Integrated Control Strategies for Freeway Bottlenecks with Vehicle Platooning

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

***Abstract*—Freeway bottlenecks caused by traffic incidents contribute significantly to large-scale traffic congestion. Traditional strategies, including variable speed limit (VSL) and ramp metering, are commonly used for freeway traffic congestion management. Recently, vehicle platooning has become a promising way to alleviate traffic bottlenecks. This work proposes a novel framework that combines VSL and vehicle platooning for freeway bottleneck, referred to as VSL-VP, in mixed traffic of connected and autonomous vehicles (CAVs) and human-driven vehicles (HDVs). First, the upstream road of a bottleneck is divided into two segments, called the former and the latter ones. VSL limits vehicle speed at the former segment, thereby reducing inflow traffic to the latter one. Then, deep reinforcement learning is employed for CAV platooning at the latter segment, where low traffic flow density and large car-following distance create conditions for smooth lane change and platoon formulation of CAVs. Simulation results demonstrate that VSL-VP significantly enhances the bottleneck throughput and reduces traffic congestion at elevated levels of CAV penetration rates.**

***Key Words*—Freeway bottleneck, variable speed limit, vehicle platooning, deep reinforcement learning.**

I. Introduction

TRAFFIC bottleneck refers to a localized constriction of traffic flow occurring when traffic demand exceeds the available capacity due to poor road design, improper traffic light time, or traffic incidents, which directly cause traffic congestion [1]. These bottlenecks are predominantly responsible for the majority of large-scale congestion on freeways [2]. Traditional traffic control methods aimed to alleviate freeway bottlenecks include ramp metering (RM) [3] and variable speed limit (VSL) [4, 5, 6]. RM prevents vehicles from entering congested areas, while VSL reduces the vehicle speed at the upstream road of the bottleneck to alleviate traffic demand. Despite the commendable performance exhibited by both methods in preventing congestion caused by bottlenecks, they possess inherent limitations [7]. RM may disrupt the nearby road traffic when the on-ramp road is temporarily closed, and VSL may adversely affect the upstream traffic of the bottleneck [8].

Considering the challenges of traditional strategies in addressing freeway bottlenecks, it is pertinent to acknowledge the significant transformation underway in traffic styles and management [9]. With the rapid advancement of vehicular automation and communication technology, connected and automated vehicles (CAVs) are progressively occupying a significant share of the automotive market [10]. In the foreseeable future, CAVs and human-driven vehicles (HDVs) are anticipated to coexist on the same roads, thereby transforming the traffic environment from predominantly HDVs to a mixture of CAVs and HDVs [11], [12]. Subsequently, the exploration of new traffic bottleneck control strategies in a mixed traffic environment becomes imperative. [13].

One of the most promising strategies in intelligent transportation systems is vehicle platooning [14]. A vehicle platoon refers to a group of CAVs driving closely together in the same lane of a road segment, maintaining a uniform, reduced car-following distance and time headway while traveling at higher speeds [15]. Consequently, vehicle platoons hold significant promise for enhancing roadway capacity. Moreover, due to minimized aerodynamic drag, the fuel consumption can be reduced [16, 17]. Numerous studies have demonstrated the advantages of vehicle platoons in mitigating traffic bottlenecks [18, 19]. Vehicle platooning is a process in which a group of adjacent CAVs form a stable platoon by joining and merging operations [20]. Zhao et al. propose a platoon formation method based on model predictive control (MPC) to optimize platoons passage through intersections on a green phase with minimal fuel consumption [21]. Smith et al. present a MPC-based approach for vehicle platooning in urban traffic settings, demonstrating the potential in enhancing urban traffic throughput [22].

Since CAVs are typically distributed randomly within mixed traffic flow, several maneuvers, including joining, leaving, merging, and splitting, need to be executed to organize neighboring CAVs into a platoon [23]. These maneuvers might have adverse impacts, such as triggering undesired congestion [24]. However, the existing studies often overlook vital decision-making details regarding these maneuvers [25]. For instance, few studies address the determination of the optimal joining time, which signify the moment when a CAV should perform

Manuscript received April 11, 2025; revised April 13, 2025; accepted April 18, 2025. This article was recommended for publication by Associate Editor Xiwang Guo 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, $tongluu2023@163.com, upomagourab@hotmail.com$).

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.

lane change and acceleration operations to ensure safe and efficient integration into the platoon. Moreover, as the platoon length increases, the dimension of the solution space expands, resulting in a significant increase in computational complexity. Recent advancements in deep reinforcement learning (DRL) have shown great promise in tackling complex control problems characterized by uncertain and high-dimensional state and action spaces [26, 27, 28]. Some studies have begun to apply DRL to control platoons. For instance, Li et al. propose a multi-agent reinforcement learning algorithm to control vehicle platoons and improve energy efficiency during traffic oscillations [29]. Another study [30] develops a DRL-based hierarchical model. It integrates platooning and coordination to enhance CAV control, consequently reducing travel time and fuel consumption at signal-free intersections. Shi et al. [31] propose a distributed CAV longitudinal control strategy based on DRL for mixed traffic comprising of CAVs and HDVs. A novel approach is constructed to significantly reduce traffic oscillation. However, to the best of our knowledge, no studies have employed DRL to determine the optimal joining time of CAVs in mixed traffic environments.

