Reinforcement Learning Based Collaborative Perception for Vehicular Networks

Hongyi Chen*, Zhiping Lin†, Yunjun Zhu†, Jieling Li†, Liang Xiao†*, Yuliang Tang†, Yanyong Zhang‡
*Institute of Artificial Intelligence, Xiamen University. Email: lxiao@xmu.edu.cn
†Dept. Information and Communication Engineering, Xiamen University. Email: lxiao@xmu.edu.cn
‡School of Computer Science and Technology, University of Science and Technology of China. Email: yanyongz@ustc.edu.cn

GLOBECOM 2024 - 2024 IEEE Global Communications Conference | 979-8-3503-5125-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/GLOBECOM52923.2024.10901016

***Abstract*—Reinforcement learning (RL)-based collaborative perception in vehicular networks chooses the sub-frame of radio channel resources for connected autonomous vehicles (CAVs) to exchange sensing data to enhance the perception performance, but leads to inaccurate detection in the light detection and ranging (LiDAR)-based object detection due to the asynchronous scan period of the LiDAR point clouds. This paper proposes a RL-based collaborative perception scheme to choose the transmit power and sub-frame to share the feature maps extracted from point clouds. Based on the estimated packet timestamp, the network topology and the channel gains among CAVs, this scheme enhances the perception accuracy and latency against path-loss and interference. The collaborative risk in the policy distribution is formulated as a weighted sum of the perception latency and packet loss rate to avoid the time asynchronization and information loss of the feature map exchange. The performance bound of the perception accuracy and latency is provided based on a Nash equilibrium of the cooperative game among CAVs. Simulation results based on five CAVs show the performance gain of the perception accuracy and latency over the benchmarks.**

***Index Terms*—Collaborative perception, vehicular networks, LiDAR, reinforcement learning.**

I. Introduction

Collaborative perception by connected autonomous vehicles (CAVs) enhances the performance to forecast motion trajectory and detect obstacles beyond the line-of-sight, thus improving traffic efficiency and road safety by detecting pedestrians and vehicles behind large occlusions [1]–[3]. For example, CAVs equipped with light detection and ranging (LiDAR) sensors (e.g., Ultra-Puck) to capture the raw sensing data such as point clouds at a frame rate of 5-20 Hz, and apply convolutional neural network (CNN) to obtain the feature map of the captured point clouds. The feature map is shared to the neighboring CAVs that perform a feature fusion model such as multi-head attention and self-attention [4], [5] to aggregate the received feature maps to enhance the perception accuracy and range over individual CAV.

CAVs share the feature maps via vehicular communications such as reservation-based and reinforcement learning (RL)-based resource allocation schemes [6]–[11]. For example, the seminal vehicle-to-vehicle (V2V) communication scheme named DARP as proposed in [6] reserves a sub-frame of radio channel resources for each CAV to exchange the feature maps with a constant rate of 10 Hz. A RL-based sensing data transmission scheme named ABFRL as proposed in [10] exploits the deep Q-network algorithm to further optimize sub-frame selection based on the bandwidth and vehicle locations to enhance the perception reliability. However, these schemes are hindered by high latency and packet loss rate (PLR) that result in the low perception accuracy due to the asynchronous LiDAR scan periods by CAVs.

In this paper, we propose a RL-based collaborative perception for LiDAR-based object detection in vehicular networks, which exploits the CNN to extract feature maps from point clouds and chooses the transmit power and sub-frame to share the feature maps with neighboring CAVs against path-loss and interference. A channel-wise self-attention-based data fusion mechanism is leveraged to aggregate the received feature maps and thus enhance the perception accuracy.

Based on the estimated packet timestamp of the feature maps, the network topology and the channel gains, this scheme formulates the RL state to maximize the expected long-term discounted utility (i.e., Q-values). In addition, the average latency and PLR are exploited to evaluate the risk of perception accuracy degradation, thus avoiding choosing the vehicular communication policy that fails the feature sharing against path-loss and interference to improve perception accuracy. The computational complexity of the proposed scheme increases linearly with the number of power levels and sub-frames.

