Integrated Control Strategies for Freeway Bottlenecks with Vehicle Platooning

Liang Qi, Lu Liang, Wenjing Luan, Tong Lu, Xiwang Guo, and Qurra Tul Ann Talukder

***Abstract*—Traffic incidents often lead to freeway bottlenecks, which are major contributors to widespread traffic congestion. Conventional approaches for managing freeway congestion, such as variable speed limits (VSL) and ramp metering, are widely implemented. In recent times, vehicle platooning has emerged as a promising strategy to mitigate traffic bottlenecks. This study introduces an innovative framework that integrates VSL with vehicle platooning to address freeway bottlenecks, named VSL-VP, specifically designed for mixed traffic environments comprising both connected and autonomous vehicles (CAVs) and human-driven vehicles (HDVs). Initially, the upstream section of a bottleneck is partitioned into two segments: the upstream and downstream portions. By imposing VSL on the upstream segment, the framework effectively curtails the volume of incoming traffic to the downstream segment. Subsequently, deep reinforcement learning is utilized to facilitate CAV platooning within the downstream segment, where reduced traffic density and increased following distances between vehicles create favorable conditions for seamless lane changes and the formation of CAV platoons. Simulation results indicate that the VSL-VP framework significantly improves bottleneck throughput and alleviates traffic congestion, particularly as the penetration rate of CAVs increases.**

***Index Terms*—bottleneck control method, variable speed limit, vehicle platooning.**

I. INTRODUCTION

A traffic bottleneck is a specific point along a road where traffic flow becomes restricted, occurring when the volume of traffic exceeds the road's capacity due to factors like inadequate road design, inefficient traffic signal timing, or traffic incidents [1]. This restriction is a primary cause of widespread congestion on freeways [2]. Traditional approaches to mitigating freeway bottlenecks include ramp metering (RM) [3] and variable speed limit (VSL) [4]-[6]. RM restricts vehicles from entering congested sections, while VSL reduces vehicle speeds upstream of the bottleneck to manage traffic demand. Although both methods are effective in preventing congestion caused by bottlenecks, they have limitations. RM can disrupt traffic on nearby roads when on-ramps are temporarily closed, and VSL can negatively impact traffic flow upstream of the bottleneck [7].

Given the challenges associated with traditional strategies for addressing freeway bottlenecks, it is essential to recognize the significant changes occurring in traffic patterns and management. The rapid development of vehicle automation and communication technologies is leading to an increasing presence of connected and automated vehicles (CAVs) in the automotive market [8]. In the near future, CAVs and human-driven vehicles (HDVs) are expected to share the roads, shifting the traffic landscape from predominantly HDVs to a mix of both [9], [10]. Consequently, there is an urgent need to explore new strategies for controlling traffic bottlenecks in this mixed traffic environment [11].

One of the most promising strategies in intelligent transportation systems is vehicle platooning [12]. A vehicle platoon consists of a group of CAVs traveling closely together in the same lane, maintaining a consistent, reduced following distance and time gap while operating at higher speeds [13]. This approach holds considerable potential for increasing road capacity. Additionally, by reducing aerodynamic drag, vehicle platoons can decrease fuel consumption [14], [15]. Several studies have highlighted the benefits of vehicle platooning in alleviating traffic bottlenecks [16], [17]. The process of vehicle platooning involves adjacent CAVs forming a stable group through joining and merging maneuvers [18]. Zhao *et al.* propose a platoon formation method using model predictive control (MPC) to optimize the passage of platoons through intersections during green phases, minimizing fuel consumption [19]. Smith *et al.* demonstrate the potential of an MPC-based approach for vehicle platooning in urban traffic settings, showing improvements in urban traffic throughput [20].

As CAVs are often randomly dispersed within mixed traffic, several actions—such as joining, leaving, merging, and splitting—are necessary to organize nearby CAVs into a platoon. These actions can potentially have negative effects, including the induction of unwanted congestion [21]. However, current research frequently overlooks crucial decision-making aspects related to these maneuvers [22]. For example, there is limited focus on identifying the optimal joining time, which is the point at which a CAV should initiate lane changes and acceleration to integrate safely and efficiently into a platoon. Furthermore, as the platoon size increases, the complexity of the solution space grows, leading to a substantial rise in computational demands. Recent developments in deep

Manuscript received March 30, 2025; revised April 10, 2025; accepted April 18, 2025. This article was recommended for publication by Associate Editor Shujin Qin upon evaluation of the reviewers' comments. This work was supported in part by the National Natural Science Foundation of China under Grant 61903229 and in part by the Natural Science Foundation of Shandong Province under Grant ZR2024MF140.

L. Qi, L. Liang, W. Luan, T. Lu, and Q. T. A. Talukder are with the College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China (e-mail: qiliangsdkd@163.com, ll18395287328@163.com, tong_lu2023@163.com, upoma_gourab@hotmail.com).

X. Guo is with the College of Computer and Communication Engineering, Liaoning Petrochemical University, Fushun 113001, China, (e-mail: x.w.guo@163.com).

Corresponding Author: Liang Qi.

reinforcement learning (DRL) have demonstrated significant potential in addressing complex control challenges marked by uncertain and high-dimensional state and action spaces [23]-[25]. Some research has started utilizing DRL for platoon control. For example, Li *et al.* introduce a multi-agent reinforcement learning algorithm to manage vehicle platoons and enhance energy efficiency during traffic fluctuations [26]. Another study [27] develops a DRL-based hierarchical model that combines platooning and coordination to improve CAV control, thereby reducing travel time and fuel consumption at intersections without traffic signals. Shi *et al.* [28] propose a distributed longitudinal control strategy for CAVs in mixed traffic environments using DRL, devising a new method to significantly mitigate traffic oscillations. However, to our knowledge, no existing studies have applied DRL to determine the optimal joining time for CAVs in mixed traffic scenarios.

In bottleneck zones, high traffic density makes it challenging for vehicles to safely change lanes for platooning. Current research primarily concentrates on platooning in low-density situations [29]-[31]. To address this, this study utilizes VSL to support vehicle platooning. Specifically, VSL controls vehicle speeds to create a stretch of road with lower traffic density and greater following distances between vehicles. This setup allows CAVs to change lanes smoothly and form platoons safely and efficiently. To our knowledge, this is the first initiative to combine VSL with vehicle platooning to mitigate bottlenecks in extensive mixed traffic scenarios. The main contributions are outlined as follows:

1) This work introduces an innovative framework that integrates VSL and vehicle platooning to tackle bottlenecks. Initially, VSL imposes speed restrictions, thereby decreasing the volume of traffic entering the bottleneck. The resulting lower traffic density and increased following distances create optimal conditions for CAVs to perform safe lane-changing maneuvers necessary for platooning.

2) To manage the high-dimensional solution space associated with CAV platooning, this study uses sliding windows to divide large-scale traffic flows into separate segments. Within each segment, deep reinforcement learning (DRL) is employed to identify the optimal joining time for CAVs in mixed traffic settings. Curriculum learning is applied during model training [32]. The model begins by learning from simpler scenarios with fewer vehicles and is subsequently refined using extensive traffic data from various conditions, including different CAV penetration rates and vehicle densities.

Furthermore, comprehensive experiments are carried out to demonstrate the effectiveness of the proposed framework in reducing bottlenecks and congestion. The study examines how varying window sizes affect bottleneck throughput and discusses the scalability of the proposed framework.

