mathematics

MDPI

Article

Investigating Autonomous Vehicle Driving Strategies in Highway Ramp Merging Zones

Zhimian Chen [1], Yizeng Wang [1], Hao Hu [2,*], Zhipeng Zhang [2], Chengwei Zhang [1] and Shukun Zhou [3]

[1] School of Ocean and Civil Engineering, Shanghai Jiaotong University, Shanghai 200240, China; chen.z.m@sjtu.edu.cn (Z.C.); wangyizeng@sjtu.edu.cn (Y.W.)
[2] State Key Laboratory of Ocean Engineering, Shanghai Jiaotong University, Shanghai 200240, China; zp.zhang@sjtu.edu.cn
[3] Shandong Hi-Speed Construction Management Group Co., Ltd., Jinan 250014, China
* Correspondence: hhu@sjtu.edu.cn

Abstract: The rapid development of autonomous driving technology is widely regarded as a potential solution to current traffic congestion challenges and the future evolution of intelligent vehicles. Effective driving strategies for autonomous vehicles should balance traffic efficiency with safety and comfort. However, the complex driving environment at highway entrance ramp merging areas presents a significant challenge. This study constructed a typical highway ramp merging scenario and utilized deep reinforcement learning (DRL) to develop and regulate autonomous vehicles with diverse driving strategies. The SUMO platform was employed as a simulation tool to conduct a series of simulations, evaluating the efficacy of various driving strategies and different autonomous vehicle penetration rates. The quantitative results and findings indicated that DRL-regulated autonomous vehicles maintain optimal speed stability during ramp merging, ensuring safe and comfortable driving. Additionally, DRL-controlled autonomous vehicles did not compromise speed during lane changes, effectively balancing efficiency, safety, and comfort. Ultimately, this study provides a comprehensive analysis of the potential applications of autonomous driving technology in highway ramp merging zones under complex traffic scenarios, offering valuable insights for addressing these challenges effectively.

Keywords: autonomous driving; ramp merging zones; mixed traffic flow; deep reinforcement learning

MSC: 93C40

check for updates

Citation: Chen, Z.; Wang, Y.; Hu, H.; Zhang, Z.; Zhang, C.; Zhou, S. Investigating Autonomous Vehicle Driving Strategies in Highway Ramp Merging Zones. *Mathematics* **2024**, *12*, 3859. https://doi.org/10.3390/math12233859

Academic Editor: Bo-Hao Chen

Received: 25 October 2024
Revised: 15 November 2024
Accepted: 19 November 2024
Published: 8 December 2024

Copyright: © 2024 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/).

1. Introduction

The automotive industry is experiencing an unprecedented transformation driven by rapid technological advancements, extending well beyond conventional vehicle design and manufacturing, and fundamentally reshaping the transportation landscape. Among these innovations, autonomous driving has emerged as a revolutionary breakthrough poised to redefine the future of road transportation by addressing critical issues related to safety, efficiency, and traffic management [1]. Many nations have identified autonomous driving technology as a cornerstone strategy for revitalizing and advancing their automotive industries. This emphasizes that advanced autonomous driving technology is now a significant driver of national competitiveness, technological progress, and societal development worldwide [2–4].

The potential of autonomous vehicles (AVs) to significantly enhance road safety by minimizing human error, alleviating congestion through optimized vehicle coordination, and improving overall passenger comfort has attracted substantial attention from both industry and academia [5]. The development of autonomous driving technology is not merely a response to transportation challenges; rather, it reflects a broader progression

toward the creation of intelligent, automated systems capable of assisting or even replacing human operators across various sectors [6]. Initially driven by the necessity for improved vehicle safety systems, foundational technologies such as adaptive cruise control (ACC), lane-keeping assistance, and collision avoidance have laid the groundwork for more sophisticated automation [7]. The evolution of autonomous systems is mirrored in other transportation domains as well. For example, Song et al. proposed an autonomous route management system for railways, aiming to enhance operational efficiency while ensuring functional safety, demonstrating similar priorities of optimizing safety and efficiency [8]. The emergence of connected and automated vehicles (CAVs) has further accelerated this evolution by incorporating V2X (vehicle-to-everything) communication, allowing vehicles to interact seamlessly with each other and surrounding infrastructure to enhance safety and operational efficiency [9]. The growing urbanization, coupled with increasing vehicle densities and escalating traffic congestion, has driven the demand for safe, efficient, and comfortable transportation modes—serving as a major impetus for the advancement of autonomous driving technologies [10]. The prevalence of road accidents—often linked to human error—and associated traffic inefficiencies, such as congestion and bottlenecks, have highlighted the urgent need for innovative solutions to substantially improve traffic safety and efficiency [11–14].

Freeway entrance ramp merging areas are among the more challenging scenarios for road transportation. These zones, where vehicles enter from on-ramps and must seamlessly integrate into the existing traffic flow, present a high degree of complexity and heightened risk [15]. These environments are characterized by variations in vehicle speeds, aggressive driver behaviors, and dense traffic, often culminating in congestion and accidents [16]. Human drivers frequently encounter difficulties in identifying optimal gaps for merging, adjusting their speeds accordingly, and predicting the behaviors of surrounding vehicles [17]. Hu et al. showed that the complexity of vehicle negotiation and interaction during merging requires drivers to complete multi-dimensional information processing and vehicle control tasks within a limited distance and time frame, significantly increasing both the workload and the probability of driver errors [18]. Zhu et al. argued that a driver's inability to accurately grasp the timing of lane changes can also lead to congestion and queue overflow in ramp merging areas [19]. Moreover, frequent "stop-and-go" traffic fluctuations resulting from unstable vehicle speeds can lead to a marked decline in occupant comfort [20,21]. Based on the above studies, it can be concluded that freeway ramp merging areas, as typical bottlenecks in freeway networks, face challenges not only due to their unique geometric design that reduces road capacity in merging areas but also because these areas require frequent lane changes and diverse driving behaviors, which contribute to increased risks of collisions, conflicts, congestion, and queuing. Therefore, mitigating these issues to eliminate the critical human factors in driving and significantly improve traffic safety, operational efficiency, and passenger comfort remains an essential objective. Autonomous vehicles (AVs), equipped with advanced sensors and sophisticated decision-making algorithms, present a promising solution to these challenges. Unlike human drivers, AVs can precisely calculate their position, speed, and state of the surrounding environment, enabling them to perform merging maneuvers with enhanced accuracy and efficiency. Li et al. proposed a control framework for autonomous vehicles that effectively coordinates both longitudinal and lateral movements during ramp merging, thereby significantly improving safety and efficiency compared to human-operated vehicles [22]. Han et al. introduced a cooperative merging strategy aimed at optimizing road space utilization and reducing traffic conflicts in merging zones [23]. Similarly, Zhao et al. proposed an on-ramp merging strategy that jointly optimized lane-keeping and lane-changing behaviors, resulting in improved travel times [24]. The Cooperative Adaptive Cruise Control (CACC) system represents another technological advancement, allowing autonomous vehicles to communicate with one another and with surrounding infrastructure, thereby enhancing coordination during merging operations. CACC helps mitigate variability in driving behavior, prevent sudden deceleration, and reduce the likelihood of traffic waves and

collisions. Zhou et al. proposed a cooperative framework based on the "follow-the-leader" strategy, developed by the PATH laboratory at the University of California, Berkeley, demonstrating that increased CACC penetration can effectively alleviate congestion and enhance traffic capacity and stability [25]. Song et al. discussed the development of train-centric communication-based autonomous train control systems, which emphasize data prediction and fusion, comparable to how V2X communication allows AVs to interact with each other and infrastructure to enhance safety and operational efficiency [26]. Weaver et al. conducted comparative experiments demonstrating that CACC can enhance road accessibility by reducing gaps between vehicles; however, smaller gaps may also hinder a human driver's ability to safely enter a CACC-controlled traffic stream [27]. In this study, instead of focusing on cooperative control within the autonomous traffic stream, we focus on single-vehicle control of each autonomous vehicle, emphasizing individual decision-making control rather than vehicle-to-vehicle communication.

Despite the potential of autonomous vehicles to alleviate the challenges associated with merging zones, it is crucial to acknowledge that AVs will not immediately dominate roadways. In the foreseeable future, road networks will feature mixed traffic environments where AVs must coexist with human-driven vehicles (HDVs). The unpredictability of human drivers, often influenced by stress, distractions, or personal driving preferences, adds complexity to these interactions, making it challenging for AVs to seamlessly coexist with HDVs [28–30]. Thus, optimizing the interaction between AVs and HDVs is vital to ensuring smooth integration during the transition period, especially when AV penetration rates are still low. Imbsweiler et al. investigated the interaction between AVs and HDVs in narrow road scenarios, proposing a series of interaction behaviors to mitigate conflicts between the two under complex traffic conditions [31]. Cui et al. utilized the SUMO simulation platform to model mixed traffic environments comprising AVs, HDVs, and intelligent network-connected vehicles at varying ratios, deriving the optimal penetration rate for AVs in urban intersection scenarios by analyzing relevant traffic parameters [32]. Ramezan et al. explored improvements in traffic performance through dedicated lanes for AVs or HDVs, suggesting designs for various lane combinations depending on AV penetration levels [33]. Zhang et al. proposed a multi-robot learning approach based on proximal policy optimization (PPO), which significantly enhances the success rate of decision-making during high-speed on-ramp merging involving AVs and HDVs [23]. These studies underscore the need for sophisticated coordination strategies for AVs to manage the unpredictability of human drivers, particularly in merging scenarios where rapid decision-making is crucial to maintaining traffic stability and safety within mixed traffic flows.

Deep reinforcement learning (DRL) has emerged as a promising solution to address the complexities of mixed traffic environments involving autonomous vehicles (AVs) and human-driven vehicles (HDVs) [34]. By integrating reinforcement learning with deep neural networks, DRL enables agents to learn optimal actions through interaction with their environment, adapting strategies based on real-time feedback [35–37]. This approach is particularly advantageous in autonomous driving scenarios characterized by dynamic and uncertain conditions involving multiple agents. Li et al. proposed a DRL method that carefully incorporates map elements to address the challenges of following and obstacle avoidance for AVs, introducing an obstacle representation to handle the non-fixed number and state of external obstacles [38]. Bin et al. developed a model combining Faster R-CNN with Double Deep Q Learning and demonstrated the efficacy of DRL in enhancing AV efficiency in a simulated real-world environment [39]. Chen et al. introduced a multi-agent reinforcement learning (MARL) framework for highway on-ramp merging, wherein AVs collaboratively learn policies to adapt to HDVs and optimize overall traffic throughput [40]. Their framework utilized parameter sharing among agents to promote cooperation and scalability, improving AV navigation in mixed traffic conditions. Lee et al. also explored DRL-based strategies for on-ramp merging, emphasizing the potential of these methods to outperform traditional control approaches in unpredictable and dynamic traffic scenar-

ios [41]. The key advantage of DRL is its capability to manage the inherent complexity and uncertainty of mixed traffic environments, allowing AVs to make decisions that balance multiple objectives, such as minimizing travel time, ensuring safety, and maintaining passenger comfort [5]. This is especially critical in freeway merging areas, where AVs must simultaneously manage acceleration, lane changes, and interactions with surrounding HDVs. Zhao et al. presented a DRL-based control framework that integrated lane-keeping and lane-changing behaviors, demonstrating improved travel time efficiency and merging success rates compared to traditional methods [24]. The integration of DRL in autonomous driving offers significant insights into how AVs can adaptively handle dynamic traffic scenarios, contributing to safer and more efficient road networks.

