IEEE Open Journal of Vehicular Technology

Received 5 May 2025; accepted 19 May 2025. Date of publication 27 May 2025;
date of current version 21 July 2025. The review of this article was coordinated by the Guest Editors for the Special Section on Orthogonal Time Frequency Space Modulation and Delay-Doppler Signal Processing.

Digital Object Identifier 10.1109/OJVT.2025.3574223

OTFS-Assisted Sensing Adaptive Cruise Control for Highways: A Reinforcement Learning Approach

YULIN LIU [1], XIAOQI ZHANG [2] (Student Member, IEEE), JUN WU [3], AND QINGQING CHENG [4] (Member, IEEE)

[1]Department of Electrical, Electronic Engineering, Southern University of Science, Technology, Shenzhen 518055, China
[2]School of Electrical, Data Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia
[3]School of Automation and Intelligent Manufacturing, Southern University of Science, Technology, Shenzhen 518055, China
[4]School of Electrical Engineering and Robotics, Queensland University of Technology, Brisbane, QLD 4000, Australia

CORRESPONDING AUTHOR: QINGQING CHENG (e-mail: qingqing.cheng@qut.edu.au).

ABSTRACT In this paper, we propose a novel channel estimation approach and driving decision method for adaptive cruise control (ACC) systems for vehicular networks, leveraging the properties of deep learning, reinforcement learning, and orthogonal time frequency space (OTFS) modulation. To achieve that, we propose to leverage deep learning (DL) to estimate motion parameters. Subsequently, we develop a reinforcement learning method to process the obtained target motion information to enable adaptive vehicle-following strategies. This ensures robust decision-making and precise control under dynamic and uncertain driving conditions, achieving superior performance in terms of both accuracy and reliability.

INDEX TERMS ACC, deep learning, OTFS, reinforcement learning.

I. INTRODUCTION

Emerging vehicular networks are anticipated to integrate intelligent capabilities essential for supporting autonomous driving and reducing traffic congestion. Adaptive cruise control (ACC) serves as a cornerstone in implementing intelligent vehicular networks (IVNs). It enables IVN participants to acquire essential information, e.g., lane details, pedestrians, and unexpected traffic incidents, ensuring safe following distances from preceding vehicles and minimizing collision risks. Furthermore, ACC can effectively manage vehicle acceleration and braking, alleviating driver fatigue, especially during long-distance and highway driving. Therefore, ACC is recognized as a transformative component of next-generation wireless systems, attracting increasing attention from both academia and industry.

Cruise control systems originated in the mid-20th century [1], with early implementations using mechanical mechanisms to sustain a fixed throttle position, offering limited functional capabilities. However, conventional cruise control (CCC) systems were unable to ensure consistent vehicle velocity and provided minimal enhancement in safety or driving comfort. To address these limitations, the first generation of ACC systems was developed to dynamically adjust vehicle velocity and maintain a specified following distance from the preceding vehicle through integrated throttle and brake control. Initially, ACC systems were deployed in luxury vehicles by automakers and suppliers, focusing on improving driving comfort and convenience while offering secondary safety benefits. Today, ACC systems are commonly integrated into vehicles across various market segments, from high-end to mid-range models. Specifically, when no vehicle is detected ahead, an ACC-equipped vehicle operates similarly to a traditional CCC system, maintaining a driver-defined velocity via throttle control. Upon detecting a preceding vehicle, the ACC system determines the safety of maintaining the preset velocity. If the vehicle ahead is too close or moving at a much lower speed, the system shifts from maintaining the diver-defined velocity to controlling a user-defined time headway, adjusting both the throttle and the brake accordingly. Modern ACC systems typically operate at a velocity range of 40 km/h to 160 km/h, with a maximum braking deceleration of approximately 0.5 g. It is important to note that the

driver maintains full control of the vehicle, enabling manual intervention whenever required.

To accomplish successful tasks, ACC system needs to measure the distance and relative velocity between adjacent vehicles using diverse sensors, e.g., radar, lidar, or a video camera. This acquired data is then transmitted to the controller to determine and perform appropriate control actions. Towards that end, the study in [2] developed a vision-based ACC system by utilizing a camera for both range and velocity measurements. Despite the progress, a single camera falls short in detecting vehicles in adjacent lanes, and its performance is vulnerable to weather conditions, such as fog or rain, compromising safe cruise control. For instance, the study in [2] examined a vision-based ACC system that used a single camera to compute both range and relative velocity. However, vision-based ACC systems are vulnerable to weather-related issues, such as fog, where a following vehicle may fail to detect the leading vehicle, thereby compromising safe cruise control. Given that, [3] proposed to integrate frequency-modulated continuous wave (FMCW) radar with constant false alarm rate (CFAR) clutter rejection, enhancing detection reliability in adverse weather conditions, such as rain. This integration can reduce the impact of clutter, ensuring accurate range and velocity measurements [4]. With the same purpose, [5] proposed a lidar-based ACC framework to detect both distance and velocity information of nearby vehicles. To further improve performance, [6] introduced a multi-dimensional sensor-based ACC system, which exploits the Kalman filter algorithm for obstacle detection and tracking. However, the use of multi-dimensional data in these systems poses significant computational challenges, especially for vehicles within large-scale IVNs, restricting its implementation in real-time scenarios. Therefore, the development of a low-complexity approach for determining range and relative velocity in ACC systems tailored for IVNs remains a critical challenge.

Beyond precise sensing, reliable communication is another essential enabler for IVNs, allowing participants to exchange critical information for timely decision-making and efficient resource allocation. Traditionally, orthogonal frequency division multiplexing (OFDM) has been adopted for vehicle-to-vehicle (V2V) communications due to its strengths in scalable multi-user access and robustness against multipath fading. For instance, the work in [7], [8], [9] developed an OFDM-based pre-coded chaos shift keying (PC-CSK) transceiver to enhance V2V communication reliability. However, OFDM suffers from a high peak-to-average power ratio (PAPR), increasing transmitter complexity and reducing power efficiency. Additionally, it experiences significant performance degradation in high-mobility environments due to large Doppler shifts and inter-carrier interference, particularly in high-speed vehicle scenarios such as highways [10], [11], [12]. Consequently, OFDM may not be well-suited for high-mobility IVN applications.

Recognizing the above challenges, a novel communication paradigm known as orthogonal time frequency space (OTFS) has been developed. Unlike OFDM modulating data symbols in the time-frequency (TF) domain, OTFS operates in the delay-Doppler (DD) domain. This shift enables OTFS to convert fluctuating and dense channel responses in the TF domain to quasi-static and sparse ones in the DD domain, thereby improving robustness to Doppler spreads and enhancing performance in high-mobility scenarios. Inspired by these benefits, several studies have explored OTFS-based V2V communications [13], [14], [15], [16], [17], where results confirmed the superior performance of OTFS for high-velocity scenarios. Another important finding in these studies is that the DD domain channel parameters are essential for realizing the sensing functionality by converting the delay and Doppler shifts into vehicle ranges and velocities. Consequently, OTFS can effectively serve both sensing and communication tasks using unified signal processing techniques, offering a promising solution for designing low-complexity ACC systems in IVNs.