The rest of this paper is structured as follows. Section II describes the problem at hand. Section III details the approach for integrating VSL and vehicle platooning to address freeway bottlenecks. Section IV outlines the experiments conducted and the analysis of the results. Lastly, Section V concludes the paper and suggests potential avenues for future research.

II. Problem Description and Preliminaries

A. Problem Description

This study focuses on a typical three-lane freeway, which is a prevalent configuration, featuring a bottleneck resulting from a lane obstruction, as illustrated in Fig. 1. The scenario includes a mixed traffic flow consisting of both connected and automated vehicles (CAVs) equipped with Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication systems, and human-driven vehicles (HDVs) without such communication capabilities. Additional supporting elements include road-side units (RSUs) and variable speed limit (VSL) controllers. RSUs are equipped with devices such as traffic detectors, cameras, and routers to gather vehicle data and enable the transmission of vehicle status through V2I communication. The VSL controller dynamically adjusts speed limits by changing the speed-limit signs on a section of the freeway. The Connected Vehicle Center (CVC) manages all connected devices via communication networks. In this setting, a CAV platoon is defined as a group of two or more adjacent CAVs traveling in the same lane at high speeds, maintaining a consistent, short following distance and time gap.

When a bottleneck occurs randomly in a certain area, an RSU collects such information and relays it to CVC. Subsequently, the CVC dynamically divides the upstream road of the bottleneck into a former road segment (S-1) for VSL implementation and a latter segment (S-2) for vehicle platooning in a dedicated lane, as illustrated in Fig. 1.

The goal of this research is to alleviate freeway bottlenecks by integrating VSL with vehicle platooning. However, there are two significant challenges that need to be tackled:

1) How to decrease vehicle density to facilitate platooning. Bottlenecks often lead to traffic congestion. When vehicle density is excessively high, lane-changing maneuvers become unfeasible at the bottleneck due to insufficient following distances, impeding the formation of CAV platoons.

2) How to coordinate individual CAVs to form platoons within a mixed traffic environment. This entails selecting a lead CAV and identifying the most suitable time for other CAVs to join the platoon.

B. Basic Definitions and Assumptions

Definition 1 (CAV penetration rate). A CAV penetration rate p_l refers to the proportion of CAVs within a road lane l

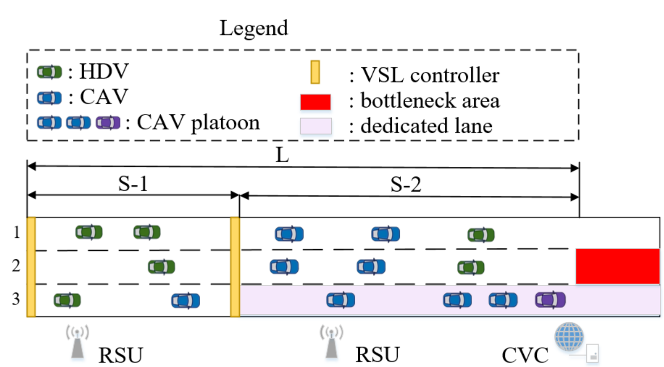


Fig. 1. A bottleneck in a three-lane freeway with mixed traffic.

calculated by

$$p_l = \frac{n_{l,c}}{n_{l,c} + n_{l,h}} \tag{1}$$

where $n_{l,c}$ and $n_{l,h}$ denote the count of CAVs and HDVs in l, respectively.

Given a penetration rate, the distribution of CAVs may vary across different spatial areas. Platoon intensity is used to measure the distribution of CAVs[33] and provide the following mathematical definition.

Definition 2 (Platoon intensity) [34]**.** A platoon intensity, denoted as I_l, represents the ratio of the actual number of CAVs in platoons to the total number of CAVs in the mixed traffic flow within the same lane. It is calculated by

$$I_l = \frac{\sum_{k=2}^{Z} k \times m_k}{n_l^c} \tag{2}$$

where k is the number of vehicles (called size) of a platoon, m_k is the number of platoons with size k, $m_k \geq 2$, and Z is the maximum size of a platoon.

Furthermore, the following assumptions are considered:

1) Within the communication range, the status of a CAV can be captured by RSUs and transmitted to the CVC. This scenario does not account for any communication delays or detection errors.
2) Each CAV receives and precisely follows the instructions from the CVC, including adjustments to speed and lane-changing actions.
3) To maintain the safety and effectiveness of the approach, HDVs are restricted from entering lanes designated for CAVs, whereas CAVs are allowed to use any lane on S-2 to form platoons. It is important to note that some previous studies have employed dedicated lanes for CAVs [35]-[38]. This strategy aims to separate CAVs from mixed traffic and reduce the adverse impacts of erratic driving behaviors by HDVs on CAVs within mixed traffic flow.

III. Methodology

This section presents a framework that integrates VSL and vehicle platooning (VSL-VP) to reduce freeway bottlenecks in a mixed traffic setting with both CAVs and HDVs. Initially, an overview of the framework is provided, succeeded by the technical specifics, where a RL-based approach for platooning is introduced. The notations used throughout the subsequent sections are defined in Table I.

A. Framework of VSL-VP

A schematic representation of the proposed VSL-VP is depicted in Fig. 2. The framework divides the upstream road of the bottleneck into two distinct segments, denoted as S-1 and S-2. Vehicle platooning occurs in S-2. In order to ensure sufficient car-following distance for lane-changing maneuvers to form platoons, VSL is adopted in S-1. As vehicles approach S-1, VSL controllers transmit speed limits, thereby slowing down traffic flow and decreasing the inflow of S-2. RSUs positioned within S-1 collect vehicle data and relay it to CVC. Subsequently, CVC calculates the CAV penetration rate and

TABLE I
NOTATIONS

Notations	Meaning
p_l	CAV penetration rate of lane l.
I_l	Platoon intensity of lane l.
F_{in}	Total inflow of the bottleneck.
F_{out}	Total outflow of the bottleneck.
$n_{l,c}$	Number of CAVs in lane l.
$n_{l,h}$	Number of HDVs in lane l.

platoon intensity for each lane and chooses a dedicated lane for vehicle platooning. A coefficient, denoted as D_l, is defined to determine lane l as a dedicated lane, calculated by:

$$D_l = (\omega_1 * p_l + \omega_2 * I_l) * b_l \tag{3}$$

where ω_1 and ω_2 represent the weights assigned to the CAV penetration rate and platoon intensity, respectively. Additionally, b_l indicates whether lane l has a bottleneck: b_l=0 if l is blocked, and otherwise, b_l=1. The lane with the highest value of D_l is selected as a dedicated lane for CAVs to formulate platoons.

When vehicles leave S-1 and enter S-2, speed limits are lifted, and vehicles begin to accelerate, thereby increasing the car-following distance from the vehicles behind them. A large car-following distance ensures smooth lane-changing maneuvers of CAVs for formulating platoons safely and efficiently. However, coordinating all CAVs on segment S-2 to start platooning simultaneously poses a challenge, and the size of the platoons also influences the traffic flow stability and road capacity [39]. Therefore, sliding-windows are employed to divide the large-scale traffic flow into individual windows, each comprising of a road section containing a finite number of vehicles. CAVs within each window independently participate in platooning based on deep reinforcement learning (DRL).

