Edge-Assisted Collaborative Perception Against Jamming and Interference in Vehicular Networks

Zhiping Lin, Liang Xiao, *Senior Member, IEEE*, Hongyi Chen, Zefang Lv, Yunjun Zhu, Yanyong Zhang, *Fellow, IEEE*, Yong-Jin Liu, *Senior Member, IEEE*

***Abstract*—Collaborative perception of connected autonomous vehicles (CAVs) that offload the sensing data, such as the feature map extracted from light detection and ranging (LiDAR) point clouds, to an edge device such as the roadside unit (RSU) to detect traffic objects has severe performance degradation due to the offloading latency and packet loss rate (PLR) under jamming and interference. In this paper, we propose an edge-assisted reinforcement learning (RL)-based collaborative perception scheme for CAVs to enhance the accuracy and speed against jamming and interference in LiDAR-based object detection. Based on the spatial confidence score of the feature map, the data size, the channel gains, the received jamming power and interference level, this scheme chooses the critical regions of the feature map, radio channel and transmit power with the hierarchical structure to enhance the learning efficiency. The risk level of the selected policy evaluates the time asynchronization and information loss of the shared feature map using the multi-level risk function based on multiple thresholds of the offloading latency and PLR, with assigning different penalties to mitigate the selection of high-risk policies that degrade perception performance. The upper performance bound in terms of the perception accuracy, latency and utility is provided based on the Stackelberg equilibrium of the game between the jammer and CAVs. Experimental results based on the Robosense RS-LiDAR-16 sensors and the Raspberry Pi to detect 10 vehicles in an $8.5\times4\times3.5$ m^3 area show the performance gain with 22.4% higher perception accuracy and 41.3% less latency compared with the benchmark against a smart jammer.**

***Index Terms*—Edge-assisted collaborative perception, LiDAR, vehicular networks, reinforcement learning, jamming.**

I. Introduction

Edge-assisted collaborative perception enables connected autonomous vehicles (CAVs) to exploit the computation resources of roadside units (RSUs) to process the sensing data, such as point clouds captured periodically by the light detection and ranging (LiDAR) sensors, to upgrade perception capabilities such as overcoming sparse observation and limited sensing range [1]–[3]. For example, CAVs equipped with Velodyne HDL32E LiDAR sensor to capture point cloud frames at 5-20 Hz, use the deep neural network (DNN) such as PointPillar to extract feature maps from point clouds, and offload the feature maps to the RSU that performs a data fusion model to create a global view for improving the perception range, such as detecting pedestrians or vehicles beyond the line-of-sight distance [4], [5].

Jammers in vehicular networks cause performance degradation and connection loss for the sensing data transmission (i.e., feature map or point clouds offloading), which potentially fails the vehicle control and leads to severe traffic accidents and road fatality. In addition, further attacks such as man-in-middle or denial of service can be launched by sending fake or replayed messages such as manipulated feature maps to the RSU or victim vehicles, which compromises the integrity, availability and confidentiality of the sensing data exchange in the collaborative perception system [6]–[8].

CAVs partition the sensing data based on the spatial confidence scores of feature maps or the vehicle locations, and only offload the important portions such as point clouds with high resolution or feature maps with higher confidence scores to balance the communication overhead and data quality [9]–[11]. For example, CAVs exploit the detection head module in [10] to divide the critical regions of feature maps based on the spatial confidence scores of objects and only share the critical regions among neighboring CAVs to improve the detection accuracy with less bandwidth consumption.

Reinforcement learning (RL)-based collaborative perception scheme in vehicular networks chooses the transmit power, the collaborative CAVs, the radio channel and the region of interest of the sensing data to improve the detection accuracy for the objects beyond the line-of-sight distance of individual sensing range [12]–[15]. For example, the seminal work in [12] applies the branching dueling Q-network to choose the region of interest of the sensing data and radio channels for collaborative CAVs based on the estimated locations and moving speeds of CAVs and the available network bandwidth to enhance the sensing range and accuracy. However, the severe channel fading, the interference and the jammers that block the offloading links lead to long latency and high packet loss rate (PLR) to degrade the perception performance.

In this paper, we propose an edge-assisted collaborative perception scheme in vehicular networks for CAVs to im-

This work was supported in part by the National Natural Science Foundation of China under Grant U21A20444, Grant 62332016, Grant 62332019 and Grant U2336214; and in part by the National Key Research and Development Program of China under Grant 2023YFB3107603. An earlier version of this paper was presented in part at IEEE GLOBECOM 2024 [14].(Corresponding author: Liang Xiao)

Zhiping Lin, Liang Xiao, and Yunjun Zhu are with the Department of Information and Communication Engineering, Xiamen University, Xiamen, China. (E-mail: lxiao@xmu.edu.cn)

Zefang Lv is with the Department of Information and Communication Engineering, Xiamen University, Xiamen, and also with the School of Computer Science, Fudan University, Shanghai, China

Hongyi Chen is with the Institute of Artificial Intelligence, Xiamen University, Xiamen 361005, China

Yanyong Zhang is with the School of Computer Science and Technology, University of Science and Technology of China, Hefei, China. (Email: yanyongz@ustc.edu.cn)

Yong-Jin Liu is with the Department of Computer Science and Technology, Tsinghua University, Beijing, China. (Email: liuyongjin@tsinghua.edu.cn)

prove the perception accuracy and speed against jamming and interference in LiDAR-based object detection. Based on the spatial confidence score and data size of the feature map, the channel gains between the CAV and RSU, the jamming signal strength and interference level to formulate the RL state, this scheme chooses the critic regions of each feature map, the radio channel and the transmit power to offload the selected regions to the RSU that performs an attention-based data fusion model to aggregate the received feature maps. Since the future state observed by each CAV is independent of the previous state if given the current state and policy, the repeated policy optimization for collaborative perception can be formulated as a Markov decision process (MDP), which is addressed by RL to maximize the expected long-term utility that formulated by a weighted sum of perception accuracy, latency and PLR.

The risk of time-asynchronization and information loss of the shared feature map is evaluated using the multi-level risk function based on multiple thresholds of the offloading latency and PLR, with assigning different penalties to mitigate the selection of high-risk policies that degrade perception performance during the RL learning process. The policy selection for collaborative perception is decomposed into three sub-policies based on the hierarchical RL structure with three decision layers for reducing the sample complexity of the action space to enhance the learning efficiency, in which each layer consists of a Q-table that evaluates the long-term utility and an E-table that evaluates the long-term risk to jointly formulate the policy distribution.

As the state space increases with the number of radio channels and spatial regions of the feature map, a deep RL version is also designed to extract the state features for the find-grained feature maps in large-scale vehicular networks. More specifically, neural networks are exploited to estimate the long-term rewards and risk levels for the policy selection, as well as to address the quantization error of the spatial confidence scores, the channel gains and the jamming and interference power for faster learning speed and higher perception performance.

The performance upper bound of collaborative perception is provided based on the Stackelberg equilibrium (SE) of the game between CAVs and jammer, in which CAVs as the leader choose regions of the feature map, radio channel and transmit power to enhance the perception accuracy and speed and the jammer as the follower determines the jamming channel and power to degrade the perception performance with less jamming energy consumption. The results show that the perception accuracy increases with the number of feature map regions if the channel gains between the RSU and CAVs exceed the lower bound depending on the jamming power and minimum signal-to-interference-plus-noise (SINR), which indicates that the selection of feature map regions fine-tunes the perception performance, bandwidth consumption and computational complexity.

Simulations are performed based on the feature fusion model in [16] and the V2X-Sim dataset in [17] show that our proposed schemes improve the perception accuracy and speed in the LiDAR-based object detection compared with the benchmark in [12]. In addition, we further conduct the experiments based on five CAVs equipped with Robosense RS-LiDAR-16 and Raspberry Pi-4B and an edge server with Intel i9-13900 CPU, NVIDIA GeForce RTX 3090 GPU to collaboratively detect vehicles, and provide a graphic user interface (GUI) to show the captured point cloud frames and detection results. A smart jammer equipped with a universal software radio peripheral (USRP) applies Q-learning to choose the jamming power up to 30 mW based on the estimated legitimate signal strength. Experimental results show that our proposed schemes outperform 22.4% perception accuracy and 41.3% less latency compared with the benchmark.

The main contributions are summarized as follows:

1) We propose the RL-based collaborative perception scheme to enhance the perception accuracy and speed for CAVs, with the hierarchical structure to reduce the sample complexity of the action space and the safe exploration to avoid the risky collaborative perception policies.
2) We provide the upper performance bound in terms of the perception accuracy, latency and utility based on the SE of the game between N CAVs and a jammer.
3) Experiments based on the Raspberry Pi and Robosense RS-LiDAR-16 sensors verify the performance gain of our proposed schemes compared with the benchmark.

This paper is organized as follows: we begin with a comprehensive review of previous literature in Section II, followed by an introduction of the system model in Section III. In Sections IV and V, we present the RL-based policy selection for collaborative perception. The upper bound of the perception performance is provided in Section VI, followed by the simulation and experimental results in Sections VII and VIII. Finally, the conclusion is summarized in Section IX.

II. Related Work

Cooperative perception in vehicular networks enables autonomous vehicles to overcome the physical limitations of onboard sensors such as line-of-sight sensing range by sharing the perceived information with nearby vehicles [18]–[23]. For instance, the collaborative perception scheme based on vehicle-to-vehicle (V2V) communications in [18] designs a spatial graph neural network to fuse the intermediate features received from neighboring vehicles to improve the motion forecasting performance and meet bandwidth requirements. In addition, an infrastructure-assisted object detection scheme proposed in [20] exploits the tracked motion trajectories of moving vehicles and pedestrians to align the multi-perspective representations of objects and mitigate the impact of asynchronous data frames in edge-assisted perception system.

Vehicular communications for collaborative perception have been studied to improve the communication efficiency in perceptual information sharing for enhancing the detection accuracy and robustness in road safety or autonomous driving applications [24]–[28]. In [24], the message generation rule and transmission frequency for vehicular collaborative perception are designed to reduce the communication load and overhead of sensing data sharing, and thus improving perception reliability. The lossy communication-aware feature

fusion model is proposed in [26] to address the negative effects of V2V communications such as packet loss in feature sharing, which designs an inter-vehicle attention module to capture the uncertainties among collaborative vehicles caused by latency to improve the accuracy.

The performance of collaborative perception depends on the bandwidth, channel gains and interference of the communication links for sensing data exchanging, which results in network congestion and high latency to decrease the perception performance [25], [29], [30]. For example, a three handshake-based collaborative perception scheme is proposed in [29] to enable the target robot to determine the most valuable collaborative robots based on the compressed request messages, and thus improving the perception capabilities and reducing bandwidth requirements. To reduce the transmission data volume, the vehicle detection results are integrated into the bird-eyes-view of point clouds, which is shared among nearby vehicles to improve accuracy with less communication overhead [30].

RL-based mobile edge computing against jamming as presented in [31]–[33] chooses the computational nodes, the offloading rate, the radio channel and the transmit power based on the channel gains, bandwidth and service requirements to enhance the offloading performance. For example, in [31], the actor-critic algorithm is applied to choose the offloading rate, the edge nodes and the transmit power based on the channel gains, the received signal strength indicator (RSSI) and SINR to reduce the offloading latency and energy consumption. However, these schemes have performance degradation in LiDAR-based object detection due to the lack of perception features such as the data quality (i.e., confidence scores of the feature maps or point density of the point clouds).

Furthermore, our previous work in [14] proposes an RL-based sub-frame selection and power control scheme for each CAV to broadcast feature maps to improve perception accuracy in LiDAR-based object detection. Compared with the previous work, this paper proposes the RL-based collaborative perception schemes to choose the regions of feature map, radio channels and transmit power against jamming and interference. The deep RL version is also proposed to address the state quantization errors of the spatial confidence scores and channel gains, which further improves the perception accuracy and speed. The performance upper bound is also provided based on the SE of the game between N CAVs and the jammer. Experiments based on 5 CAVs equipped with LiDAR sensors and Raspberry Pi to offload feature maps to an edge server are conducted to show the efficacy of our proposed schemes against a smart jammer.