Despite considerable advancements in applying deep reinforcement learning (DRL) to autonomous driving, there are still notable gaps in effectively addressing the complexities and real-time decision-making requirements of real-world traffic scenarios. Several scholars have highlighted the need for efficient vehicle coordination and real-time decision-making in dynamic conditions, particularly during freeway on-ramp merging—an area that remains inadequately addressed [1,7–10,15,42]. This study aims to bridge these gaps by proposing a novel DRL-based control model specifically designed for freeway entrance ramp merging scenarios, with the primary objective of enhancing safety, efficiency, and passenger comfort for AV operations in this challenging environment. Unlike previous research that tends to focus on isolated aspects of AV control, this study adopts an integrated approach, combining both longitudinal and lateral control into a unified framework. Leveraging DRL, the proposed model learns the optimal driving strategies required to adapt to the inherent unpredictability of freeway merging areas, and its effectiveness is validated through simulation experiments conducted on the SUMO platform.

This paper aims to make the following contributions:

- Investigation of Autonomous Driving Strategies in Realistic Highway Ramp Merging Scenarios: A highway ramp merging scenario was developed on the SUMO simulation platform, replicating real-world conditions. Two autonomous driving traffic flows—rule-based and deep reinforcement learning—were constructed to ensure that AVs exhibit the behavior of experienced drivers during ramp merging. Theoretically, this provides a valuable benchmark for autonomous vehicle control and traffic simulation research. Practically, it demonstrates the feasibility of deploying AVs in complex traffic environments.
- Development of a Deep Reinforcement Learning-Based Control Model: This research proposes a novel DRL-based control framework that integrates both longitudinal and lateral strategies for AVs in highway ramp merging areas. The model addresses the challenges posed by complex interactions between AVs and HDVs and supports rapid decision-making under dynamic traffic conditions.
- Comprehensive Traffic Performance Analysis: This research provides a holistic evaluation of traffic performance by assessing multiple metrics, such as congestion reduction, safety enhancement, and travel time efficiency. Using the SUMO platform, the study explores the impact of AV deployment in mixed traffic, offering valuable insights into AV integration and its influence on traffic flow under various penetration rates.

The purpose of this paper is to examine the diverse driving strategies employed by autonomous vehicles at highway entrance ramp merging areas. The remainder of this paper is organized as follows. Section 2 outlines the methodology employed. Section 3 describes the configuration of the simulation environment and the design of the simulation experiment. Section 4 presents a comprehensive review and analysis of the results obtained from the simulation experiment. Section 5 offers an extensive and in-depth discussion of the entire research design, simulation process, and results. Section 6 provides a summary of the paper.

2. Methods

The following sections explain the relevant methods used in the study. Vehicle control typically encompasses behaviors, such as following other vehicles and changing lanes.

The IDM(Intelligent Driver Model) microscopic car-following model is employed to simulate the longitudinal behavior of human-driven vehicles [43,44], while "LC2013" as the default lane-changing model of the open-source traffic microsimulator SUMO is utilized to simulate lateral behavior [45].

In the context of autonomous vehicles, two distinct methodologies are employed for simulation and control: namely, deep reinforcement learning (DRL) and SUMO simulation. Accordingly, the vehicle control models are distinct, delineating two categories of autonomous vehicles, as follows:

- SUMO-Simulated Autonomous Vehicle Flow (ZD1): The IDM model is used to simulate the longitudinal behavior, and the LC2013 model is used to simulate the lateral behavior of ZD1 vehicles.
- Deep Reinforcement Learning Controlled Autonomous Vehicle Flow (ZD2): The longitudinal following and lateral lane-changing behaviors of ZD2 vehicles are regulated by the deep reinforcement learning algorithm.

These models enable a comprehensive evaluation of the driving strategies of both human-driven and autonomous vehicles in the complex merging scenarios at highway entrance ramps.

2.1. Vehicle Basic Control Model

2.1.1. Car-Following Model

The car-following model is a mathematical representation of the behavior of vehicles following a lead vehicle in a lane [46]. In vehicle control, this model serves as an algorithmic framework to infer the driving behavior of the trailing vehicle based on the driving state of the lead vehicle within a single-lane's traffic scenario. The IDM is utilized to describe vehicle behavior under both free-flow and congested-flow conditions. This unified expression captures the distinct behaviors observed in each traffic flow state, as shown in Equations (1)–(4).

$$a = a_{free} + a_{break} \tag{1}$$

$$a_{free} = a_{\max}^{(n)} \left[1 - \left(\frac{v_n}{v_{\max}^{(n)}} \right)^{\delta} \right] \tag{2}$$

$$a_{break} = -a_{\max}^{(n)} \left(\frac{s^*}{s_n} \right)^2 \tag{3}$$

In the equation, a is the vehicle acceleration, a_{free} is the acceleration of the vehicle in a free-flowing condition, a_{break} is the deceleration of the vehicle when braking, n represents the vehicle number, $a_{\max}^{(n)}$ represents the maximum acceleration of vehicle n in free-flow conditions, $v_{\max}^{(n)}$ denotes the maximum velocity of vehicle n in free-flowing conditions, δ is the acceleration exponent, v_n represents the current velocity of vehicle n, s_n denotes the distance to the leading vehicle. s^* represents the desired minimum gap, calculated as follows:

$$s^*(v_n, \Delta v_n) = s_{\min}^{(n)} + \max\{T_n v_n + \frac{v_n \Delta v_n}{2\sqrt{a_{\max}^{(n)} b_{\max}^{(n)}}}, 0\} \tag{4}$$

In the equation, $s_{\min}^{(n)}$ represents the minimum safe distance, T_n represents the reaction time of the driver, $b_{\max}^{(n)}$ represents the maximum comfortable deceleration, Δv_n represents the relative velocity between the current vehicle and the leading vehicle. The purpose of $\max\{T_n v_n + \frac{v_n \Delta v_n}{2\sqrt{a_{\max}^{(n)} b_{\max}^{(n)}}}, 0\}$ in Equation (4) is to ensure that the desired minimum gap between vehicles remains adequate under all driving conditions. Specifically, it ensures

that the greater of either the safe minimum gap or the additional gap derived from reaction time and velocity conditions are maintained. This prevents unrealistic or negative gaps, ensuring both the numerical stability and physical feasibility of the model.

These equations collectively describe the vehicle's acceleration and deceleration behaviors, providing a comprehensive framework for simulating longitudinal vehicle dynamics in both free-flow and congested traffic conditions.

2.1.2. Lane-Change Model

In this study, we employed the ramp vehicle lane change model, designated "LC2013", which was proposed by the Institute of Transportation Technology of the German Aerospace Center (DLR) and integrated into the open-source traffic simulation platform SUMO, developed by DLR. In each simulation cycle, the model divides the lane change process into four stages. Initially, it assesses and identifies the optimal target lane. Subsequently, it calculates the safe velocity for maintaining the original lane based on the velocity target in the previous cycle. Thirdly, it determines the necessary conditions for completing the lane change. Finally, it decides whether the vehicle should execute the lane change operation or plan for the lane change demand in the next cycle if the current state does not meet the lane change conditions. The extent to which the vehicle modifies its velocity is contingent upon the urgency of the lane change.

The LC2013 model introduces a novel categorization of lane change motivations, divided into four distinct categories [45], as follows:

- Strategic Lane Change: Undertaken when it is not possible to reach the subsequent section of the route by continuing in the original lane.
- Cooperative Lane Change: A lateral movement in the same direction as the prevailing traffic flow, facilitating the smooth entry of a vehicle from an adjacent lane.
- Tactical Lane Change: Employed by a driver to increase velocity and overall efficiency when the driving state of the preceding vehicle restricts the driver's ability to achieve the desired velocity. The objective is to expedite and optimize traffic flow by changing lanes.
- Obligatory Lane Change: Previously defined as "mandatory lane changing". The driver must adhere to established traffic regulations and promptly resume their original lane, typically the left lane, to guarantee optimal utilization of the overtaking lane (default position for the left lane).

In this study, the aforementioned two models are employed to generate the human-driven vehicle, designated as RL, and the autonomous vehicle, identified as ZD1. The two vehicles are distinguished solely by the parameters indicated in Table 1.

Table 1. Parameter settings for RL and ZD1 in the simulation.

Attribute Category	Specific Parameters	RL	ZD1	Definition
Vehicle type	Vehicle-type	Human-driven vehicles	Autonomous vehicles under the SUMO strategy	Different vehicle types
Vehicle physical attributes	Length (m)	5.0	5.0	Vehicle clear length
	Width (m)	1.8	1.8	Vehicle clear width
	Height (m)	1.5	1.5	Vehicle clear height
	Color	(1,0,0)	(1,1,0)	Red (RL) and Yellow (ZD1)
Car-following	Accel (m/s^2)	3.0	3.0	Maximum acceleration
	Dccel (m/s^2)	7.5	7.5	Maximum deceleration capacity
	maxVelocity (km/h)	100	100	Maximum velocity
	tau	1.5	0.5	Driver reaction time
	sigma	0.5	0	Driving proficiency level
	minGap (m)	2.5	1.0	Minimum headway distance
Lane-change	lcStrategic	0.5	0.5	Readiness to implement strategic change
	lcCooperative	0.5	0.5	Readiness to implement cooperative change
	lcVelocityGain	0.1	1	Willingness to change lanes for higher velocity

2.2. Deep Reinforcement Learning

Deep reinforcement learning (DRL) represents the convergence of reinforcement learning and deep learning, enabling autonomous vehicles to make behavioral decisions and process high-risk data effectively. The objective of this study is to utilize the DRL to develop efficient and safe driving strategies for autonomous vehicles in mixed traffic scenarios at highway ramp merging areas. In DRL, an agent encounters tasks within real-world scenarios and abstracts these tasks into models that can be solved. The essence of DRL is a typical Markov decision process (MDP). Initially, the agent executes the required actions based on the environmental conditions and its current state, thereby obtaining immediate reward values and modifying the environment. The agent then selects subsequent actions based on the prevailing state and strategy, repeating this process iteratively. Ultimately, the agent identifies a strategy that maximizes the cumulative reward value, representing the optimal approach for task completion (see Figure 1).