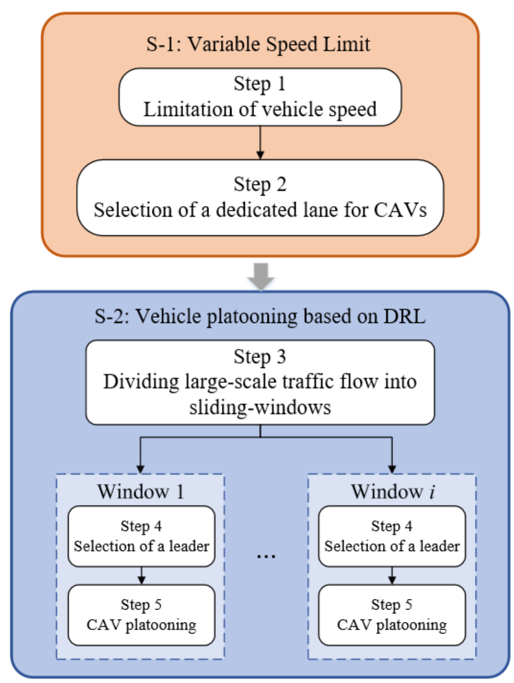


Fig. 2. Hierarchical framework of the VSL-VP.

Upon entering S-2, vehicles are separated into several windows according to their position and the maximum communication distance between two CAVs. Here, the window size represents the maximum number of vehicles in the window.

In each window, the CAV nearest to the bottleneck is usually chosen as the leader. This leader then shifts to the designated lane. Subsequently, DRL is employed to identify the best time for the remaining CAVs to join a platoon. Windows with only one CAV do not need a leader. The system checks if there is another platoon within the communication range of this CAV. If a platoon is found, the CAV joins it; if not, it stays independent. Additional details about DRL are provided in the subsequent subsections.

B. Reinforcement Learning for Platooning

Reinforcement learning (RL) is an algorithm that describes how an agent takes actions to maximize expected benefits in an unknown environment [40]. As shown in Fig. 3, an RL process mainly consists of the interaction between agent G and the environment O. G first takes the current state s as the input and learns a policy to choose an action a. After a is performed, a reward r is obtained. Subsequently, environment O changes and the state is transformed into a new one s'. The agent dynamically interacts with the environment and updates its strategy to maximize the accumulated rewards [41]. This procedure is regarded as a Markov Decision Process (MDP), which can be described as a four-tuple $\{S, A, P, R\}$, where:

S denotes a traffic state space, where $s \in S$ is a specific state;

A denotes an action space, where $a \in A$ is a specific action;

$P=S \times A \times S$ denotes the transmission probability among states; and

R denotes a reward space, where $r \in R$ is a specific reward.

To progressively enhance the environment towards an optimal state, reinforcement learning (RL) agents choose actions based on an optimal policy function represented by π. The objective of π is to maximize the cumulative expected rewards starting from the initial state. If agents know the optimal cumulative reward at a given state, they can choose actions that provide the maximum reward [42]. The cumulative reward can be determined recursively using the Bellman equation [43]. For example, when an agent in a particular state s takes an action a to transition to the next state s' and receives a reward r, represented as a tuple (s, a, r, s'), the cumulative reward denoted by $Q^{\pi}(s, a)$ under policy π can be calculated by

$$Q^{\pi}(s,a) = E_{s'}\left[r + \gamma \max_{a'} Q^{\pi}(s',a') \mid s,a\right] \tag{4}$$

where $a' \sim \pi(s')$ is an action selected according to policy π at state s', γ is a discount factor, and a' is the best possible action.

Deep Q-network (DQN) [44] is proposed to estimate the $Q^{\pi}(s, a)$, i.e., determining the optimal time for a CAV to join a platoon. CAV platooning is formalized as a MDP with RL settings, which involves trial-error interaction with the environment, as illustrated in Fig. 3. The design of the state, action, transition probability, and reward is given next.

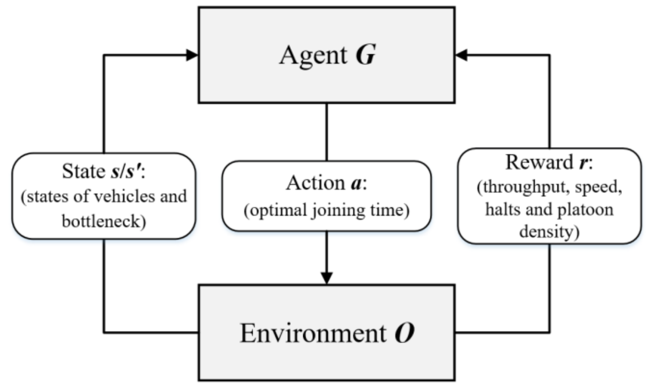


Fig. 3. RL for vehicle platooning.

Agent: A set of vehicles in a sliding-window is treated as an agent. Notice that each window conducts platooning independently, thus minimizing the interference among agents. The goal of each agent is to enhance traffic throughput at the bottleneck and reduce congestion.

State: s_i is a state of window i, denoted by:

$$s_i = (d_{i,l}, d_{i,c}, d_{i,h}, d_b) \tag{5}$$

where $d_{i,l}$, $d_{i,c}$, and $d_{i,h}$ respectively represent the vehicle state of the leader, CAVs, and HDVs within window i, including their position, speed, and lane allocation, and d_b denotes the state of the bottleneck, including its position and the blocked lane.

Note that the number of vehicles (up to a maximum of L) and CAV penetration rate in each window may vary. To ensure a uniform dimension of state s across different windows, windows with fewer than L vehicles are padded with zero vectors. The state information is refreshed after each time step.

Action: At each time step t, DRL receives a state s_i and selects an action a_i, which comprises the actions of all CAVs within window i. The size of the action a_i is determined by the number of CAVs present. A discrete action space $\{a_{i,j}\}$ is utilized, where $a_{i,j}$=1 signifies that CAV j within window i takes an action to join the platoon by changing lanes, accelerating to reach the leading CAV, and merging into the platoon. Conversely, $a_{i,j}$=0 indicates that CAV j maintains its current state.

Reward: An agent receives rewards designed to promote future positive actions aimed at improving the traffic throughput of the bottleneck. Four rewards are employed to target specific aspects of traffic flow dynamics within the bottleneck.

1) The first reward, denoted as r_1, is linked to traffic throughput. It is awarded only at the conclusion of the simulation and is defined as:

$$r_1 = F_{in} - F_{out} \tag{6}$$

where F_{in} and F_{out} represent the total inflow and outflow of the bottleneck, respectively.

2) The second reward, r_2, reflects the average vehicle speed at the bottleneck:

$$r_2 = \bar{v} \tag{7}$$

where $\bar{v}$ represents the average vehicle speed.

3) The third reward, r_3, relates to the number of stops made by vehicles in the bottleneck area. To discourage prolonged traffic congestion before the bottleneck, a penalty term is introduced:

$$r_3 = -n_s \quad (8)$$

where n_s is the total number of stops made by vehicles. In the simulation, a vehicle traveling at a speed lower than 3 m/s is considered to have stopped.

4) The fourth reward, r_4, is associated with platoon density, which is the ratio of CAVs participating in platooning to the total number of CAVs in each window. A small positive value of r_4 is assigned to windows with high platoon density at each time step to incentivize early platoon formation:

$$r_4 = \frac{n_{i,p}}{n_{i,c}} \quad (9)$$

where $n_{i,c}$ and $n_{i,p}$ are the total number of CAVs and the number of CAVs participating in platooning within window i, respectively.