III. System Model

A. Task model

As shown in Fig. 1, both the RSU and N CAVs denoted by $\mathcal{N} = \{1, 2, \cdots, N\}$ equipped with LiDAR sensors capture point clouds for collaboratively detecting traffic objects such as pedestrians and vehicles, in which the RSU performs the object detection based on the received feature maps and provides perception results to CAVs for improving traffic efficiency and road safety. At time slot k, RSU transforms the point clouds into a bird's-eye-view (BEV) map via the voxelization operation that quantizes the point clouds into regular cells, and further exploits the anchor-based PointPillar backbone network to extract feature map $\boldsymbol{F}^{(k)} \in \mathbb{R}^{H \times W \times C}$ from the BEV map, where H, W and C denote the height, width and the number of feature channels [4], [5].

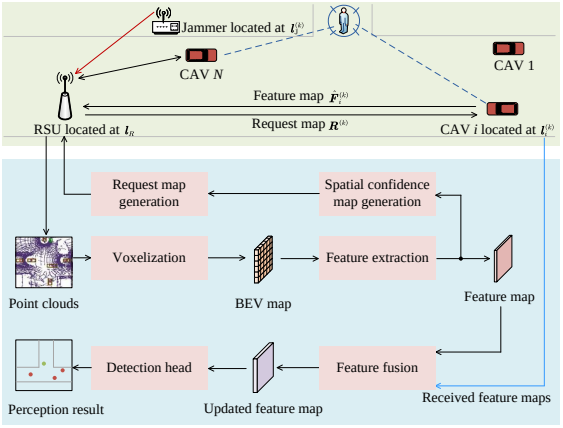


Fig. 1. Overview of collaborative perception in vehicular networks, in which the RSU aggregates feature maps via a self-attention-based feature fusion mechanism to enhance accuracy.

The spatial confidence map $\boldsymbol{C}^{(k)} \in \mathbb{R}^{H \times W}$ is generated based on the detection head module in [10] that indicates the confidence for correctly detecting objects at each spatial region of the feature map $\boldsymbol{F}^{(k)}$, which in turn is used to calculate request map $\boldsymbol{R}^{(k)} = \mathbf{1} - \boldsymbol{C}^{(k)}$ to represent the uncertainty for detecting objects based on the individual observations, i.e., the region with a higher value in the request map represents the lower probability to accurately detect the object. RSU located at $\boldsymbol{l}_{\mathrm{R}}$ broadcasts the request map $\boldsymbol{R}^{(k)}$ to neighboring CAVs as the basis for evaluating the importance of feature maps.

CAV $i \in \mathcal{N}$ extracts the feature map $\boldsymbol{F}_i^{(k)}$ and divides the feature map into M regions denoted by $\mathcal{M}$ each with $\varpi_{i,m}^{(k)}$-bits based on the confidence score thresholds [10] or feature selective module [11]. For example, CAV i uses the anchor-based PointPillar backbone network to extract a feature map with the dimension of 64×64×256 that each element is denoted by a float32 data type, which yields an approximate 1-MB data for each region if the feature map is equally divided into four regions. According to [10], the spatial confidence score for each region of the feature map denoted by $\boldsymbol{o}_i^{(k)} = \left[o_{i,m}^{(k)}\right]_{1 \le m \le M}$ that indicates the importance level for the collaborative object detection is evaluated by the element-wise dot-product of the m-th region of the request map $\boldsymbol{R}_m^{(k)}$ and the spatial confidence map $\boldsymbol{C}_{i,m}^{(k)}$.

RSU exploits a channel-wise self-attention mechanism as presented in [16] to aggregate the extracted feature map $\boldsymbol{F}^{(k)}$ and the received feature maps $\hat{\boldsymbol{F}}_i^{(k)}$ based on the trainable collaboration attention weights $\boldsymbol{W}_i^{(k)}$ and $\boldsymbol{W}^{(k)}$ given by

$$\boldsymbol{H}^{(k)} = \sum_{i=1}^{N} \boldsymbol{W}_i^{(k)} \odot \hat{\boldsymbol{F}}_i^{(k)} + \boldsymbol{F}^{(k)} \odot \boldsymbol{W}^{(k)}, \tag{1}$$

where $\odot$ represents the element-wise dot-product operation. Finally, the detection head module up-samples the updated feature map $\boldsymbol{H}^{(k)}$, and classifies the foreground-background categories and regresses the bounding boxes to generate perception results $\boldsymbol{z}^{(k)}$ that consists of location, orientation, size and confidence score of detected objects. The perception tasks consist of the bounding box detection that detects the existence of objects at the specific regions and the object classification that estimates the probabilities of each class for the detected objects [34]. According to [35], the perception accuracy can be represented by the averaged confidence scores of the objects, which are obtained by the calibrated detection DNN model via the train-time auxiliary loss formulation to indicate the probability of the presence of the objects.

B. Communication Model

CAV $i \in \mathcal{N}$ located at $\boldsymbol{l}_i^{(k)}$ chooses the region of the feature map $\boldsymbol{x}_i^{(k)} = [x_{i,m}^{(k)}]_{m\in\mathcal{M}} \in \{0,1\}^M$, the uplink radio channel $a_i^{(k)}$ out of the F channels each with B hertz and the transmit power $p_i^{(k)}$ up to P_{V} mW with L_1 levels for offloading the selected regions of the feature map against jamming and interference. The RSU performs the feature fusion model given by Eq. (1) to obtain detection results $\boldsymbol{z}^{(k)}$, and further evaluates perception accuracy $\rho^{(k)}$, offloading latency $\tau_i^{(k)}$ and PLR $b_i^{(k)}$, which are sent back to CAVs.

In addition, the offloading latency of the feature maps from CAV i is evaluated based on the labeled timestamps on each transmitted packet, thus, the perception latency $t^{(k)}$ is evaluated based on the offloading and inference latency $\hat{\tau}^{(k)}$ given by $t^{(k)} = \max_{i\in\mathcal{N}} \tau_i^{(k)} + \hat{\tau}^{(k)}$, where the inference latency depends on the computational complexity of the feature fusion model and the computational capability of the edge device. The packet loss rate is evaluated based on the ratio of received packets in each time slot, in which the SINR of each packet usually should be greater than a predefined threshold γ_0 corresponding to modulation modes for successfully decoding.

According to [36], the channel gain $h_i^{(k)}$ between the RSU and CAV i depends on the reference path-loss h_0, the path-loss exponent α, the communication distance $d_i^{(k)}$ and the shadow fading X_σ that modeled by a zero mean normal distribution with a standard deviation of σ given by

$$h_i^{(k)} = h_0 \left(d_i^{(k)}\right)^{\alpha} X_\sigma \tag{2}$$

C. Jamming and Interference Model

Jamming attacks in vehicular networks lead to the performance degradation and connection loss for the CAVs' transmission, in which the jamming policy can be chosen randomly (e.g., random jammer) or based on the features of the vehicular links (e.g., reactive jammer) [6]. In the collaborative perception system, the jammer located at $\boldsymbol{l}_{\mathrm{J}}^{(k)}$ tracks the communication links of the feature map offloading, and estimates the RSSI of each link $\boldsymbol{\psi}^{(k)} = [\psi_l^{(k)}]_{1\le l\le F}$ via energy detection technique.

As an advanced type of reactive jamming in [8], the RL-based smart jammer applies the Q-learning algorithm to select the jamming power $y_1^{(k)} \in \{jP_{\mathrm{J}}/L_2 | 0 \le j \le L_2\}$, the center frequency $y_2^{(k)} \in \left[\underline{f}, \bar{f}\right]$ and the bandwidth $y_3^{(k)} \in (0, \hat{B}]$ based on the estimated RSSI $\boldsymbol{\psi}^{(k)}$ to formulate the RL state, where P_{J}, L_2, $\underline{f}$, $\bar{f}$ and $\hat{B}$ refer to the maximum and quantization level of the jamming power, the frequency range and maximum bandwidth of the jammer, respectively. The jamming utility u_{J} aims at decreasing the SINR $\gamma_i^{(k)}$ of each offloading link with less energy cost given by

$$u_{\mathrm{J}}^{(k)} = -\sum_{i=1}^{N} \gamma_i^{(k)} - c_{\mathrm{J}} y_1^{(k)}, \tag{3}$$

where c_{J} weighs the importance of energy consumption. The channel gain between the jammer and RSU is denoted by $g^{(k)}$.

According to [37], the interference level of each offloading link depends on the neighboring CAVs $\mathcal{N}_{-i}$, the transmit power $p_j^{(k)}$ and the interference factor $\varsigma_j^{(k)}$ and the channel gains $h_j^{(k)}$ given by

$$I_i^{(k)} = \sum_{j\in\mathcal{N}_{-i}} \varsigma_j^{(k)} p_j^{(k)} h_j^{(k)}, \tag{4}$$

where the interference factor depends on the channel selection of neighboring CAVs and the collision avoidance schemes of vehicular communication protocols. For example, the dedicated short-range communication protocol as presented in [38] specifies the binary backoff-based collision avoidance scheme, in which the interference factor depends on the number of backoff stages and the size of backoff windows. Without confusion, the superscript k will be omitted in the following.

D. Problem Formulation

For the collaborative perception system, the accuracy depends on the data quality of feature maps received from CAVs and the DNN models such as the feature fusion and inference model for LiDAR-based object detection, in which the data quality increases with the confidence scores of the feature maps $\boldsymbol{O} = [\boldsymbol{o}_i]_{i\in\mathcal{N}}$ and decreases with the latency t according to [10], [19]. Each feature map is divided into M regions each with ϖ_m-bits, and thus the data size for a transmitted packet is given by $\sum_{m=1}^{M} \varpi_m x_{i,m}$. According to [39], the inference latency $\hat{\tau}$ depends on the computational loads, the CPU clock rate ζ and the CPU cycles ψ for per-bit processing, as well as the offloading latency $\boldsymbol{\tau} = [\tau_i]_{i\in\mathcal{N}}$ depending on the data size of the selected regions and the data rate of each CAV.

CAVs optimize the selection of feature map regions $\boldsymbol{\chi} = [x_{i,m}]_{i\in\mathcal{N},m\in\mathcal{M}}$, the transmit power $\boldsymbol{p} = [p_i]_{i\in\mathcal{N}}$ and the radio channel $\boldsymbol{a} = [a_i]_{i\in\mathcal{N}}$ to maximize the object function for collaborative perception that is modeled by the weighted sum of the accuracy ρ, latency t and PLR b_i given by

$$\max_{p_i,a_i,x_{i,m}} \rho - c_{\mathrm{t}} t - \frac{c_{\mathrm{b}}}{N} \sum_{i=1}^{N} \mathcal{I}\left(\gamma_i < \gamma_0\right) \tag{5a}$$

$$\text{s.t.} \quad \rho = \mathcal{F}\left(\boldsymbol{O}, \boldsymbol{\tau}, \hat{\tau}\right) \tag{5b}$$

$$t = \max_{i\in\mathcal{N}} \frac{\sum_{m=1}^{M} \varpi_m x_{i,m}}{B\log\left(1+\gamma_i\right)} + \sum_{i=1}^{N}\sum_{m=1}^{M} \frac{\psi \varpi_m x_{i,m}}{\zeta} \tag{5c}$$

$$\gamma_i = \frac{p_i h_i}{\sum\limits_{j\in\mathcal{N}_{-i}} \varsigma_j p_j h_j + \kappa y_1 g + N_0} \tag{5d}$$