While OTFS excels in sensing and communication, it necessitates advanced decision-making strategies to fully capitalize on its advantages. Notably, distance and relative velocity to the preceding vehicle can be estimated at present time step, however, the preceding vehicle's driving status such as velocity at next time step can not be accurately predicted. This unknown information causes challenges for the following vehicle to maintain within desired system constraints, motivating the use of reinforcement learning (RL) that can iteratively learn the vehicle dynamics by the following vehicle's interaction with the preceding vehicle and improve driving strategy. Q-learning is a popular choice among many researchers due to its simplicity and effectiveness in environments with discrete, well-defined state and action spaces. For instance, [18] innovatively applied a critic-only Q-learning method for designing optimal tracking control. Additionally, [19] introduced a model-based approximate Q-learning algorithm that trains the control laws to maximize the expected time before system constrains, such as safety, comfort, and fuel economy are violated. Further extending the utility of Q-learning, [20], [21] employed it in an autonomous car-following system that processes video frames, thereby enhancing the adaptability of ACC systems to complex traffic conditions. Previous studies have validated the effectiveness of Q-learning in complex driving environments. In this work, we adapt Q-learning specifically for highway scenarios, optimizing it to leverage the unique capabilities of OTFS for enhanced safety and operational efficiency.

Motivated by the above, in this paper, we investigate the design of OTFS sensing-assisted ACC system for IVNs. Specifically, the roadside infrastructure units (RSUs) transmit OTFS signals to the mobile vehicles for downlink communication. Meanwhile, the received signals at the vehicles can be further radiated to their adjacent vehicles. In such a context, the following vehicle captures the reflected signal from the preceding vehicle and extracts the delays and the Doppler shifts to determine the corresponding distance and velocity. Subsequently, the optimal driving policy for the following

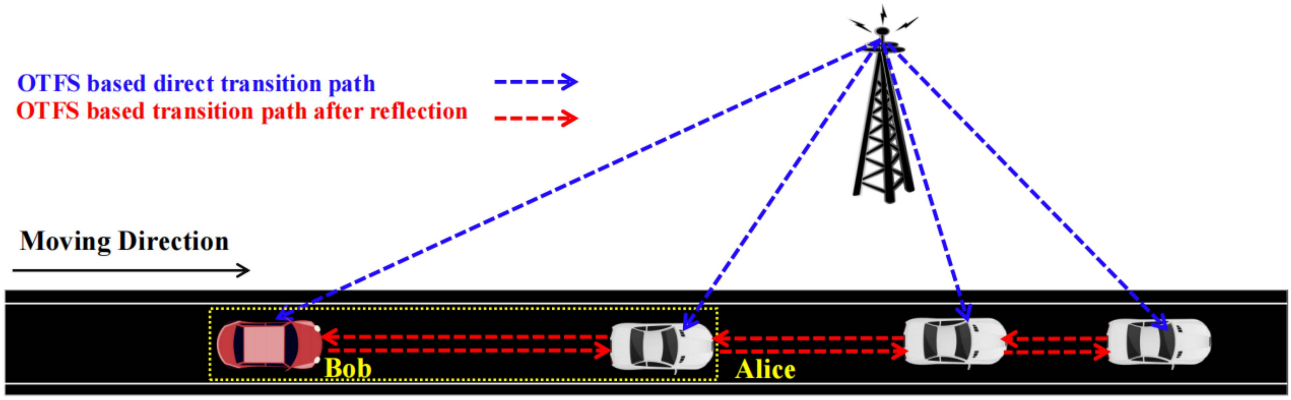


FIGURE 1. Physical scenario.

vehicle can be obtained, enabling a low-overhead ACC system design for joint sensing and communication functionalities. To the best of authors' knowledge, this is the first attempt to exploit OTFS for designing the ACC system in IVNs. The main contributions are listed as follows:

- We propose a novel adaptive cruise control system for vehicular networks, by leveraging the properties of OTFS signals and RL. This integration enhances the accuracy of vehicle sensing under varying conditions but also significantly boosts the adaptability and performance of cruise control functions across diverse traffic scenarios.
- To perform sensing tasks, we first formulate the parameter estimation as a regression problem and then develop a dedicated parameter estimation network (PEnet) by exploiting the power of ResNet. PEnet can capture intricate signal features and mitigate noise and interference, enabling efficient and accurate feature extraction for sensing tasks. Additionally, PEnet is able to reduce computational complexity significantly, demonstrating its potential in real-world implementations.
- With extracted target information, we develop an adaptive vehicle-following strategy by leveraging RL to learn optimal control policies through interactions with dynamic environments. The designed strategy can ensure accurate and robust decision-making by accounting for real-time environmental changes and complex vehicular interactions. Additionally, it is capable of adapting to varying and uncertain driving conditions, enhancing performance in responsiveness and reliability, even under challenging scenarios.

The remainder of this paper is organized as follows: Section II introduces system models, and Section III derives proposed methods. Our simulation results are provided in Section IV, while Section V concludes this paper.

Notations: Boldface capital letters are used to represent matrices, such as $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{v}$. Boldface lowercase letters are used to represent vectors, such as $\mathbf{x}$, $\mathbf{y}$, $\mathbf{h}$, $\mathbf{k}$, $\mathbf{l}$, and $\mathbf{s}$. Blackboard bold letters denote complex number vector spaces; for instance, $\mathbb{C}^{MN\times 1}$ represents a complex number vector space of dimensions $MN \times 1$. Additionally, $\mathbb{E}$ denotes expectation in probability and statistics. The set of states and the action space in RL are defined in uppercase calligraphic letters, such as $\mathcal{S}$ and $\mathcal{A}$. Random variables are denoted by italic letters, like x, y, etc.

II. SYSTEM MODELS

We consider a vehicular system supported by an RSU, as shown in Fig. 1, including k vehicles in the traffic flow, indexed from 0 to $k-1$. The following autonomous vehicle, indexed as $k-1$, is denoted as Bob. Each vehicle receives both the direct transition signal from the RSU and the reflected signal from its preceding vehicle. All vehicles share the same frequency band. To simplify the problem, the height of the RSU is ignored. In the following, we first introduce our kinetic model for the vehicle system, followed by a detailed explanation of the OTFS-assisted sensing model.

A. VEHICLE KINETIC MODEL

In this section, we focus on the straight-line motion of vehicles for simplicity, then extend the single-vehicle kinematic model to a multi-vehicle system.

Specifically, the single vehicle kinetic model can be characterized by its position, velocity, and acceleration along the longitudinal axis. For a specific vehicle j, its position update process can be expressed as

$$x_j(t+\Delta t) = x_j(t) + v_j(t)\cdot\Delta t + \frac{1}{2}a_j(t)\cdot(\Delta t)^2, \quad (1)$$

where $x_j(t)$, $v_j(t)$, and $a_j(t)$ represent the position, velocity, and acceleration of vehicle j at time t, respectively, and Δt denotes the time step. Also, its velocity update process is given by

$$v_j(t+\Delta t) = v_j(t) + a_j(t)\cdot\Delta t. \quad (2)$$

Next, we extend this kinetic model to a multi-vehicle system. To model dynamics between vehicles, we define the distance $d(t)$ and relative velocity $v(t)$ between vehicle $i+1$ (following vehicle) and vehicle i (preceding vehicle) as

$$d_{i,i+1}(t) = x_i(t) - x_{i+1}(t),$$

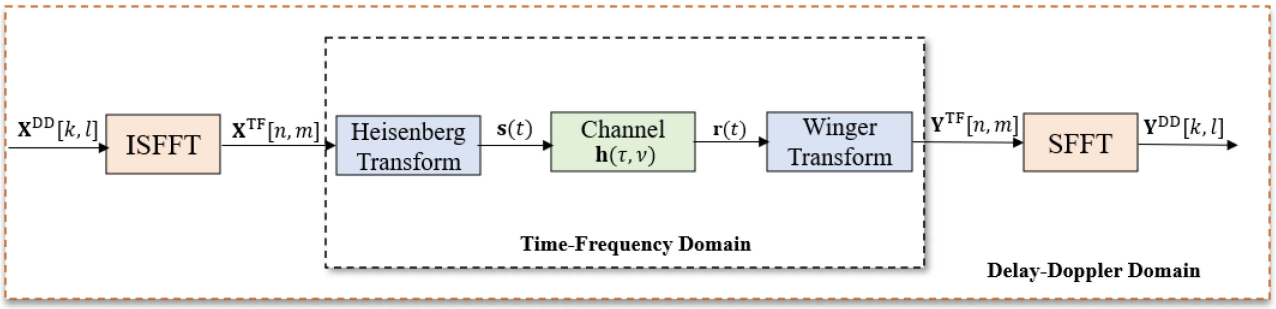


FIGURE 2. OTFS modulation and demodulation process.

