HDV. Secondly, we use the modified DDPG algorithm to learn the lane change decision of the CAV. Finally, we verify that the modified DDPG converges faster. It is worth noting that the modified

DDPG is more adaptive to the environment and can find more lane change opportunities compared to the rule-based lane change decision.

In the future, obtaining the actual driving data of HDV in low visibility and mixed traffic conditions can be further considered. It is also possible to further explore the issues of lane change decisions and overtaking decisions for multiple lanes in low visibility and mixed traffic conditions. On the basis of vehicle micro-control, the traffic operation rule and vehicle scheduling in this environment can continue to be studied.

**Author Contributions:** Conceptualization, B.G., R.W., T.W. and C.L.; methodology, Z.X. and P.G.; software, R.W.; validation, Z.X. and R.W.; formal analysis, B.G., Z.X. and T.W.; investigation, Z.X.; resources, T.W. and P.G.; data curation, Z.X. and P.G.; writing—original draft preparation, B.G., Z.X. and R.W.; writing—review and editing, Z.X. and R.W.; visualization, B.G., T.W., R.W. and C.L.; supervision, T.W. and C.L.; project administration, T.W., P.G. and C.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors acknowledge the Scientific Research Project of the Education Department of Jilin Province (Grant No. JJKH20221020KJ), the Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, China Academy of Transportation Science (Grant No. 2021B1206) and the Qingdao Social Science Planning Research Project (Grant No. 2022-389), who are partly supporting this work.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Li, G.; Yang, Y.; Zhang, T.; Qu, X.; Cao, D.; Cheng, B.; Li, K. Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios. *Transp. Res. Part C Emerg. Technol.* **2021**, *122*, 102820. [[CrossRef](#)]
- Deng, J.-H.; Feng, H.-H. A multilane cellular automaton multi-attribute lane-changing decision model. *Phys. A Stat. Mech. Its Appl.* **2019**, *529*, 121545. [[CrossRef](#)]
- Hruszczak, M.; Lowe, B.T.; Schrodell, F.; Freese, M.; Bajcinca, N. Game Theoretical Decision Making Approach for a Cooperative Lane Change. *IFAC-PapersOnLine* **2020**, *53*, 15247–15252. [[CrossRef](#)]
- Das, A.; Khan, M.N.; Ahmed, M.M. Detecting lane change maneuvers using SHRP2 naturalistic driving data: A comparative study machine learning techniques. *Accid. Anal. Prev.* **2020**, *142*, 105578. [[CrossRef](#)]
- Xu, T.; Zhang, Z.; Wu, X.; Qi, L.; Han, Y. Recognition of lane-changing behaviour with machine learning methods at freeway off-ramps. *Phys. A Stat. Mech. Its Appl.* **2021**, *567*, 125691. [[CrossRef](#)]
- Yoneda, K.; Suganuma, N.; Yanase, R.; Aldibaja, M. Automated driving recognition technologies for adverse weather conditions. *IATSS Res.* **2019**, *43*, 253–262. [[CrossRef](#)]
- Furda, A.; Vlacic, L. Multiple Criteria-Based Real-Time Decision Making by Autonomous City Vehicles. *IFAC Proc. Vol.* **2010**, *43*, 97–102. [[CrossRef](#)]
- Chen, D.; Srivastava, A.; Ahn, S. Harnessing connected and automated vehicle technologies to control lane changes at freeway merge bottlenecks in mixed traffic. *Transp. Res. Part C Emerg. Technol.* **2021**, *123*, 102950. [[CrossRef](#)]
- Karimi, M.; Roncoli, C.; Alecsandru, C.; Papageorgiou, M. Cooperative merging control via trajectory optimization in mixed vehicular traffic. *Transp. Res. Part C Emerg. Technol.* **2020**, *116*, 102663. [[CrossRef](#)]
- Yu, H.; Tseng, H.E.; Langari, R. A human-like game theory-based controller for automatic lane changing. *Transp. Res. Part C Emerg. Technol.* **2018**, *88*, 140–158. [[CrossRef](#)]
- Chen, K.; Pei, X.; Okuda, H.; Zhu, M.; Guo, X.; Guo, K.; Suzuki, T. A hierarchical hybrid system of integrated longitudinal and lateral control for intelligent vehicles. *ISA Trans.* **2020**, *106*, 200–212. [[CrossRef](#)] [[PubMed](#)]
- Yi, H.; Edara, P.; Sun, C. Modeling Mandatory Lane Changing Using Bayes Classifier and Decision Trees. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 647–655. [[CrossRef](#)]
- Jin, C.J.; Knoop, V.L.; Li, D.W.; Meng, L.Y.; Wang, H. Discretionary lane-changing behavior: Empirical validation for one realistic rule-based model. *Transp. A* **2019**, *15*, 244–262. [[CrossRef](#)]
- Xi, C.; Shi, T.; Wu, Y.; Sun, L. Efficient Motion Planning for Automated Lane Change based on Imitation Learning and Mixed-Integer Optimization. In Proceedings of the 23rd IEEE International Conference on Intelligent Transportation Systems, Rhodes, Greece, 20–23 September 2020.
- Jin, H.; Duan, C.; Liu, Y.; Lu, P. Gauss mixture hidden Markov model to characterise and model discretionary lane-change behaviours for autonomous vehicles. *IET Intell. Transp. Syst.* **2020**, *14*, 401–411. [[CrossRef](#)]
- Tang, J.J.; Liu, F.; Zhang, W.H.; Ke, R.M.; Zou, Y.J. Lane-changes prediction based on adaptive fuzzy neural network. *Expert Syst. Appl.* **2018**, *91*, 452–463. [[CrossRef](#)]