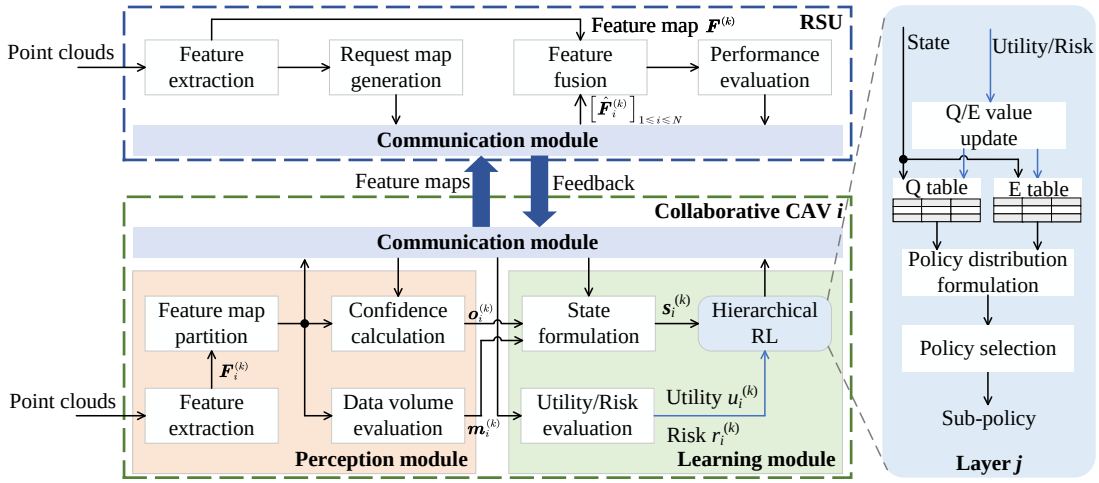


Fig. 2. Illustration of RL-based collaborative perception in vehicular networks, in which CAV i exploits a hierarchical RL in the learning module to choose the feature map regions, radio channel and transmit power against jamming and interference.

$$x_{i,m} \in \{0,1\}, \forall i \in \mathcal{N}, \forall m \in \mathcal{M} \tag{5e}$$

$$a_i \in \{1, \cdots, F\} \tag{5f}$$

$$p_i \in \{lP_{\mathrm{V}}/L_1 | 1 \le l \le L_1\} \tag{5g}$$

where $\mathcal{I}(\cdot)$ is the indicator function that takes value 1 if the argument is true.

The implicit function $\mathcal{F}(\cdot)$ in Eq. (5b) depends on the DNN model for the LiDAR-based object detection with the inputs of received feature maps, in which the accuracy in turn is determined by the collaborative perception policies among N CAVs. The latency t given by Eq. (5c) depends on the maximum offloading latency and the inference latency, which in turn is determined by the SINR γ_i of each packet given by Eq. (5d) that decreases with interference level from neighboring CAVs and the received jamming signal strength. The binary parameter κ takes value 1 if the jammer chooses the same radio channel as CAV i and 0 otherwise.

The optimal selection of the feature map, radio channel and transmit power against jamming and interference for N CAVs in Eq. (5) relies on the confidence score of feature maps, the received jamming power, interference level and channel gains of each offloading link, which are seldom accurately known in time by each CAV in the practical collaborative perception system. The RL-based scheme enhances the perception accuracy and speed via trial-and-error since the collaborative perception can be modeled as an MDP, in which the future state observed by each CAV is independent of the previous state if given the current perception policy and state.

IV. RL-based Collaborative Perception Against Jamming and Interference

The distributed RL-based collaborative perception scheme (RLCP) is proposed for each CAV to enhance the perception accuracy and speed of LiDAR-based object detection against jamming and interference. Based on the spatial confidence scores, the channel gains and the estimated jamming power and interference level of each channel, this scheme chooses the regions of feature maps, the radio channel and transmit power with the hierarchical structure to reduce the sample complexity of the action space. In addition, the risk level of the selected policies is evaluated using the multi-level risk function based on multiple thresholds of the offloading latency and packet loss rate, with assigning different penalties to mitigate the selection of high-risk policies that degrade perception performance.

As shown in Fig. 2, the collaborative perception system consists of three modules for performing the perception, learning and communications, in which the RSU extracts feature map $\boldsymbol{F} \in \mathbb{R}^{H \times W \times C}$ from captured LiDAR point clouds via the anchor-based PointPillar backbone network, and further calculates the request map $\boldsymbol{R}$ that indicates the uncertainty of detecting an object at each spatial region. CAV $i \in \mathcal{N}$ extracts and further divides the feature map $\boldsymbol{F}_i$ into M regions each with $\varpi_{i,m}$-bits, and calculates the spatial confidence score $\boldsymbol{o}_i$ based on the received request map to indicate the importance level of each region for the collaborative object detection.

Based on the data size $\boldsymbol{\varpi} = [\varpi_{i,m}]_{1\le i\le N, 1\le m\le M}$ and spatial confidence score $\boldsymbol{O} = [o_{i,m}]_{1\le i\le N, 1\le m\le M}$ of feature maps, the channel gains h_i between the CAV and RSU via the pilot-based channel estimation scheme, the received jamming power and interference level $\boldsymbol{J} = [J_l]_{1\le l\le F}$ of each offloading link via the energy detection technique such as [40], the current state $\boldsymbol{s}_i^{(k)}$ is formulated by,

$$\boldsymbol{s}_i^{(k)} = \left[\boldsymbol{O}, \boldsymbol{\varpi}, h_i, \boldsymbol{J}, \rho\right] \tag{6}$$

where the confidence score and the data size of the feature map, the channel gain, the received jamming signal strength and interference level and the perception accuracy are quantized into D_1, D_2, D_3, D_4 and D_5 levels, respectively.

Unlike the centralized RL that the state and action spaces increase exponentially with the network scale, this scheme designs the distributed RL with the three-layer hierarchical structure as shown in Fig. 2 to reduce the sample complexity of the action space, which decomposes the selection of feature map regions, radio channel and transmit power into three sub-policies to improve the learning efficiency. Each policy decision layer consists of a Q-table to evaluate the long-term

reward for each state-action pair and an E-table to evaluate the long-term risk of the selected policy.

More specifically, based on the state $\boldsymbol{s}_i^{(k)}$ as the input, the layer $j \in \{1,2,3\}$ of hierarchical RL outputs $Q\left(\boldsymbol{s}_i^{(k)}, \boldsymbol{x}_j'\right)$ and $E\left(\boldsymbol{s}_i^{(k)}, \boldsymbol{x}_j'\right)$ to formulate the policy distribution for selecting the sub-policies given by

$$\pi\left(x_j'|\boldsymbol{s}_i^{(k)}, \varphi_{j-1}\right) = \frac{\exp\left(\frac{Q\left(\boldsymbol{s}_i^{(k)}, \varphi_{j-1}, x_j'\right)}{\xi_j E\left(\boldsymbol{s}_i^{(k)}, \varphi_{j-1}, x_j'\right)+1}\right)}{\sum\limits_{\hat{x}\in\boldsymbol{X}_j} \exp\left(\frac{Q\left(\boldsymbol{s}_i^{(k)}, \varphi_{j-1}, \hat{x}\right)}{\xi_j E\left(\boldsymbol{s}_i^{(k)}, \varphi_{j-1}, \hat{x}\right)+1}\right)}, \quad (7)$$

where φ_j represents the selected sub-policy from the layer j of hierarchical RL, and thus $\varphi_0 = \varnothing$, $\varphi_1 = \boldsymbol{x}_i$ and $\varphi_2 = a_i$, respectively. Therefore, the action space $\boldsymbol{X}_j$ for each sub-policy is given by $\boldsymbol{X}_1 = \{0,1\}^M$, $\boldsymbol{X}_2 = \{1, \cdots, F\}$ and $\boldsymbol{X}_3 = \{lP_{\mathrm{V}}/L_1 | 1 \le l \le L_1\}$, respectively. Instead of the traditional ϵ-greedy-based policy exploration method, the modified Boltzmann distribution is formulated based on the proportional weighting factor $\boldsymbol{\xi} = [\xi_j]_{1\le j\le 3}$ to tradeoff the importance of policy reward and risk level. The policy distribution in Eq. (7) tends to choose the sub-policy with a higher Q-value and a lower risk for reducing the perception performance degradation, in which the risk level of each policy is evaluated based on the offloading latency and PLR.

After receiving the feature maps, RSU performs the self-attention-based feature fusion model to obtain the detection result $\boldsymbol{z}$, and further evaluates the perception accuracy ρ, offloading latency τ_i and PLR b_i, which in turn are sent back to each CAV. In particular, the accuracy can be estimated based on the average confidence score that indicates the probability of the presence of the objects at each spatial region [34], [35], as well as the perception latency is calculated based on the offloading and inference latency given by Eq. (5c). The PLR is calculated based on the rate of received packets over the total transmitted packets in each time slot. Note that the SINR of the received packets usually should be greater than a predefined threshold γ_0 corresponding to the specific modulation mode to successfully decode the packet. The utility u_i is given by

$$u_i = \rho - c_{\mathrm{t}}\tau_i - c_{\mathrm{b}} b_i, \quad (8)$$

where c_{t} and c_{b} represent the importance of perception latency and PLR. The Q-values in each layer $j \in \{1,2,3\}$ of the hierarchical RL are updated iteratively via the Bellman equation as follows,

$$\begin{aligned} Q\left(\boldsymbol{s}^{(k)}, \varphi_{j-1}, x_j\right) = (1-\vartheta) Q\left(\boldsymbol{s}^{(k)}, \varphi_{j-1}, x_j\right) + \\ \vartheta\left(u_i + \delta \max_{x_j'\in\boldsymbol{X}_j} Q\left(\boldsymbol{s}^{(k+1)}, \varphi_{j-1}, x_j'\right)\right), \end{aligned} \quad (9)$$

where the learning rate $\vartheta \in (0,1]$ weighs the current communication experience and the discount factor $\delta \in (0,1]$ evaluates the importance of the future utility [41].

In addition, the policy risk r_i to result in the time asynchronization and information loss of the feature maps is evaluated with the weighted sum of the offloading latency τ_i and PLR b_i, which are quantized into C_1 and C_2 levels with the different penalty factors $c_{l,1}$ and $c_{l,2}$ associating with multiple thresholds $\mu_{l,1}$ and $\mu_{l,2}$ given by

$$r_i = \sum_{l=0}^{C_1} c_{l,1} \mathcal{I}\left(\tau_i > \mu_{l,1}\right) + \sum_{l=0}^{C_2} c_{l,2} \mathcal{I}\left(b_i > \mu_{l,2}\right), \quad (10)$$

The E-values in each layer of the hierarchical structure are updated based on the risk function to maximize the expected long-term discounted risks, that is, the risk levels from previous L steps are used to evaluate the long-term risks in policy exploration given by

$$E\left(\boldsymbol{s}^{(k)}, \varphi_{j-1}, x_j'\right) = \sum_{l=0}^{L} \varrho^l r_i^{(k-l)}, \quad (11)$$

where ϱ weighs the discounted previous risk levels.

Algorithm 1: RL-based collaborative perception against jamming and interference for CAV i

Initialize ϑ, β, δ, $\tau^{(0)}$, $\rho^{(0)}$ and $b^{(0)}$;
for $k = 1, 2, ..., K$ **do**
  Calculate the spatial confidence scores $o_{i,m}$ and data size $\varpi_{i,m}$ of the feature map;
  Estimate the received jamming signal strength and interference level J_l of each radio channel;
  Formulate $\boldsymbol{s}_i^{(k)}$ via Eq. (6);
  Obtain $Q\left(\boldsymbol{s}_i^{(k)}, \boldsymbol{x}'\right)$ and $E\left(\boldsymbol{s}_i^{(k)}, \boldsymbol{x}'\right)$ from the first layer;
  Select regions of the feature map $\boldsymbol{x}_i$ via Eq. (7)to formulate $\hat{\boldsymbol{F}}_i$;
  Formulate $[\boldsymbol{s}_i^{(k)}, \boldsymbol{x}_i]$ to choose the radio channel a_i;
  Formulate $[\boldsymbol{s}_i^{(k)}, a_i]$ to choose the transmit power p_i;
  Offload the feature map $\hat{\boldsymbol{F}}_i$ to the RSU;
  Receive $\boldsymbol{z}$, τ_i, b_i and ρ from the RSU;
  Calculate u_i via Eq. (8);
  Evaluate r_i via Eq. (10);
  for $j = 1, 2, 3$ **do**
    Update Q-values via Eq. (9);
    Update E-values via Eq. (11);
  end
end

V. DRL-based Collaborative Perception Against Jamming and Interference

We also propose a deep RL version, named DRLCP, for CAVs to choose the feature map regions, radio channels and transmit power against jamming and interference for the fine-grained regions of feature maps in large-scale vehicular networks. More specifically, neural networks are exploited to extract state features that increase with the number of collaborative CAVs, regions of feature maps and radio channels, as well as to address the quantization errors for the spatial confidence scores and the channel gains, thus accelerating the learning efficiency and improving the detection accuracy.