C. DQN

The architecture of DRL with a DQN is illustrated in Fig. 4, featuring two distinct tiers. The upper tier handles decision-making processes, including neural network training and action refinement. Specifically, the training involves three main elements [44]. A Q-network with the current parameters θ plays a pivotal role in determining the policy governing the agent's actions. A target Q-network with the previous parameters $\hat{\theta}$ is employed to generate Q values crucial for the loss function during the training process. A replay memory serves as a storage system for saving and retrieving training samples. During training, data is randomly sampled from the replay memory. The parameters θ are updated multiple times per time step and are periodically copied to $\hat{\theta}$ after every κ iterations. In the Q-network, $Q(s, a; \theta)$ represents the output estimating the value gained by the agent taking action a in state s. Meanwhile, $\hat{Q}(s', a'; \hat{\theta})$ represents the output of the target network. At each iteration ζ, the parameter θ is updated to minimize the loss function defined as:

$$L(\theta) = E\left[\left(r + \gamma \max_{a'} \hat{Q}(s', a'; \hat{\theta}) - Q(s, a; \theta)\right)^2\right] \quad (10)$$

Once training is complete, the Q values produced by the Q network require further processing to be converted into actions, including the elimination of invalid actions. For instance, once a CAV successfully joins a platoon, it is prohibited from performing additional joining maneuvers.

The low level is responsible for executing decisions. Following the actions determined by the upper level, RSUs transmit these actions to each CAV. Based on the car-following and lane-changing models of CAVs, these actions are converted into speed and acceleration commands, which the CAVs then carry out.

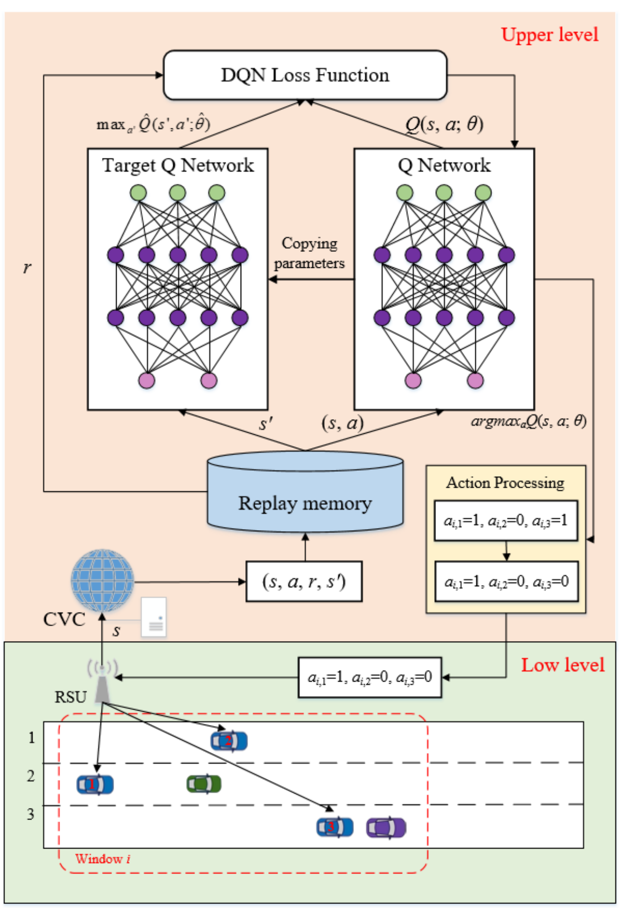


Fig. 4. Framework for DRL with a DQN.

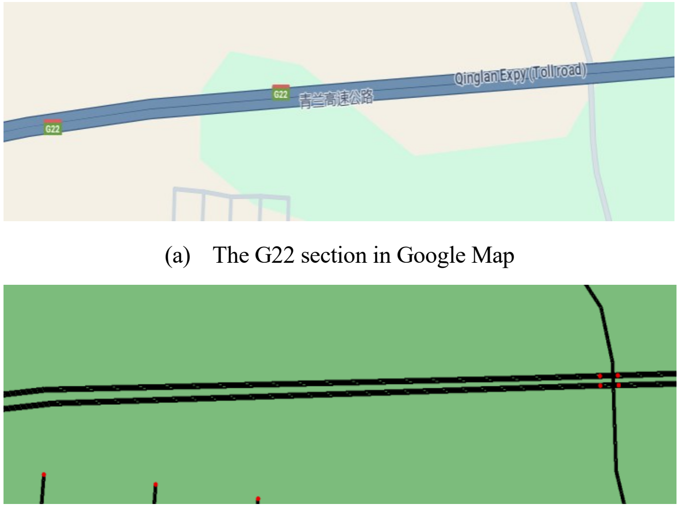
IV. Experiments

This section outlines the experiments conducted on platooning and bottleneck control. It encompasses both the training and testing phases of the platooning model, as well as an evaluation of the VSL-VP approach in managing bottlenecks under varying CAV penetration rates. Additionally, the influence of window size on the performance of VSL-VP is examined.

A. Simulation Scenario

The simulation platform utilized is PLEXE-SUMO [45], an OMNeT++ framework that offers platooning f capabilities through the Plexe models integrated into SUMO [46]. The simulation scenario for the experiment is based on a segment of QinglanExpy, G22, located in Qingdao, China, as depicted in Fig. 5.

In simulation, SUMO plays a crucial role due to its provision of various classical car-following and lane-changing models. For the longitudinal movement of HDVs, the Intelligent Driver Model (IDM) is employed [47] because of its ability to realistically mimic the acceleration and deceleration patterns of HDVs. Meanwhile, a Cooperative Adaptive Cruise Control (CACC) model [48], [49] is chosen to manage the longitudinal motion of CAVs. Known for its cooperative and adaptive features, CACC is well-suited for capturing the autonomous and interconnected behavior of CAVs. Regarding lateral movement, both HDVs and CAVs are governed by SUMO's default lane change model LC2013 [50], which incorporates realistic considerations such as safety distances and traffic conditions to ensure an accurate representation of lateral dynamics. Furthermore, the detailed simulation parameters are listed in Table II [36]. The source codes developed for this study are available at https://github.com/TongLu0223/VSL-VP.



(a) The G22 section in Google Map

(b) The G22 section in SUMO

Fig. 5. Simulation scenario.