17. Sheikh, M.S.; Wang, J.; Regan, A. A game theory-based controller approach for identifying incidents caused by aberrant lane changing behavior. *Phys. A Stat. Mech. Its Appl.* **2021**, *580*, 126162. [[CrossRef](#)]
18. Lowe, R.; Wu, Y.; Tamar, A.; Harb, J.; Abbeel, P.; Mordatch, I. Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments. In Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA, 4–9 December 2017.
19. Peng, H.X.; Shen, X.M. Multi-Agent Reinforcement Learning Based Resource Management in MEC- and UAV-Assisted Vehicular Networks. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 131–141. [[CrossRef](#)]
20. Zheng, Z.D. Recent developments and research needs in modeling lane changing. *Transp. Res. Part B—Methodol.* **2014**, *60*, 16–32. [[CrossRef](#)]
21. Shalev-Shwartz, S.; Shammah, S.; Shashua, A. On a formal model of safe and scalable self-driving cars. *arXiv* **2017**, arXiv:1708.06374.
22. Gipps, P.G. A model for the structure of lane-changing decisions. *Transp. Res. Part B Methodol.* **1986**, *20*, 403–414. [[CrossRef](#)]
23. Rickert, M.; Nagel, K.; Schreckenberg, M.; Latour, A. Two lane traffic simulations using cellular automata. *Phys. A* **1996**, *231*, 534–550. [[CrossRef](#)]
24. Kesting, A.; Treiber, M.; Helbing, D. General lane-changing model MOBIL for car-following models. *Transp. Res. Rec.* **2007**, *1999*, 86–94. [[CrossRef](#)]
25. Wang, Z.; Zhao, X.M.; Xu, Z.G.; Li, X.P.; Qu, X.B. Modeling and field experiments on autonomous vehicle lane changing with surrounding human-driven vehicles. *Comput. Aided Civ. Infrastruct. Eng.* **2021**, *36*, 877–889. [[CrossRef](#)]
26. Sun, D.; Kondyli, A. Modeling Vehicle Interactions during Lane-Changing Behavior on Arterial Streets. *Comput. Aided Civ. Infrastruct. Eng.* **2010**, *25*, 557–571. [[CrossRef](#)]
27. Peng, J.S.; Guo, Y.S.; Fu, R.; Yuan, W.; Wang, C. Multi-parameter prediction of drivers' lane-changing behaviour with neural network model. *Appl. Ergon.* **2015**, *50*, 207–217. [[CrossRef](#)]
28. Zhao, C.; Li, Z.H.; Li, L.; Wu, X.B.; Wang, F.Y. A negotiation-based right-of-way assignment strategy to ensure traffic safety and efficiency in lane changes. *IET Intell. Transp. Syst.* **2021**, *15*, 1345–1358. [[CrossRef](#)]
29. Hidas, P. Modelling lane changing and merging in microscopic traffic simulation. *Transp. Res. Part C Emerg. Technol.* **2002**, *10*, 351–371. [[CrossRef](#)]
30. Hidas, P. Modelling vehicle interactions in microscopic simulation of merging and weaving. *Transp. Res. Part C Emerg. Technol.* **2005**, *13*, 37–62. [[CrossRef](#)]
31. Yoo, J.H.; Langari, R. Stackelberg Game Based Model of Highway Driving. In Proceedings of the 5th Annual Dynamic Systems and Control Division Conference/11th JSME Motion and Vibration Conference, Fort Lauderdale, FL, USA, 17–19 October 2012; pp. 499–508.
32. Chen, T.Y.; Shi, X.P.; Wong, Y.D. Key feature selection and risk prediction for lane-changing behaviors based on vehicles' trajectory data. *Accid. Anal. Prev.* **2019**, *129*, 156–169. [[CrossRef](#)]
33. Gindele, T.; Brechtel, S.; Dillmann, R. A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments. In Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems, Funchal, Portugal, 19–22 September 2010; pp. 1625–1631.
34. Chen, Y.L.; Dong, C.Y.; Palanisamy, P.; Mudalige, P.; Muelling, K.; Dolan, J.M. Attention-based Hierarchical Deep Reinforcement Learning for Lane Change Behaviors in Autonomous Driving. In Proceedings of the 32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; pp. 1326–1334.
35. Jiang, S.H.; Chen, J.Y.; Shen, M.C. An Interactive Lane Change Decision Making Model With Deep Reinforcement Learning. In Proceedings of the 7th IEEE International Conference on Control, Mechatronics and Automation (ICCMA), Delft, The Netherlands, 6–8 November 2019; pp. 370–376.
36. Wang, P.; Li, H.H.; Chan, C.Y. Continuous Control for Automated Lane Change Behavior Based on Deep Deterministic Policy Gradient Algorithm. In Proceedings of the 30th IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; pp. 1454–1460.
37. Lv, K.; Pei, X.; Chen, C.; Xu, J. A Safe and Efficient Lane Change Decision-Making Strategy of Autonomous Driving Based on Deep Reinforcement Learning. *Mathematics* **2022**, *10*, 1551. [[CrossRef](#)]
38. Kim, M.-S.; Eoh, G.; Park, T. Reinforcement learning with data augmentation for lane change decision-making. *J. Inst. Control Robot. Syst.* **2021**, *27*, 572–577. [[CrossRef](#)]
39. Wang, G.; Hu, J.; Li, Z.; Li, L. Harmonious Lane Changing via Deep Reinforcement Learning. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 4642–4650. [[CrossRef](#)]
40. Ammourah, R.; Talebpour, A. Deep Reinforcement Learning Approach for Automated Vehicle Mandatory Lane Changing. *Transp. Res. Rec.* **2022**, *2677*, 712–724. [[CrossRef](#)]
41. He, X.; Yang, H.; Hu, Z.; Lv, C. Robust Lane Change Decision Making for Autonomous Vehicles: An Observation Adversarial Reinforcement Learning Approach. *IEEE Trans. Intell. Veh.* **2022**, *8*, 184–193. [[CrossRef](#)]
42. Wallace, A.M.; Halimi, A.; Buller, G.S. Full Waveform LiDAR for Adverse Weather Conditions. *IEEE Trans. Veh. Technol.* **2020**, *69*, 7064–7077. [[CrossRef](#)]
43. Heinzler, R.; Schindler, P.; Seekircher, J.; Ritter, W.; Stork, W.; IEEE. Weather Influence and Classification with Automotive Lidar Sensors. In Proceedings of the 30th IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; pp. 1527–1534.