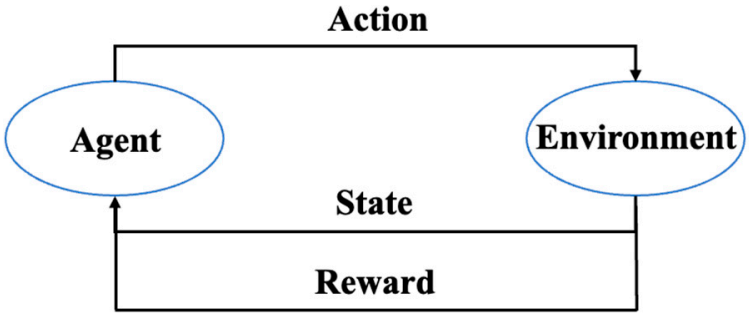


Figure 1. The Fundamental Process of Reinforcement Learning.

The Markov decision process is defined by a six-tuple (S, A, P, R, γ, T), where S represents the state value space, A represents the action value space, which contains all possible behaviors of the agent, P represents the state transition probability matrix, which is the probability of selecting an action from a state to the next state, R represents the value of the reward obtained, γ is a discount factor between 0 and 1, which indicates the importance of future returns compared to current returns, and T is the time span. The objective is to execute a sequence of actions commencing at time t in order to maximize the expected reward [47].

$$R_t = E_\pi(\sum_{k=0}^{T} \gamma^t r_{t+k}) \tag{5}$$

where r_{t+k} is the reward value at a time $t + k$, and π is the policy. The optimal policy of an MDP problem is expressed as:

$$\pi^* = \arg\max_\pi R_t \tag{6}$$

Deep neural networks, employing multiple hidden layers, generate a strategy by extracting features from a set of inputs. In continuous control problems, the policy gradient method, utilizing the gradient descent algorithm, enables the modification of the random policy parameters associated with the action. The Proximal Policy Optimization (PPO) algorithm is employed for the continuous control of autonomous vehicles in this study.

2.3. Problem Description

This section employs the Traci interface between SUMO and Python to control the second-generation autonomous vehicle "ZD2" using the Proximal Policy Optimization (PPO) algorithm. The autonomous vehicle inputs characteristic information about the road and surrounding vehicles observed from the road network into the training model. Through iterative training and reward learning, the model can eventually output driving actions that are executable and meet the required constraints. The specific process is illustrated in Figure 2. The Traci interface facilitates real-time interaction between the SUMO simulation

environment and the Python-based control algorithm. By continuously gathering data on the vehicle's surroundings and its own state, the ZD2 vehicle updates its strategy to optimize performance under varying traffic conditions, including the following:

- Data Collection: The autonomous vehicle gathers real-time data, including its own speed, position, acceleration, and the relative positions and velocities of surrounding vehicles.
- State Representation: These data are processed to form a comprehensive state representation, which serves as the input for the DRL model.
- Action Selection: Based on the current state and the policy learned through PPO, the vehicle selects the optimal action (e.g., accelerating, decelerating, lane changing).
- Reward Calculation: The selected action is executed, and the resultant state change is evaluated to calculate a reward. The reward function considers factors such as safety, efficiency, and comfort.
- Policy Update: Using the calculated reward, the PPO algorithm updates the policy to improve future decision-making.

The diagram in Figure 2 illustrates the full cycle of model implementation, beginning with the interaction between the SUMO simulation environment and the Python-based control algorithm using the Traci interface. It shows how the data are collected from the vehicle environment, processed into state representations, and subsequently used to determine the next action via the DRL framework. These actions are executed in real-time, and the resulting impact on the environment is assessed using a reward mechanism. The policy is iteratively updated to enhance decision-making for future scenarios. This cycle allows the ZD2 vehicle to continuously learn and refine its driving strategy, ultimately achieving optimal performance under complex traffic conditions.

This iterative process enables the ZD2 vehicle to learn and refine its driving strategy, achieving optimal performance in complex ramp merging scenarios.

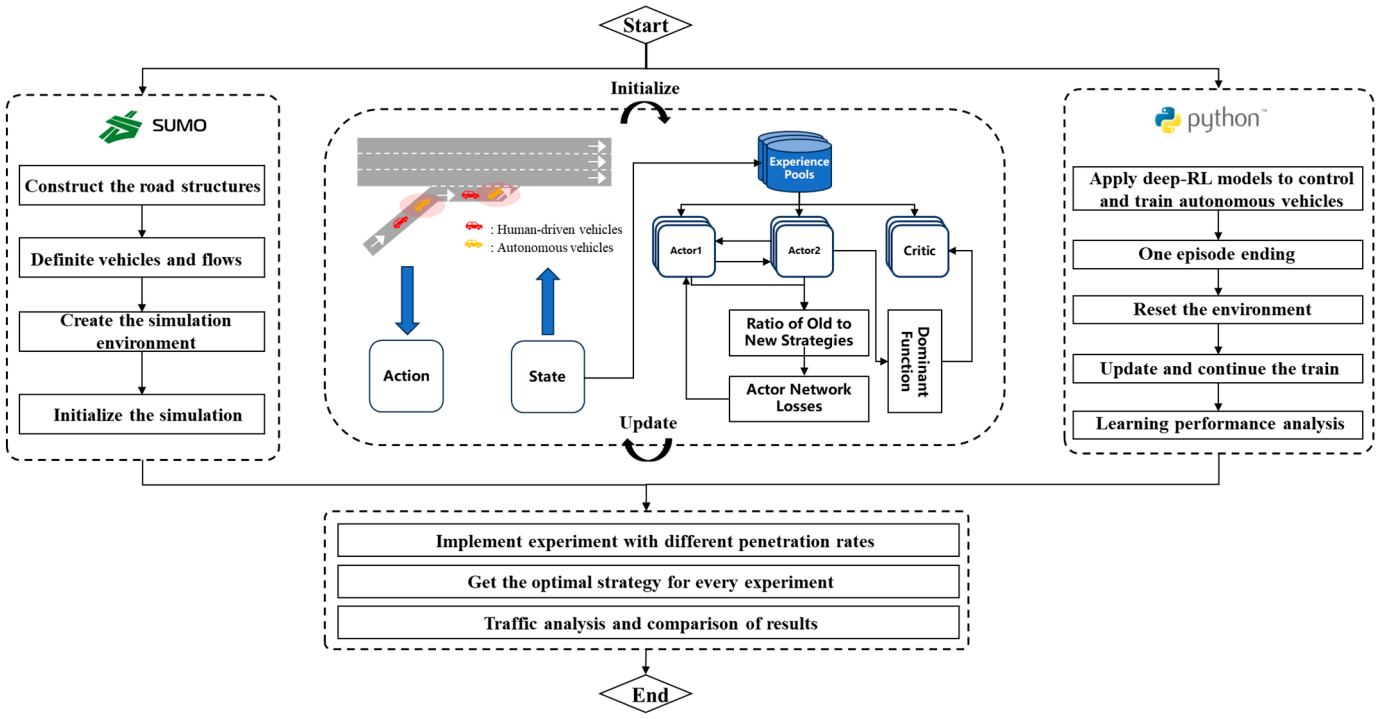


Figure 2. Diagram for model implementation.

2.3.1. State Space

This study provides a comprehensive examination of the ramp merging scenario and the design of the associated reward and punishment mechanism. To ensure the security

of vehicle operation and the requisite velocity and position adjustments for the vehicle on the ramp to complete the lane change, the control vehicle is designed to obtain additional relevant status information about the preceding and following vehicles in the current lane and the preceding and following vehicles in the adjacent lanes, in addition to its own driving status.

As illustrated in Figure 3, a 21-dimensional state observation feature set has been devised. This feature set encompasses the position, velocity, and acceleration of the controlled vehicle itself, as well as the position, velocity, and acceleration of the six surrounding vehicles. The controlled vehicle's position provides critical information for determining lane-changing strategies, while its velocity and acceleration are key factors in timing these maneuvers and ensuring smooth driving. The surrounding vehicles' relative distance, velocity, and acceleration help the controlled vehicle assess collision risk, choose safe gaps for merging, and adapt to changes in the behavior of other vehicles. These features collectively ensure that the controlled vehicle can make informed decisions to enhance safety and efficiency during ramp merging. During the training phase, each controlled vehicle acquires its own observation state and inputs it into the training network concurrently.

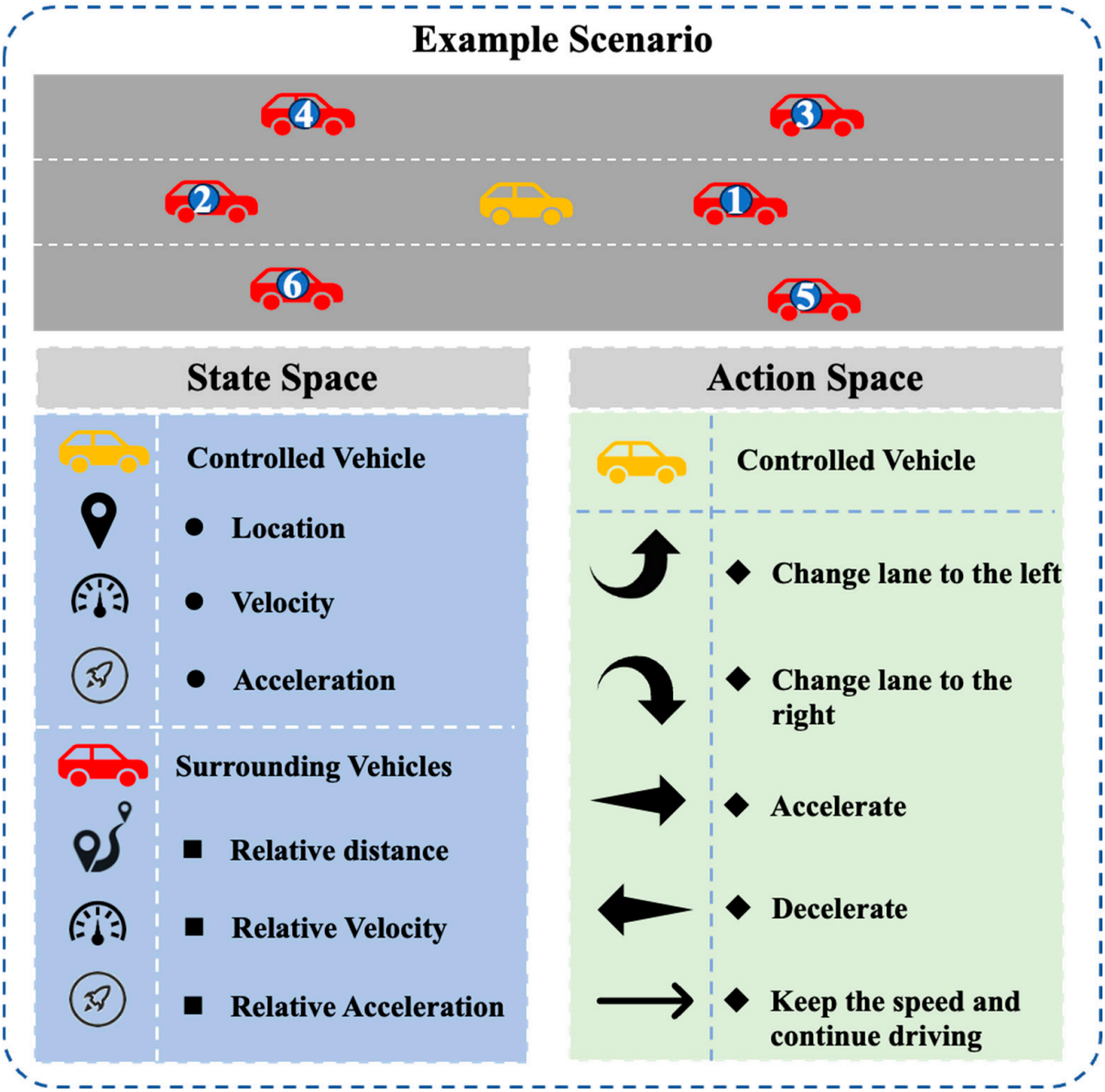


Figure 3. Schematic diagram of state space and action space.