TABLE II
PARAMETER SETTING OF EXPERIMENTS

Parameter	**Value**
SUMO	
Simulation time per episode (*s*)	6000
Vehicle size (*m*)	4.8×1.8
Simulation step (*s*)	0.1
Length of S-1 (*m*)	3000
Length of S-2 (*m*)	500
Lane-change model	LC2013
Window size *L* (*veh*)	5
Max window length (*m*)	100
V2V communication distance (*m*)	100
Maximum speed limit in S-1 (*m*/*s*)	25
HDV	
Maximum speed (*m*/*s*)	33.33
Desired speed (*m*/*s*)	30.55
Maximum acceleration (m/s^2)	3.5
Minimum acceleration (m/s^2)	-2.8
Car-following model	IDM
CAV	
Maximum speed (*m*/*s*)	33.33
Maximum acceleration (m/s^2)	3.5
Minimum acceleration (m/s^2)	-4
Car-following model	CACC
Model training	
Discount factor	0.95
Learning rate	0.001
Training batch size	256
Memory length	20000
Epsilon decay	0.98

B. Results on Platooning

Initially, the DRL-based vehicle platooning model is trained within an individual window located in S-2. A window size *L*=5 is initialized. In this sample scenario, five vehicles, comprising both CAVs and HDVs, are randomly generated on the three-lane freeway, with a lane blocked ahead due to a traffic incident, creating a bottleneck. During training within a single window, the traffic flow remains stable, with no significant congestion. Therefore, only vehicle platooning is employed for control. When vehicles within the window enter S-2, a leader is designated, and subsequently, DQN determines the optimal time for CAVs to join the platoon. The training utilizes only rewards r_2 and r_4 only. The reward function is defined as

$$r = \omega_3 * r_2 + \omega_4 * r_4 \tag{11}$$

where ω_3=0.9 and ω_4=0.1. A higher weight is assigned to r_2 since r_2 is the most critical optimization objective. The result of 700 episodes of training is shown in Fig. 6. It can be observed that the average reward gradually increases and eventually stabilizes, indicating a clear convergence trend. After successfully converging, the model controls the CAVs within the window to form a stable platoon before reaching the bottleneck area. A video of this scenario can be found at https://youtu.be/I4CYQyf-TmE.

Fig. 7 depicts the changes in the average vehicle speed within the given scenario. Vehicles traverse the bottleneck at around 30 seconds. Due to the need for vehicles in the blocked lane to slow down and change lanes in advance, the average vehicle speed is notably reduced. The findings show that CAVs display smoother speed variations and attain higher speeds through the bottleneck compared to vehicles not utilizing platooning.

Subsequently, curriculum learning is utilized to accelerate the learning process. Curriculum learning is especially advantageous for safety-critical tasks, such as autonomous driving, because beginning with a proficient model can significantly decrease the number of hazardous blind explorations [51]. Specifically, the model's training proceeds in

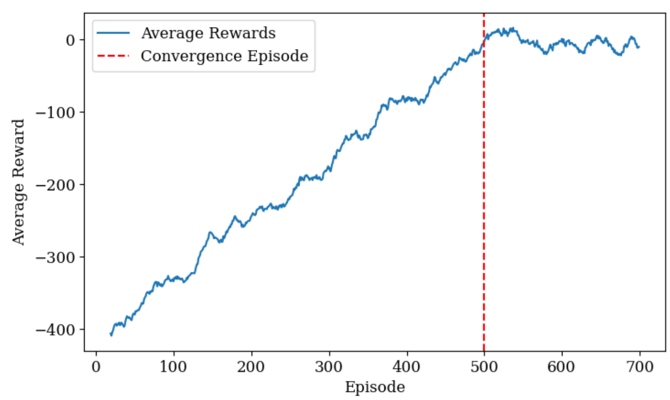


Fig. 6. Reward in training for platooning.

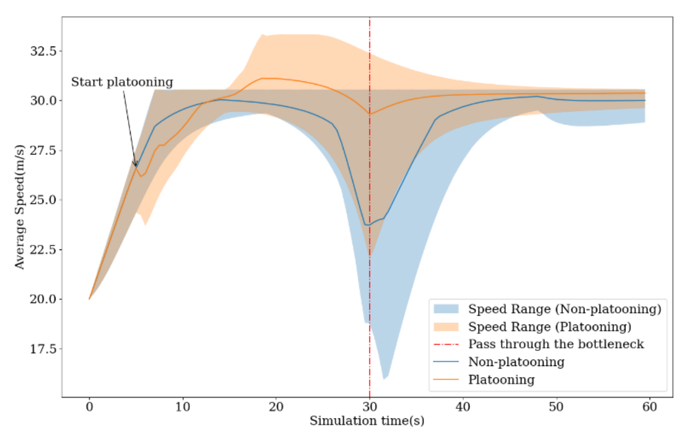


Fig. 7. Average vehicle speed.

a large-scale traffic flow with 100 vehicles entering the road per minute.

To improve the model's performance in a large-scale traffic flow, additional reward constraints are introduced to account for the complexity of traffic flow dynamics. The training process includes various traffic scenarios with different CAV penetration rates. During the initial training phase, the vehicle density is set at 40 vehicles per kilometer. All rewards r_1-r_4 are taken into account during the training process. The agent is trained using a weighted sum of all rewards:

$$r = \omega_5 * r_1 + \omega_6 * r_2 + \omega_7 * r_3 + \omega_8 * r_4 \tag{12}$$

where ω_5=0.68, ω_6=0.1, ω_7=0.2 and ω_8=0.02. The reasoning behind these weight values is as follows. In a large-scale traffic flow, the throughput of the bottleneck, represented by r_1, is the most crucial optimization objective. Hence, it is assigned the highest weight. To prevent congestion and excessive jams before the bottleneck, rewards r_2 and r_3 are also given weights. Reward r_4 encourages CAVs to form platoons as quickly as possible, thus speeding up the model's convergence.

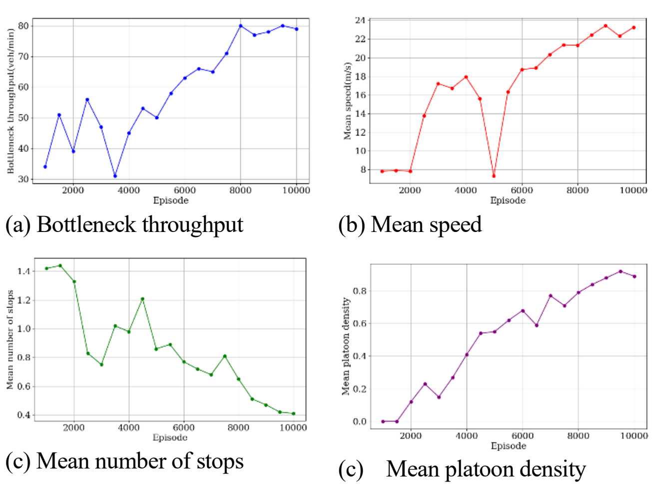
Fig. 8 presents the performance across four key metrics: bottleneck throughput, mean vehicle speed, mean number of stops, and mean platoon density at the bottleneck. It illustrates the convergence of each reward throughout the training process. Specifically, the bottleneck throughput, mean speed, and platoon density within the bottleneck area show a gradual increase as training progresses. This indicates that the model effectively learns to improve traffic throughput and overall system efficiency. In contrast, there is a significant decrease in the mean number of stops for vehicles, suggesting a reduction in traffic oscillations. The results highlight the effectiveness of the proposed method in optimizing traffic performance.

C. Results on Bottleneck Control

To verify the effectiveness of VSL-VP, the following four different strategies are employed for comparative experiments:

1) a baseline entails not utilizing any additional control strategies for CAVs. Instead, it solely relies on the car-following and lane-changing models of CAVs;

2) Only VSL is adopted on road segment S-1 to limit the speed of vehicles;



(a) Bottleneck throughput (b) Mean speed

(c) Mean number of stops (c) Mean platoon density

Fig. 8. Training performance in large-scale traffic flow.

3) VSL-VP without a dedicated lane; and

4) VSL-VP with a dedicated lane.

These are implemented on six CAV penetration rates: 0, 20%, 40%, 60%, 80%, and 100%, respectively. The bottleneck throughputs are shown in Fig. 9. As the penetration rate increases, the throughput of all strategies rises, and VSL-VP demonstrates its efficacy in enhancing throughput. Specifically, at penetration rates of 60%, 80%, and 100%, VSL-VP has a significant improvement in bottleneck throughput compared with baseline and VSL.