In bottleneck areas, when traffic density is high, it becomes impractical for vehicles to safely perform lane-changing maneuvers for platooning. Existing studies mainly focus on platooning under low-density conditions [32, 33, 34]. Therefore, this work implements VSL to facilitate vehicle platooning. More specifically, VSL restricts vehicle speeds to establish a roadway segment characterized by reduced traffic flow density and increased car-following distance. This ensures a smooth lane change of CAVs for the formation of platoons in a safe and efficient manner. To the best of our knowledge, this is the first work to integrate VSL and vehicle platooning to alleviate bottlenecks in large-scale mixed traffic flows. The primary contributions are summarized as follows:

1) It proposes a novel framework that combines VSL and vehicle platooning to address bottlenecks. Firstly, VSL sets speed limits, thereby reducing the inflow to the bottleneck. The low traffic flow density and extended car-following distance form favorable conditions for CAVs to execute safe lane-changing maneuvers for platooning.
2) To address the high-dimensional solution space of CAV platooning, this work employs sliding windows to partition large-scale traffic flows into independent windows. Within each window, DRL is utilized to determine the optimal joining time for CAVs in a mixed traffic environment. Curriculum learning is applied for training the model [35]. Initially, the model is trained by learning from experiences gathered from a simple scenario with fewer vehicles. Subsequently, the model is calibrated using large-scale traffic data across diverse conditions, encompassing varying levels of CAV penetration rates and vehicle densities.

In addition, extensive experiments are conducted to demonstrate the efficacy of the proposed framework in alleviating bottlenecks and reducing congestion. The impact of the varying window sizes on bottleneck throughput is explored, and the scalability of the proposed framework is discussed.

The remainder of this paper is organized as follows. Section II provides a problem description. Section III elaborates on the methodology of combining VSL and vehicle platooning for freeway bottlenecks. Section IV presents the conducted experiments and analysis of the results. Finally, Section IV-E concludes the paper and discusses future directions for research.

II. Problem Description And Preliminaries

A. Problem Description

In this work, a typical three-lane freeway is considered, which is the most common configuration, with a bottleneck caused by a blocked lane, as depicted in Fig. 1. The scenario involves mixed traffic flow comprising of both connected and automated vehicles (CAVs) equipped with Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication devices and human-driven vehicles (HDVs) that lacks communication devices. Additional auxiliary components include roadside units (RSUs) and VSL controllers. RSU equips devices (e.g., traffic detectors, cameras, routers, etc.) to collect vehicle data and facilitates the transmission of vehicle states via V2I communication. VSL controller dynamically controls the speed limits by altering the speed-limit sign on a freeway segment. Connected Vehicle Center (CVC) oversees all connected devices through communication. In this context, a CAV platoon is defined as two or more adjacent CAVs traveling with a uniform, small car-following distance and time headway but at high speeds within the same lane.

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 objective of this work is to mitigate the freeway bottleneck by combining VSL and vehicle platooning. However, there are two critical problems to be addressed:

1) How to reduce the vehicle density for platooning. Bottlenecks typically result in traffic congestion. When the vehicle density is too high, lane-changing maneuvers become impossible at the bottleneck due to inadequate car-following distance, hindering the formation of CAV platoons..
2) How to manage individual CAVs to form platoons in a mixed traffic environment. This involves selecting a leader CAV and determining the optimal time for other CAVs to join a platoon..

B. Basic Definitions and Assumptions

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

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

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

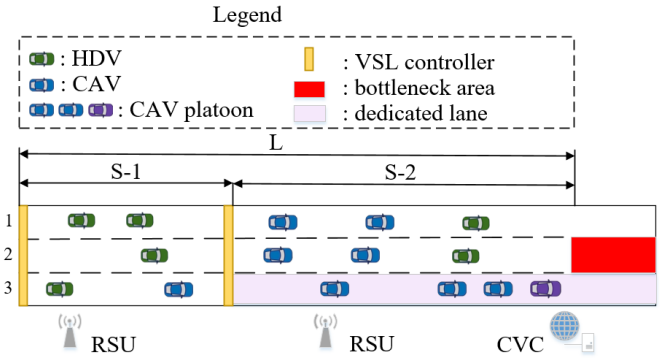


Fig. 1. A bottleneck in a three-lane freeway with mixed traffic.

Given a penetration rate, the distribution of CAVs may vary across different spatial areas. Platoon intensity is used to measure the distribution of CAVs[36] and provide the following mathematical definition.

Definition 2 (Platoon intensity)[37]. A platoon intensity, denoted as Il, 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

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

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

Additionally, the following assumptions are made:

1) Within the communication range, the state of a CAV can be acquired by RSUs and transferred to CVC. No communication delay or detection error is considered in this scenario.
2) Each CAV receives and strictly adheres to the corresponding commands from the CVC, including speed and lane-changing maneuvers.
3) To ensure the safety and efficiency of the approach, HDVs are prohibited from entering the CAV dedicated lane, while CAVs are permitted to enter any lane on S-2 to form a platoon. It is worth noting that some existing studies utilized a dedicated lane for CAVs [38, 39, 40, 41]. This serves to segregate CAVs from the heterogeneous traffic and mitigate the negative effects of random driving behaviors exhibited by HDVs on CAVs in mixed traffic flow.