2.3.2. Action Space

The deep reinforcement learning algorithm determines the actions the agent can perform based on the received state input and the set reward mechanism. Once the deep network has undergone sufficient iterative learning, the agent can execute these actions. This study proposes a set of five actions for the autonomous driving vehicle, as illustrated in Figure 3, to be used throughout the entire highway ramp merging area. The five actions are: (1) accelerate, (2) decelerate, (3) change lane to the left, (4) change lane to the right, and (5) maintain the current velocity and continue driving in the current lane. These actions enable the vehicle to effectively navigate the ramp merging area, ensuring a balance between safety, efficiency, and passenger comfort.

2.3.3. Reward

In the context of a highway ramp merging scenario, it is anticipated that an autonomous vehicle will attain a high velocity safely and proceed to complete a lane change in a timely and appropriate manner, maintaining an adequate distance between vehicles and merging into the flow of traffic on the main road at a velocity commensurate with prevailing traffic conditions. In light of these objectives, this study proposes the following reward function:

- Safety reward

In the field of transportation, safety is of paramount importance. Safety assessment and risk prevention at source are very common in the field of traffic safety [48,49]. Consequently, the development of autonomous driving control strategies must prioritize safety and establish measures to address potential risks. In road traffic, particularly at high velocities, vehicle collisions represent a significant risk factor for severe accidents. To circumvent such incidents, this paper proposes the implementation of a safety incentive based on the "Time to Collision" (TTC) metric [50,51], as illustrated in Equations (7) and (8). The proposed control strategy explicitly incorporates a collision risk factor within the reward function, using the TTC metric as an essential part of the deep reinforcement learning model. This inclusion ensures that safety considerations are integral to the control strategy, particularly during ramp merging scenarios, where the risk of collision is high.

$$TTC(t) = -\frac{\Delta X_{n-1,n}(t)}{\Delta V_{n-1,n}(t)} \tag{7}$$

$$R_{safe\ (TTC)} = \begin{cases} \log \frac{TTC}{4}, & 0 \le TTC \le 4 \\ 0, & Else \end{cases} \tag{8}$$

In the equation, $TTC(t)$ represents the estimated time to collision between vehicle $n-1$ and vehicle n, $\Delta X_{n-1,n}(t)$ denotes the distance between vehicle $n-1$ and vehicle n at a given time t, $\Delta V_{n-1,n}(t)$ represents the velocity disparity between vehicle $n-1$ and vehicle n at a given time t.

- Efficiency reward

In the scenario of merging onto a highway via an acceleration lane, vehicles on the main road typically maintain a higher velocity to meet the posted speed limit, while vehicles on the acceleration lane travel at a lower velocity before entering the parallel road section [52,53]. Upon reaching the acceleration lane, vehicles accelerate to match the speed of the main road and seek a suitable distance to complete the lane change. It is necessary to implement a system that rewards the velocity change of autonomous vehicles in the merging area, guiding them to accelerate to the main road speed as quickly as possible under safe conditions. This will increase the probability of successful lane changes, speed up vehicle traffic, and simultaneously reduce the interference of merging vehicles on the main line. This can be achieved through the application of a penalty system, as shown in Equation (9). Concurrently, to circumvent the accumulation of autonomous vehicles at a low velocity during the action exploration phase, which would impair the efficiency of the

simulation and training, a penalty is imposed for a lower driving velocity, as illustrated in Equation (10).

$$R_{v1} = 1 - \left| \frac{v_t - v_{op}}{v_{op}} \right| \tag{9}$$

$$R_{v2} = \begin{cases} -1, & v_t \leq 5\ \mathrm{m/s} \\ 0, & Else \end{cases} \tag{10}$$

In the equation, v_t is the velocity of the controlled vehicle at the present moment in time, v_{op} is the optimal velocity of the controlled vehicle on the current road network.

- Comfort reward

In addition to ensuring the safety and efficiency of the driving process, autonomous driving must also provide the driver and passengers with an optimal driving experience. The most intuitive sensation for humans during a journey is sudden acceleration or deceleration, which necessitates keeping the vehicle's acceleration value to a minimum. The objective of this study is to ensure the comfort of the vehicle during the driving process by constraining the acceleration value. Given that both rapid acceleration and rapid deceleration result in a suboptimal driving experience, and that the corresponding accelerations are positive and negative values, the design of the reward function employs the square term of the relative value of the maximum acceleration to penalize both situations. In particular, the greater the discrepancy between the square term of the relative value and zero, the greater the associated penalty value [54,55]. That is to say,

$$R_c = -\left(\frac{a_t}{a_{\max}} \right)^2 \tag{11}$$

In the equation, a_t is the acceleration of the controlled vehicle at the present moment in time, $a_{\max}$ is the maximum acceleration of the controlled vehicle.

- Low-disturbance reward

The act of changing lanes by an autonomous vehicle, whether from a ramp to the main road or from one lane on the main road to another, will inevitably result in some degree of interference with the flow of traffic on the main road. This is particularly true in situations where other vehicles are present on the main road. Accordingly, the original objective of the reward function is to discourage frequent lane changes while encouraging vehicles to successfully change lanes from the parallel acceleration lane to the main road. Furthermore, penalties are set for other lane changes based on the spacing between the preceding and following vehicles.

$$R_{lc1} = \begin{cases} 1, & Change\ lane\ successfully \\ -1, & Failed\ to\ change\ lane \\ 0, & Else \end{cases} \tag{12}$$

$$R_{lc2} = -\left| \frac{h_{lead} - h_{follow}}{h_{lead} + h_{follow}} \right| \tag{13}$$

In the equation, h_{lead} is the distance between the autonomous vehicle and the leading vehicle in the newly designated lane, h_{follow} is the distance between the autonomous vehicle and the following vehicle in the newly designated lane.

- Total reward

During the experiment, the initial weight is set to 1, and the test weight parameters are adjusted according to the experimental performance until the weight parameters that maximize the experimental effect are obtained.

$$R_{Total} = w_1 R_{safe\ (TTC)} + w_2 R_{v1} + w_3 R_{v2} + w_4 R_c + w_5 R_{lc1} + w_6 R_{lc2} \tag{14}$$

In the equation, w_i represents the weight parameter of every reward. After extensive training and testing, the final weight coefficients were determined as follows: $w_1 = 2$, $w_2 = 1$, $w_3 = 1.5$, $w_4 = 1.5$, $w_5 = 1$, $w_6 = 1$. These weight coefficients were selected to achieve an optimal balance between safety, velocity, lane-change efficiency, and comfort, thereby maximizing overall experimental performance.

3. Modeling and Experiment

This section provides a detailed account of the scenarios and experimental details incorporated into the SUMO simulation software.

3.1. Configuration of Road Structures

The principal forms of ramp merging areas are parallel and straight configurations. In the parallel configuration, the ramp is connected to the main road by a section of the same width as the ramp. This design provides an optimal view, allowing the vehicle exiting the ramp to fully observe the traffic conditions on the main road, which is conducive to merging and suitable for sections with high traffic flow rates. In the straight configuration, the ramp is directly connected to the main road at an acute angle. Given that the independent acceleration section of the parallel layout is conducive to observing the simulation state, this study primarily focuses on parallel ramp merging.

As illustrated in Figure 4, our investigation entailed the analysis and conceptualization of a one-way, three-lane highway with a total length of 850 m. The merging area is equipped with parallel ramps, with a total length of 384 m.

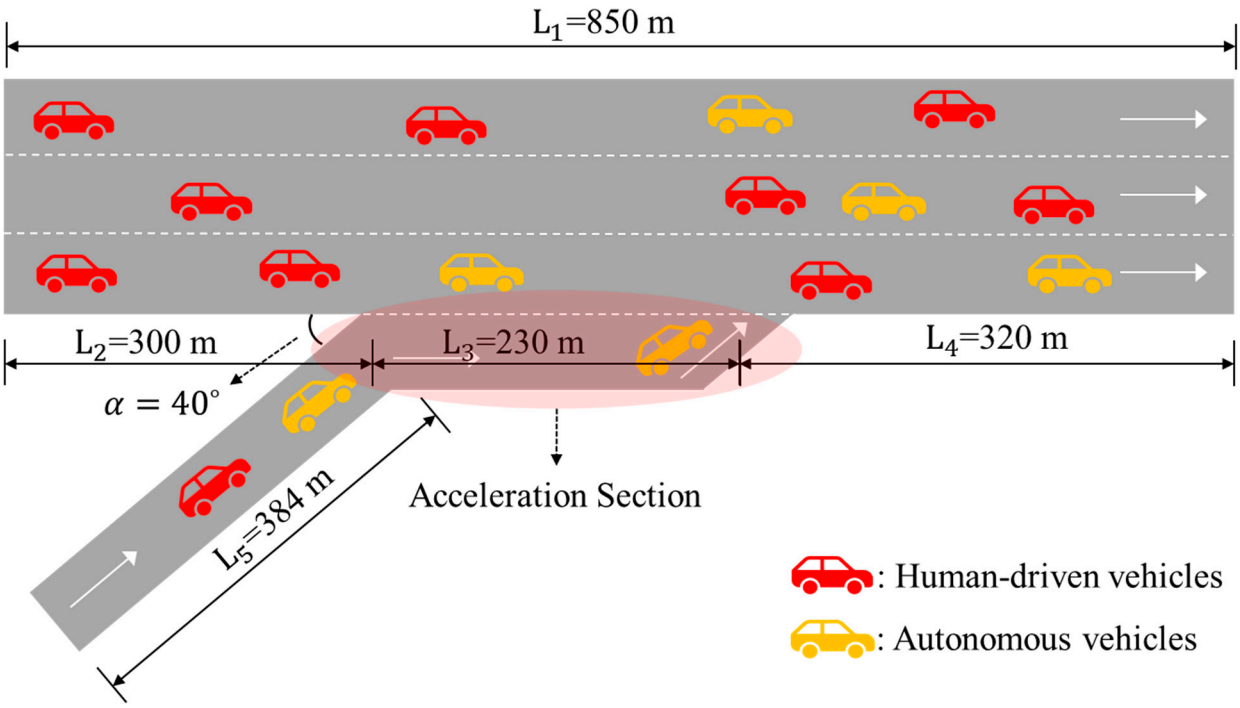


Figure 4. Visualization of road structures.