Then, the effects of the dedicated lane on the performance of the proposed method are analyzed. At penetration rates of 20%, 40%, and 60%, the introduction of the dedicated lane decreases bottleneck throughput because a dedicated lane restricts the entry of HDVs although enhancing CAV platooning. As the penetration rate gradually increases to over 80%, CAV platoons can take full advantage of the dedicated lane, thereby increasing bottleneck throughput. The impact of the dedicated lane is dependent on the CAV penetration rate, which has also been confirmed by recent studies [37], [38].

The results of different metrics under six penetration rates are shown in Table III. At penetration rates of 20% and 40%, VSL-VP has lower mean speed but higher mean time loss, mean stops per vehicle, mean stop duration, mean jam length, and max jam length compared to both baseline and VSL. At such penetration rates, a large number of HDVs are present on the road. This makes it hard for CAVs to form platoons. As a result, the advantages of CAV platoons are not realized, and the platooning process has a negative impact on the traffic flow, leading to a low mean speed and more congestion. Nevertheless, when the penetration rate surpasses 60%, CAVs can readily form platoons. These CAV platoons can leverage their benefits, swiftly navigating through the bottleneck. Consequently, VSL-VP achieves a higher mean speed and a lower mean time loss in comparison to both the baseline and VSL approaches. It effectively mitigates congestion, bringing the values of mean stops, mean stopping duration, mean jam length, and related metrics down to zero. Videos showcasing these experiments are available at https://www.youtube.com/playlist?list=PLzq7Vw-HmU2fLOuS4oGOkpTEyLHkFjV8CF.

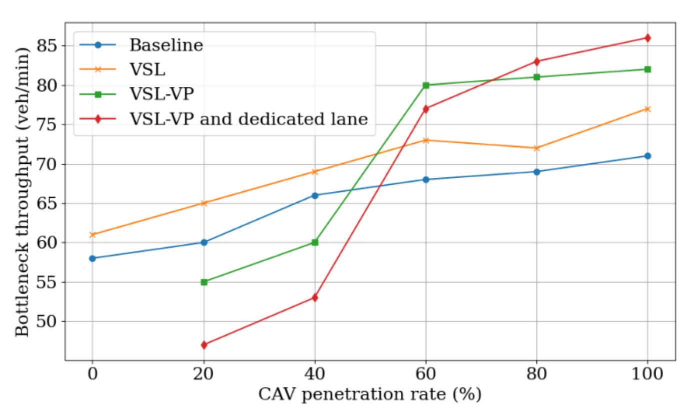


Fig. 9. Performance at different CAV penetration rates.

TABLE III
COMPARISON RESULTS

Strategy	Performance evaluations						
	Throughput (*veh*/*min*)	Mean speed (*m*/*s*)	Mean time loss (*s*)	Mean stops per vehicle	Mean stopping duration (*s*)	Mean jam length (*m*)	Max jam length (*m*)
(a) 0% CAV penetration rate							
Baseline	58	6.97	31.36	1.17	65.52	255.14	347.91
VSL	**61**	**12.17**	**24.76**	**0.89**	**49.31**	**231.54**	**308.04**
(b) 20% CAV penetration rate							
Baseline	60	6.62	31.98	1.97	60.99	226.51	287.24
VSL	**65**	**14.52**	**21.39**	**0.86**	**39.09**	**151.66**	**207.88**
VSL-VP	55	11.01	28.47	2.35	72.49	227.06	320.57
VSL-VP with a dedicated lane	47	10.38	29.83	2.56	74.61	317.87	404.61
(c) 40% CAV penetration rate							
Baseline	66	6.11	33.27	2.24	38.01	227.64	309.81
VSL	**69**	**15.09**	**19.04**	**0.91**	**20.42**	**150.99**	**220.92**
VSL-VP	60	13.45	25.01	2.49	39.11	298.54	378.45
VSL-VP with a dedicated lane	53	10.32	29.65	2.72	43.98	301.99	398.21
(d) 60% CAV penetration rate							
Baseline	68	6.46	32.3	3.03	17.6	154.66	246.51
VSL	73	13.26	25.71	0.82	14.17	106.75	208.23
VSL-VP	**80**	**30.4**	**0.22**	**0.00**	**0.00**	**0.00**	**0.00**
VSL-VP with a dedicated lane	77	24.99	4.78	0.45	6.35	12.48	33.27
(e) 80% CAV penetration rate							
Baseline	69	7.43	27.1	2.88	12.67	107.95	217.86
VSL	72	12.13	26.2	0.72	9.69	82.73	181.67
VSL-VP	81	30.43	0.1	0.00	0.00	0.00	0.00
VSL-VP with a dedicated lane	**83**	**32.44**	**-0.31**	**0.00**	**0.00**	**0.00**	**0.00**
(f) 100% CAV penetration rate							
Baseline	71	15.16	19.26	0.83	8.25	49.73	94.61
VSL	77	23.15	4.1	0.00	0.00	0.00	0.00
VSL-VP	82	30.55	0.00	0.00	0.00	0.00	0.00
VSL-VP with a dedicated lane	**86**	**33.33**	**-0.79**	**0.00**	**0.00**	**0.00**	**0.00**

The outcomes of the experiments illustrate the efficacy of VSL-VP in easing freeway bottlenecks. This is accomplished by improving the throughput at bottlenecks, minimizing traffic fluctuations, and decreasing congestion, especially in situations where the CAV penetration rate is above 60%.

D. Window Size Test

In this study, faced with the intricacies of large-scale traffic flow, sliding-windows are employed to efficiently manage the traffic. Specifically, a constant window size of L=5 is maintained, thereby setting a maximum limit of 5 vehicles for platoon size. The selection of platoon size is vital in shaping the behavior of vehicle platoons and their influence on traffic conditions [39]. Our experiments investigate various window sizes across six scenarios with differing CAV penetration rates and vehicle densities. The aim is to evaluate how changes in window size affect the performance of VSL-VP under varying conditions. The results are presented in Table IV.

The effect of window size on bottleneck throughput differs across scenarios. In scenario 1, an initial increase in window size enhances bottleneck throughput. However, further enlargement of the window size results in a decrease. In scenario 6, as the window size grows, the bottleneck throughput continues to rise. This can be explained as follows: In scenario 1, when the window size initially increases, the number of CAVs within the window grows, promoting the formation of longer and more numerous platoons. As the window size expands further, longer CAV platoons may restrict other HDVs within the window, limiting their lane-changing maneuvers. Moreover, with an excessive number of CAVs, platoons may not fully form by the time they reach the bottleneck, leading to a decrease in bottleneck throughput. In scenario 6, since the CAV penetration rate is 100%, there are no HDVs on the road. CAVs can form platoons more rapidly and securely. Thus, as the window size increases, CAVs can create longer platoons, enhance traffic efficiency, and thereby boost bottleneck throughput. However, considering other aspects, such as the model's time complexity, a larger window size is not always beneficial.