We derive the upper performance bound of the collaborative perception based on the Nash equilibrium of the cooperative game, which shows that the collaborative CAV tends to share the feature map in the nearest sub-frame with the maximum transmit power to minimize the latency to enhance the perception accuracy. Simulations were performed based on the feature fusion model in [12] and a simulated multi-vehicle perception dataset for autonomous driving in [13]. The results show that our proposed scheme enhances 17.1% average perception accuracy and decreases 67.3% latency in the LiDAR-based object detection compared with the benchmark in [10].

The structure of this paper is organized as follows: Section II reviews related work. Section III introduces the system

This work was supported in part by the National Natural Science Foundation of China under Grant U21A20444, 62332016 and 91638204 and the National Key Research and Development Program of China under Grant 2023YFB3107603. The corresponding author is Liang Xiao.

model. The RL-based collaborative perception scheme is detailed in Section IV, with performance analysis in Section V. Section VI discusses simulation results, followed by the conclusion in Section VII.

II. Related Work

Collaborative perception in vehicular networks enables vehicles to overcome the physical limitations of onboard sensors such as line-of-sight sensing range by sharing the perceived information with neighboring vehicles [14]–[16]. For instance, the collaborative perception scheme based on V2V communications in [14] designs a spatial graph neural network to fuse the feature maps received from neighboring vehicles to improve the motion forecasting performance. An infrastructure-assisted object detection scheme in [15] exploits the motion trajectories of moving vehicles to mitigate the impact of asynchronous data frames of the point clouds due to the transmission latency to improve perception accuracy.

The performance of multi-agent collaborative perception in large-scale networks is restricted to the scarce transmission bandwidth that results in network congestion and high message latency. In [17], a three handshake-based collaborative perception scheme is proposed to enable the target robot to determine the most valuable collaborative robots based on the compressed request messages, thus improving the perception capabilities and reducing bandwidth requirements. A spatial confidence map is built in [4] to compress the redundancy of feature maps, which enables agents to share spatially critical information to save bandwidth usage and improve robustness in collaborative object detection tasks.

V2V communication protocols have been studied to reduce transmission latency and improve communication efficiency [6]–[9] for vehicular latency-sensitive applications. For example, a positioning error threshold-based vehicular broadcasting protocol is proposed in [7], in which each vehicle chooses a specific error threshold to determine whether to broadcast the message based on the position estimation accuracy to enhance reliability and scalability. In addition, a reinforcement learning-based resource allocation scheme in [9] chooses transmit power and channel based on channel gains and interference level to reduce transmission latency.

III. System Model

A. Task Model

CAVs share feature maps of point clouds to upgrade the perception capacities such as expanding sensing range and enhancing detection accuracy to support autonomous driving applications, such as collision avoidance and path planning. As shown in Fig. 1, collaborative CAV $i \in \mathcal{N} = \{1, 2, ..., N\}$ equipped with LiDAR captures point clouds with f_i-Hz, and further quantizes the point clouds into a bird-eye view $\boldsymbol{X}_i^{(k)}$ via voxelization operation at time slot k. The feature map $\boldsymbol{F}_i^{(k)}$ with dimension $H \times W \times C$ is extracted by the CNN that is composed of four blocks with convolutions, rectified linear units activation function and batch normalization, in which

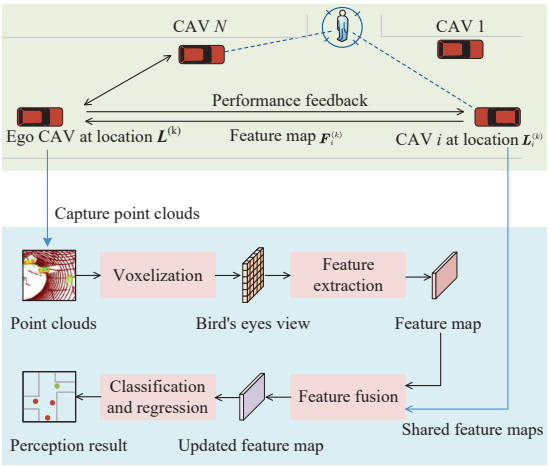


Fig. 1. Overview of collaborative perception, in which each CAV performs object detection by aggregating the received feature maps to enhance perception accuracy.

H, W and C are the height, width and number of feature channels.