As shown in Fig. 3, the current state $\boldsymbol{s}_i^{(k)}$ is formulated via Eq. (6) to input the Q-network (QNN) and E-network

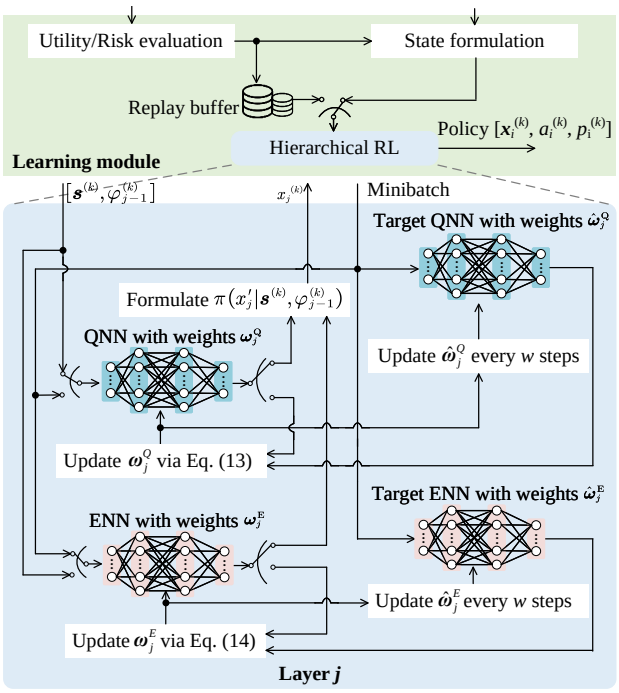


Fig. 3. Illustration of deep RL-based collaborative perception scheme for each CAV, which designs neuron networks to address the quantization error of state space.

Algorithm 2: DRL-based vehicular communication scheme for collaborative perception

Initialize ϑ, β, δ, $\tau^{(0)}$, $\rho^{(0)}$, $b^{(0)}$;
for $k = 1, 2, ..., K$ **do**
  Same as Lines 3-5 in Algorithm 1;
  Obtain $Q\left(\boldsymbol{s}_i^{(k)}, \boldsymbol{x}'; \boldsymbol{\omega}_1^{\mathrm{Q}}\right)$ and $E\left(\boldsymbol{s}_i^{(k)}, \boldsymbol{x}'; \boldsymbol{\omega}_1^{\mathrm{E}}\right)$;
  Formulate $\pi\left(\boldsymbol{x}'|\boldsymbol{s}_i^{(k)}\right)$ via Eq. (7) to select $\boldsymbol{x}_i$;
  Formulate $\left[\boldsymbol{s}_i^{(k)}, \boldsymbol{x}_i\right]$ to choose the radio channel a_i via Eq. (7);
  Formulate $\left[\boldsymbol{s}_i^{(k)}, a_i\right]$ to choose the transmit power p_i via Eq. (7);
  Same as Lines 10-13 in Algorithm 1;
  Formulate the collaborative perception experience $\boldsymbol{e}^{(k)}$ and store it in the memory pool;
  Sample a minibatch from the replay buffer $\mathcal{D}$;
  for $j = 1, 2, 3$ **do**
    Update $\boldsymbol{\omega}_j^{\mathrm{Q}}$ via Eq. (12);
    Update $\boldsymbol{\omega}_j^{\mathrm{E}}$ via Eq. (13);
  end
  if $k \bmod w = 0$ **then**
    for $j = 1, 2, 3$ **do**
      Update $\hat{\boldsymbol{\omega}}_j^{\mathrm{Q}}$ by copying QNN weights;
      Update $\hat{\boldsymbol{\omega}}_j^{\mathrm{E}}$ by copying ENN weights;
    end
  end
end

(ENN) with full-connected (FC) layers in the first level of hierarchical RL. The input layer of QNN in the first level consists of $\Omega_1 = 2M + 2F + 2$ neurons that depend on the number of neighboring CAVs, feature map regions and radio channels, and the output layer consists of 2^M neurons to obtain the estimated long-term utility $Q\left(\boldsymbol{s}_i^{(k)}, \boldsymbol{x}'; \boldsymbol{\omega}_1^{\mathrm{Q}}\right)$ and policy risk $E\left(\boldsymbol{s}_i^{(k)}, \boldsymbol{x}'; \boldsymbol{\omega}_1^{\mathrm{E}}\right)$ for the region selection. Two FC layers that includes $f_{1,1}$ and $f_{2,1}$ ReLU activation neurons extract the state features for less sample computational complexity compared to deep Q-learning presented in [41] with several convolutional layers.

Similarly, the inputs of the second and third layers of the hierarchical structure have $\Omega_2 = 3M + 2F + 2$ and $\Omega_3 = 2M + 2F + 3$ neurons for the state $\left[\boldsymbol{s}_i^{(k)}, \boldsymbol{x}_i\right]$ and $\left[\boldsymbol{s}_i^{(k)}, a_i\right]$ as the inputs, respectively. The radio channel a_i and transmit power p_i are chosen based on the outputs with F and L_1 neurons to formulate the policy distribution via Eq. (7). Each CAV evaluates the utility u_i via Eq. (8) and risk level r_i via Eq. (10) based on the latency τ_i, PLR b_i and accuracy ρ.

The collaborative perception experiences are formulated by $\boldsymbol{e}^{(k)} = \left[\boldsymbol{s}_i^{(k)}, \boldsymbol{x}_i, a_i, p_i, u_i, r_i, \boldsymbol{s}_i^{(k+1)}\right]$, which is stored in the replay buffer, i.e., $\mathcal{D} \leftarrow \mathcal{D} \cup \boldsymbol{e}^{(k)}$. A minibatch with Z experiences is uniformly sampled from the replay buffer for updating the weights of the neural network, i.e., $\mathcal{B} = \left\{\boldsymbol{e}^{(z)} | 0 \le z \le Z\right\}$, which is the input of the neural networks in three layers of hierarchical RL. Specifically, the target QNN outputs the maximum Q-value for the collaborative perception policy under the current state to update the weights of QNN $\boldsymbol{\omega}_{1\le j\le 3}^{\mathrm{Q}}$ that minimizes the mean-square error between the estimated and the target Q-values according to the stochastic gradient descent method given by

$$\boldsymbol{\omega}_j^{\mathrm{Q}} \leftarrow \arg\min_{\boldsymbol{\omega}} \frac{1}{Z} \sum_{z=1}^{Z} \left(u^{(z)} - Q\left(\boldsymbol{s}^{(z)}, \varphi_{j-1}^{(z)}, x_j^{(z)}; \boldsymbol{\omega}\right) + \delta Q\left(\boldsymbol{s}^{(z+1)}, \underset{x \in \boldsymbol{X}_j}{\operatorname{argmax}} Q\left(\boldsymbol{s}^{(z)}, \varphi_{j-1}^{(i)}, x; \hat{\boldsymbol{\omega}}_j^{\mathrm{Q}}\right); \boldsymbol{\omega}_j^{\mathrm{Q}}\right)\right)^2 \quad (12)$$

Similarly, ENNs with weights $\boldsymbol{\omega}_{1\le j\le 3}^{\mathrm{E}}$ that evaluates the risk levels of the selected policies are updated by

$$\boldsymbol{\omega}_j^{\mathrm{E}} \leftarrow \arg\min_{\boldsymbol{\omega}} \frac{1}{Z} \sum_{z=1}^{Z} \left(r^{(z)} - E\left(\boldsymbol{s}^{(z)}, \varphi_{j-1}^{(z)}, x_j^{(z)}; \boldsymbol{\omega}\right) + \delta E\left(\boldsymbol{s}^{(z+1)}, \underset{x \in \boldsymbol{X}_j}{\operatorname{argmax}} E\left(\boldsymbol{s}^{(z)}, \varphi_{j-1}^{(z)}, x; \hat{\boldsymbol{\omega}}_j^{\mathrm{E}}\right); \boldsymbol{\omega}_j^{\mathrm{E}}\right)\right)^2, \quad (13)$$

where the discount factor δ weighs the importance of future rewards and risk values. Both the weights of target neural networks $\hat{\omega}_j^{\mathrm{Q}}$ and $\hat{\omega}_j^{\mathrm{E}}$ are updated every w steps by copying the weights of QNN and ENN for stable learning.

VI. Performance Evaluation

We provide the upper performance bound of perception accuracy, latency and utility based on the SE of the game between N CAVs and the jammer, as well as the computational complexity of our proposed schemes.

A. Game Model

N CAVs $\mathcal{N}$ as the leader firstly chooses the collaborative perception policies $\boldsymbol{A} = [\boldsymbol{\chi}, \boldsymbol{p}, \boldsymbol{a}] \in \boldsymbol{X} = \boldsymbol{X}_1 \times \boldsymbol{X}_2 \times \boldsymbol{X}_3$, which consists of the region of feature maps $\boldsymbol{\chi} = [x_{i,m}]_{i\in\mathcal{N},m\in\mathcal{M}}$, the transmit power $\boldsymbol{p} = [p_i]_{i\in\mathcal{N}}$ and the radio channels $\boldsymbol{a} = [a_i]_{i\in\mathcal{N}}$, to maximize the overall perception utility given by

$$\bar{u} = \frac{1}{N}\sum_{i=1}^{N} u_i \tag{14}$$

The jammer $\mathcal{J}$ as the follower chooses the jamming power $y_1 \in [0, P_{\mathrm{J}}]$, the center frequency $y_2 \in [\underline{f}, \bar{f}]$ and the bandwidth $y_3 \in (0, \hat{B}]$ to maximize the jamming utility u_{J} given by Eq. (3). Without loss generality, we assume the jamming channel y_2 and the bandwidth y_3 are chosen from the F vehicular communication channels, thus the jamming strategy $\boldsymbol{y}$ can be simplified to allocate the power y_l to the l-th vehicular channel, which is given by $\boldsymbol{y} = [y_l]_{l\in\mathcal{F}} \in \boldsymbol{Y} = [0, P_{\mathrm{J}}]^F$ with the maximum power constraint $\sum_{l=1}^{F} y_l \le P_{\mathrm{J}}$. The Stackelberg game between the CAVs and the jammer is given by $\mathcal{G} = \{\mathcal{N}, \mathcal{J}, \boldsymbol{A}, \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Y}, \bar{u}, u_{\mathrm{J}}\}$.