III. Methodology

This section introduces a framework that combines VSL and vehicle platooning (VSL-VP) to alleviate the freeway bottlenecks in a mixed traffic environment consisting of CAVs and HDVs. Firstly, a framework is outlined, followed by the technical details where a RL-based method for platooning is proposed. Some notations that will be used in the following content are defined in Table I.

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.

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 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:

$$D_l = (\omega_1 * p_l + \omega_2 * I_l) * b_l$$

where 1 and 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 [42]. 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). 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.

Within each window, the CAV closest to the bottleneck is typically selected as a leader. The leader moves to the dedicated lane. Then, DRL is used to determine the optimal time for CAVs, excluding the leader, to join a platoon. Windows containing only one CAV do not require a leader. It verifies if there is another platoon within the communication range of this CAV. If a platoon is detected, the CAV joins it; otherwise, it remains independent. Further details of DRL are presented in the following subsections.

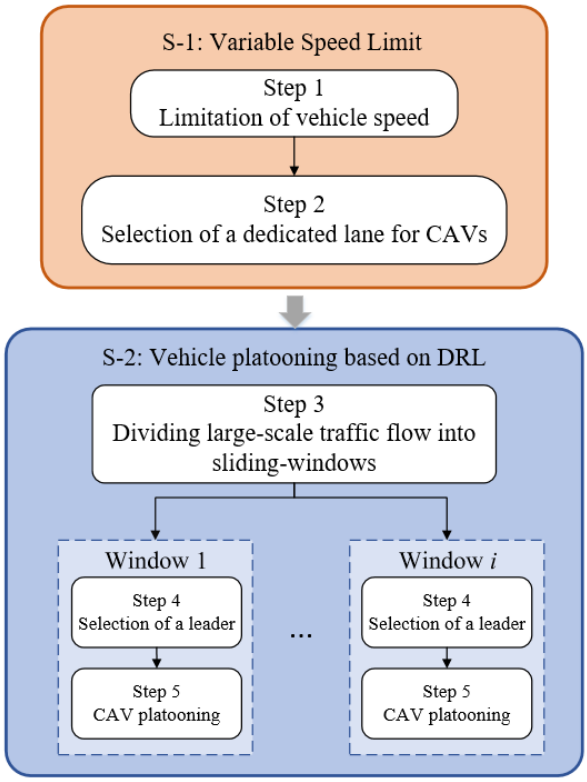


Fig. 2. Hierarchical framework of the VSL-VP.

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 [43]. 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 [44]. This procedure is regarded as a Markov Decision Process (MDP), which can be described as a four-tuple $\{S, A, P, R\}$, where:

1) S denotes a traffic state space, where $s \in S$ is a specific state;
2) A denotes an action space, where $a \in A$ is a specific action;
3) $P = SAS$ denotes the transmission probability among states; and
4) R denotes a reward space, where $r \in R$ is a specific reward.

In order to gradually improve the environment to the ideal state, RL agents select actions according to an optimal policy function denoted by π. The goal of π is to maximize the cumulative expected rewards originating from the initial state. If agents are aware of the optimal cumulative reward at a specific state, they can select actions that yield the highest reward [45]. The cumulative reward can be recursively determined by using

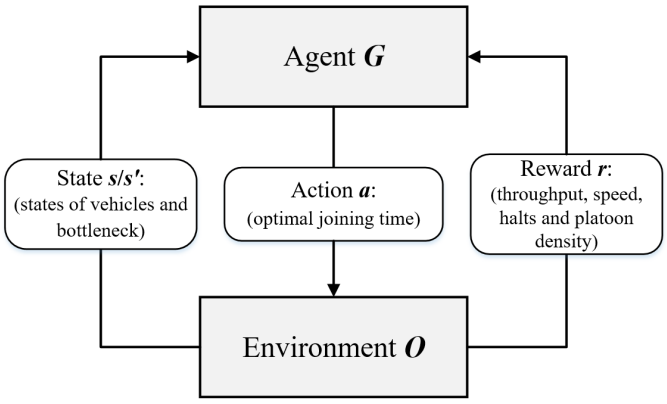


Fig. 3. RL for vehicle platooning.

the Bellman equation [46]. For instance, the agent at a certain state s takes an action a to reach the next state s' and gets a reward r, which is denoted by a tuple (s, a, r, s'), and then the cumulative reward denoted by $Q\pi(s, a)$ under policy π can be calculated by

$$Q^{\pi}(s,a) = E_{s'} r + \gamma [max]_{(}a')[Q^{\pi}](s', a')|s, a$$

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) [47] 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.

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:

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

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 time step t, DRL receives a state s_i and takes an action a_i, which is a set of actions of all CAVs within window i. The size of the action ai depends on the number of CAVs. A discrete action space $\{a_{i,j}\}$ is used, where $a_{ij} = 1$ indicates CAV j within window i takes an action to join the platoon. It changes the lane, accelerates to catch the leading CAV, and

joins the platoon, and $a_{ij} = 0$ indicates that j maintains the current state unchanged.

Reward: An agent receives rewards that encourage future positive behaviors with the goal of enhancing traffic throughput of the bottleneck. Four rewards are utilized to address specific aspects of traffic flow dynamics within the bottleneck.

1) The first reward, denoted as r_1, is associated with traffic throughput. It is only granted at the end of the simulation, which is defined as:

$$r_1 = F_{in} - F_{out}$$

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

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

$$r_2 = v$$

where represents the average vehicle speed.