$$v_{i,i+1}(t) = v_i(t) - v_{i+1}(t), \tag{3}$$

where $d(t)$ represents the longitudinal distance between two vehicles, and $v(t)$ stands for the relative velocity. And after Δt, it can be written as

$$\begin{aligned} d_{i,i+1}(t+\Delta t) &= x_i(t+\Delta t) - x_{i+1}(t+\Delta t), \\ v_{i,i+1}(t+\Delta t) &= v_i(t+\Delta t) - v_{i+1}(t+\Delta t), \end{aligned} \tag{4}$$

where $i+1$ denotes any following vehicle up to $k-1$.

These parameters are crucial for ACC, where the following vehicle adjusts its acceleration $a_{i+1}(t)$ to maintain a safe distance and avoid collisions [22]. Note that, the proposed kinematic model for the multi-vehicle system provides an effective way to simulate longitudinal motion and evaluate control strategies under straight-line driving conditions.

B. OTFS-ASSISTED SENSING MODEL

In the study of OTFS modulation, a data grid in the DD domain is considered, represented by a matrix of size $N \times M$, where N and M denote the number of slots and subcarriers, respectively. The data of an OTFS frame in the DD domain is expressed as $\mathbf{X}^{\mathrm{DD}}[k,l]$, where integers $l \in \{0,\ldots,M-1\}$ and $k \in \{0,\ldots,N-1\}$ denote the DD indices, respectively [23], [24]. As shown in Fig. 2, data denoted as $\mathbf{X}^{\mathrm{DD}}[k,l]$, undergoes a transformation to the TF domain through an inverse symplectic finite Fourier transform (ISFFT) [25], [26], given by

$$\mathbf{X}^{\mathrm{TF}}[n,m] = \frac{1}{\sqrt{MN}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} \mathbf{X}^{\mathrm{DD}}[k,l] e^{j2\pi\left(\frac{nk}{N} - \frac{ml}{M}\right)}, \tag{5}$$

where time and frequency indices are denoted as n and m, respectively. Then the transmitted OTFS waveform from RSU in the time domain can be derived using the Heisenberg transform, where the modulation symbols $\mathbf{X}^{\mathrm{TF}}[n,m]$ are transformed into $s(t)$. These signals are shaped by $g_{\mathrm{tx}}(t)$ and modulated onto orthogonal subcarriers, which is [27], [28]

$$\mathbf{s}(t) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{X}^{\mathrm{TF}}[n,m] e^{j2\pi m\Delta f(t-nT)} g_{\mathrm{tx}}(t-nT), \tag{6}$$

where $g_{\mathrm{tx}}(t)$ is the transmit shaping function that localizes the signal in the time domain. After traveling through the wireless channels, $\mathbf{s}(t)$ is collected by the receiver side, denoted as $\mathbf{r}(t)$, expressed as

$$\mathbf{r}(t) = \iint h(\tau,\nu)\mathbf{s}(t-\tau) e^{j2\pi\nu(t-\tau)} d\tau d\nu, \tag{7}$$

where $h(\tau,\nu)$ is channel response in the DD domain, given by

$$h(\tau,\nu) = \sum_{i=1}^{P} h_i \delta(\tau-\tau_i)\delta(\nu-\nu_i), \tag{8}$$

where h_i is the channel gain of the ith path, τ_i and ν_i are the delay and the Doppler spread of the ith path, respectively.

Then, Wigner Transform is performed on $\mathbf{r}(t)$ to obtain the TF domain response, i.e.,$\mathbf{Y}^{\mathrm{TF}}[n,m]$, by

$$\begin{aligned} &\mathbf{Y}^{\mathrm{TF}}[n,m] \\ &= \left[\int g_{\mathrm{rx}}(t-\tau)\mathbf{r}(t) e^{-j2\pi\nu(t-\tau)} dt\right]\Big|_{\tau=nT,\nu=m\Delta f}, \end{aligned} \tag{9}$$

where g_{rx} represents the transmit pulse. Then we map the signal from the TF domain to the DD domain using symplectic finite Fourier transform (SFFT), given by

$$\mathbf{Y}^{\mathrm{DD}}[k,l] = \frac{1}{\sqrt{MN}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \mathbf{Y}^{\mathrm{FT}}[n,m] e^{2j\pi\left(\frac{ml}{M} - \frac{nk}{N}\right)}. \tag{10}$$

After the Wigner Transform, the noise $v[k,l]$ is added to $\mathbf{Y}^{\mathrm{DD}}[k,l]$, and $h_w[k-k',l-l']$ represents the equivalent channel in DD domain, which is

$$\begin{aligned} &h_w\left[k-k',l-l'\right] \\ &= \sum_{i=1}^{P} h_i e^{-j2\pi\nu_i l_{\tau_i}} w\left(k-k'-k_{\nu_i}, l-l'-l_{\tau_i}\right), \end{aligned} \tag{11}$$

where $k_{\nu_i} = \nu_i NT$, $l_{\tau_i} = \tau_i M\Delta f$, and the sampling function $w(k-k'-k_{\nu_i}, l-l'-l_{\tau_i})$ can be written as

$$\begin{aligned} &w\left(k-k'-k_{\nu_i}, l-l'-l_{\tau_i}\right) \\ &= \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} e^{-j2\pi(\nu nT - \tau m\Delta f)}\Big|_{\nu=k-k'-k_{\nu_i},\tau=l-l'-l_{\tau_i}}. \end{aligned} \tag{12}$$

The relationship between the input signal in DD domain, $\mathbf{X}^{\mathrm{DD}}[k,l]$ and the output signal in DD domain, $\mathbf{Y}^{\mathrm{DD}}[k,l]$ is

$$\mathbf{Y}^{\mathrm{DD}}[k,l] = \sum_{k'=0}^{N-1}\sum_{l'=0}^{M-1} \mathbf{X}^{\mathrm{DD}}[k,l] h_{\mathrm{w}}\left[k-k', l-l'\right] + v[k,l]. \tag{13}$$

For simplicity, the relationship between the received signal vector and the channel coefficients in the delay-Doppler domain can be expressed as a linear system in matrix form, which is

$$\mathbf{y} = \mathbf{\Phi}_x \mathbf{h}_w + \mathbf{v}, \tag{14}$$

where $\mathbf{Y} \in \mathbb{C}^{MN\times 1}$ is the received signal vector, $\mathbf{v} \in \mathbb{C}^{MN\times 1}$ represents noise, and $\mathbf{h}_w \in \mathbb{C}^{MN\times 1}$ denotes the vector of effective channel coefficients given by equation (11). The sensing matrix is derived from the pilot symbols. The channel estimation problem of OTFS systems involves obtaining channel vector $\mathbf{h}_w$, which contains the motion parameter $\mathbf{k}_\tau = [k_{\tau_1}, \ldots, k_{\tau_P}]$ and $\mathbf{l}_\tau = [l_{\tau_1}, \ldots, l_{\tau_P}]$. The traditional radar parameter estimation methods often rely on techniques such as matched filtering. However, these approaches are sensitive to variations in target properties, such as rotation, translation, and scaling, and they require precise prior knowledge of the target signal waveform, which is always challenging to achieve. To address these challenges, we propose a deep learning-based estimation algorithm to effectively estimate Doppler frequency shifts and time delays, offering enhanced accuracy and robustness.