B. Performance Bound

Each feature map is divided into M regions each with ϖ-bits. According to [39], the inference latency $\hat{\tau}$ depends on the computational load, the CPU clock rate ζ and the CPU cycles ψ for per-bit processing given by

$$\hat{\tau} = \frac{\psi\varpi}{\zeta}\sum_{i=1}^{N}\sum_{m=1}^{M} x_{i,m} \tag{15}$$

The perception accuracy of LiDAR-based object detection depends on the DNN models for the feature fusion and inference, the data quality such as the confidence score of the feature map and the offloading performance such as latency and PLR, which in turn depends on the collaborative perception policies of CAVs given by Eq. (5). According to [10], [19], the accuracy ρ increases with the spatial confidence scores $\boldsymbol{O} = [o_{i,m}]_{1\le i\le N, 1\le m\le M}$ and decreases with the offloading latency $\boldsymbol{\tau} = [\tau_i]_{1\le i\le N}$ and inference latency $\hat{\tau}$ denoted by $\rho = \mathcal{F}(\boldsymbol{O}, \boldsymbol{\tau}, \hat{\tau})$ if given the specific DNN models. For simplicity, the accuracy is modeled by a linear function of the spatial confidence scores and offloading latency with the regression parameters $\boldsymbol{\alpha} = [\alpha_{i,m}]_{1\le i\le N, 1\le m\le M}$, $\boldsymbol{\beta} = [\beta_i]_{1\le i\le N}$ and β_0 given by

$$\rho = \sum_{i=1}^{N}\sum_{m=1}^{M} x_{i,m}\left(\alpha_{i,m}o_{i,m} - \frac{\psi\varpi\beta_0}{\zeta} - \frac{\beta_i\varpi}{r_i}\right), \tag{16}$$

where $r_i = B\log(1+\gamma_i)$ is the data rate of CAV i. The accuracy given by Eq. (16) is a simplified expression for analyzing the impacts of collaborative perception policies on the perception accuracy, which can be easily developed to the more practical models by collecting a large amount of sensing data to fit the accuracy curve, and we leave it in the future work. Without loss generality, the channel gains between the CAVs and the RSU are assumed to follow $h_1 \le \cdots \le h_N$.

Theorem 1. The upper performance bound of the collaborative perception scheme is given by

$$\rho^* = \sum_{i=1}^{N}\sum_{m=1}^{M}\alpha_{i,m}o_{i,m} - \frac{\beta_0\psi\varpi MN}{\zeta} - \sum_{i=1}^{N}\frac{\varpi M\beta_i}{B\log\left(1+\frac{P_{\mathrm{V}}h_i}{\phi_i P_{\mathrm{J}}g}\right)} \tag{17}$$

$$t^* = \frac{\psi\varpi MN}{\zeta} + \frac{\varpi M}{B\log\left(1+\frac{P_{\mathrm{V}}h_N}{\phi_i P_{\mathrm{J}}g}\right)} \tag{18}$$

$$\bar{u}^* = \sum_{i=1}^{N}\sum_{m=1}^{M}\alpha_{i,m}o_{i,m} - \frac{\beta_0\psi\varpi MN}{\zeta} - \frac{\varpi M}{B}\sum_{i=1}^{N}\frac{N\beta_i + c_{\mathrm{t}}}{N\log\left(1+\frac{P_{\mathrm{V}}h_i}{\phi_i P_{\mathrm{J}}g}\right)} \tag{19}$$

where

$$\phi_i = \frac{\sqrt{h_i}}{\sum_{i=1}^{N}\sqrt{h_i}} \tag{20}$$

if $N \le F$, $P_{\mathrm{J}}g \gg N_0$

$$\alpha_{i,m}o_{i,m} \ge \frac{\varpi\beta_i}{B\log\left(1+\frac{P_{\mathrm{V}}h_N}{P_{\mathrm{J}}g}\right)} + \frac{\beta_0\psi\varpi}{\zeta} \tag{21}$$

$$h_N \ge \max\left\{\frac{\gamma_0 L_1 P_{\mathrm{J}}g}{P_{\mathrm{V}}}, \frac{c_{\mathrm{J}}P_{\mathrm{J}}^2 g^2}{P_{\mathrm{V}}g}\right\} \tag{22}$$

Proof. See Appendix A. □

Remark 1: Each CAV offloads M regions of the feature map to achieve the optimal perception accuracy given by Eq. (17) if the minimum confidence score for the detection accuracy gain is greater than the offloading latency for M feature map regions given by Eq. (21) and the channel gains exceed the lower bound given by Eq. (22) that depends on the maximum jamming power and the SINR threshold γ_0 for decoding packets. The jammer allocates the jamming power to N CAV-RSU offloading links with the power factor ϕ_i for each link given by Eq. (20) to maximize the jamming utility under the power constraint P_{J}. The perception latency given by Eq. (18) increases with the data size of feature maps and the worse channel gains h_N between the CAV and RSU, yielding the upper bound of utility given by Eq. (19).

C. Computational Complexity

According to [41], the computational complexity Γ_1 of RLCP depends on the total execution steps K and the process of searching the policy with the maximum Q-value at each layer in each time slot given by $\mathcal{O}\left(\sum_{j=1}^{3} K|\boldsymbol{X}_j|\right)$, which is determined by the number of feature map regions, radio channels and power levels, and thus can be further simplified to $\mathcal{O}\left(2^M K\right)$ if $2^M \gg F \approx L_1$. In addition, the sample complexity Γ_1' of RLCP for achieving the optimal policies depends on the size of state and action spaces in each layer of the hierarchical structure given by $\mathcal{O}\left(\sum_{j=1}^{3}|\boldsymbol{S}_j||\boldsymbol{X}_j|\right)$, and can be simplified to $\mathcal{O}\left(2^M F D_1^M D_2^M D_3 D_4^F D_5\right)$ if omitting

the lower terms [42]. The sample complexity shows that the total execution steps K should be much larger than the size of the joint state-action space for sufficient exploration, and thus the computational complexity Γ_1 increases with the quantization level of the state and action spaces. In particular, the performance of the RL-based scheme decreases with the quantization error of the state space, but has to balance the performance gain and computational complexity.

Furthermore, the computational complexity of the deep version, denoted by Γ_2, depends on the multiplication operations for the forward-propagation to obtain the estimated long-term utility and risk for policy selection, as well as the back-propagation for the gradient descent to update the weights of neural networks as shown in Fig. 3. More specifically, with the same architecture of four neural networks in level j of hierarchical RL, we have the dimension of input size Ω_j, the number of neurons in the full-connected layers $f_{1,j}$ and $f_{2,j}$ and the output size $|\boldsymbol{X}_j|$. In the policy selection process, the number of multiplications in the forward-propagation of QNN and ENN in the three levels is given by

$$L_j^{\mathrm{F}} = \Omega_j f_{1,j} + (f_{1,j}+1) f_{2,j} + (f_{2,j}+1)|\boldsymbol{X}_j| \tag{23}$$

Similarly, in the weight update process given by Eqs. (12) and (13), all the four neural networks in the three levels perform L_j^{F} multiplications for the forward-propagation to choose collaborative perception policies, and the QNN and ENN also perform L_j^{B} multiplications in the back-propagation for weights update given by

$$L_j^{\mathrm{B}} = 2\Omega_j f_{1,j} + 2(f_{1,j}+1) f_{2,j} + 3(f_{2,j}+1)|\boldsymbol{X}_j| \tag{24}$$

DRLCP samples Z experiences at each time slot for weight updating, yielding the total number of multiplications given by

$$L^{\mathrm{T}} = (4Z+2)\sum_{j=1}^{3} L_j^{\mathrm{F}} + 4Z\sum_{j=1}^{3} L_j^{\mathrm{B}} \tag{25}$$

According to [8], the number of neurons on the FC layers relies on the total learning steps K of each episode and the dimension of the output layers in each level given by

$$f_{1,j} = \sqrt{K|\boldsymbol{X}_j|} + 2\sqrt{\frac{K}{|\boldsymbol{X}_j|}} \tag{26}$$

$$f_{2,j} = \sqrt{K|\boldsymbol{X}_j|} \tag{27}$$

Theorem 2. The computational complexities of the proposed RLCP and DRLCP schemes are given by $\mathcal{O}\left(2^M K\right)$ and $\mathcal{O}\left(2^M K Z\right)$ if $2^M \gg F \approx L_1$.

Proof. See Appendix B. □

Remark 2: For both the RLCP and DRLCP schemes, the computational complexity increases exponentially with the number of feature map regions, which indicates that the feature map should be divided into a moderate number of regions to fine-tune a trade-off among the perception performance, the communication overhead such as latency and bandwidth consumption, and the computational complexity of the collaborative perception schemes. In particular, the sample complexity of RLCP for achieving the optimal policies depends on the dimension of state and action spaces, which has to balance the performance gain and the convergence speed.

TABLE I: Parameter settings in the simulation

Parameter	Value	Parameter	Value
$B/\hat{B}$	10 MHz	ϖ	1 MB
$P_{\mathrm{V}}/P_{\mathrm{J}}$	100 mW	N_0	-174 dBm/Hz
L_1/L_2	10/10	F	6
M	4	N	5
C_1/C_2	2/2	h_0	68.83
f_1/f_2	128/128	σ	5.5
γ_0	10	α	2.75
ϑ	0.4	δ	0.3

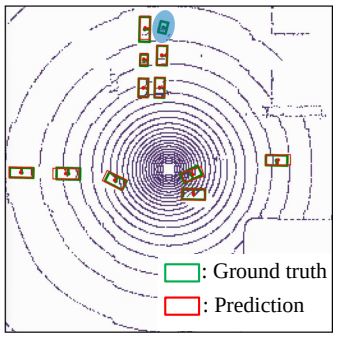
VII. Simulation Results

Simulations were performed on a server with Intel i9-13900 CPU, NVIDIA GeForce RTX 3090 GPU and 128G memory, which applies the feature fusion model in [16] and the autonomous driving dataset containing 47,200 samples in [17] to evaluate the performances of collaborative perception in terms of perception accuracy, latency and utility for the LiDAR-based object detection. The perception accuracy is evaluated based on the intersection-over-union (IoU) threshold of 0.5 that represents the proportion of the overlapping area of bounding boxes between the prediction and ground truth (e.g., a vehicle is detected if the proportion of the overlapping area is greater than 0.5).

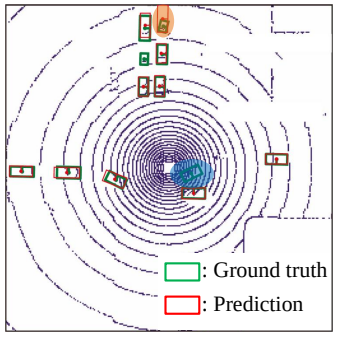
Each CAV partitions the feature map into four regions each with 1-MB, one out of the six radio channels and the transmit power up to $P_V = 100$ mW that is quantified into $L_1 = 10$ levels [38]. The smart jammer is away from the RSU by 200 m estimates the RSSI of vehicular uplink channels, and applies a Q-learning algorithm to choose one out of the six transmit channels and power up to 100 mW to maximize the jamming utility given by Eq. (3). Based on the vehicular channel model as proposed in [36], the reference path-loss h_0 is set to 68.83 at reference distance $d_0 = 1$ m, the path-loss exponent α is 2.75 and the standard deviation of shadowing σ is 5.5. The received packet is successfully decoded if the SINR is greater than the predefined threshold $\gamma_0 = 10$.

The vehicular collaborative perception policy is optimized with the learning rate $\vartheta = 0.4$ and discounted factor $\delta = 0.3$ to maximize the utility function given by Eq. (8) with the coefficients $c_{\mathrm{t}} = 0.001$ and $c_{\mathrm{b}} = 0.01$. The modified Boltzmann distributions for the policy selection make a trade-off between the policy quality and exploration risk with the weights $\xi_{\mathrm{T}} = \xi_{\mathrm{L}} = 10$ in the learning process, and the risk function in Eq. (10) takes $c_{l,1} = c_{l,2} = 0.5$ to evaluate the risk level based on both the offloading latency and PLR that are quantified into two levels with the threshold $\mu_{l,1} = 50$ and $\mu_{l,2} = 0.1$, respectively. The parameters of simulation settings are summarized in Table I for ease of reference.