Therefore, in each unique scenario, the optimal window size varies depending on factors like CAV penetration rate and vehicle density. The most effective approach involves dynamically adjusting the window size in real-time, allowing for adaptive responses to different traffic flows. This dynamic method of determining window size is a key area of focus for our future work,

TABLE IV
IMPACT OF DIFFERENT WINDOW SIZE

Scenario	CAV penetration	Vehicle density (*veh*/*km*)	Window size	Throughput (*veh*/*min*)
1	60%	30	5	67
			10	68
			15	66
			20	63
2	60%	40	5	80
			10	80
			15	79
			20	76
3	80%	30	5	68
			10	69
			15	70
			20	68
4	80%	40	5	81
			10	82
			15	79
			20	78
5	100%	30	5	70
			10	71
			15	72
			20	74
6	100%	40	5	82
			10	83
			15	83
			20	84

as it has the potential to improve the performance of our vehicle platooning model across various traffic situations.

V. CONCLUSION

In this research, we introduce VSL-VP, an innovative framework that integrates VSL and vehicle platooning to address freeway bottlenecks. The approach divides the road into two segments preceding the bottleneck. In the first segment, VSL is used to decrease vehicle density and increase the following distance between cars in the second segment by restricting vehicle speeds. This creates favorable conditions for platooning. In the second segment, sliding-windows are initially used to divide the large-scale traffic flow into separate windows. Within each window, DQN is employed to identify the best time for CAVs to join a platoon, allowing them to form a stable platoon and navigate through the bottleneck area, thereby improving bottleneck throughput and reducing traffic congestion. We evaluate various traffic performance metrics of VSL-VP under different CAV penetration rate scenarios. The results show that VSL-VP effectively enhances bottleneck throughput and alleviates traffic congestion at CAV penetration rates exceeding 60%. We also analyze the effectiveness of dedicated lanes for CAVs and the influence of different window sizes on bottleneck throughput.

In future research, we plan to propose further optimizations to enhance the efficiency of VSL-VP. Specifically, the window size, which determines the maximum platoon size, affects bottleneck throughput. To refine and optimize this component of the framework, we will explore new methods, such as Multi-Agent Reinforcement Learning (MARL) [52], will be tried. Integrating MARL introduces a more advanced decision-making structure, facilitating a cooperative approach among multiple intelligent agents. By utilizing MARL, we aim to dynamically determine the optimal window size for vehicle platoons in real-time, improving adaptability to changing traffic conditions and boosting the overall efficiency and effectiveness of the VSL-VP framework in freeway bottleneck scenarios.

REFERENCES

[1] W. S. Vickrey, “Congestion theory and transport investment,” *Am. Econ. Rev.*, vol. 59, no. 2, pp. 251-260, May 1969.
[2] T. T. Nguyen, S. C. Calvert, H. L. Vu, and H. Van-Lint, “An automated detection framework for multiple highway bottleneck activations,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 5678-5692, Jun. 2022.
[3] F. Tajdari, C. Roncoli, and M. Papageorgiou, “Feedback-based ramp metering and lane-changing control with connected and automated vehicles,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 939-951, Feb. 2022.
[4] P. Mao, X. Ji, X. Qu, L. Li, and B. Ran, “A variable speed limit control based on variable cell transmission model in the connecting traffic environment,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 17632-17643, Oct. 2022.
[5] Y. Zhang and P. A. loannou, “Combined variable speed limit and lane change control for highway traffic,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp.1812-1823, Jul. 2017.
[6] Z. Li, P. Liu, C. Xu, H. Duan, and W. Wang, “Reinforcement learning-based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 3204-3217, Nov. 2017.
[7] C. Wang, Y. Xu, J. Zhang, and B. Ran, “Integrated traffic control for freeway recurrent bottleneck based on deep reinforcement learning,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 15522-15535, Sept. 2022.
[8] P. Bansal and K. M. Kockelman, “Forecasting americans' long-term adoption of connected and autonomous vehicle technologies,” *Transp. Res. A, Policy Pract.*, vol. 95, pp. 49-63, 2017.
[9] H. Shi *et al.*, “Connected automated vehicle cooperative control with a deep reinforcement learning approach in a mixed traffic environment,” *Transp. Res. C, Emerg. Technol.*, vol. 133, 2021, Art. no. 103421.
[10] X. Qu *et al.*, “Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: a reinforcement learning based approach,” *Appl. Energy*, vol. 257, 2020, Art. no. 114030.
[11] J. Lin, C. Yu, L. Wang, G. Liu, J. Wang, and W. Ma, “Optimization of lane-changing advisory in mixed traffic of connected vehicles and human-driven vehicles at expressway bottlenecks,” *IEEE Trans. Intell. Veh.*, early access, Oct. 2023, doi: 10.1109/TIV.2023.3326565.
[12] P. Kasture and H. Nishimura, “Platoon definitions and analysis of correlation between number of platoons and jamming in traffic system,” *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 1, pp.319-328, Jan. 2021.
[13] Z. Wang, Y. Bian, S. E. Shladover, G. Wu, S. E. Li, and M. J. Barth, “A survey on cooperative longitudinal motion control of multiple connected and automated vehicles,” *IEEE Trans. Intell. Transp. Syst. Mag.*, vol. 12, no. 1, pp. 4-24, Dec. 2020.
[14] D. Cao *et al.*, “A platoon regulation algorithm to improve the traffic performance of highway work zones,” *Comput-Aided Civ. Inf.*, vol. 36, no. 7, pp. 941-956, Apr. 2021.
[15] C. Earnhardt *et al.*, “Cooperative exchange-based platooning using predicted fuel-optimal operation of heavy-duty vehicles,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 17312-17324, Oct. 2022.
[16] M. Čičić, X. Xiong, L. Jin, and K. H. Johansson, “Coordinating vehicle platoons for highway bottleneck decongestion and throughput improvement,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8959-8971, Jul. 2022.
[17] C. Chen, J. Wang, and Q. Xu, “Mixed platoon control of automated and human-driven vehicles at a signalized intersection: dynamical analysis