III. PROPOSED METHODS

In this section, we first develop a deep learning (DL)-based channel estimation approach to estimate Doppler shifts and delays for vehicle sensing. Then, a reinforcement learning (RL)-based approach is proposed to optimize the adaptive vehicle-following strategy.

A. ACC PROBLEM FORMULATION

As illustrated in Fig. 1, we focus on Bob (depicted in red) following its preceding vehicle (Alice) along the highway. Let $x_b(t)$, $v_b(t)$, $a_b(t)$ denote the position, velocity, and acceleration of Bob at time t, respectively. Similarly, $x_a(t)$, $v_a(t)$, $a_a(t)$ represent the corresponding information of the Alice. The distance between the two vehicles is defined as $d_{a,b}(t) = x_a(t) - x_b(t)$.

In our formulation, the primary objective of the ACC system is to maintain a safe following distance, defined as d_{safe}. Then the allowable distance error ε_d is explicitly defined as the tolerance within which the actual following distance $d_{a,b}(t)$ can be derived from d_{safe}, i.e., $e_d(t) = |d_{a,b}(t) - d_{\mathrm{safe}}| \le \varepsilon_d$. Similarly, we define the allowable velocity difference ε_v to limit the speed discrepancy between Alice and Bob, i.e., $e_v(t) = |v_a(t) - v_b(t)| \le \varepsilon_v$. To ensure a safe and comfortable driving experience, we impose the following constraints: $a_{\max}$ and $a_{\min}$ are defined as the maximum and minimum of the allowable accelerations, respectively. Besides, $v_{\min}$ and $v_{\max}$ are defined as the speed constraints on the highway. These constraints and limits are summarized in Table 1.

TABLE 1. System Constraints

Constraint Type	Mathematical Expression
Speed limits	$v_{\min} \le v_a(t) \le v_{\max}$, $v_{\min} \le v_b(t) \le v_{\max}$
Acceleration bounds	$a_{\min} \le a_a(t) \le a_{\max}$, $a_{\min} \le a_b(t) \le a_{\max}$
Allowable errors	$e_d(t) = \lvert d_{a,b}(t) - d_{\mathrm{safe}}\rvert \le \varepsilon_d$, $e_v(t) = \lvert v_a(t) - v_b(t)\rvert \le \varepsilon_v$

The ACC control problem can be formulated as finding the optimal driving action that minimizes:

$$J = \int \left[w_1 e_d^2(t) + w_2 e_v^2(t) + w_3 a_h^2(t)\right] dt, \tag{15}$$

where w_1, w_2, and w_3 are weighting factors that balance the trade-off between distance keeping, velocity matching, and ride comfort, respectively. In the following RL section, these objectives are incorporated into the reward function to guide driving decision-making process.

B. DL-BASED CHANNEL ESTIMATION

1) PILOT ARRANGEMENT

In this work, we adopt the pilot arrangement as designed in [29], where the single pilot is embedded in the DD domain with zero guard space. We consider $x^{\mathrm{DD}}[k,l]$ corresponds to the individual elements of the DD domain data matrix $X^{\mathrm{DD}}[k,l]$, which represents the entire OTFS frame in the delay-Doppler domain as mentioned above [12], [26], [30]. The pilot arrangement is given by

$$x^{\mathrm{DD}}[k,l] = \begin{cases} x_p & k = k_p, l = l_p, \\ 0 & l \in [l_p - l_{\max}, l_p + l_{\max}], \\ & k \in [k_p - 2k_{\max}, k_p + 2k_{\max}], \\ x_d & \text{otherwise.} \end{cases} \tag{16}$$

where x_p is the pilot located at $[l_p, l_p]$, and x_d is the data symbol. Then, the received symbol at the guard space without the noise term is

$$y^{\mathrm{DD}}[k,l] = h_w[k-k_p, l-l_p]x_p + \bar{w}[k,l], \tag{17}$$

where h_w is the effective channel in the DD domain demonstrated by (11), and $\bar{w}[k,l]$ is the additive Gaussian noise. For motion parameters estimation, the above equation can be rewritten (17) as

$$\tilde{y}^{\mathrm{DD}}[k',l'] = y^{\mathrm{DD}}[k,l]/x_p = h_w[k',l'] + \tilde{w}[k',l'], \tag{18}$$

where $\tilde{w}[k',l'] = \bar{w}[k,l]/x_p$ denotes the additive Gaussian noise. $h_w[k',l']$ is the shifting results of $h_w[k,l]$ based on the pilot position. As a result, we can estimate the delay index and the Doppler index directly from $\tilde{y}_{\mathrm{DD}}[k',l']$.

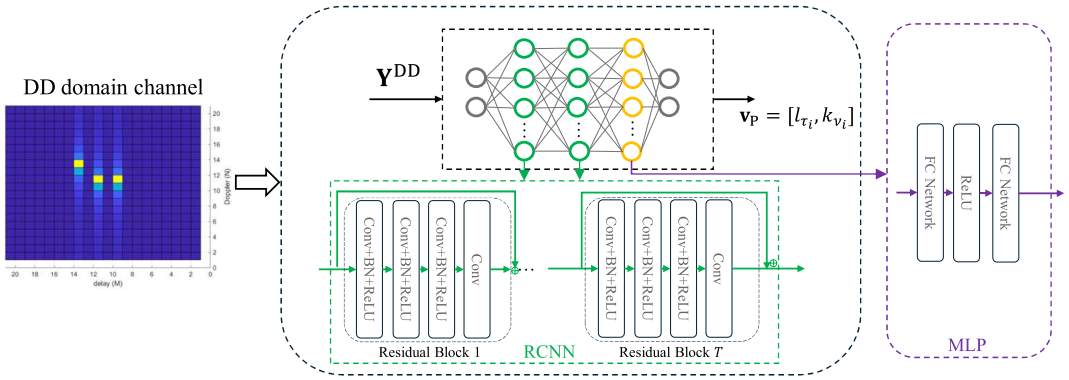


FIGURE 3. The structure of the proposed network.

Additionally, the above received signal in the guard space area can be expressed in matrix form as

$$\mathbf{Y}^{\mathrm{DD}} = \mathbf{h}_{\omega} + \mathbf{v}^{\mathrm{DD}}, \tag{19}$$

where $\mathbf{h}_{\omega}$ represents the matrix form of the channel response in the DD domain, and $\mathbf{v}^{\mathrm{DD}}$ denotes the Gaussian noise matrix. Specifically, the element in the k-th row and l-th column of $\mathbf{h}_{\omega}$ is defined as $\mathrm{H}_{\omega}[k, l] = h_w[k, l]$. Given this, we propose to employ DL to estimate the off-grid delay indices l_{τ_i} and Doppler shifts $k_{\nu_i} = k_i + \kappa_i$ directly from the received signals. Note that the off-grid delay indices correspond to the spatial position of the target in the DD domain, e.g., yellow squares in Fig. 3. The proposed DL-based approach can directly learn and extract underlying channel characteristics for sensing purposes, overcoming the limitations of traditional non-DL methods that heavily rely on assumptions of guide alignment and perfect channel sparsity.