3.2. Configuration of Vehicles

Once the road network has been constructed, the subsequent phase is to determine the types of vehicles that will utilize it and the manner in which they will traverse it. It is essential that the vehicle types that are incorporated into the traffic flow are clearly defined and recognized, and that they are stored by the system. It thus follows that the definition of vehicle types is a prerequisite for the definition of traffic flows.

The experiment defines three categories of vehicles: manually driven vehicles, autonomous vehicles utilizing the SUMO strategy, and autonomous vehicles utilizing the

DRL strategy. It is not feasible to include two distinct types of autonomous vehicles in a single experiment. The parameters of each vehicle type are presented in Table 2.

Table 2. Model Parameter for Different Vehicle Types.

Number	ID	Length (m)	Color	Car-Following Model	Lane-Change Model	Parameters of Car-Following Model	Parameters of Lane-Change Model
1	RL	5	Red	IDM	LC2013	Modify the IDM model parameters	Default
2	ZD1	5	Yellow	IDM	LC2013	Modify the IDM model parameters	Default
3	ZD2	5	Yellow	DRL	DRL	DRL	DRL

3.3. Details of the Experiments

In this experiment, the road network is defined, and traffic flow is generated by writing and modifying XML files. For vehicle control, considering compatibility with SUMO, both the human-driven vehicle and the first autonomous vehicle, designated as "ZD1," are implemented by adjusting parameters within the XML file. For the second autonomous vehicle, "ZD2", which is controlled using deep reinforcement learning (DRL), all codes are written in Python using the PyCharm IDE. The DRL control algorithm for "ZD2" is implemented using the stable-baseline3 framework, while the action and state spaces are established with the help of NumPy and other relevant function libraries. The programming language, framework, and main libraries utilized in the experiment are summarized in Table 3.

Table 3. Experimental Program and Corresponding Function Library Versions.

Name	Version
Python	3.7.4
SUMO	1.19.0
PyCharm	2021.3
Traci	1.19.0
numpy	1.21.6
gymnasium	0.28.1
stable-baseline3	2.0.0

This study employs a multi-faceted experimental approach to elucidate the operational characteristics of traffic flow in mixed traffic scenarios. The investigation entails the implementation of diverse driving strategies and penetration rates to ascertain their impact on traffic dynamics. The fundamental experiment group considers the traffic situation in the absence of autonomous vehicles. Consequently, both the vehicles departing from the main road (flow rate 3000 veh/h) and the vehicles departing from the ramp (flow rate 600 veh/h) are human-driven vehicles. The control group is composed of two distinct groups. The vehicles departing from the main road are all human-driven vehicles, with a flow rate of 3000 veh/h. In contrast, the vehicles departing from the ramp comprise a mixed traffic flow of human-driven vehicles and autonomous vehicles, with a total flow rate of 600 veh/h. The first control group comprises a mixed traffic flow of RL and ZD1, while the second control group is a mixed traffic flow of RL and ZD2. Given that the proportion of autonomous vehicles in a realistic mixed traffic flow scenario remains relatively low, each control group considers a penetration rate of 10% and 20%. The flow rates were chosen to represent a moderate-to-high traffic volume, which allows for the evaluation of traffic performance under more challenging conditions, while still being within the

operational capacity of a 3-lane mainline and single-lane ramp. The specific arrangement of the experiment is illustrated in Figure 5.

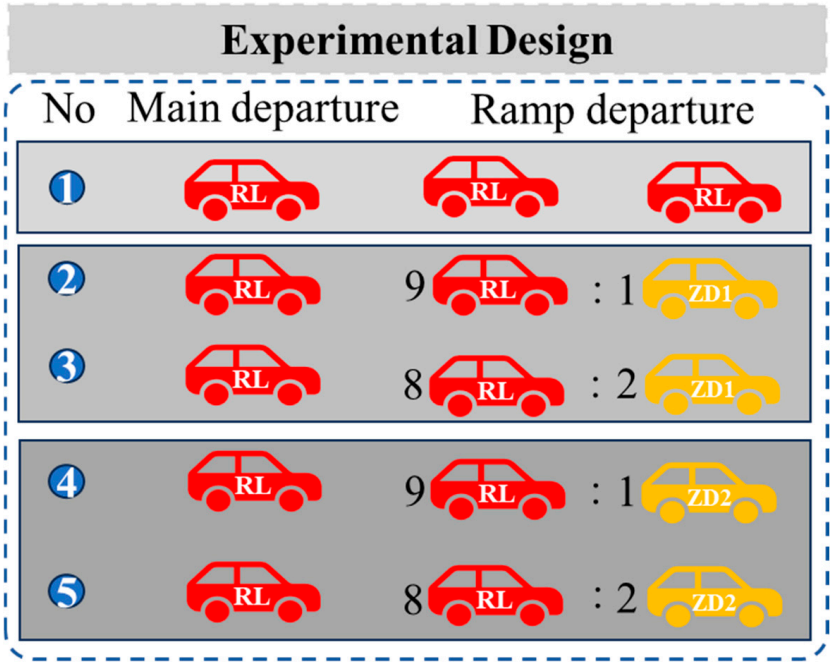


Figure 5. Description of the specific experimental conditions.

The initial three groups of experiments were conducted in SUMO software for a total of 1200 s, with a simulation time step of 0.2 s for each experiment. To assess the influence of randomness inherent to the simulation process on the resulting data, each experiment was repeated ten times, with the random number generator seeded with a different value in each instance. The final two groups of experiments were conducted in SUMO for 1200 s per training round, with a learning rate of 1×10^{-4}. The discount rate was 0.999, and the total number of training episodes was 800. The data utilized for the analysis of the results is the mean of each group of experiments, thereby eliminating the influence of randomness and enhancing the reliability of the results. To select the optimal hyperparameters for the deep reinforcement learning model, a combination of manual tuning and iterative experimentation was employed. The learning rate was adjusted gradually, starting from a larger value of 1×10^{-3} and then decreasing to ensure stable convergence. The batch size was set to 32 to balance training efficiency and stability. The policy network architecture was designed with three layers containing 128, 64, and 32 units, respectively, which allowed for sufficient model complexity while mitigating the risk of overfitting. These parameters were fine-tuned based on the performance during training, where metrics such as total reward and episode length were monitored to determine the final settings. After extensive testing, the hyperparameters that yielded the best overall performance were selected for the final experiments.

4. Results

4.1. Learning Performance

To evaluate the performance of the autonomous vehicle controlled by deep reinforcement learning, ramp departure scenarios with penetration rates of 10% and 20% were set. The relevant reward function, as mentioned above, was used to train the vehicle, corresponding to Experiments 4 and 5, respectively. Figure 6 illustrates the learning outcomes of the two sets of experiments. The learning curves for both experiments demonstrate that the reward value increased rapidly during the initial 500 training iterations and then exhibited a convergent trend. This indicates that the reward function was set with a reasonable

degree of precision and could converge rapidly. These results also demonstrate that an autonomous vehicle controlled by a deep reinforcement learning algorithm is capable of maintaining an optimal driving strategy in the merging area of a highway ramp.

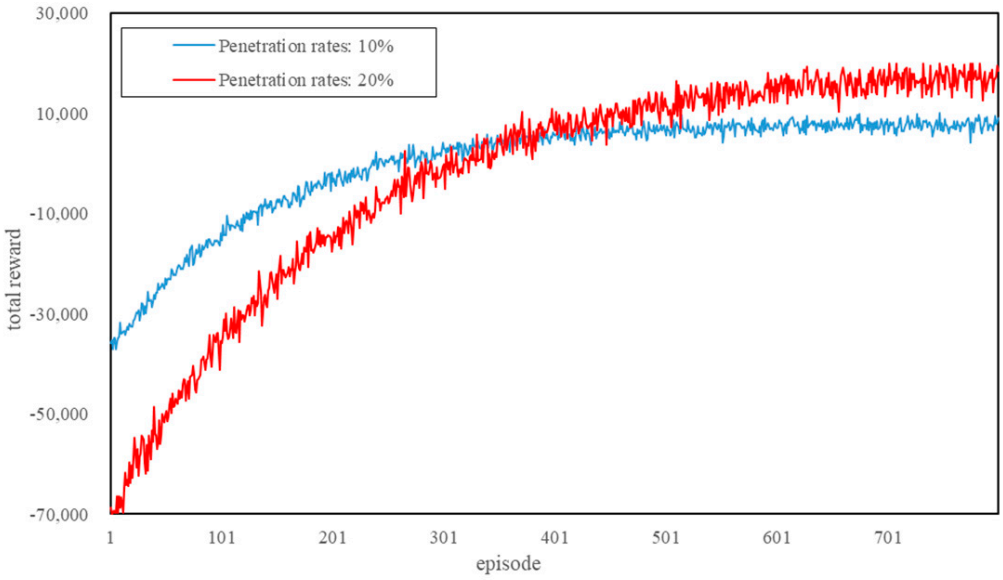


Figure 6. Learning performance in experiments 4 and 5 (10% and 20% penetration rates of autonomous vehicles).