and optimal control," *Transp. Res. C, Emerg. Technol.*, vol. 127, 2021, Art. no. 103138.
[18] K. Li, J. Wang, and Y. Zheng, "Cooperative formation of autonomous vehicles in mixed traffic flow: Beyond platooning," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 15951-15966, Sept. 2022.
[19] W. Zhao *et al.*, "A platoon based cooperative eco-driving model for mixed automated and human-driven vehicles at a signalized intersection," *Transp. Res. C, Emerg. Technol.*, vol. 95, pp. 802-821, 2018.
[20] S. W. Smith *et al*., "Improving urban traffic throughput with vehicle platooning: theory and experiments," *IEEE Access*, vol. 8, pp. 141208-141223, July. 2020.
[21] Q. Deng, "A general simulation framework for modeling and analysis of heavy-duty vehicle platooning," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3252-3262, Nov. 2016.
[22] Q. Li, Z. Chen, and X. Li, "A review of connected and automated vehicle Platoon merging and splitting operations," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 22790-22806, Dec. 2022.
[23] B. R. Kiran *et al*., "Deep reinforcement learning for autonomous driving: a survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 4909-4926, Jun. 2022.
[24] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26-28, Nov. 2017.
[25] O. Vinyals *el al.*, "Grandmaster level in StarCraft II using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350-354, Oct. 2019.
[26] M. Li, Z. Cao, and Z. Li, "A reinforcement learning-based vehicle platoon control strategy for reducing energy consumption in traffic oscillations," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5309-5322, Dec. 2021.
[27] D. Li *et al.*, "COOR-PLT: A hierarchical control model for coordinating adaptive platoons of connected and autonomous vehicles at signal-free intersections based on deep reinforcement learning," *Transp. Res. C, Emerg. Technol.*, vol. 146, 2023, Art. no. 103933.
[28] H. Shi *et al.*, "A deep reinforcement learning based distributed control strategy for connected automated vehicles in mixed traffic platoon," *Transp. Res. C, Emerg. Technol.*, vol. 148, 2023, Art. no. 104019.
[29] O. Pauca, A. Maxim, and C. F. Caruntu, "Cooperative platoons merging for obstacle avoidance on highways," in *Proc. Int. Conf. Syst. Theory, Contr. Comp. (ICSTCC)*, Lasi, Romania, 2021, pp. 25-30.
[30] S. B. Prathiba, G. Raja, K. Dev, N. Kumar, and M. Guizani, "A hybrid deep reinforcement learning for autonomous vehicles smart-platooning," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 13340-13350, Dec. 2021.
[31] A. Irshayyid and J. Chen, "Comparative study of cooperative platoon merging control based on reinforcement learning," *Sensors*, vol. 23, no. 2, 2023, Art. no. 990.
[32] Y. Bengio *et al.*, "Curriculum learning," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, Montreal, Canada, pp. 41-48.
[33] J. Zhou and F. Zhu, "Modeling the fundamental diagram of mixed human-driven and connected automated vehicles," *Transp. Res. C, Emerg. Technol.*, vol. 115, 2020, Art. no. 102614.
[34] Y. Jiang *et al.*, "Platoon intensity of connected automated vehicles: Definition, formulas, examples, and applications," *J. Adv. Transp.*, vol. 2023, 2023.
[35] S. S. McDonald and C. Rodier, "Envisioning automated vehicles within the built environment: 2020, 2035, and 2050," in *Proc. Road Veh. Autom. 2*, Springer, Cham, 2015, pp. 225-233.
[36] X. Huang, P. Lin, M. Pei, B. Ran, and M. Tan, "Reservation-based cooperative ecodriving model for mixed autonomous and manual vehicles at intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 9, pp. 9501-9517, Sept. 2023.
[37] L. Ye and T. Yamamoto, "Impact of dedicated lanes for connected and autonomous vehicle on traffic flow throughput," *Physica A*, vol. 512, pp. 588-597, Dec. 2018.
[38] D. Chen *et al.*, "Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles," *Transp. Res. B, Meth.*, vol. 100, pp. 196-221, 2017.
[39] J. Zhou and F. Zhu, "Analytical analysis of the effect of maximum platoon size of connected and automated vehicles," *Transp. Rec. C, Emerg. Technol.*, vol. 122, 2021, Art. no. 102882.
[40] R. S. Sutton and A. G. Barto, *Introduction to reinforcement learning*. Cambridge: MIT press, 1998, pp. 223-260.
[41] D. Ma, B. Zhou, X. Song, and H. Dai, "A deep reinforcement learning approach to traffic signal control with temporal traffic pattern mining," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 11798-11800, Aug. 2022.
[42] X. Liang, X. Du, G. Wang, and Z. Han, "A deep reinforcement learning network for traffic light cycle control," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1243-1253, Feb. 2019.
[43] B. O'Donoghue, I. Osband, R. Munos, and V. Mnih, "The uncertainty bellman equation and exploration," in *Proc. Intern. Conf. Mach. Learn.*, Stockholm, Sweden, 2018, pp. 3836-3845.
[44] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529-533, Feb. 2015.
[45] M. Segata *et al.*, "Plexe: A platooning extension for veins," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Paderborn, GER, 2014, pp. 53-60.
[46] P. A. Lopez *et al.*, "Microscopic traffic simulation using SUMO," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Maui, HI, USA, 2018, pp. 2575-2582.
[47] M. A. S. Kamal, T. Hayakawa, and J. -i. Imura, "Development and evaluation of an adaptive traffic signal control scheme under a mixed-automated traffic scenario," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 2, pp. 590-602, Feb. 2020.
[48] G. J. L. Naus *et al.*, "String-stable CACC design and experimental validation: A frequency-domain approach," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4268-4279, Nov. 2010.
[49] V. Milanés *et al.*, "Cooperative adaptive cruise control in real traffic situations," *IEEE Trans. Veh. Technol.*, vol. 15, no. 1, pp. 296-305, Feb. 2014.
[50] J. Erdmann, "SUMO's lane-changing model," in *Proc. Model. Mob. Op. Dat., 2nd SUMO Conf.*, Berlin, Germany, 2015, pp. 105-123.
[51] D. Chen *et al.*, "Deep multi-agent reinforcement learning for highway on-ramp merging in mixed traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 11, pp. 11623-11638, Nov. 2023.
[52] T. Chu, J. Wang, L. Codecà and Z. Li, "Multi-Agent Deep Reinforcement Learning for Large-Scale Traffic Signal Control," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1086-1095, Mar. 2020.

Liang Qi received his Ph.D. degree in Computer Software and Theory from Tongji University, Shanghai, China in 2017. From 2015 to 2017, he was a Visiting Doctoral Student at New Jersey Institute of Technology. He is currently an Associate Professor in the Shandong University of Science and Technology. He has published 190+ papers in journals and conference proceedings (45 in IEEE Transactions/Journal). He received the Best Student Paper Award-Finalist in the 15th IEEE International Conference on Networking, Sensing and Control (ICNSC'2018) and the Best Paper Award-Finalist in the 3rd IEEE International Conference on Automation in Manufacturing, Transportation and Logistics (iCaMaL2023). His current research interests include Petri nets, optimization, machine learning, and intelligent transportation systems. He is currently serving as an Associate Editor for IEEE Transactions on Intelligent Transportation Systems.

Lu Liang received her B.S. degrees in Intelligence Science and Technology from Shandong University of Science and Technology, Qingdao, China, in 2024, where is currently pursuing the M.S. degree. Her current research interests include reinforcement learning, connected and autonomous vehicles, and intelligent transportation systems.

Wenjing Luan received her Ph.D. degree in Computer Software and Theory from Tongji University, Shanghai, China in 2018. She is currently with Shandong University of Science and Technology, Qingdao, China. From May to July, 2017, she was a visiting student in the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, USA. Her current research interests include recommender systems and intelligent transportation systems.

Tong Lu received his B.S. degrees in Intelligence Science and Technology from Shandong University of Science and Technology, Qingdao, China, in 2022, where is currently pursuing the M.S. degree. His current research interests include reinforcement learning, connected and autonomous vehicles, and intelligent transportation systems.

Xiwang Guo received his Ph.D. degree in System Engineering from Northeastern University, Shenyang, China, in 2015. He is currently an Associate Professor in the College of Computer and Communication Engineering at Liaoning Petrochemical University. From 2016 to 2018, he was a visiting scholar of Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, USA. He has published 150+ papers in journals and conference proceedings, including the IEEE TITS, IEEE/CAA JAS, IEEE TSMCS, IEEE TCSS, IEEE TASE, IEEE TCYB, IEEE IoT-J, IEEE TAI, and IEEE RA-L. His current research interests include optimization algorithms, reinforcement learning, manufacturing, and intelligent transportation.

Qurra Tul Ann Talukder received her B.S. degree in Software Engineering from Shandong University of Science and Technology, Qingdao, China, in 2022, where she is currently pursuing the M.S. degree. Her current research interests include optimization algorithms, language models, intelligent guidance systems, and intelligent transportation systems.