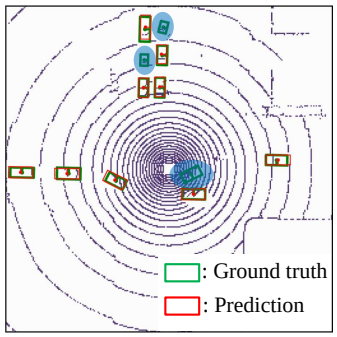
An example result of LiDAR-based collaborative vehicle detection at an intersection in town areas as shown in Fig. 4



(a) DRLCP

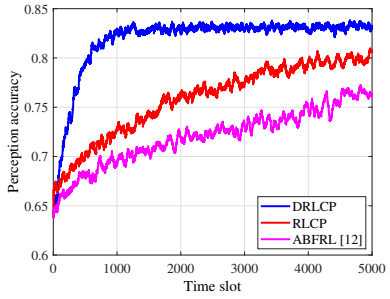


(b) RLCP

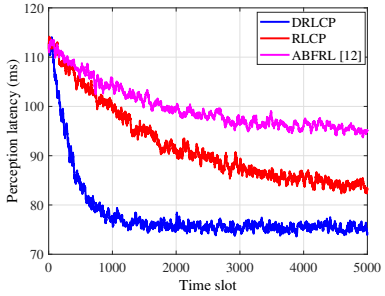


(c) ABFRL [12]

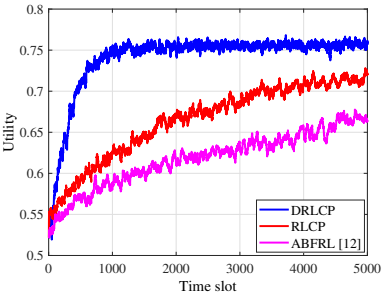
Fig. 4. Snapshot of detection results for LiDAR-based collaborative perception, in which the orange and blue annotations represent the inaccurate and miss detection, respectively.



(a) Perception accuracy



(b) Perception latency



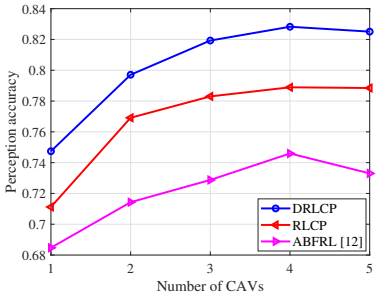
(c) Utility

Fig. 5. Performance of the RL-based collaborative perception scheme in LiDAR-based object detection, in which the perception accuracy, latency and utility are provided over 200 runs each with 5000 time slots.

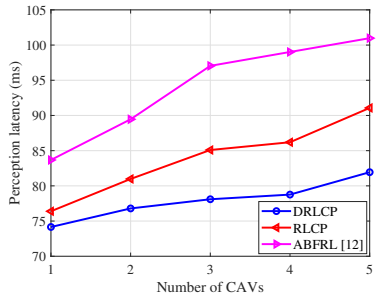
shows that our proposed schemes detect surrounding vehicles more accurately compared with the benchmark ABFRL in [12] due to the lower offloading latency and PLR against smart jamming and interference. For example, the prediction bounding boxes (red) for the vehicle locations of the DRLCP scheme are well-aligned with the ground truth (green), while ABFRL fails to detect 3 vehicles out of the line-of-sight sensing range caused by the information loss of feature maps from neighboring CAVs.

The performance of collaborative perception in vehicular networks as shown in Fig. 5 averaged by 200 runs each with 5000 time slots shows that the perception accuracy, speed and utility of our proposed schemes converge to the optimal performance after 1500 and 4000 time slots, respectively. For example, the perception accuracy of DRLCP increases from 65.4% to 83.2% and the perception latency decreases from 112.4 to 75.5 ms due to the safe exploration to avoid choosing the risk policies to reduce the PLR of the feature map offloading. In addition, our proposed schemes outperform the benchmark ABFRL [12] with 8.3% higher perception accuracy and 21.1% less latency after 4000 time slots. The reason is that both the confidence scores and the estimated received jamming power are exploited to choose the feature map regions and radio resources against smart jamming, thus reducing the offloading latency and PLR to improve the perception performance.

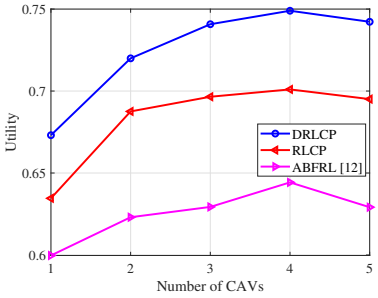
The perception performance averaged over 200 runs each with 5000 time slots in Fig. 6 shows that our proposed schemes outperform benchmark with the increasing number of CAVs. For example, DRLCP improves 9.2% perception accuracy and reduces 18.9% perception latency compared with ABFRL in [12] if the number of CAVs is 5. The perception latency of our proposed schemes is more robust compared with the benchmark against the increasing number of CAVs under



(a) Perception accuracy



(b) Perception latency



(c) Utility

Fig. 6. Average perception performance in terms of accuracy, latency and utility, which are provided over 200 runs with the number of CAVs increasing from 1 to 5.

interference and jamming, in which the deep version improves 10.4% performance gain of the perception latency compared with ABFRL if the number of CAVs increases from 1 to 5.

VIII. Experimental Results

Experiments were performed based on an edge device and five CAVs to detect 10 objects moving at 1 m/s in an $8.5\times4\times3.5$ m^3 area as shown in Fig. 7, in which CAVs capture point clouds via Robosense RS-LiDAR-16 sensors and share feature maps via the Raspberry Pi-4B with the Cortex-A72 processor. The GUI running on the edge device provides the data visualization for the captured point cloud frames and the detection results of the collaborative perception systems as shown in Fig. 8.

Following IEEE 802.11ac protocol, CAVs choose the transmit power with 10 levels ranging from 0 to 100 mW and the radio channels at 5170-5330 MHz each with 20 MHz bandwidth for offloading feature maps that is equally divided

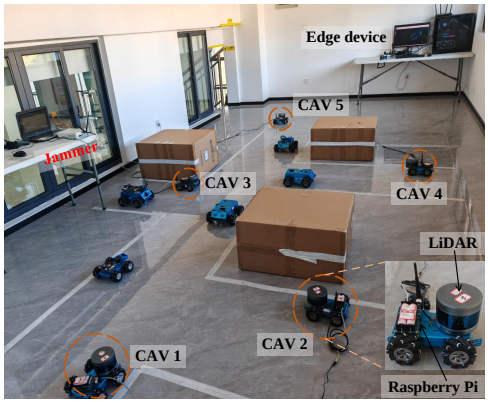


Fig. 7. Experimental setting of the collaborative perception against a smart jammer in an $8.5\times4\times3.5$ m^3 area.

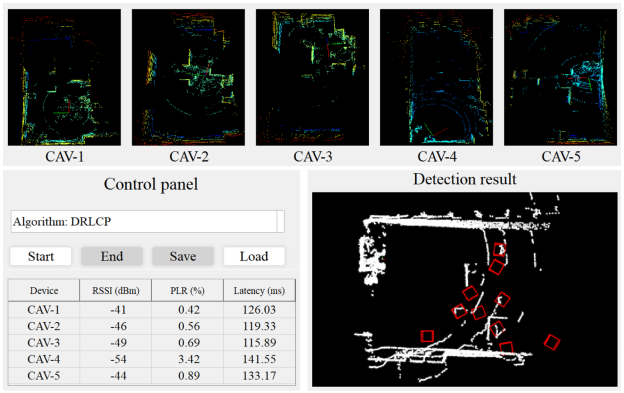
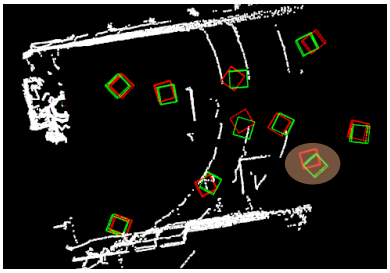


Fig. 8. GUI of the vehicular perception system running in the edge device.

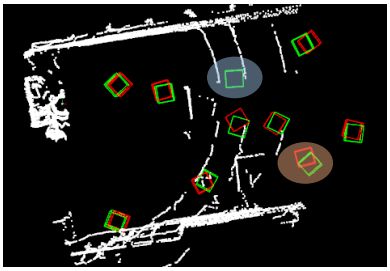
into four regions each with 1-MB data size. The edge device located at (0, 0, 0.8) m that equips with Intel I9-13900 CPU, NVDIA GeForce RTX 3090 GPU and 128G memory to perform the feature fusion and object detection after receiving all the feature maps from neighboring CAVs, which assumes the packet loss if the feature maps are not successfully received within 400 ms. The smart jammer located at (4, 5, 0.8) m equips with USRP N210 controlled by a laptop to emit Gaussian signals, which applies Q-learning algorithm to select the jamming power up to 30 mW and one out of the six channels based on the estimated uplink RSSI.

A snapshot of detection result as shown in Fig. 9 illustrates the inaccurate and miss detection of objects due to the packet loss and long latency of sensing data transmission. Our proposed schemes show the performance gain of perception accuracy over the benchmark in [12] with the lower location offset and miss detection rate. For example, ABFRL fails to detect three vehicles due to the loss of critical perceptual information, as well as having position-misalignment of one object caused by the time asynchronous of feature maps due to the long offloading latency. Our proposed DRLCP scheme detects all the 10 objects with one of them having a slight location offset as marked by inaccurate detection due to the lower PLR and offloading latency.

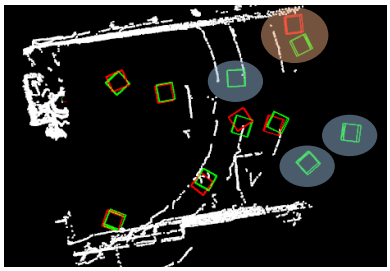
Experimental results in Fig. 10 averaged by 20 runs each



(a) DRLCP

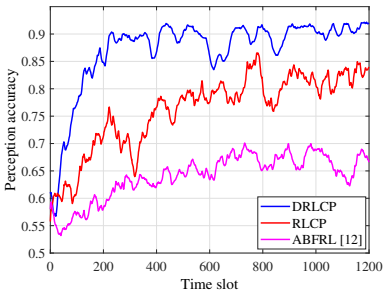


(b) RLCP

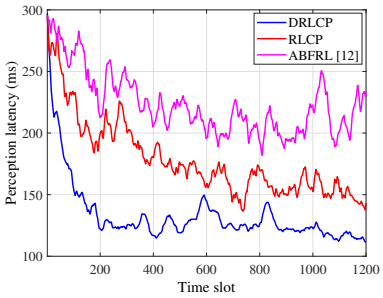


(c) ABFRL [12]

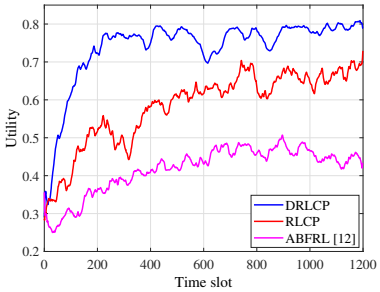
Fig. 9. Snapshot of detection results with the experimental setting in Fig. 7, in which the orange and blue annotations represent the inaccurate and miss detection, respectively.



(a) Perception accuracy



(b) Perception latency



(c) Utility

Fig. 10. Experimental performance provided over 20 runs each with 1200 time slots.

with 1200 steps show that the proposed schemes converge to the optimal performance after 200 and 600 time slots, respectively. The perception accuracy of DRLCP increases from 58.6% to 91.2% and the perception latency decreases from 295.2 to 120.3 ms due to the lower PLR and offloading latency against smart jamming. Our proposed schemes outperform ABFRL in [12] in terms of the perception accuracy, latency and utility. For example, the RLCP scheme is a low-complexity algorithm that enhances 15.1% accuracy with 28.5% less latency compared with ABFRL. In addition, the deep version further enhances 22.4% accuracy and 41.3% less latency compared with the benchmark.