2) THE PROPOSED DL-BASED METHOD FOR PARAMETERS ESTIMATION

The received signal $\mathbf{Y}^{\mathrm{DD}}$ is a complex-valued matrix, where we separate the real and imaginary components and use them as input features. Then the input of the proposed DL model is defined as

$$\mathbf{X}_{\mathrm{in}} = \left\{\mathrm{Re}\left\{\mathbf{Y}^{\mathrm{DD}}\right\}, \mathrm{Im}\left\{\mathbf{Y}^{\mathrm{DD}}\right\}\right\}, \tag{20}$$

where $\mathrm{Re}\{\mathbf{Y}^{\mathrm{DD}}\}$ and $\mathrm{Im}\{\mathbf{Y}^{\mathrm{DD}}\}$ indicate the real and imaginary parts of $\mathbf{Y}^{\mathrm{DD}}$, respectively. Subsequently, $\mathbf{X}_{\mathrm{in}}$ is fed into the DL model for training. To this end, we propose a PEnet consisting of two sub-modules: a residual convolutional neural network (RCNN) and a multi-layer perceptron (MLP). Figure. 3 illustrates the structure of our proposed model. The processed features are fed into the RCNN, as shown in the green dashed box. The RCNN comprises T residual blocks to extract spatial features from the 2D input. Each residual block contains N_L layers, where the operations in the 1st to $(N_L - 1)$th layers are identical and include "Conv", "'BN" and "ReLU" operations. Here, "Conv" represents the convolutional layer, and "BN" is a batch normalization layer to accelerate the training process and mitigate gradient vanishing. The ReLU function is employed as the activation function throughout the RCNN. The flow of data through the residual blocks is indicated by green arrows, highlighting the hierarchical feature extraction process within the RCNN module. After feature extraction through the RCNN, to generate the final output from the extracted features, the MLP processes the features to produce a $P_{\max} \times 2$ vector $\mathbf{v}_{\mathrm{p}} = [l_{\tau_1}, \ldots, l_{\tau_P}, k_{\nu_1}, \ldots, k_{\nu_P}]$, where $P_{\max}$ denotes the maximum number of paths, which constrains the shape of the output. The transition of data from the RCNN to the MLP is indicated by the purple arrow, emphasizing how the learned representations are processed in the final stage. We rewrite the output as

$$\mathbf{v}_{\mathrm{p}} = f_{\boldsymbol{\theta}}(\mathbf{X}_{\mathrm{in}}), \tag{21}$$

where $\boldsymbol{\theta}$ represents the trainable network parameters. Thus, the PEnet can adopt the stochastic gradient descent algorithm (e.g., Adam) to refine the parameters $\boldsymbol{\theta}$ to realize a well-trained model, which is capable of radar parameters.

Upon estimating delay and Doppler indices, the corresponding estimated distance d_i and relative velocity v_i between the $i-1$-th and i-th vehicle can be obtained by

$$d_i = \frac{\hat{l}_{\tau_i}}{M\Delta f} \times c, \quad v_i = \frac{\hat{k}_{\nu_i}}{NT} \times \frac{c}{f_c}, \tag{22}$$

where c represents the speed of light, and f_c is the carrier frequency.

C. RL-BASED DRIVING DECISIONS

As discussed previously, distance and relative velocity are crucial for ACC problems, providing Bob with essential information for its driving behavior. Towards that end, we propose to leverage RL to enable Bob to make optimal decisions based on a Markov decision process (MDP) principle. To be specific, as depicted in Fig. 4, at each time step t, the agent observes the current state s_t from the state space $\mathcal{S}$, which typically refers to the set of all possible states that the agent can be in. Then, it selects an action a_t from the action space $\mathcal{A}$ based on a specific rule or current policy π. The performed action will

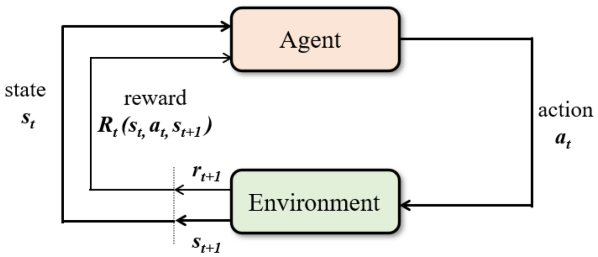


FIGURE 4. Markov decision process.

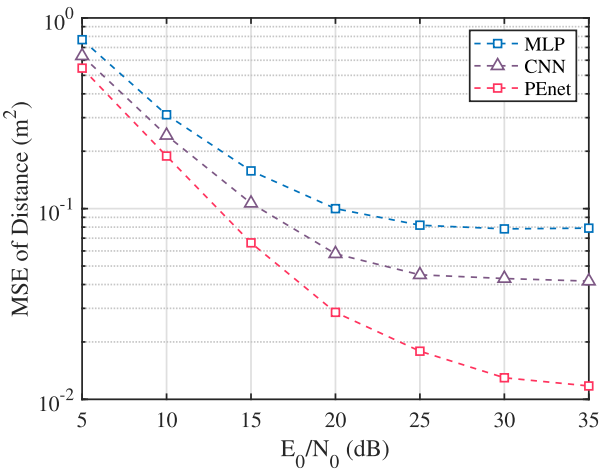


FIGURE 5. The MSE of distance with DL-based baselines.

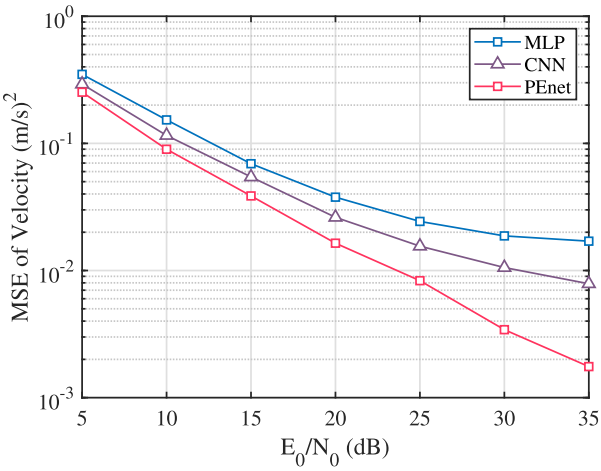


FIGURE 6. The MSE of velocity with DL-based baselines.

result in an immediate reward r_t from the environment and a transition to a new state s_{t+1}.

Specifically, Bob learns the value function for a given policy through interaction with Alice to determine an optimal solution. Also, based on the value function, Bob continuously develops and learns the optimal policy π^* to maximize the expectation of cumulative reward, $\mathbb{E}[\sum_{t=0}^{T-1} \gamma^t r_t]$, where γ is the discount factor that controls the weighting of future rewards on the cumulative return. The value function

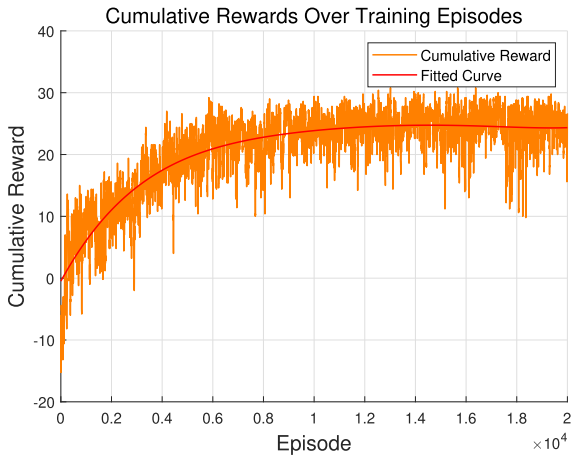


FIGURE 7. Cumulative rewards over training episodes with a fitted curve.

is the expected return (i.e., cumulative reward), which is a sum of the reward subject to Bob's states and is affected by Bob's current policy, given by $V^\pi(s) = \mathbb{E}_\pi[\sum_{t=0}^{T-1} \gamma^t r_t \mid s_t = s]$. Similarly, the corresponding action value function is $Q^\pi(s,a) = \mathbb{E}_\pi[\sum_{t=0}^{T-1} \gamma^t r_t \mid s_t = s, a_t = a]$. The relationship between $V(s)$ and $Q(s,a)$ can be represented using a state-transition-matrix, denoted as $P(s'|s,a)$, representing the probability of reaching the next state s' after taking the action a at state s. Then we have

$$Q^\pi(s,a) = \sum_{s' \in \mathcal{S}} P\left(s' \mid s, a\right)\left[R\left(s,a,s'\right) + \gamma V^\pi\left(s'\right)\right]. \quad (23)$$

To achieve an optimal decision, we investigate the Bellman Optimality Equation for the action-value function $Q^*(s,a)$, which is

$$Q^*(s,a) = \sum_{s' \in \mathcal{S}} P\left(s' \mid s, a\right)\left[R\left(s,a,s'\right) + \gamma \max_{a'} Q^*\left(s',a'\right)\right], \quad (24)$$

where $\max_{a'} Q^*(s',a')$ represents the maximum expected future reward for state s' considering all possible actions [31], [32].