3) The third reward, r_3, corresponds to the number of stops made by vehicles in the bottleneck area. To discourage excessively long traffic jams before a bottleneck, a penalty term is designed:

$$r_3 = -n_s$$

(8) where ns is the total number of stops made by vehicles. In the simulation, a vehicle with a speed lower than 3 m/s is counted as a stop.

4) The fourth reward, r_4, is associated with platoon density, which refers to the proportion of CAVs participating in platooning to the total CAVs in each window. A small positive value of r_4 is assigned to windows with high platoon density at each time step to encourage the achievement of platooning earlier:

$$r_4 = \frac{n_{i,p}}{n_{i,c}}$$

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 framework for DRL with a DQN is depicted in Fig. 4, encompassing two distinct levels. The upper level is responsible for decision making, including neural network training and action processing. Specifically, the training consists of three key components [47]. A Q-network with the current parameters plays a pivotal role in determining the policy governing the agent's actions. A target Q-network with the previous parameters is employed to generate Q values crucial for the loss function during the training process. A replay memory is served as a repository for storing and recalling training samples. During the training process, the data are randomly extracted from the replay memory. Then, the parameters are updated multiple times per time step and are periodically copied to ̂after every iterations. In the Q-network, $\hat{Q}(s, a;)$ represents the output estimating the value obtained by the agent taking action a under state s. Simultaneously,

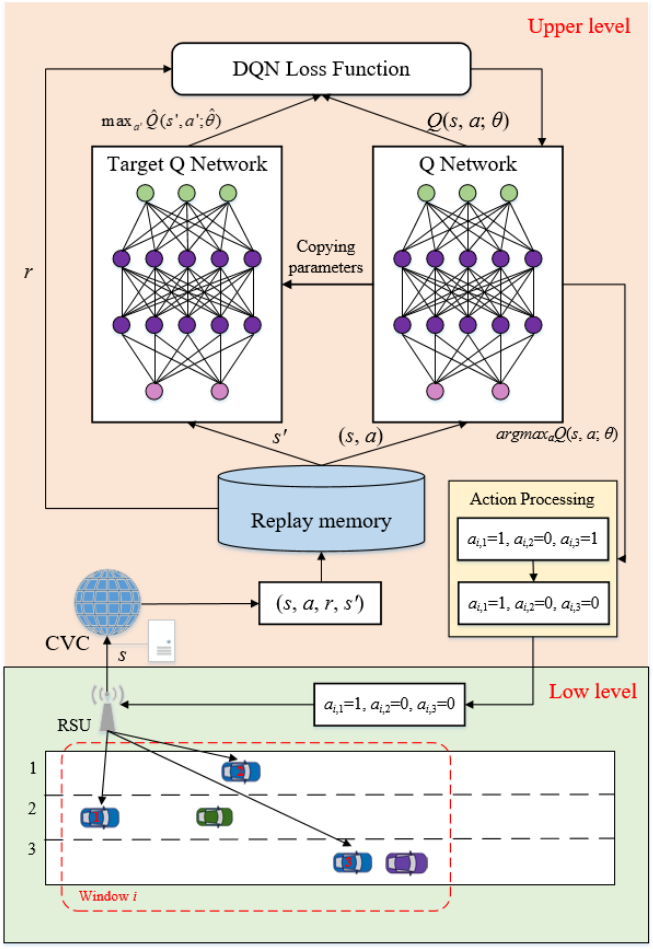


Fig. 4. Framework for DRL with a DQN.

represents the output of the target network. At each iteration , the parameter is updated to minimize the loss function given by:

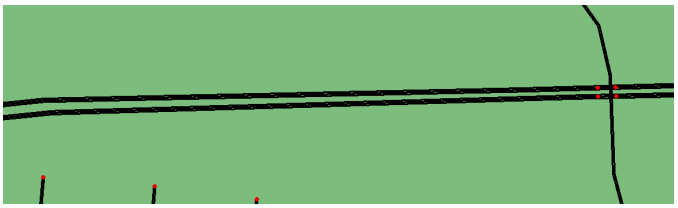
$$L(\theta) = E[(r + \gamma max_{a'} \hat{Q}(s', a'; \hat{\theta}) - Q(s, a; \theta))^2]$$

When the training is completed, the Q values output by the Q network need to be further processed to convert into actions. This includes filtering out illegal actions. For example, once a CAV successfully joins a platoon, it is restricted from executing additional joining maneuvers. The low level is responsible for decision execution. Following the actions determined by the upper level, RSUs relay these actions to each CAV. Based on the car-following model and lane-changing model of CAVs, these actions are translated into speed and acceleration commands, which the CAVs then execute.

IV. Experiments

This section presents the experiments on platooning and bottleneck control. It includes the training and testing phases of the platooning model, as well as the evaluation of VSL-VP in bottleneck control under varying CAV penetration rates. Additionally, the impact of the window size on the performance of VSL-VP is assessed.

(a) The G22 section in Google Map.



(b) The G22 section in SUMO.

Fig. 5. Simulation scenario.

A. Simulation Scenario

The simulation platform is PLEXE-SUMO [48], which is an OMNeT++ framework that provides platooning functionalities through the Plexe models implemented in SUMO [49]. The simulation scenario for the experiment is based on a section of QinglanExpy, G22, which is located at Qingdao, China, as illustrated in Fig. 5.