4.2. Traffic Analysis

4.2.1. Velocity Aspect

The velocity of vehicles is an important indicator of traffic efficiency. Accordingly, the data obtained from the simulation experiment were utilized to extract the instantaneous velocity of the vehicles at each time step at the beginning of each ramp. Subsequently, the velocity was plotted as a function of position (x-coordinate), as illustrated in Figure 7. Experiment 1 illustrates the impact of the absence of autonomous vehicles on the velocity of RL vehicles (Figure 7b). It can be observed that the velocity of RL vehicles at the ramp (approximately 300 m) declines significantly, and then, after entering the main road (approximately 400 m), the velocity recovers and remains stable. This deceleration indicates that human-driven vehicles will initially reduce their velocity before merging into the main road. In Experiment 2 (Figure 7c) and Experiment 3 (Figure 7d), the impact of varying penetration rates of autonomous vehicles on the velocity of RL and ZD1 vehicles is illustrated. The results of the two experiments are comparable. Both RL and ZD1 vehicles exhibit a notable decline in velocity at the entrance to the ramp. However, ZD1 vehicles demonstrate a more rapid and consistent recovery in velocity after entering the main road, whereas RL vehicles exhibit a slower and more erratic recovery. As the penetration rate of autonomous vehicles increases (Experiment 3), the overall traffic flow recovers more smoothly. However, fluctuations in the velocity of RL vehicles still occur. The enhanced stability of autonomous vehicles may be attributed to their superior capacity to navigate and adapt to varying traffic conditions. In Experiments 4 (Figure 7e) and 5 (Figure 7f), the impact of varying penetration rates of autonomous vehicles on the velocity of RL and ZD2 vehicles is illustrated. The findings indicate that both RL and ZD2 vehicles experience a decline in velocity at the ramp. However, ZD2 vehicles demonstrate a more pronounced and less pronounced recovery, respectively, after entering the main road, whereas RL vehicles exhibit a slower and more pronounced recovery. As the penetration rate of ZD2 vehicles increases (Experiment 5), the velocity recovery of the overall traffic flow becomes more stable. The deep reinforcement learning-controlled autonomous vehicle demonstrated superior stability in terms of velocity recovery, indicating that the deep reinforcement learning control algorithm enables the vehicle to respond to traffic conditions with greater

intelligence, optimize acceleration and deceleration processes, reduce velocity fluctuations, and enhance driving efficiency and safety. Regarding the chattering phenomenon observed in Figure 7, particularly in Experiments 2 and 3, the high instability can be attributed to the higher fluctuation levels of RL and ZD1 vehicles. This instability is likely caused by frequent lane changes and inconsistent acceleration patterns during merging, which lead to speed variations and reduced stability. The ZD1 vehicles exhibited faster lane changes but were prone to abrupt speed changes, which further contributed to this instability. In contrast, the deep reinforcement learning approach used for ZD2 vehicles allowed for more stable lane-changing behavior and smoother acceleration, thereby reducing chattering and enhancing stability in Experiments 4 and 5.

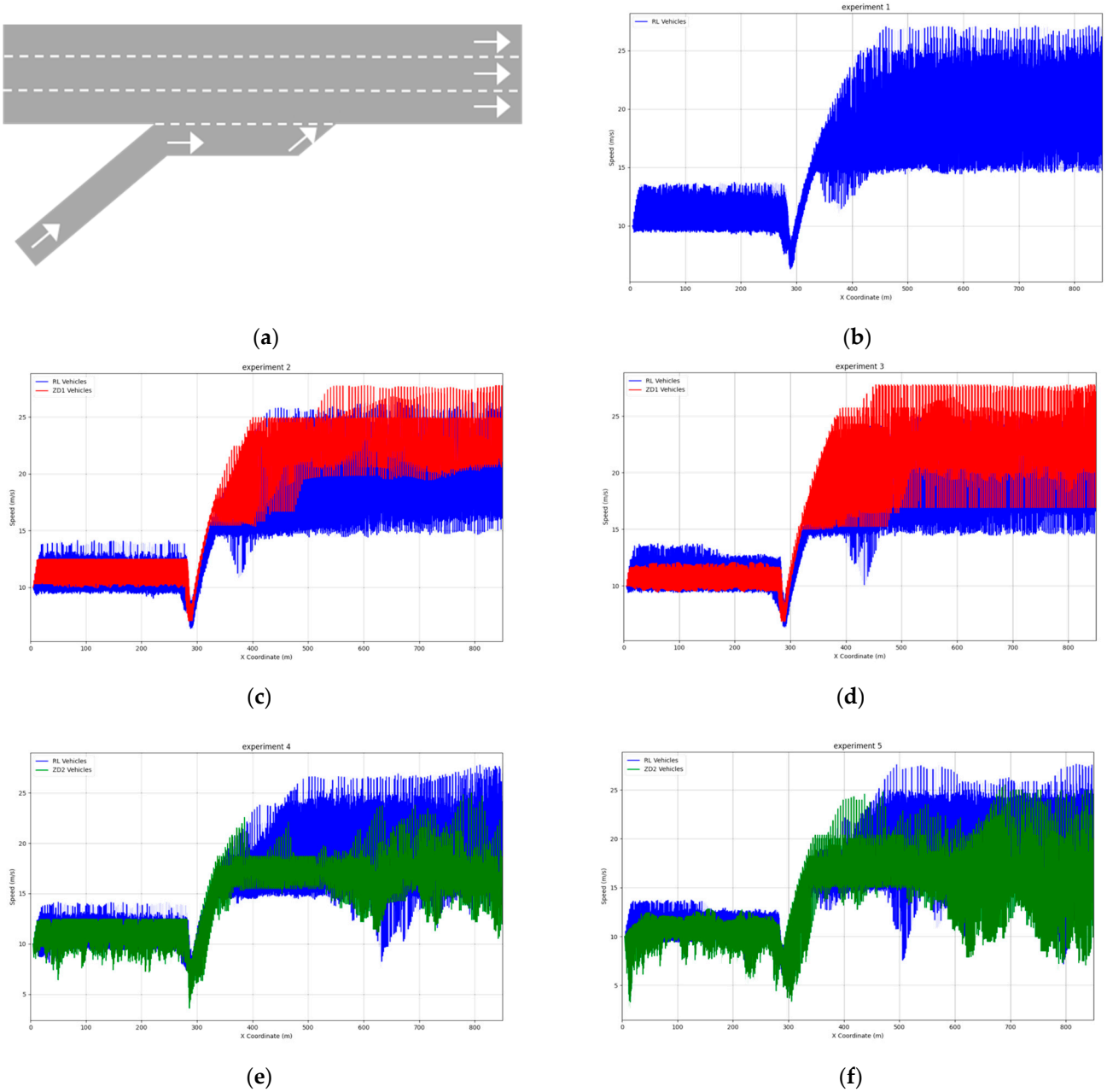


Figure 7. This is the velocity versus time curves for different types of vehicles in each experimental scenario. (**a**) traffic scenario; (**b**) Experiment 1 correspondence; (**c**) Experiment 2 correspondence; (**d**) Experiment 3 correspondence; (**e**) Experiment 4 correspondence; (**f**) Experiment 5 correspondence.

4.2.2. Acceleration Aspect

The magnitude of the acceleration value has a direct impact on the level of comfort experienced by drivers and passengers. A reduction in the magnitude of the acceleration fluctuations results in a more comfortable driving experience for the driver and passengers. Consequently, based on the simulation experimental data, the acceleration values of each vehicle type at each time step from the ramp are extracted and averaged, as illustrated in Figure 8. As illustrated in the figure, the average acceleration of the RL vehicle in Experiment 1 was 0.174 m/s^2, which represents the performance of a human-driven vehicle under baseline conditions. In Experiments 2 and 3, the RL and ZD1 vehicles exhibited acceleration values of 0.177 m/s^2 and 0.167 m/s^2, respectively. This indicates that the human-driven vehicles demonstrated greater consistency in their performance in mixed traffic conditions. In contrast, the ZD1 vehicles exhibited acceleration values of 0.226 m/s^2 and 0.212 m/s^2, respectively. These higher acceleration values suggest that the ZD1 vehicles exhibited a more pronounced change in speed.

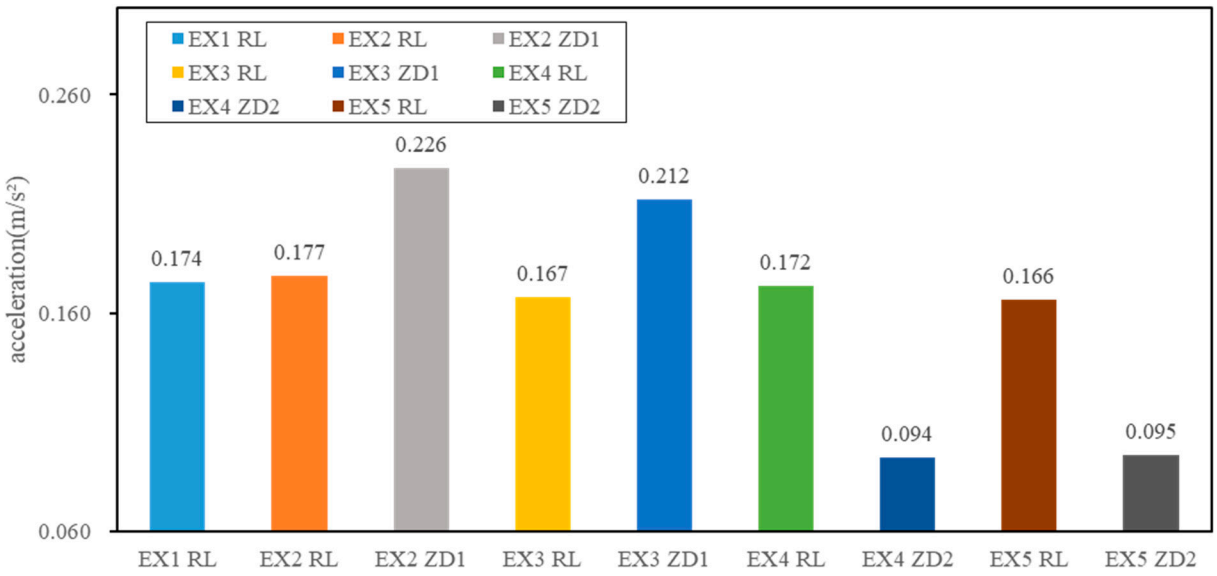


Figure 8. Average acceleration of various types of vehicles.

In contrast, the mean acceleration values of the autonomous vehicles in Experiments 4 and 5 demonstrate stability. In Experiments 4 and 5, the acceleration values of the RL vehicle were 0.172 m/s^2 and 0.166 m/s^2, respectively, exhibiting fluctuations analogous to those observed in the preceding experiments. In contrast, the ZD2 vehicles exhibited markedly lower acceleration values of 0.094 m/s^2 and 0.095 m/s^2, respectively. This suggests that autonomous vehicles controlled by deep reinforcement learning demonstrate enhanced stability during speed changes. The markedly reduced acceleration values indicate that the ZD2 vehicles are more proficient in optimizing acceleration and deceleration during driving, thereby reducing the incidence of abrupt acceleration and deceleration. This enhances driving comfort and stability.

4.2.3. Lane-Change Aspect

Lane-changing behavior is a key research focus for autonomous vehicles in complex scenarios, especially for the merging area at the entrance to the highway ramp in this scenario, where the timing of lane-changing needs to be accurately determined. This is an exciting area of research. Figure 9 shows the impressive relative changes in lane-changing speed, lead time, and follow time of different types of vehicles in the merging area at the entrance to the highway ramp relative to the baseline human-driven vehicle (EX1 RL). The baseline data are change_speed 15.42 m/s, leaderGap 67.06 m, and followGap 136.30 m. The experimental data show some fascinating insights. It seems that the ZD1 vehicles

in EX2 and EX3 exhibited higher lane-changing speeds, reaching 1.10 and 1.08 times the baseline, respectively. This suggests that the ZD1 vehicles underwent a significant change in speed during lane-changing, which could potentially lead to some interesting speed fluctuations and instability. The deep reinforcement learning autonomous vehicle (ZD2) in EX4 and EX5 has an impressive lane-change speed of 0.95 and 0.92, respectively. While it is slightly lower than the baseline, it is much more stable, which is great news for driving stability and comfort. In terms of front-vehicle distance, the ZD2 vehicles in EX4 and EX5 are doing an amazing job. They are able to maintain a safe and optimal distance from the preceding vehicle during the merging process, optimize the driving path, and improve traffic efficiency. In contrast, the relative values of the front distance of the ZD1 vehicles in EX2 and EX3 were 1.07 and 1.05, respectively. This indicates that the front distance was larger during lane changes, which presents an exciting opportunity to optimize traffic efficiency. In terms of the rear distance, the relative values of the ZD2 vehicles in EX4 and EX5 were 0.97 and 0.74, respectively, which were significantly lower than the baseline. This is great news. It means that the ZD2 vehicles were able to more accurately adjust the distance from the rear vehicle when merging, reducing traffic congestion and improving the stability of the traffic flow. The relative values of the rear distance in EX2 and EX3 for ZD1 vehicles were 0.98 and 1.03, respectively. While these values did not change much compared to the baseline, they did show some fluctuation during lane changes, which is an exciting sign of progress.