Ego CAV leverages a feature channel-wise self-attention data fusion mechanism based on the normalized matrix-valued weights as presented in [12] to aggregate both the ego and received feature maps to produce the updated representation $\boldsymbol{H}^{(k)}$ taking the time duration δ_3, i.e.,

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

where $\odot$ symbolizes the dot-product operation, $\boldsymbol{W}_i^{(k)}$ and $\boldsymbol{W}^{(k)}$ denote the collaboration and self-attention weights, which are the trainable matrix for representing the importance of each spatial region among the ego and received feature maps. Note that in perception tasks, all the CAVs play dual roles as both ego and collaborative agents, each taking on different responsibilities, i.e., share or aggregate the feature maps to enhance the perception performance. The output CNN up-samples the feature representation $\boldsymbol{H}^{(k)}$, classifies the foreground-background categories and regresses the bounding boxes to generate perception results $z^{(k)}$.

As shown in Fig. 2, each time slot with T seconds for the collaborative object detection consists of a policy selection period with duration δ_1 to choose the transmit power and sub-frame, the feature sharing period with duration δ_2 that consists of M_2 sub-frames and the data fusion period with duration δ_3 to aggregate all the feature maps via Eq. (1). CAVs in the policy selection period estimate the packet timestamp $\hat{t}_i^{(k)}$ of the feature maps based on the length of each time slot T and the LiDAR scan period f_i, given by

$$\hat{t}_i^{(k)} = \frac{\left\lfloor f_i \left(kT - t_i^{(k-1)} \right) \right\rfloor}{f_i} + t_i^{(k-1)}, i \in \mathcal{N}, \tag{2}$$

where $\lfloor \cdot \rfloor$ is the floor function.

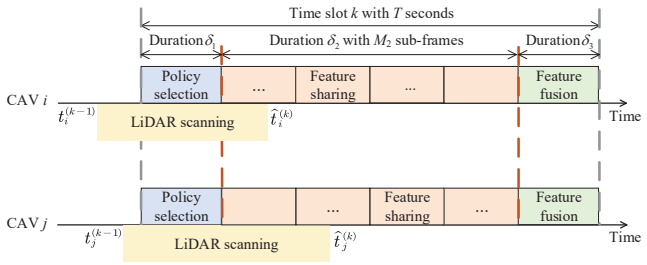


Fig. 2. Timeline of collaborative perception, in which each time slot with T seconds consists of a policy selection period for N CAVs to choose vehicular communication policy based on RL, a feature sharing period with M_2 sub-frames to share the feature map of point clouds and a feature fusion period to aggregate all the feature maps.

B. Communication Model

The communication channel time is divided into frames with duration δ_2 for feature sharing that consists of M_2 sub-frames with equal duration as shown in Fig. 2. Collaborative CAV i located at $\boldsymbol{L}_i^{(k)}$ chooses the transmit power $p_i^{(k)} \in \mathbf{A}_1 = \{jP/M_1 | 0 \le j \le M_1\}$ and one of the sub-frames $m_i^{(k)} \in \mathbf{A}_2 = \{1, 2, ..., M_2\}$ to share a D-bits packet containing the feature map $\boldsymbol{F}_i^{(k)}$ and the packet timestamp $t_i^{(k)}$. Upon receiving the packets, ego CAV aggregates the feature maps via a channel-wise self-attention mechanism to evaluate the perception accuracy $\rho^{(k)}$ based on the perception results $z^{(k)}$, calculates the latency $\tau_i^{(k)}$ of the packet and the PLR $b_i^{(k)}$ within n frames, and sends back these performances to collaborative CAV i via acknowledgment packet.

The channel gain $h_i^{(k)}$ between the ego CAV and collaborative CAV i depends on the path-loss and the shadow fading X_σ given by

$$h_i^{(k)} = h_0 + 10\lambda \lg d_i^{(k)} + X_\sigma, \tag{3}$$

where λ and h_0 are the path-loss exponent corresponding to the communication distance $d_i^{(k)}$ and reference path-loss, and X_σ follows a zero mean Gaussian distribution with standard deviation σ. To successfully decode the received packets, the signal-to-interference-plus-noise-ratio (SINR) needs to exceed the threshold γ_0 corresponding to the specific modulation mode. Without confusion, we will omit superscript k and subscript i in the following for simplicity.