44. Chen, C.; Zhao, X.; Liu, H.; Ren, G.; Liu, X. Influence of adverse weather on drivers' perceived risk during car following based on driving simulations. *J. Mod. Transp.* **2019**, *27*, 282–292. [[CrossRef](#)]
45. Caro, S.; Cavallo, V.; Marendaz, C.; Boer, E.R.; Vienne, F. Can headway reduction in fog be explained by impaired perception of relative motion? *Hum. Factors* **2009**, *51*, 378–392. [[CrossRef](#)]
46. Saffarian, M.; Happee, R.; de Winter, J.C.F. Why do drivers maintain short headways in fog? A driving-simulator study evaluating feeling of risk and lateral control during automated and manual car following. *Ergonomics* **2012**, *55*, 971–985. [[CrossRef](#)]
47. Huang, H.; Xu, H.; Chen, F.; Zhang, C.; Mohammadzadeh, A. An Applied Type-3 Fuzzy Logic System: Practical Matlab Simulink and M-Files for Robotic, Control, and Modeling Applications. *Symmetry* **2023**, *15*, 475. [[CrossRef](#)]
48. Jomaa, M.; Abbes, M.; Tadeo, F.; Mami, A. Greenhouse Modeling, Validation and Climate Control based on Fuzzy Logic. *Eng. Technol. Appl. Sci. Res.* **2019**, *9*, 4405–4410. [[CrossRef](#)]
49. Huang, C.; Farooq, U.; Liu, H.; Gu, J.; Luo, J. A Pso-Tuned Fuzzy Logic System for Position Tracking of Mobile Robot. *Int. J. Robot. Autom.* **2019**, *34*, 84–94. [[CrossRef](#)]
50. Rajagiri, A.K.; Mn, S.R.; Nawaz, S.S.; Suresh Kumar, T. Speed control of DC motor using fuzzy logic controller by PCI 6221 with MATLAB (Conference Paper). *E3S Web Conf.* **2019**, *87*, 01004. [[CrossRef](#)]
51. Zhao, X.; Wang, Z.; Xu, Z.; Wang, Y.; Li, X.; Qu, X. Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle. *Transp. Res. Part C Emerg. Technol.* **2020**, *114*, 205–224. [[CrossRef](#)]
52. Gao, K.; Tu, H.; Shi, H.; Li, Z. Effect of low-visibility in haze weather condition on longitudinal driving behavior in different car following stages. *J. Jilin Univ.* **2017**, *47*, 1716–1727.
53. Jiang, R.; Wu, Q.; Zhu, Z. Full velocity difference model for a car-following theory. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **2001**, *64*, 017101. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.