To address the above problem, a variety of RL architectures have been employed, e.g., [33]. Among them, Q-learning demonstrates its promise due to its flexibility in reward structures and simple implementation. Consequently, this work utilizes Q-learning to seek the optimal action-value function $Q^*(s,a)$ by iteratively updating its estimate based on observed state transitions. The update rule can be expressed as

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[R\left(s,a,s'\right) + \gamma \max_{a'} Q\left(s',a'\right) - Q(s,a)\right], \quad (25)$$

where α is the learning rate that controls how much the newly observed value updates the existing q-values. By following

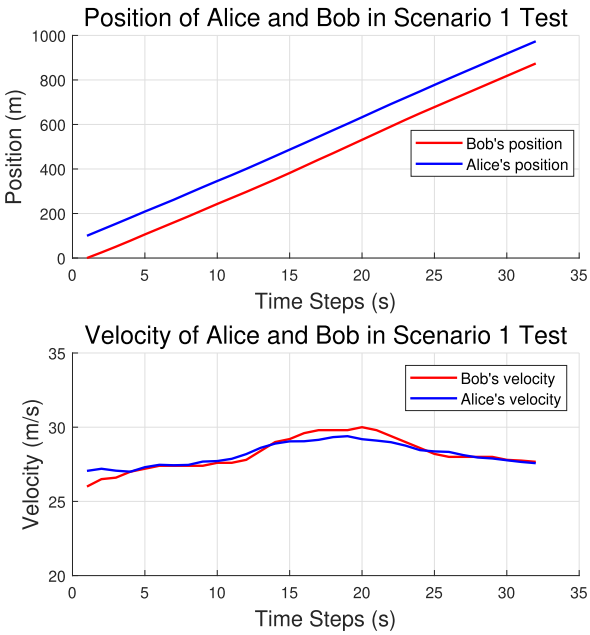


FIGURE 8. Distance and velocity of two vehicles in a scenario 1 test.

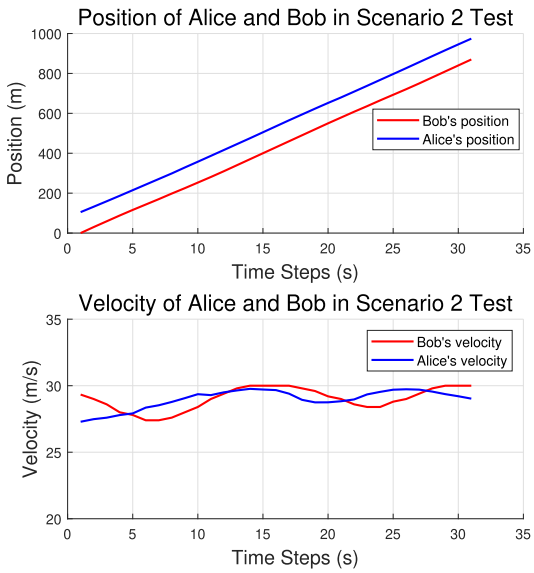


FIGURE 9. Distance and velocity of two vehicles in a scenario 2 test.

the iterative process, Q-learning can find the optimal policy by continuously refining its estimates of $Q^*(s, a)$.

The objective is to incentivize Bob to maintain a safe distance from Alice, while ensuring that the reward is updated based on the relative velocity and proximity to the ideal safe following distance. By continuously refining its estimates of $Q^*(s, a)$, Bob is trained to optimize his driving behavior based on the distance and relative velocity to Alice, ensuring safe following within the desired system constraints. Upon obtaining optimal $Q^*(s, a)$, Bob can effectively adapt to dynamic traffic conditions, maintaining a balance between safety, efficiency, and achieving integration with the OTFS-based ACC framework, ensuring reliable decision-making, even under variable vehicular dynamics, ultimately improving system performance.

In highway scenarios, the state is defined as a vector $\mathbf{s} = [d_{a,b}, v_{a,b}]$, where $d_{a,b}$ represents the distance between Alice and Bob, and $v_{a,b}$ denotes their relative velocity. The road length is constrained by $d_{\max}$, while the velocities of both vehicles are limited within the range $[v_{\min}, v_{\max}]$. To facilitate efficient learning and decision-making, the state space $\mathcal{S}$ is constructed by discretizing both the relative distance $d_{a,b}$ and the relative velocity $v_{a,b}$ into a finite set of intervals. This discretization ensures that Bob can effectively evaluate different driving conditions while maintaining computational feasibility. The action space $\mathcal{A}$ consists of three discrete acceleration choices: braking, maintaining a constant speed, and accelerating. These actions allow the agent to adaptively adjust Bob's motion in response to Alice's behavior while ensuring safe and efficient driving. In this setup, a reward function is designed to encourage Bob to maintain a safe distance from Alice, with penalties for both excessive proximity (risk of collision) and large deviations from the ideal safe distance. The reward is structured as follows. Bob receives a penalty, i.e., -2, if the distance between Alice and Bob is either less than a minimum safe distance or greater than a maximum following distance, indicating a risk of collision or excessive following distance. On the other hand, if the distance between Alice and Bob is within the safe range, and the relative velocity between them is small, Bob receives a positive reward, i.e., +1. Otherwise, the reward is 0.

IV. SIMULATION RESULTS

In this section, extensive simulations are conducted to verify the effectiveness of the proposed DL-based channel estimation and Q-learning algorithm for the ACC system in highway scenarios.

For the OTFS system, the size of each OTFS frame is $M \times N$, where $M = 12$ is the subcarrier numbers, and $N = 8$ is the time slots. The path number is 4, and the channel coefficients follow $\mathcal{CN}(0, \frac{1}{P})$ distribution. The maximum doppler index is $k_{\max} = 3$, and maximum delay index is $l_{\max} = 4$. The Δf is set to 7.5 kHz, and f_c is 3 GHz. The number of training samples is 10^4. The block number T is set to 4, and the layer number N_L is set to 12. The proposed approach is implemented on a desktop computer equipped with a Xeon Gold 6226R CPU operating at 2.9 GHz and an NVIDIA GeForce RTX 3090 GPU. Network training is performed using the Adam optimizer, with the learning rate initialized at 0.01. Additionally, we employ two metrics to evaluate the performance of the proposed parameter estimation approach, which are given by $\text{MSE}(\mathbf{d}) = \mathbb{E}[|\mathbf{d} - \hat{\mathbf{d}}|^2]$ and $\text{MSE}(\mathbf{v}) = \mathbb{E}[|\mathbf{v} - \hat{\mathbf{v}}|^2]$. Here, $\text{MSE}(\mathbf{d})$ and $\text{MSE}(\mathbf{v})$ are the MSE of the parameters distance and velocity. In Q-learning setting, training parameters such as learning rate (α), exploration rate (ε), and discount