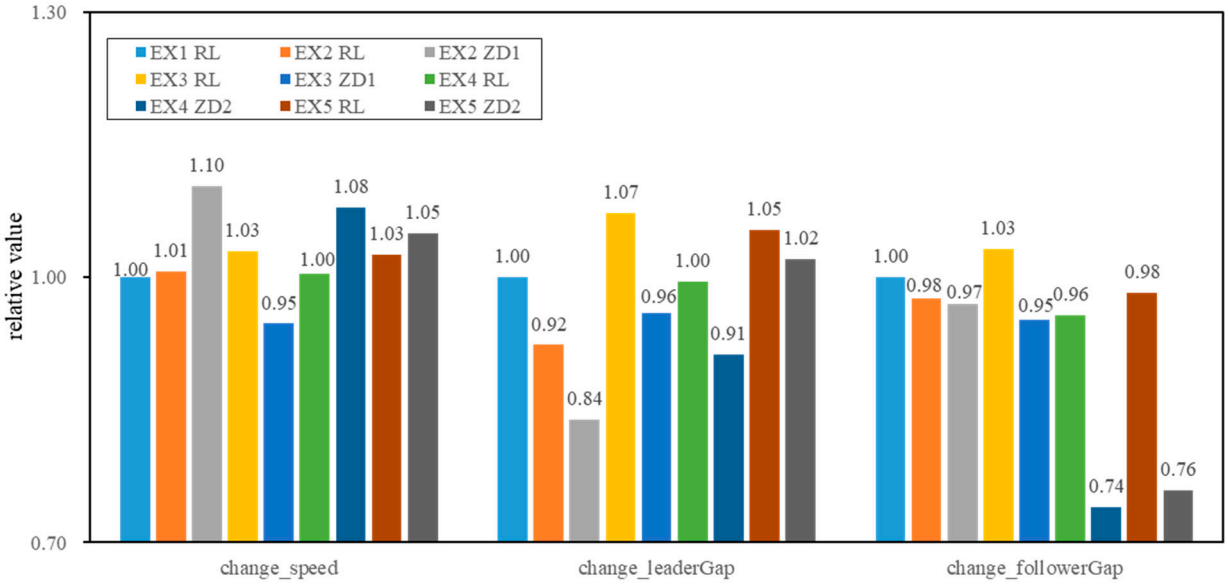


Figure 9. Relative data for lane changes for various types of vehicles.

4.2.4. Comparison

In this section, we summarize the comparative analysis of three vehicle types—human-driven (RL), rule-based autonomous vehicles (ZD1, using IDM), and deep reinforcement learning-controlled autonomous vehicles (ZD2, using DRL). The comparison focuses on velocity recovery, acceleration consistency, lane-changing efficiency, and safety aspects in terms of leader and follower gaps. Table 4 highlights the advantages and disadvantages of each approach.

Table 4. Comparative Analysis of Ramp Departing Vehicles.

Experiment No.	Vehicle Type	Model	Velocity Recovery	Acceleration Consistency	Lane Change and Gap Behavior
1	RL	Default	Moderate, unstable	Moderate fluctuations	Erratic lane changes, inconsistent gaps
2	RL	Default	Slower, erratic	High fluctuation levels	Delayed lane changes, inconsistent gaps
	ZD1	IDM	Rapid, consistent	High acceleration, prone to abrupt changes	Fast lane changes, moderate gaps
3	RL	Default	Slower, erratic	Moderate fluctuations	Delayed lane changes, inconsistent gaps
	ZD1	IDM	Rapid, consistent	High but slightly reduced fluctuations	Improved lane-changing, less optimal gaps
4	RL	Default	Slower, erratic	Moderate fluctuations	Erratic changes, frequent corrections
	ZD2	DRL	Pronounced, stable	Low and stable acceleration	Smooth changes, optimal gaps
5	RL	Default	Slower, erratic	Moderate fluctuations	Erratic changes, slow response
	ZD2	DRL	Pronounced, stable	Low and stable acceleration	Consistent changes, well-managed gaps

5. Discussions

This study examines the efficacy of deep reinforcement learning (DRL) in enhancing the driving strategy of autonomous vehicles (AVs) in highway ramp merging scenarios. The simulation results reveal several pivotal findings pertaining to traffic efficiency, safety, and passenger comfort.

Firstly, the implementation of a DRL-based driving strategy (ZD2) markedly enhances traffic flow stability and efficiency. The ZD2 vehicles exhibited a superior capacity to maintain a constant velocity and mitigate fluctuations in speed compared to both human-operated vehicles (RL) and rule-based autonomous vehicles (ZD1). Such stability is crucial for preventing traffic congestion and ensuring a smooth merging process.

Secondly, the level of lane changing significantly improved with the implementation of DRL. During lane changes, the ZD2 vehicles maintained a high lane change speed while keeping a reasonable distance from the vehicle in front, thereby achieving efficiency without compromising safety. The ZD2 vehicles demonstrated the capacity to change lanes with greater precision and timeliness, without undue disruption to the flow of traffic on the main road. This outcome evinces the efficacy of the DRL algorithm in navigating complex driving scenarios.

However, the study also identified areas for further research and improvement. While the DRL-based ZD2 vehicle exhibited superior performance in numerous areas, its average travel speed was slightly lower than that of the baseline human-driven vehicle, suggesting scope for optimizing the balance between safety and efficiency. Future research should investigate the impact of increased autonomous vehicle penetration and more complex traffic scenarios, such as highways with varying densities in each lane. These extensions will facilitate a more comprehensive understanding of the scalability and robustness of DRL-based driving strategies.

6. Conclusions

This study illustrates the considerable potential of deep reinforcement learning (DRL) in optimizing the driving strategy of autonomous vehicles in highway ramp merging scenarios. The DRL-based autonomous vehicle (ZD2) demonstrates superior performance

in terms of traffic flow stability, safety, and passenger comfort when compared to both the traditional rule-based autonomous vehicle (ZD1) and the human-driven vehicle (RL).

The primary findings are as follows: ZD2 vehicles exhibit superior stability in maintaining a consistent speed, which is instrumental in preventing traffic congestion and ensuring seamless merging. They also demonstrate superior control when changing lanes, maintaining a reasonable distance from other vehicles while performing lane changes at high speeds, thereby reducing the risk of collision. Additionally, the more stable acceleration mode of ZD2 vehicles provides a more comfortable driving experience for passengers. In conclusion, the findings of this research are anticipated to provide guidance for the future advancement of AV technology and traffic management systems, thereby facilitating the development of autonomous driving in real-world scenarios.

Nevertheless, further research is required to achieve an optimal balance between safety and efficiency, particularly in more complex and higher-density traffic scenarios. Moreover, the scalability of DRL-based strategies must be investigated to ensure their efficacy with a significant increase in the number of autonomous vehicles on the road.

Author Contributions: The authors confirm their contribution to the paper as follows: study conception and design: H.H., Z.Z. and Z.C.; data collection: Z.C., Y.W. and S.Z.; analysis and interpretation of results: Z.C. and C.Z.; manuscript preparation: Z.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The authors will supply the relevant data in response to reasonable requests.

Acknowledgments: The authors employed ChatGPT for language polishing assistance to enhance the clarity and quality of the manuscript.

Conflicts of Interest: Author Shukun Zhou was employed by Shandong Hi-Speed Construction Management Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Yurtsever, E.; Lambert, J.; Carballo, A.; Takeda, K. A Survey of Autonomous Driving: Common Practices and Emerging Technologies. *IEEE Access* **2020**, *8*, 58443–58469. [CrossRef]
2. Bose, A.; Ioannou, P. Mixed manual/semi-automated traffic: A macroscopic analysis. *Transp. Res. Part C Emerg. Technol.* **2003**, *11*, 439–462. [CrossRef]
3. Fagnant, D.J.; Kockelman, K. Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transp. Res. Part A Policy Pract.* **2015**, *77*, 167–181. [CrossRef]
4. Qin, Y.Y.; Wang, H.; Ran, B. Impact of Connected and Automated Vehicles on Passenger Comfort of Traffic Flow with Vehicle-to-vehicle Communications. *KSCE J. Civ. Eng.* **2019**, *23*, 821–832. [CrossRef]
5. Grigorescu, S.; Trasnea, B.; Cocias, T.; Macesanu, G. A Survey of Deep Learning Techniques for Autonomous Driving. *J. Field Robot.* **2020**, *37*, 362–386. [CrossRef]
6. Kiran, B.R.; Sobh, I.; Talpaert, V.; Mannion, P.; Al Sallab, A.A.; Yogamani, S.; Pérez, P. Deep Reinforcement Learning for Autonomous Driving: A Survey. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 4909–4926. [CrossRef]
7. Talebpour, A.; Mahmassani, H.S. Influence of Connected and Autonomous Vehicles on Traffic Flow Stability and Throughput. *Transp. Res. Part C Emerg. Technol.* **2016**, *71*, 143–163. [CrossRef]
8. Song, H.F.; Li, L.L.; Li, Y.; Tan, L.G.; Dong, H.R. Functional Safety and Performance Analysis of Autonomous Route Management for Autonomous Train Control System. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 13291–13304. [CrossRef]
9. Zhou, S.; Zhuang, W.; Yin, G.; Liu, H.; Qiu, C. Cooperative On-Ramp Merging Control of Connected and Automated Vehicles: Distributed Multi-Agent Deep Reinforcement Learning Approach. In Proceedings of the 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), Macau, China, 8–12 October 2022; pp. 402–408. [CrossRef]
10. Paden, B.; Cap, M.; Yong, S.Z.; Yershov, D.; Frazzoli, E. A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles. *IEEE Trans. Intell. Veh.* **2016**, *1*, 33–55. [CrossRef]
11. Jiang, W.; Huang, Z.; Wu, Z.; Su, H.; Zhou, X. Quantitative Study on Human Error in Emergency Activities of Road Transportation Leakage Accidents of Hazardous Chemicals. *Int. J. Environ. Res. Public Health* **2022**, *19*, 14662. [CrossRef]
12. Jin, Z.; RuiChun, H.E. Research on Road Traffic Safety Based on Hammer's Human Error Theory. *China Saf. Sci. J.* **2008**, *18*, 53–58.