IV. RL-BASED COLLABORATIVE PERCEPTION IN VEHICULAR NETWORKS

We propose a RL-based collaborative perception scheme (RLCP) for each CAV to choose the transmit power and sub-frame to share the feature map and further exploit a self-attention-based data fusion mechanism to aggregate the received feature maps to enhance the perception accuracy in LiDAR-based object detection. A weighted sum of average latency and PLR is exploited to evaluate the potential policy risk to avoid perception performance degradation.

Algorithm 1: RL-based collaborative perception in vehicular networks

Initialize β, δ, c_1, c_2, ξ_1 and ξ_2.
for k = 1, 2, ..., K **do**
  Extract feature map $\boldsymbol{F}$ from LiDAR point clouds
  Estimate the packet timestamp $\left[\hat{t}_i\right]_{i\in\mathcal{N}}$ via (2)
  Obtain locations $[\boldsymbol{L}_i]_{i\in\mathcal{N}}$ and channel gains $[h_i]_{i\in\mathcal{N}}$
  Formulate $\boldsymbol{s}^{(k)}$ via (4)
  Input $\boldsymbol{s}^{(k)}$ to the top level
  Obtain $Q_{\mathrm{T}}\left(\boldsymbol{s}^{(k)}, p'\right)$ and $E_{\mathrm{T}}\left(\boldsymbol{s}^{(k)}, p'\right)$
  Formulate policy distribution $\boldsymbol{\pi}_{\mathrm{T}}$ via (5)
  Select p based on π_{T}
  Formulate $\hat{\boldsymbol{s}}^{(k)} = \left[\boldsymbol{s}^{(k)}, p\right]$ for the lower level
  Formulate π_{L} via (5)
  Select m based on the policy distribution $\boldsymbol{\pi}_{\mathrm{L}}$
  Share $\boldsymbol{F}$ with the transmit power p at sub-frame m
  Receive τ, b and ρ from ego CAV
  Evaluate $u^{(k)}$ via (6)
  Evaluate risk level $r^{(k)}$ via (7)
  Update Q-value via Bellman equation
  Update E-value via (8)
end for

As shown in Fig. 3, in the policy selection period, a collaborative CAV estimates the packet timestamp $\left[\hat{t}_i\right]_{i\in\mathcal{N}}$ of the feature maps via Eq. (2), and obtains the network topology $[\boldsymbol{L}_i]_{i\in\mathcal{N}}$ and channel gains $[h_i]_{i\in\mathcal{N}}$ to formulate the state $\boldsymbol{s}^{(k)}$ as follows,

$$\boldsymbol{s}^{(k)} = \left[[\boldsymbol{L}_i]_{i\in\mathcal{N}}, [\hat{t}_i]_{i\in\mathcal{N}}, [h_i]_{i\in\mathcal{N}}, \bar{\rho}, \bar{\tau}, \bar{b} \right], \tag{4}$$

where $\bar{\rho}$, $\bar{\tau}$ and $\bar{b}$ are the average perception accuracy, latency and PLR, respectively.

Unlike the centralized RL that the state and action spaces increase exponentially with the number of CAVs, the distributed two-level hierarchical RL decomposes the policy into two sub-policies to improve learning efficiency. More specifically, based on the state $\boldsymbol{s}^{(k)}$ as the input, the top level of the hierarchical RL outputs the expected long-term reward $Q_{\mathrm{T}}\left(\boldsymbol{s}^{(k)}, p'\right)$ for each state-action pair and the long-term risk $E_{\mathrm{T}}\left(\boldsymbol{s}^{(k)}, p'\right)$ for evaluating the policy risk. A modified Boltzmann distribution π_{T} given by Eq. (5) is then formulated to choose the transmit power $p \in \{jP/M_1 | 0 \le j \le M_1\}$.