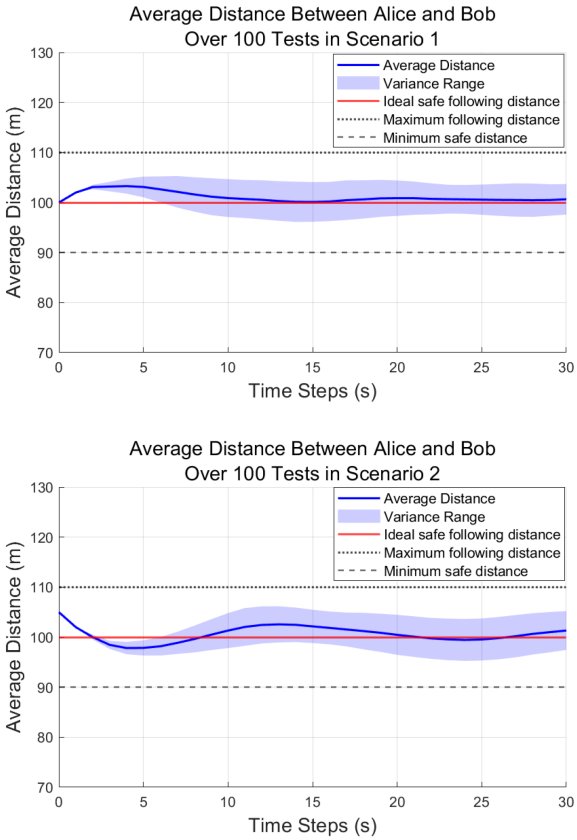


FIGURE 10. Average distance between Alice and Bob over 100 tests in two scenarios.

factor (γ) are set as 0.1, 0.01, and 0.9, respectively. Bob is trained over 20,000 episodes, where it iteratively learns the optimal policy based on the observed rewards. The cumulative reward during training provides insight into how well the agent improves its behavior over time.

As shown in Fig. 5, we demonstrate the MSE of distances of our proposed Resnet with another two DL-based baselines: MLP and CNN. The results illustrate that the MSE of the estimated delay and Doppler parameters decreases as the signal-to-noise ratio (SNR) increases. This behavior is expected, as higher SNR levels reduce the influence of noise. The MLP- and CNN-based methods exhibit suboptimal performance due to their limited ability to capture spatial features from the input data. In contrast, the proposed Resnet integrates the strengths of both CNN and MLP, by leveraging spatial feature extraction through RCNN and parameter regression through MLP. As a result, the proposed Resnet achieves significantly lower estimation errors, particularly at higher SNR values. Similar conclusions can be obtained for the velocity estimation, as depicted in Fig. 6.

Fig. 7 demonstrates the cumulative reward of our proposed Q-learning. As training progresses, the cumulative reward increases and stabilizes, achieving convergence around the 10,000 episode mark. This is evident as the fluctuations in reward decrease and the fitted curve levels out, indicating that Bob's performance is stabilizing. It is learning a more consistent strategy for maintaining safe following distances and making efficient speed adjustments in the simulated driving environment.

To further validate the effectiveness of our proposed Resnet, we consider two testing scenarios. In test scenario 1, Bob starts with a velocity significantly lower than Alice's, meaning Bob has to accelerate and adapt to Alice's pace. As shown in Fig. 8, although Bob's velocity is lower, he gradually speeds up to close the gap and ensures a safe following distance. In other words, Bob can adjust the velocity to safely maintain a distance from Alice while avoiding collisions. In test scenario 2, the initial distance between Alice and Bob is larger than the safe distance. Bob aims to follow Alice as closely as possible while avoiding excessive deviation from the ideal safe distance. As shown in Fig. 9, Bob reduces the distance between Alice without violating the safe distance threshold.

These two scenarios highlight how Bob handles different driving conditions, either adjusting to a slower pace or maintaining a safe following distance when starting farther apart. As shown in figure 10, the average distance between Alice and Bob stabilizes around the ideal safe following distance (i.e., 100 m) over time. The variance range indicates some fluctuations, but the distance stays within the safe range (i.e., [90 m, 110 m]).

V. CONCLUSION

In this paper, we proposed a DL-based channel estimation and a RL-based driving decision method to enhance the ACC system in IVNs. The developed DL approach can ensure communication reliability between vehicles, significantly reducing estimation errors in dynamic environments. The proposed RL method is able to optimize the decision-making process to maintain a safe following distance and adapt to varying traffic conditions. Our extensive simulations have demonstrated the feasibility and effectiveness of the proposed methods.

This work indicates a promising direction for more intelligent, adaptive, and safe automotive technologies, marking a significant step into fully autonomous vehicles. Future studies could explore the scalability of the proposed methods in more congested traffic scenarios with more complex and practical conditions such as pedestrian recognition and multi-lane navigation.

REFERENCES

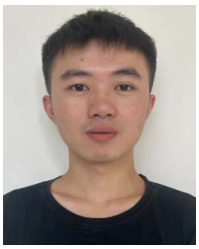
[1] L. Xiao and F. Gao, "A comprehensive review of the development of adaptive cruise control systems," *Veh. Syst. Dyn.*, vol. 48, no. 10, pp. 1167–1192, Oct. 2010, doi: 10.1080/00423110903365910.

[2] G. Stein, O. Mano, and A. Shashua, "Vision-based ACC with a single camera: Bounds on range and range rate accuracy," in *Proc. IEEE IV Intell. Vehicles Symp. Proc.*, Columbus, OH, USA, Jul. 2003, pp. 120–125.