In simulation, the utilization of SUMO is pivotal, given its provision of an array of classical car-following and lane-changing models. For the longitudinal motion of HDVs, an Intelligent Driver Model (IDM) is employed [50] because of its ability to realistically replicate the behavior of HDVs in terms of acceleration and deceleration patterns. Meanwhile, a Cooperative Adaptive Cruise Control (CACC) model [51, 52] is selected to control the longitudinal motion of CAVs. CACC, renowned for its cooperative and adaptive characteristics, is well-suited for capturing the autonomous and interconnected behavior of CAVs. Regarding lateral motion, both HDVs and CAVs are governed by SUMO's default lane change model LC2013 [53], which incorporates realistic considerations such as safety distances and traffic conditions to ensure a good representation of lateral dynamics. In addition, the detailed simulation parameters are shown in Table II [39]. Our developed source codes are accessible at https://github.com/TongLu0223/VSL-VP.

B. Result on Platooning

Firstly, the DRL-based vehicle platooning model is trained within an individual window situated in S-2. The initialization of a window size $L = 5$ is conducted. In this sample scenario, five vehicles, including CAVs and HDVs, are randomly generated on the three-lane freeway, and a lane is blocked ahead by a traffic incident, resulting in a bottleneck. During training within a single window, the traffic flow is stable, and there is 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

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

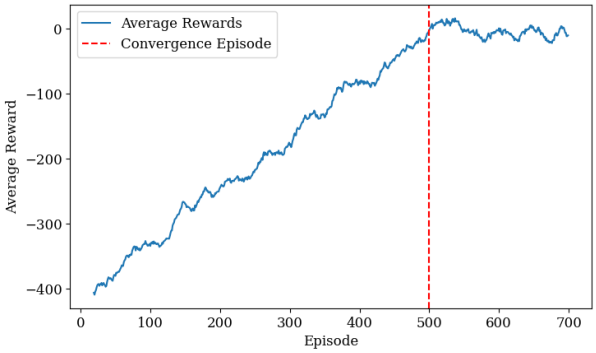


Fig. 6. Reward in training for platooning.

the optimal time for CAVs to join the platoon. The training utilizes rewards r_2 and r_4 only. The reward function is

$$r_{=}\omega_3 * r_2 + \omega_4 * r_4$$

where 3 = 0.9 and 4 = 0.1. A higher weight is assigned to r_2 since r_2 is the most significant optimization objective. The result of 700 episodes of training is shown in Fig. 6. It can be observed that the average reward gradually rises and eventually stabilizes, demonstrating 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 in https://youtu.be/I4CYQyf-TmE.

Fig. 7 depicts the variation of the average vehicle speed within this sample scenario. Vehicles pass through the bot-

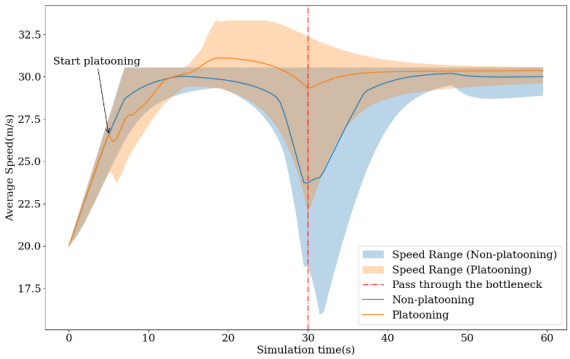


Fig. 7. Average vehicle speed.

tleneck at approximately 30 seconds. Since vehicles in the blocked lane need to decelerate and change lanes in advance, the average vehicle speed is significantly reduced. The results indicate that the CAVs exhibit smooth speed variations and achieve high speeds through the bottleneck compared to vehicles without platooning.

Then, curriculum learning is adopted to speed up the learning process. Curriculum learning is especially preferable for safety-critical tasks (e.g., autonomous driving), as starting from a proficient model can greatly reduce the number of blind explorations that may pose risks [54]. Specifically, the training of the model is continued in a large-scale traffic flow with 100 vehicles entering the road per minute. To enhance the performance of the model in a large-scale traffic flow, additional reward constraints are introduced to capture the complexity of traffic flow dynamics. The training process incorporates various traffic scenarios with different CAV penetration rates. During the initial training phase, the vehicle density is set to 40 vehicles per kilometer. All rewards $r_1 - r_4$ are considered in the training process. The agent is trained by a weighted sum of all rewards:

$$r = \omega_5 * r_1 + \omega_6 * r_2 + \omega_7 * r_3 + \omega_8 * r_4$$

where $5 = 0.68$, $6 = 0.1$, $7 = 0.2$ and $8 = 0.02$. The rationales for determining the values of weights are as follows. In the context of a large-scale traffic flow, throughput of the bottleneck represented by r_1 is the most critical optimization objective. Therefore, it receives the highest weight. To prevent congestion and excessive jams before the bottleneck, rewards r_2 and r_3 are also assigned weights. Reward r_4 encourages CAVs to form platoons as soon as possible, thereby augmenting the convergence speed of the model.