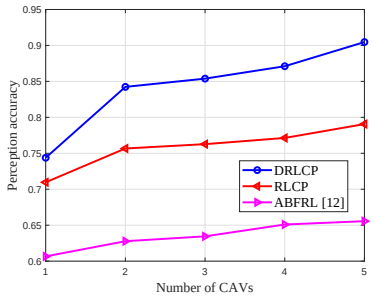
As shown in Fig. 11, our proposed schemes are scalable to the increasing number of collaborative CAVs, which shows the performance gain of 15.3% perception accuracy for the DRLCP scheme if the number of collaborative CAVs increases from 1 to 5. In addition, our scheme enhances at least 14.4% perception accuracy with less 16.1% latency compared with the benchmark in [12]. The scalability of our proposed schemes can be further evaluated with faster CAV speed and wider variety of traffic objects in a larger experimental area, and we will leave it in the future work.

IX. Conclusion

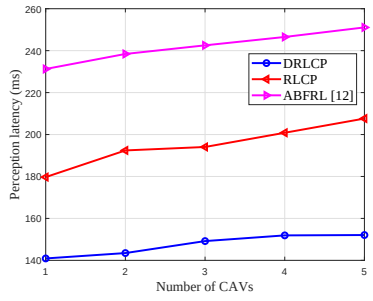
This paper proposed the RL-based collaborative perception schemes for LiDAR-based object detection against jamming and interference, which chooses the regions of the feature map, the radio channel and the transmit power with the hierarchical structure to enhance learning efficiency and the multi-risk levels to evaluate the performance degradation. An upper performance bound based on the Stackelberg equilibrium of the game between N CAVs and the jammer shows that the perception accuracy increases with the number of selected regions, the bandwidth and the maximum transmit power. Experiments were conducted based on five CAVs equipped with LiDAR and Raspberry Pi to collaboratively detect 10 objects in an $8.5\times4\times3.5$ m^3 area against a smart jammer that chooses jamming power up to 30 mW. Experimental results show the performance gains of 22.4% higher perception accuracy and 41.3% less latency compared with the benchmark.

The potential limitations of our proposed system lie in faster CAV speed, larger network size and data fabrication

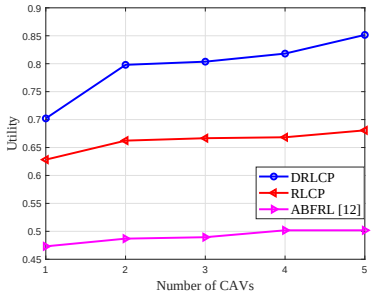
attacks from collaborative CAVs in a larger experimental area, which requires lower transmission latency and data verification schemes to avoid perception errors. In future work, we will further exploit the spatial consistency checking schemes to validate received sensing data to enhance the perception performance, apply the adaptive modulation and coding and the data compression schemes to further decrease the transmission latency in bandwidth-constraint vehicular networks, and exploit the motion trajectories and speed of traffic objects to mitigate the misalignment of object locations caused by both the inconsistent frame rates of LiDAR sensors and the long transmission latency.



(a) Perception accuracy



(b) Perception latency



(c) Utility

Fig. 11. Average performance provided over 20 runs with the number of CAVs increasing from 1 to 5.

Appendix A
Proof of Theorem 1

Proof: By Eqs. (16) and (21), we have

$$\alpha_{i,m}o_{i,m} - \frac{x_{i,m}\psi\varpi\beta_0}{\zeta} - \frac{x_{i,m}\beta_i\varpi}{r_i} \geq 0 \tag{28}$$

Thus, by Eqs. (8), (5d), (14), (16) and (22) and (28), $\forall \boldsymbol{x}_i \in \{0,1\}^M$, $\forall a_i \in \{1,\cdots,F\}$, $\forall p_i \in (0, P_{\mathrm{V}}]$ and $\forall \boldsymbol{y} \in [0, P_{\mathrm{J}}]^F$, we have

$$\begin{aligned}
\bar{u}\Big(\boldsymbol{A},\boldsymbol{y}\Big) &= \sum_{i=1}^{N}\sum_{m=1}^{M} x_{i,m}\alpha_{i,m}o_{i,m} \\
&- \frac{x_{i,m}\psi\varpi\beta_0}{\zeta} - \frac{x_{i,m}\varpi(\beta_i + c_{\mathrm{t}})}{B\log\left(1+\frac{p_i h_i}{\kappa y_i g + N_0}\right)} \\
&\leq \sum_{i=1}^{N}\sum_{m=1}^{M}\alpha_{i,m}o_{i,m} - \frac{\beta_0\psi\varpi MN}{\zeta} \\
&- \frac{\varpi M}{B}\sum_{i=1}^{N}\frac{N\beta_i + c_{\mathrm{t}}}{N\log\left(1+\frac{P_{\mathrm{V}}h_i}{\kappa y_i g + N_0}\right)} \\
&= \bar{u}\left(\boldsymbol{A}^*,\boldsymbol{y}\right),
\end{aligned} \tag{29}$$

where A^* is given by

$$\boldsymbol{A}^* = \Big[\underbrace{[1,\cdots,1]}_{MN}, [1,\cdots,N], \underbrace{[P_{\mathrm{V}},\cdots,P_{\mathrm{V}}]}_{N}\Big] \tag{30}$$

By Eqs. (3), (5d), (22) and (29), $\forall \boldsymbol{y} \in (0, P_{\mathrm{J}}]^F$, $N \leq F$ and $P_{\mathrm{J}} g \gg N_0$ we have

$$\begin{aligned}
u_{\mathrm{J}}(\boldsymbol{A}^*,\boldsymbol{y}) &= -\sum_{i=1}^{N}\frac{P_{\mathrm{V}}h_i}{P_{\mathrm{J}}g} - c_{\mathrm{J}}\sum_{l=1}^{F} y_l \\
&\leq -\sum_{i=1}^{N}\frac{P_{\mathrm{V}}h_i}{P_{\mathrm{J}}g} - c_{\mathrm{J}}\sum_{l=1}^{N} y_i = u_{\mathrm{J}}(\boldsymbol{A}^*,\boldsymbol{y}')
\end{aligned} \tag{31}$$

As $\sum_{i=1}^{N} y_i \leq P_{\mathrm{J}}$, we have

$$\mathcal{L}\left(\boldsymbol{y}',\lambda\right) = -\sum_{i=1}^{N}\frac{P_{\mathrm{V}}h_i}{P_{\mathrm{J}}g} - c_{\mathrm{J}}\sum_{l=1}^{N} y_i + \lambda\left(P_{\mathrm{J}} - \sum_{i=1}^{N} y_i\right) \tag{32}$$

$$\frac{\partial\mathcal{L}\left(\boldsymbol{y}',\lambda\right)}{\partial y_i} = \frac{P_{\mathrm{V}}h_i}{y_i^2 g} - c_{\mathrm{J}} + \lambda = 0 \tag{33}$$

$$\frac{\partial\mathcal{L}\left(\boldsymbol{y}',\lambda\right)}{\partial\lambda} = P_{\mathrm{J}} - \sum_{i=1}^{N} y_i = 0 \tag{34}$$

By Eqs. (29), (31), (33) and (34), $\forall \boldsymbol{y} \in (0, P_{\mathrm{J}}]^F$, we have

$$u_{\mathrm{J}}\left(\boldsymbol{A}^*,\boldsymbol{y}^*\right) \geq u_{\mathrm{J}}\left(\boldsymbol{A},\boldsymbol{y}\right) \tag{35}$$

where $\boldsymbol{y}^* = [y_i^*]_{1\leq i\leq F}$ is given by

$$\boldsymbol{y}^* = \Big[y_1^*,\cdots,y_N^*,\underbrace{0,\cdots,0}_{F-N}\Big], \quad y_i^* = \frac{\sqrt{h_i}P_{\mathrm{J}}}{\sum_{i=1}^{N}\sqrt{h_i}} \tag{36}$$

Thus, $\left[\boldsymbol{A}^*,\boldsymbol{y}^*\right]$ is a Stackelberg equilibrium of the game for the collaborative perception against jamming and interference, yielding the performance upper bound given by Eqs. (17)-(19).

Appendix B
Proof of Theorem 2

Proof: According to [42], as $|\boldsymbol{X}_1| = 2^M$, $|\boldsymbol{X}_2| = F$, $|\boldsymbol{X}_3| = L_1$, $|\boldsymbol{S}_1| = D_1^M D_2^M D_3 D_4^F D_5$, $|\boldsymbol{S}_2| = |\boldsymbol{S}_1||\boldsymbol{X}_1|$ and $|\boldsymbol{S}_3| =$

$|\boldsymbol{S}_1||\boldsymbol{X}_2|$, we have Γ_1' given by

$$\Gamma_1' = \mathcal{O}\left(\sum_{j=1}^{3}|\boldsymbol{S}_j||\boldsymbol{X}_j|\right) \tag{37}$$

$$= \mathcal{O}\Big(2^M D_1^M D_2^M D_3 D_4^F D_5 + 2^M F D_1^M D_2^M D_3 D_4^F D_5 + F L_1 D_1^M D_2^M D_3 D_4^F D_5\Big) \tag{38}$$

$$= 2^M F D_1^M D_2^M D_3 D_4^F D_5 \tag{39}$$

where Eq. (38) is obtained by submitting the $|\boldsymbol{S}_j|$ and $|\boldsymbol{X}_j|$, and Eq. (39) is obtained by omitting the lower terms.

According to [41], as $\Omega_1 = 2M+2F+2$, $\Omega_2 = 3M+2F+2$ and $\Omega_3 = 2M+2F+3$, we have

$$\Gamma_2 = \mathcal{O}\Big(4Z\sum_{j=1}^{3}\Omega_j f_{1,j} + (f_{1,j}+1)f_{j,2} + (f_{2,j}+1)|\boldsymbol{X}_j| + \Omega_j f_{1,j} + 2(f_{1,j}+1)f_{j,2} + 3(f_{2,j}+1)|\boldsymbol{X}_j|\Big) \tag{40}$$

$$= \mathcal{O}\left(Z\sum_{j=1}^{3}\sqrt{K|\boldsymbol{X}_j|} + K|\boldsymbol{X}_j| + \sqrt{K|\boldsymbol{X}_j|^3}\right) \tag{41}$$

$$= \mathcal{O}\Big(KZ\Big(2^M + F + L_1\Big)\Big) \tag{42}$$

$$= \mathcal{O}\Big(2^M KZ\Big), \tag{43}$$

where Eq. (40) is obtained by Eqs. (23)-(25), and Eq. (41) is derived by submitting Eqs. (26) and (27). Eq. (42) is obtained by submitting $|\boldsymbol{X}|_1 = 2^M$, $|\boldsymbol{X}|_2 = F$ and $|\boldsymbol{X}|_3 = L_1$, and is further simplified to Eq. (43) as $2^M \gg F \approx L_1$.

REFERENCES

[1] Q. Yang, S. Fu, H. Wang *et al.*, "Machine-Learning-Enabled Cooperative Perception for Connected Autonomous Vehicles: Challenges and Opportunities," *IEEE Network*, vol. 35, no. 3, pp. 96–101, May 2021.

[2] X. Zheng, S. Li, Y. Li *et al.*, "Confidence Evaluation for Machine Learning Schemes in Vehicular Sensor Networks," *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2833–2846, April 2023.

[3] H. Zhu, J. Deng, Y. Zhang *et al.*, "VPFNET: Improving 3d Object Detection with Virtual Point Based LiDAR and Stereo Data Fusion," *IEEE Trans. Multimedia*, vol. 25, pp. 5291–5304, 2023.

[4] R. Xu, H. Xiang, Z. Tu *et al.*, "V2X-ViT: Vehicle-to-Everything Cooperative Perception with Vision Transformer," in *Proc. European Conf. Comput. Vision (ECCV)*, Tel Aviv, Israel, 2022, pp. 107–124.

[5] B. Yang, W. Luo, and R. Urtasun, "Pixor: Real-Time 3D Object Detection from Point Clouds," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit. (CVPR)*, Salt Lake City, UT, 2018, pp. 7652–7660.

[6] X. Lu, L. Xiao, P. Li *et al.*, "Reinforcement Learning Based Physical Cross-Layer Security and Privacy in 6G," *IEEE Commun. Surveys Tut.*, vol. 25, no. 1, pp. 425–466, Firstquarter 2023.