13. Hirst, W.M.; Mountain, L.J.; Maher, M.J. Sources of error in road safety scheme evaluation: A method to deal with outdated accident prediction models. *Accid. Anal. Prev.* **2004**, *36*, 717–727. [CrossRef] [PubMed]
14. McCorry, B.; Murray, W. Reducing commercial vehicle road accident costs. *Int. J. Phys. Distrib. Logist. Manag.* **1993**, *23*, 35–41. [CrossRef]
15. Xu, D.; Zhang, B.; Qiu, Q.; Li, H.; Guo, H.; Wang, B. Graph-Based Multi-Agent Reinforcement Learning for On-Ramp Merging in Mixed Traffic. *Appl. Intell.* **2024**, *54*, 6400–6414. [CrossRef]
16. Irshayyid, A.; Chen, J.; Xiong, G. A Review on Reinforcement Learning-Based Highway Autonomous Vehicle Control. *Green Energy Intell. Transp.* **2024**, *3*, 100156. [CrossRef]
17. Liu, K.W.; Li, N.; Tseng, H.E.; Kolmanovsky, I.; Girard, A. Interaction-Aware Trajectory Prediction and Planning for Autonomous Vehicles in Forced Merge Scenarios. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 474–488. [CrossRef]
18. Hu, J.B.; He, L.C.; Wang, R.H. Safety evaluation of freeway interchange merging areas based on driver workload theory. *Sci. Prog.* **2020**, *103*, 0036850420940878. [CrossRef]
19. Zhu, J.; Tasic, I. Safety analysis of freeway on-ramp merging with the presence of autonomous vehicles. *Accid. Anal. Prev.* **2021**, *152*, 105966. [CrossRef]
20. Ntousakis, I.A.; Nikolos, L.K.; Papageorgiou, M. Optimal vehicle trajectory planning in the context of cooperative merging on highways. *Transp. Res. Part C-Emerg. Technol.* **2016**, *71*, 464–488. [CrossRef]
21. Sarvi, M.; Kuwahara, M. Microsimulation of freeway ramp merging processes under congested traffic conditions. *IEEE Trans. Intell. Transp. Syst.* **2007**, *8*, 470–479. [CrossRef]
22. Li, W.; Zhao, Z.; Liang, K.; Zhao, K. *Coordinated Longitudinal and Lateral Motions Control of Automated Vehicles Based on Multi-Agent Deep Reinforcement Learning for On-Ramp Merging*; SAE Technical Paper, 2024-01-2560; SAE International: Warrendale, PA, USA, 2024. [CrossRef]
23. Zhang, X.; Wu, L.; Liu, H.; Wang, Y.; Li, H.; Xu, B. High-Speed Ramp Merging Behavior Decision for Autonomous Vehicles Based on Multiagent Reinforcement Learning. *IEEE Internet Things J.* **2023**, *10*, 22664–22672. [CrossRef]
24. Zhao, R.; Sun, Z.; Ji, A. A Deep Reinforcement Learning Approach for Automated OnRamp Merging. In Proceedings of the 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), Macau, China, 8–12 October 2022; pp. 3800–3806. [CrossRef]
25. Zhou, Y.J.; Zhu, H.B.; Guo, M.M.; Zhou, J.L. Impact of CACC vehicles' cooperative driving strategy on mixed four-lane highway traffic flow. *Physica A* **2020**, *540*, 122721. [CrossRef]
26. Song, H.F.; Gao, S.G.; Li, Y.D.; Liu, L.; Dong, H.R. Train-Centric Communication Based Autonomous Train Control System. *IEEE Trans. Intell. Veh.* **2023**, *8*, 721–731. [CrossRef]
27. Weaver, S.M.; Balk, S.A.; Philips, B.H. Merging into strings of cooperative-adaptive cruise-control vehicles. *J. Intell. Transport. Syst.* **2021**, *25*, 401–411. [CrossRef]
28. Pei, Y.; Chi, B.; Lyu, J.; Yue, Z. An Overview of Traffic Management in "Automatic+Manual" Driving Environment. *J. Transp. Inf. Saf.* **2021**, *39*, 1–11.
29. Ozioko, E.F.; Kunkel, J.; Stahl, F. Road Intersection Coordination Scheme for Mixed Traffic (Human-Driven and Driverless Vehicles): A Systematic Review. *J. Adv. Transp.* **2022**, *2022*, 15. [CrossRef]
30. Guo, L.X.; Jia, Y.Y. Predictive Control of Connected Mixed Traffic under Random Communication Constraints. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2020; IEEE: New York, NY, USA, 2020; pp. 11817–11823. [CrossRef]
31. Imbsweiler, J.; Palyafári, R.; León, F.P.; Deml, B. Investigation of decision-making behavior in cooperative traffic situations using the example of a narrow passage. *AT-Autom.* **2017**, *65*, 477–488. [CrossRef]
32. Cui, X.T.; Li, X.S.; Zheng, X.L.; Zhang, X.Y.; Zhao, L.; Zhang, J.H. Research on Traffic Characteristics of Signal Intersections with Mixed Traffic Flow. In Proceedings of the IEEE 5th International Conference on Intelligent Transportation Engineering (ICITE), Beijing, China, 11–13 September 2020; IEEE: New York, NY, USA, 2020; pp. 306–312. [CrossRef]
33. Ramezani, M.; Machado, J.A.; Skabardonis, A.; Geroliminis, N. Capacity and Delay Analysis of Arterials with Mixed Autonomous and Human-Driven Vehicles. In Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Napoli, Italy, 26–28 June 2017; IEEE: New York, NY, USA, 2017; pp. 280–284.
34. Wang, X.P.; Wu, C.Z.; Xue, J.; Chen, Z.J. A Method of Personalized Driving Decision for Smart Car Based on Deep Reinforcement Learning. *Information* **2020**, *11*, 295. [CrossRef]
35. Silver, D.; Huang, A.; Maddison, C.J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **2016**, *529*, 484–489. [CrossRef]
36. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
37. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; et al. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529–533. [CrossRef] [PubMed]
38. Li, Z.; Yuan, S.H.; Yin, X.F.; Li, X.Y.; Tang, S.X. Research into Autonomous Vehicles Following and Obstacle Avoidance Based on Deep Reinforcement Learning Method under Map Constraints. *Sensors* **2023**, *23*, 844. [CrossRef] [PubMed]
39. Bin Issa, R.; Das, M.; Rahman, M.S.; Barua, M.; Rhaman, M.K.; Ripon, K.S.N.; Alam, M.G.R. Double Deep Q-Learning and Faster R-CNN-Based Autonomous Vehicle Navigation and Obstacle Avoidance in Dynamic Environment. *Sensors* **2021**, *21*, 1468. [CrossRef] [PubMed]

40. Chen, D.; Hajidavalloo, M.R.; Li, Z.; Chen, K.; Wang, Y.; Jiang, L.; Wang, Y. Deep Multi-Agent Reinforcement Learning for Highway On-Ramp Merging in Mixed Traffic. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 11623–11638. [CrossRef]
41. Lee, J.; Eom, C.; Lee, D.; Kwon, M. Deep Reinforcement Learning-Based Autonomous Driving Strategy under On-Ramp Merge Scenario. *Trans. KSAE* **2024**, *32*, 569–581. [CrossRef]
42. Qiao, L.; Bao, H.; Xuan, Z.; Liang, J.; Pan, F. Autonomous Driving Ramp Merging Model Based on Reinforcement Learning. *Comput. Eng.* **2018**, *44*, 20–24. [CrossRef]
43. Treiber, M.; Hennecke, A.; Helbing, D. Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. E* **2000**, *62*, 1805–1824. [CrossRef]
44. Krajzewicz, D.; Erdmann, J.; Behrisch, M.; Bieker, L. Recent development and applications of SUMO—Simulation of urban mobility. *Int. J. Adv. Syst. Meas.* **2012**, *5*, 128–138.
45. Erdmann, J. SUMO's Lane-Changing Model. In Proceedings of the 2nd SUMO User Conference (SUMO), Berlin, Germany, 15–16 May 2014; Springer: Berlin/Heidelberg, Germany, 2015; pp. 105–123. [CrossRef]
46. Helbing, D.; Hennecke, A.; Shvetsov, V.; Treiber, M. Micro- and macro-simulation of freeway traffic. *Math. Comput. Model.* **2002**, *35*, 517–547. [CrossRef]
47. Zhu, L.; Lu, L.J.; Wang, X.N.; Jiang, C.M.; Ye, N.F. Operational Characteristics of Mixed-Autonomy Traffic Flow on the Freeway With On- and Off-Ramps and Weaving Sections: An RL-Based Approach. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 13512–13525. [CrossRef]
48. Zeng, X.; Lin, H.; Wang, Y.; Yuan, T.; He, Q.; Huang, J. Safety Risk Identification of Rail Transit Signaling System Based on Accident Data. *J. Tongji University Nat. Sci.* **2022**, *50*, 418–424.
49. Zeng, X.; Lin, H.; Fang, Y.; Wang, Y.; Liu, Y.; Ma, Z. Safety Target Assignment Analysis Method Based on Complexity Algorithm. *J. Tongji University Nat. Sci.* **2022**, *50*, 1–5.
50. Pu, Z.Y.; Li, Z.B.; Jiang, Y.; Wang, Y.H. Full Bayesian Before-After Analysis of Safety Effects of Variable Speed Limit System. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 964–976. [CrossRef]
51. Vogel, K. A comparison of headway and time to collision as safety indicators. *Accid. Anal. Prev.* **2003**, *35*, 427–433. [CrossRef]
52. Vinitsky, E.; Kreidieh, A.; Flem, L.L.; Kheterpal, N.; Jang, K.; Wu, C.; Wu, F.; Liaw, R.; Liang, E.; Bayen, A.M. Benchmarks for reinforcement learning in mixed-autonomy traffic. In Proceedings of the 2nd Conference on Robot Learning, Zürich, Switzerland, 29–31 October 2019.
53. Zhang, G.H.; Wang, Y.H.; Wei, H.; Chen, Y.Y. Examining headway distribution models with urban freeway loop event data. *Transp. Res. Rec.* **2007**, *1999*, 141–149. [CrossRef]
54. Zhu, M.X.; Wang, Y.H.; Pu, Z.Y.; Hu, J.Y.; Wang, X.S.; Ke, R.M. Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving. *Transp. Res. Part C Emerg. Technol.* **2020**, *117*, 102662. [CrossRef]
55. Belletti, F.; Haziza, D.; Gomes, G.; Bayen, A.M. Expert Level Control of Ramp Metering Based on Multi-Task Deep Reinforcement Learning. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 1198–1207. [CrossRef]

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.