Fig. 8 displays the performance on four major metrics, respectively: bottleneck throughput, mean vehicle speed, mean number of stops, and mean platoon density at the bottleneck. It reveals the convergence of each reward over the process of training. Specifically, the bottleneck throughput, mean speed, and platoon density within the bottleneck area exhibit a gradual increase trend with the training. This suggests that the model effectively learns to enhance traffic throughput and overall system efficiency. Contrarily, there is a notable

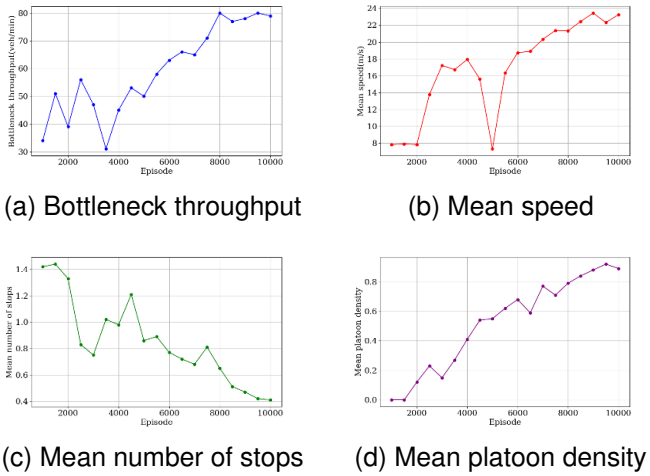


Fig. 8. Training performance in large-scale traffic flow.

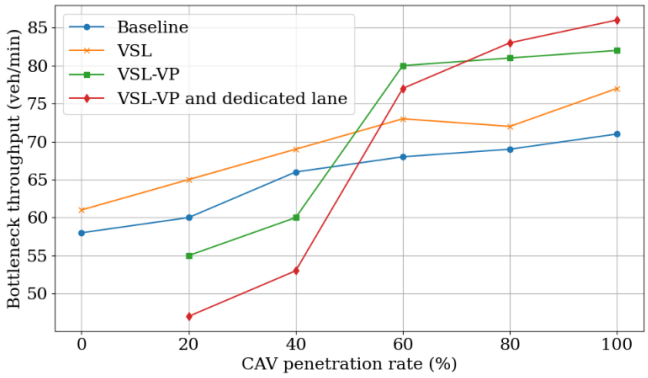


Fig. 9. Performance at different CAV penetration rates.

reduction in the mean number of stops for vehicles, signifying a decrease in traffic oscillations. The results demonstrate the effectiveness of the proposed method in achieving traffic performance optimization.

C. Result 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;
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.

TABLE III Comparison Results under Different CAV Penetration Rates

Strategy	Throughput (veh/min)	Mean speed (m/s)	Mean time loss (s)	Mean stops per vehicle	Mean stopping duration (s)	Mean / 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	32.33	3.03	17.60	154.66 / 246.51
VSL	73	13.15	25.71	1.25	14.17	106.75 / 208.23
VSL-VP	**80**	**30.40**	**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	27.11	2.88	12.67	107.95 / 217.86
VSL	72	12.13	26.02	0.72	9.69	82.73 / 181.67
VSL-VP	81	32.45	0.00	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	78	23.15	4.10	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**

Then, the effects of the dedicated lane on the performance of the proposed method are analyzed. At penetration rates of 20%, 40%, and 6%, 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 [40], [41].

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. However, as the penetration rate increases to over 60%, CAVs can form platoons easily. CAV platoons can play their advantages, passing through the bottleneck quickly. Therefore, VSL-VP has a higher mean speed but lower mean time loss compared with both baseline and VSL. It can successfully alleviate congestion, reducing the values of mean stops, mean stopping duration, mean jam length, and mean jam length to zero. Videos of these experiments can be found in https://www.youtube.com/playlist?list=PLzq7Vw-HmU2fLOuS4oGOkpTEyLHkFjV8CF.

The experimental results demonstrate the effectiveness of VSL-VP in alleviating freeway bottlenecks. This is achieved by enhancing bottleneck throughput, mitigating traffic oscillations, and reducing congestion, particularly in scenarios with the CAV penetration rate exceeding 60%.

TABLE IV Impact of Different Window Size (Transposed)

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

D. Window Size Test

In this work, confronted with the complexities of large-scale traffic flow, sliding-windows are used to manage the traffic flow effectively. Specifically, a fixed window size of $L = 5$ is maintained, thereby imposing a maximum platoon size constraint of 5 vehicles. The choice of platoon size is crucial in determining the dynamics of vehicle platoons and their impact on traffic conditions [42]. Our experiments delve into the exploration of various window sizes, conducted in six scenarios with different CAV penetration rates and vehicle densities. The objective is to assess the impact of altering the window size on the performance of VSL-VP under varying conditions. The results are shown in Table IV.

The impact of window size on bottleneck throughput varies in different scenarios. In scenario 1, the initial increase in the window size contributes to an improvement in bottleneck throughput. However, a further increase in the window size leads to a decline. In scenario 6, as the window size increases, the bottleneck throughput keeps increasing. This phenomenon can be attributed to the following reasons: In scenario 1, when the window size initially increases, the number of CAVs within the window rises, facilitating the formation of longer and more platoons. As the window size continues to increase, longer CAV platoons may impose restrictions on other HDVs within the window, limiting their lane-changing operations.

Additionally, with an excessive number of CAVs, platoons may not have completed formation by the time they reach the bottleneck, resulting in a reduction in bottleneck throughput. In scenario 6, because the CAV penetration rate is 100%, there are no HDVs on the road. CAVs can form platoons more quickly and safely. Therefore, as the window size increases, CAVs can form longer platoons, improve traffic efficiency, and thus increase bottleneck throughput. Nevertheless, considering other factors, such as the time complexity of the model, a larger window size is not always good.