[3] M. M. Zaman, I. H. Shibly, M. S. Hossain, and M. S. Rahman Shanto, "Advancements in radar technology for enhanced adaptive cruise control systems," in *Proc. 2024 IEEE Int. Conf. Power, Elect., Electron. Ind. Appl.*, Dhaka, Bangladesh, Sep. 2024, pp. 1–6.
[4] A. Abosekeen, T. B. Karamat, A. Noureldin, and M. J. Korenberg, "Adaptive cruise control radar-based positioning in GNSS challenging environment," *IET Radar Sonar Navig.*, vol. 13, no. 10, pp. 1666–1677, Oct. 2019. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-rsn.2019.0004
[5] G. R. Widmann et al., "Comparison of lidar-based and radar-based adaptive cruise control systems," *SAE Trans.*, vol. 109, pp. 126–139, 2000. [Online]. Available: http://www.jstor.org/stable/44699119
[6] P. Wei, Y. Zeng, W. Ouyang, and J. Zhou, "Multi-sensor environmental perception and adaptive cruise control of intelligent vehicles using kalman filter," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 3, pp. 3098–3107, Mar. 2024.
[7] Z. Chen, L. Zhang, J. Zhang, Z. Wu, and D. Luobu, "An OFDM-based pre-coded chaos shift keying transceiver for reliable V2V transmission," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6710–6715, Jun. 2022.
[8] W. Yuan, S. Li, Z. Wei, Y. Li, and P. Fan, "On hybrid detection of wireless communications over interference channels: A generalized framework," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 4, pp. 1214–1229, Apr. 2025.
[9] F. Xia et al., "Symbiotic sensing and communication: Framework and beamforming design," *IEEE Trans. Wireless Commun.*, vol. 24, no. 3, pp. 2417–2434, Mar. 2025.
[10] S. Arslan and M. Saritas, "The effects of OFDM design parameters on the V2X communication performance: A survey," *Veh. Commun.*, vol. 7, pp. 1–6, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214209616300286
[11] J. Wu, W. Yuan, Z. Wei, K. Zhang, F. Liu, and D. Wing Kwan Ng, "Low-complexity minimum BER precoder design for ISAC systems: A delay-doppler perspective," *IEEE Trans. Wireless Commun.*, vol. 24, no. 2, pp. 1526–1540, Feb. 2025.
[12] B. Li, W. Yuan, F. Liu, N. Wu, and S. Jin, "OTFS-based ISAC: How delay doppler channel estimation assists environment sensing?," *IEEE Wireless Commun. Lett.*, vol. 13, no. 12, pp. 3563–3567, Dec. 2024.
[13] S. Lu et al., "Integrated sensing and communications: Recent advances and ten open challenges," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19094–19120, Jun. 2024.
[14] J. Wu, W. Yuan, F. Liu, Y. Cui, X. Meng, and H. Huang, "UAV-based target tracking: Integrating sensing into communication signals," in *Proc. 2022 IEEE/CIC Int. Conf. Commun.*, Oct. 2022, pp. 309–313. [Online]. Available: https://api.semanticscholar.org/CorpusID
[15] K. Zhang et al., "Radar sensing via OTFS signaling: A delay doppler signal processing perspective," in *Proc. ICC 2023-IEEE Int. Conf. Commun.*, Oct. 2023, pp. 6429–6434. [Online]. Available: https://api.semanticscholar.org/CorpusID:256194058
[16] X. Wei, L. Zhang, W. Yuan, F. Liu, S. Li, and Z. Wei, "SDR System Design and Implementation on Delay-Doppler Communications and Sensing," in *Proc. IEEE Wireless Commun. Netw. Conf.*, May 2023, pp. 1–6.
[17] Q. Cheng, Z. Shi, D. N. Nguyen, and E. Dutkiewicz, "Sensing OFDM signal: A deep learning approach," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7785–7798, Nov. 2019.
[18] B. Luo, D. Liu, T. Huang, and D. Wang, "Model-free optimal tracking control via critic-only Q-learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 10, pp. 2134–2144, Oct. 2016.
[19] Z. Li, T. Chu, I. V. Kolmanovsky, and X. Yin, "Training drift counteraction optimal control policies using reinforcement learning: An adaptive cruise control example," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 9, pp. 2903–2912, Sep. 2018.
[20] M. Masmoudi, H. Ghazzai, M. Frikha, and Y. Massoud, "Autonomous car-following approach based on real-time video frames processing," in *Proc. 2019 IEEE Int. Conf. Veh. Electron. Saf.*, Cairo, Egypt, 2019, pp. 1–6.
[21] M. Masmoudi, H. Friji, H. Ghazzai, and Y. Massoud, "A reinforcement learning framework for video frame-based autonomous car-following," *IEEE Open J. Intell. Transp. Syst.*, vol. 2, pp. 111–127, May 2021.
[22] Y. Liu, Y. Shi, X. Zhang, J. Wu, and S. Yang, "Reinforcement learning-based car-following control for autonomous vehicles with OTFS," in *Proc. 2024 IEEE Wireless Commun. Netw. Conf.*, Dubai, UAE, Jul. 2024, pp. 1–6.
[23] W. Yuan, Z. Wei, S. Li, J. Yuan, and D. W. K. Ng, "Integrated sensing and communication-assisted orthogonal time frequency space transmission for vehicular networks," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1515–1528, Nov. 2021.
[24] X. Zhang, C. Liu, W. Yuan, J. A. Zhang, and D. W. K. Ng, "Sparse prior-guided deep learning for OTFS channel estimation," *IEEE Trans. Veh. Technol.*, vol. 73, no. 12, pp. 19913–19918, Dec. 2024.
[25] Y. Shi, B. Li, X. Wei, and H. Huang, "Experimental implementation of an OTFS-based ISAC system," in *Proc. 2024 IEEE/CIC Int. Conf. Commun.*, Hangzhou, China, Oct. 2024, pp. 587–592.
[26] X. He, W. Yuan, and P. Fan, "On the pilot-aided channel estimation for windowed OTFS with data interference in rapidly time-varying channels," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 16359–16374, Nov. 2024.
[27] X. Zhang, H. Wen, Z. Yan, W. Yuan, J. Wu, and Z. Li, "A novel joint channel estimation and symbol detection receiver for orthogonal time frequency space in vehicular networks," *Entropy*, vol. 25, no. 9, Sep. 2023, Art. no. 1358. [Online]. Available: https://www.mdpi.com/1099-4300/25/9/1358
[28] X. Zhang, W. Yuan, C. Liu, J. Wu, and D. W. K. Ng, "Predictive beamforming for vehicles with complex behaviors in ISAC systems: A deep learning approach," *IEEE J. Sel. Topics Signal Process.*, vol. 18, no. 5, pp. 828–841, Jul. 2024.
[29] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in delay–doppler channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4906–4917, May 2019.
[30] L. Tan, W. Yuan, X. Zhang, K. Zhang, Z. Li, and Y. Li, "DNN-based radar target detection with OTFS," *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 15786–15791, Oct. 2024.
[31] Z. Gao, T. Sun, and H. Xiao, "Decision-making method for vehicle longitudinal automatic driving based on reinforcement Q-learning," *Int. J. Adv. Robot. Syst.*, vol. 16, Jun. 2019, Art. no. 1729881419853185. [Online]. Available: https://api.semanticscholar.org/CorpusID
[32] J. Clifton and E. B. Laber, "Q-learning: Theory and applications," *Annu. Rev. Statist. Appl.*, vol. 7, pp. 279–301, Mar. 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID
[33] R. S. Sutton and A. G. Barto, "The reinforcement learning problem," in *Reinforcement Learning: An Introduction*, Cambridge, MA, USA: MIT Press, 1998, pp. 51–85.

YULIN LIU is currently working toward the B.E. degree with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China. Her research interests include integrated sensing and communications, orthogonal time frequency space, and deep reinforcement learning for wireless communications.

XIAOQI ZHANG (Student Member, IEEE) received the B.E. degree from the School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China, in 2021, and the M.S. degree from the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China, in 2023. He is currently working toward the Ph.D. degree with the Department of Electrical and Data Engineering, University of Technology Sydney, Sydney, Australia. His research interests include the area of integrated sensing and communications, orthogonal time frequency space, and machine learning for communication.



JUN WU received the B.E. degree from the School of Electronic Engineering, Southwest Jiaotong University, Chengdu, China, in 2021. He is currently working toward the Ph.D. degree with the School of Automation and Intelligent Manufacturing, Southern University of Science and Technology, Shenzhen, China. His research interests include the area of integrated sensing and communications, UAV communications, and convex optimization.

QINGQING CHENG (Member, IEEE) received the M.E. degree from the Harbin Institute of Technology, Harbin, China, in 2015, the Master of Research (M.Res.) degree from Macquarie University, Sydney, Australia, in 2016, and the Ph.D. degree from the University of Technology Sydney, Ultimo, Australia, in 2020. From 2020 to 2024, she was a Postdoctoral Research Fellow with the University of New South Wales, Sydney, Australia. She is currently a Lecturer (equivalent to Assistant Professor) with the Queensland University of Technology, Brisbane, Australia. Her research interests include deep learning for wireless communications, integrated sensing and communications, 5G/6G systems, high-mobility networks, privacy preservation, cognitive radio, and massive MIMO. She was the recipient of the Best Paper Award at IEEE Globecom 2023 and 2025 ARC Discovery Early Career Researcher Award (DECRA). Dr. Cheng is also an Associate Editor for IEEE TRANSACTIONS ON MOBILE COMPUTING.