[7] H. Pirayesh and H. Zeng, "Jamming Attacks and Anti-Jamming Strategies in Wireless Networks: A Comprehensive Survey," *IEEE Commun. Surveys Tutor.*, vol. 24, no. 2, pp. 767–809, 2022.

[8] Z. Lv, L. Xiao, Y. Du *et al.*, "Multi-Agent Reinforcement Learning based UAV Swarm Communications Against Jamming," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9063–9075, 2023.

[9] X. Zhang, A. Zhang, J. Sun *et al.*, "EMP: Edge-Assisted Multi-Vehicle Perception," in *Proc. Annual Int. Conf. Mobile Comput. Netw. (MOBICOM)*, New Orleans, LA, 2021, pp. 545–558.

[10] Y. Hu, S. Fang, Z. Lei *et al.*, "Where2com: Communication-Efficient Collaborative Perception via Spatial Confidence Maps," in *Proc. Advances in Neural Inf. Process. Syst. (NeurIPS)*, New Orleans, LA, 2022.

[11] Y. Zhai, J. Fu, Y. Lu *et al.*, "Feature Selective Networks for Object Detection," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit. (CVPR)*, Salt Lake City, UT, 2018, pp. 4139–4147.

[12] M. K. Abdel-Aziz, C. Perfecto, S. Samarakoon *et al.*, "Vehicular Cooperative Perception Through Action Branching and Federated Reinforcement Learning," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 891–903, Feb. 2022.

[13] Y. Jia, R. Mao, Y. Sun *et al.*, "MASS: Mobility-Aware Sensor Scheduling of Cooperative Perception for Connected Automated Driving," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 14 962–14 977, Nov. 2023.

[14] H. Chen, Z. Lin, Y. Zhu *et al.*, "Reinforcement Learning Based Collaborative Perception for Vehicular Networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Cape Town, South Africa, 2024, pp. 1–6.

[15] Q. Xie, X. Zhou, T. Qiu *et al.*, "Soft Actor-Critic-Based Multilevel Cooperative Perception for Connected Autonomous Vehicles," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21 370–21 381, Nov. 2022.

[16] Y. Li, S. Ren, P. Wu *et al.*, "Learning Distilled Collaboration Graph for Multi-Agent Perception," in *Proc. Advances in Neural Inf. Process. Syst. (NeurIPS)*, Virtual, 2021, pp. 29 541–29 552.

[17] Y. Li, D. Ma, Z. An *et al.*, "V2X-Sim: Multi-Agent Collaborative Perception Dataset and Benchmark for Autonomous Driving," *IEEE Robot. Autom. Lett.*, Oct. 2022.

[18] T.-H. Wang, S. Manivasagam, M. Liang *et al.*, "V2VNET: Vehicle-to-Vehicle Communication for Joint Perception and Prediction," in *proc. European Conf. Comput. Vision (ECCV)*, Glasgow, UK, 2020, pp. 605–621.

[19] Z. Lei, S. Ren, Y. Hu *et al.*, "Latency-Aware Collaborative Perception," in *proc. European Conf. Comput. Vision (ECCV)*, Tel Aviv, Israel, 2022, pp. 316–332.

[20] S. Shi, J. Cui, Z. Jiang *et al.*, "VIPS: Real-Time Perception Fusion for Infrastructure-Assisted Autonomous Driving," in *Proc. Annual Int. Conf. Mobile Comput. Netw. (MOBICOM)*, Sydney, Australia, 2022, pp. 133–146.

[21] Y. He, L. Ma, Z. Jiang *et al.*, "VI-Eye: Semantic-Based 3D Point Cloud Registration for Infrastructure-Assisted Autonomous Driving," in *Proc. Annual Int. Conf. Mobile Comput. Netw. (MOBICOM)*, New Orleans, LA, 2021, pp. 573–586.

[22] J. Cui, H. Qiu, D. Chen *et al.*, "Coopernaut: End-to-End Driving with Cooperative Perception for Networked Vehicles," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit. (CVPR)*, New Orleans, LA, 2022, pp. 17 252–17 262.

[23] G. Luo, C. Shao, N. Cheng *et al.*, "EdgeCooper: Network-Aware Cooperative LiDAR Perception for Enhanced Vehicular Awareness," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 1, pp. 207–222, 2024.

[24] G. Thandavarayan, M. Sepulcre, and J. Gozalvez, "Generation of Cooperative Perception Messages for Connected and Automated Vehicles," *IEEE Trans. Veh. Technol.*, Dec. 2020.

[25] D. D. Yoon, B. Ayalew, and G. G. M. Nawaz Ali, "Performance of Decentralized Cooperative Perception in V2V Connected Traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6850–6863, July 2022.

[26] J. Li, R. Xu, X. Liu *et al.*, "Learning for Vehicle-to-Vehicle Cooperative Perception Under Lossy Communication," *IEEE Trans. Intell. Veh.*, vol. 8, no. 4, pp. 2650–2660, April 2023.

[27] W. Zhuang, Q. Ye, F. Lyu *et al.*, "SDN/NFV Empowered Future IoV with Enhanced Communication, Computing, and Caching," *Proc. IEEE*, vol. 108, no. 2, pp. 274–291, Feb. 2020.

[28] Y. Xiao, Q. Du, Y. Zhang *et al.*, "Secure Vehicular Communications with Varying QoS and Environments: A Unified Cross-Layer Policy-Adaptation Approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 11, pp. 13 462–13 471, 2023.

[29] Y.-C. Liu, J. Tian, C.-Y. Ma *et al.*, "Who2com: Collaborative Perception via Learnable Handshake Communication," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Paris, France, 2020, pp. 6876–6883.

[30] H. Ngo, H. Fang, and H. Wang, "Cooperative Perception with V2V Communication for Autonomous Vehicles," *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 11 122–11 131, Set. 2023.

[31] L. Xiao, X. Lu, T. Xu *et al.*, "Reinforcement Learning-Based Mobile Offloading for Edge Computing Against Jamming and Interference," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6114–6126, 2020.

[32] S. Liu, Y. Xu, G. Li *et al.*, "Multidimensional Resource Management for Distributed MEC Networks in Jamming Environment: A Hierarchical DRL Approach," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 16 859–16 872, 2024.

[33] Y. Xu, J. Chen, Y. Xu *et al.*, "Energy-Efficient Channel Access and Data Offloading Against Dynamic Jamming Attacks," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 4, pp. 1734–1746, 2021.

[34] X. Ye, K. Qu, W. Zhuang, and X. Shen, "Accuracy-Aware Cooperative Sensing and Computing for Connected Autonomous Vehicles," *IEEE Trans. Mobile Comput.*, vol. 23, no. 8, pp. 8193–8207, 2024.

[35] M. A. Munir, M. H. Khan, S. Khan, and F. S. Khan, "Bridging Precision and Confidence: A Train-Time Loss for Calibrating Object Detection," in *IEEE/CVF Conf. Comp. Vision Pattern Recog. (CVPR)*, Vancouver, Canada, 2023, pp. 11 474–11 483.
[36] L. Cheng, B. E. Henty, D. D. Stancil *et al.*, "Mobile Vehicle-to-Vehicle Narrow-Band Channel Measurement and Characterization of the 5.9 GHz Dedicated Short Range Communication (DSRC) Frequency Band," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 8, pp. 1501–1516, Oct. 2007.
[37] K. Koufos and C. P. Dettmann, "Moments of Interference in Vehicular Networks with Hardcore Headway Distance," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8330–8341, Dec. 2018.
[38] J. B. Kenney, "Dedicated Short-Range Communications (DSRC) Standards in the United States," *Proc. IEEE*, vol. 99, no. 7, pp. 1162–1182, July 2011.
[39] Y. Xiao, L. Xiao, K. Wan *et al.*, "Reinforcement Learning Based Energy-Efficient Collaborative Inference for Mobile Edge Computing," *IEEE Trans. Commun.*, vol. 71, no. 2, pp. 864–876, Feb. 2023.
[40] O. Punal, C. Pereira, A. Aguiar *et al.*, "Experimental Characterization and Modeling of RF Jamming Attacks on VANETs," *IEEE Trans. Veh. Technol.*, vol. 64, no. 2, pp. 524–540, Feb. 2015.
[41] X. Lu, L. Xiao, G. Niu *et al.*, "Safe Exploration in Wireless Security: A Safe Reinforcement Learning Algorithm with Hierarchical Structure," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 732–743, Feb. 2022.
[42] A. L. Strehl, L. Li, E. Wiewiora *et al.*, "PAC Model-Free Reinforcement Learning," in *Proc. Int. Conf. Machine Learning (ICML)*, New York, NY, USA, 2006, pp. 881–888.

Zhiping Lin (Student Member, IEEE) received his B.S. degree in communication engineering from Huaqiao University in 2017 and his M.S. degree in information and communication engineering from Xiamen University in 2020. He is currently pursuing a Ph.D. degree with the Department of Information and Communication Engineering, Xiamen University. His research interests include collaborative vehicular perception and anti-jamming communications in vehicular networks.

Liang Xiao (Senior Member, IEEE) received the B.S. degree in communication engineering from the Nanjing University of Posts and Telecommunications, China, in 2000, the M.S. degree in electrical engineering from Tsinghua University, China, in 2003, and the Ph.D. degree in electrical engineering from Rutgers University, NJ, USA, in 2009. She was a Visiting Professor with Princeton University, Virginia Tech, and the University of Maryland, College Park. She is currently a Professor with the Department of Information and Communication Engineering, Xiamen University, Xiamen, China. She was a recipient of the Best Paper Award for 2016 INFOCOM Big Security WS and 2017 ICC. She has served as an Associate Editor for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY and a Guest Editor for IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.

Hongyi Chen (Student Member, IEEE) received the B.S. degree in computer science and technology from Fujian Normal University, China, in 2023. He is currently pursuing a Master's degree with the Institute of Artificial Intelligence at Xiamen University, China. His research interests include network security and wireless communications.

Zefang Lv (Member, IEEE) received the B.S. degree in statistics from Shandong University in 2016, the M.S. degree in applied statistics from North China Electric Power University in 2020, and the Ph.D. degree in communication and information systems from Xiamen University in 2024. She is currently a Post doctor with the School of Computer Science, Fudan University, Shanghai. Her research interests include wireless communications, UAV communications and reinforcement learning.

Yunjun Zhu received the B.S. degree in Electrical Engineering and Automation from South China University of Technology, China, in 2021. He is currently pursuing a Master's degree in the Department of Informatics and Communication Engineering at Xiamen University, China. His research interests include network security and wireless communications.

Yanyong Zhang (Fellow, IEEE) received the B.S. degree from the University of Science and Technology of China (USTC) in 1997 and the Ph.D. degree from Penn State University in 2002. From 2002 and 2018, she was on the faculty of the Electrical and Computer Engineering Department at Rutgers University. Since July 2018, she has been joined the school of Computer Science and Technology at USTC. She has 21 years of research experience in the areas of sensor networks, ubiquitous computing, and high-performance computing, and has published more than 140 technical papers in these fields. She was also a member of the Wireless Information Networks Laboratory (Winlab). She received the NSF CAREER award in 2006. She has served/currently serves as the Associate Editor for several journals, including IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE TRANSACTIONS ON SERVICE COMPUTING, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, and Elsevier Smart Health. She has served on various conference TPCs including DSN, Sensys, Infocom, etc. She is the TPC Co-Chair for IPSN'22.

Yong-Jin Liu (Senior Member, IEEE) received the B.Eng. degree from Tianjin University, China, in 1998, and the Ph.D. degree from The Hong Kong University of Science and Technology, Hong Kong, China, in 2004. He is now a Full Professor with the Department of Computer Science and Technology, Tsinghua University, China. His research interests include machine learning, cognitive computation, and pattern analysis. For more information, please visit https://cg.cs.tsinghua.edu.cn/people/ Yongjin/Yongjin.htm