Hence, in each distinct scenario, the optimal window size varies based on factors such as CAV penetration rate and vehicle density. The most effective strategy involves dynamically determining the window size in real time, enabling adaptive responses to different traffic flows. This dynamic approach to window size determination is a primary focus of our future work, as it holds the potential to enhance the performance of our vehicle platooning model across a range of traffic scenarios.

E. Conclusion

In this work, VSL-VP, a novel framework that combines VSL and vehicle platooning for the freeway bottleneck, is proposed. It divides two road segments before the bottleneck area. In the former segment, VSL is employed to reduce vehicle density and increase car-following distance in the latter segment by limiting vehicle speeds. This facilitates the conditions for platooning. In the latter road segment, sliding-windows are initially applied to partition the large-scale traffic flow into independent windows. Within each window, DQN is utilized to determine the optimal time for CAVs to join the platoon, enabling CAVs to form a stable platoon and pass through the bottleneck area, consequently enhancing bottleneck throughput and reducing traffic congestion. Various traffic performance metrics of VSL-VP under scenarios with different CAV penetration rates are accessed. The results demonstrate the effectiveness of VSL-VP in improving bottleneck throughput and reducing traffic congestion at over 60% CAV penetration rate. The effectiveness of the dedicated lane for CAVs and the impact of varying window sizes on bottleneck throughput are also analyzed.

In future studies, further optimization will be proposed to improve the efficiency of VSL-VP. Specifically, the window size, determining the maximum platoon size, has an impact on the bottleneck throughput. To refine and optimize this aspect of the framework, some new approaches, such as Multi-Agent Reinforcement Learning (MARL) [55], will be tried. The incorporation of MARL introduces a more sophisticated decision-making framework, enabling a collective and collaborative approach among multiple intelligent agents. By employing MARL, the optimal window size should be dynamically determined for vehicle platoons in real-time, enhancing adaptability to evolving traffic conditions and improving the overall efficiency and effectiveness of the proposed VSL-VP framework in the bottleneck scenario of freeways.

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. 17 632–17 643, Oct. 2022.

[5] Y. Zhang and P. A. Ioannou, "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] S. Qin, S. Zhang, J. Wang, S. Liu, X. Guo, and L. Qi, "Multiobjective multiverse optimizer for multirobotic u-shaped disassembly line balancing problems," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 2, pp. 882–894, 2023.

[8] 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. 15 522–15 535, Sept. 2022.

[9] X. Wang, M. Zhou, Q. Zhao, S. Liu, X. Guo, and L. Qi, "A branch and price algorithm for crane assignment and scheduling in slab yard," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 3, pp. 1122–1133, 2020.

[10] 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.

[11] 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, p. 103421, 2021.

[12] 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, p. 114030, 2020.

[13] 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.*, Oct. 2023, early access.

[14] 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.

[15] 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.

[16] 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.

[17] 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. 17 312–17 324, Oct. 2022.

[18] 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.

[19] 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, p. 103138, 2021.

[20] 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. 15 951–15 966, Sept. 2022.

[21] 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.

[22] S. W. Smith *et al.*, "Improving urban traffic throughput with vehicle platooning: theory and experiments," *IEEE Access*, vol. 8, pp. 141 208–141 223, July 2020.

[23] G. Xu, Z. Zhang, Z. Li, X. Guo, L. Qi, and X. Liu, "Multi-objective discrete brainstorming optimizer to solve the stochastic multiple-product robotic disassembly line balancing problem subject to disassembly failures," *Mathematics*, vol. 11, no. 6, p. 1557, 2023.

[24] 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.

[25] 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. 22 790–22 806, Dec. 2022.

[26] 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.

[27] 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.

[28] O. Vinyals *et al.*, "Grandmaster level in starcraft ii using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, Oct. 2019.

[29] 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.

[30] 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, p. 103933, 2023.

[31] 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, p. 104019, 2023.

[32] 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.

[33] 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. 13 340–13 350, Dec. 2021.

[34] A. Irshayyid and J. Chen, "Comparative study of cooperative platoon merging control based on reinforcement learning," *Sensors*, vol. 23, no. 2, p. 990, 2023.

[35] Y. Bengio *et al.*, "Curriculum learning," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, Montreal, Canada, 2009, pp. 41–48.

[36] J. Zhou and F. Zhu, "Modeling the fundamental diagram of mixed human-driven and connected automated vehicles," *Transp. Res. C, Emerg. Technol.*, vol. 115, p. 102614, 2020.

[37] Y. Jiang *et al.*, "Platoon intensity of connected automated vehicles: Definition, formulas, examples, and applications," *J. Adv. Transp.*, vol. 2023, 2023.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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, p.

102882, 2021.
[43] R. S. Sutton and A. G. Barto, *Introduction to reinforcement learning*. Cambridge: MIT press, 1998.
[44] 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. 11 798–11 800, Aug. 2022.
[45] 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.
[46] 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.
[47] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, Feb. 2015.
[48] M. Segata *et al.*, "Plexe: A platooning extension for veins," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Paderborn, GER, 2014, pp. 53–60.
[49] 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.
[50] 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.
[51] 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.
[52] 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.
[53] J. Erdmann, "Sumo's lane-changing model," in *Proc. Model. Mob. Op. Dat., 2nd SUMO Conf.*, Berlin, Germany, 2015, pp. 105–123.
[54] 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. 11 623–11 638, Nov. 2023.
[55] 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.

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. 17 632–17 643, Oct. 2022.
[5] Y. Zhang and P. A. Ioannou, "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] S. Qin, S. Zhang, J. Wang, S. Liu, X. Guo, and L. Qi, "Multiobjective multiverse optimizer for multirobotic u-shaped disassembly line balancing problems," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 2, pp. 882–894, 2023.
[8] 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. 15 522–15 535, Sept. 2022.
[9] X. Wang, M. Zhou, Q. Zhao, S. Liu, X. Guo, and L. Qi, "A branch and price algorithm for crane assignment and scheduling in slab yard," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 3, pp. 1122–1133, 2020.
[10] 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.
[11] 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, p. 103421, 2021.
[12] 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, p. 114030, 2020.
[13] 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.*, Oct. 2023, early access.
[14] 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.
[15] 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.

[16] 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.

[17] 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. 17 312–17 324, Oct. 2022.

[18] 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.

[19] 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, p. 103138, 2021.

[20] 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. 15 951–15 966, Sept. 2022.

[21] 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.

[22] S. W. Smith *et al.*, "Improving urban traffic throughput with vehicle platooning: theory and experiments," *IEEE Access*, vol. 8, pp. 141 208–141 223, July 2020.

[23] G. Xu, Z. Zhang, Z. Li, X. Guo, L. Qi, and X. Liu, "Multi-objective discrete brainstorming optimizer to solve the stochastic multiple-product robotic disassembly line balancing problem subject to disassembly failures," *Mathematics*, vol. 11, no. 6, p. 1557, 2023.

[24] 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.

[25] 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. 22 790–22 806, Dec. 2022.

[26] 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.

[27] 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.

[28] O. Vinyals *et al.*, "Grandmaster level in starcraft ii using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, Oct. 2019.

[29] 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.

[30] 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, p. 103933, 2023.

[31] 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, p. 104019, 2023.

[32] 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.

[33] 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. 13 340–13 350, Dec. 2021.

[34] A. Irshayyid and J. Chen, "Comparative study of cooperative platoon merging control based on reinforcement learning," *Sensors*, vol. 23, no. 2, p. 990, 2023.

[35] Y. Bengio *et al.*, "Curriculum learning," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, Montreal, Canada, 2009, pp. 41–48.

[36] J. Zhou and F. Zhu, "Modeling the fundamental diagram of mixed human-driven and connected automated vehicles," *Transp. Res. C, Emerg. Technol.*, vol. 115, p. 102614, 2020.

[37] Y. Jiang *et al.*, "Platoon intensity of connected automated vehicles: Definition, formulas, examples, and applications," *J. Adv. Transp.*, vol. 2023, 2023.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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, p. 102882, 2021.

[43] R. S. Sutton and A. G. Barto, *Introduction to reinforcement learning*. Cambridge: MIT press, 1998.

[44] 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. 11 798–11 800, Aug. 2022.

[45] 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.
[46] 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.
[47] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, Feb. 2015.
[48] M. Segata *et al.*, "Plexe: A platooning extension for veins," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Paderborn, GER, 2014, pp. 53–60.
[49] 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.
[50] 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.
[51] 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.
[52] 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.
[53] J. Erdmann, "Sumo's lane-changing model," in *Proc. Model. Mob. Op. Dat., 2nd SUMO Conf.*, Berlin, Germany, 2015, pp. 105–123.
[54] 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. 11 623–11 638, Nov. 2023.
[55] 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 (S'16-M'18) received his B.S. degree in Information and Computing Science and M.S. degree in Computer Software and Theory from Shandong University of Science and Technology, Qingdao, China, in 2009 and 2012, respectively, and Ph.D. degree in Computer Software and Theory from Tongji University, Shanghai, China in 2017. He is currently a lecturer of Computer Science and Technology at Shandong University of Science and Technology, Qingdao, China. From 2015 to 2017, he was a visiting student in the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, USA. He has authored over 50 technical papers in journals and conference proceedings, including IEEE Transactions on System, Man and Cybernetics: Systems, IEEE Transactions on Intelligent Transportation Systems, and IEEE/CAA Journal of Automatica Sinica. He received the Best Student Paper Award-Finalist in the 15th IEEE International Conference on Networking, Sensing and Control (ICNSC'2018). His current research interests include Petri nets, discrete event systems and optimization algorithms.

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 B.S. degree in Computer Science and Technology from Shenyang Institute of Engineering, Shenyang, China, in 2006, M.S. degree in Aeronautics and Astronautics Manufacturing Engineering. from Shenyang Aerospace University, Shenyang, China, in 2009, Ph. D. degree in System Engineering from Northeastern University, Shenyang, China, in 2015. He is currently an associate professor of the College of Computer and Communication Engineering at Liaoning Shihua 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 authored 60+ technical papers in journals and conference proceedings, including IEEE Transactions on Cybernetics, IEEE Transactions on System, Man and Cybernetics: Systems, IEEE Transactions on Intelligent Transportation Systems, and IEEE/CAA Journal of Automatica Sinica. His current research interests include Petri nets, remanufacturing, recycling and reuse of automotive, intelligent optimization algorithm.

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.