Similarly, the state for the lower level is formulated by $\hat{\boldsymbol{s}}^{(k)} = \left[\boldsymbol{s}^{(k)}, p\right]$ to choose the sub-frame $m \in \{1, 2, ..., M_2\}$ based on π_{L} given by

$$\pi_{\mathrm{T/L}} = \frac{\exp\left(\xi_j Q\left(\boldsymbol{x}^{(k)}, a\right) - E\left(\boldsymbol{x}^{(k)}, a\right)\right)}{\sum\limits_{a'\in\mathbf{A}_j} \exp\left(\xi_j Q\left(\boldsymbol{x}^{(k)}, a'\right) - E\left(\boldsymbol{x}^{(k)}, a'\right)\right)}, \tag{5}$$

where $\boldsymbol{x}^{(k)} \in \left\{\boldsymbol{s}^{(k)}, \hat{\boldsymbol{s}}^{(k)}\right\}$ and $a \in \{p, m\}$, and $[\xi_j]_{1\le j\le 2}$ weighs the trade-off between the long-term reward and risk in terms of power control and sub-frame selection, respectively.

In the feature sharing period, each collaborative CAV shares the feature map $\boldsymbol{F}$ to neighboring vehicles with the chosen transmit power and sub-frame. Ego CAV exploits the channel-

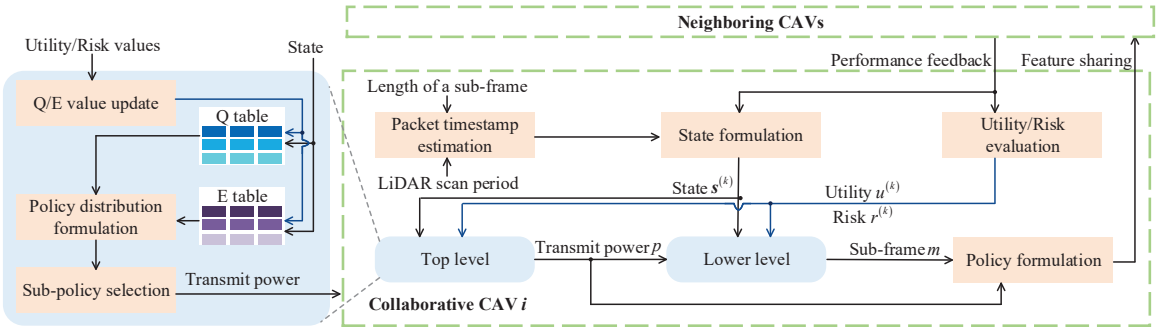


Fig. 3. Illustration of RL-based collaborative perception scheme, in which collaborative CAV i chooses the transmit power $p \in \mathbf{A}_1 = \{jP/M_1 | 0 \leq j \leq M_1\}$ and sub-frame $m \in \mathbf{A}_2 = \{1, 2, ..., M_2\}$ to share the feature map to the neighboring CAVs.

wise self-attention mechanism to aggregate the received feature maps via Eq. (1), further evaluates the perception accuracy ρ based on the perception results z, calculates the latency τ and PLR b of the feature map sharing, and finally responds these performances to neighboring CAVs in the feature fusion period.

The utility $u^{(k)}$ is evaluated based on a weighted sum of average perception accuracy $\bar{\rho}$, latency $\bar{\tau}$ and PLR $\bar{b}$ as follows,

$$u^{(k)} = \bar{\rho} - c_1\bar{\tau} - c_2\bar{b} \tag{6}$$

Similar to [18], the Q-values in both top and lower levels are updated via the Bellman equation. In addition, the average latency $\bar{\tau}$ and PLR $\bar{b}$ are exploited to formulate risk function $r^{(k)}$, which in turn are quantized into C_1 and C_2 levels with the different penalty factors $c_{l,1}$ and $c_{l,2}$ associating with each risk level $\mu_{l,1}$ and $\mu_{l,2}$ respectively, i.e.,

$$r^{(k)} = \sum_{l=0}^{C_1} c_{l,1}\mathcal{I}\left(\bar{\tau} > \mu_{l,1}\right) + \sum_{l=0}^{C_2} c_{l,2}\mathcal{I}\left(\bar{b} > \mu_{l,2}\right), \tag{7}$$

where $\mathcal{I}(\cdot)$ represents the indicator function that returns 1 if the argument is true and 0 otherwise. Finally, the E-values of both top and lower levels are updated based on the previous L steps of risk levels given by

$$E_{\mathrm{T/L}} = \sum_{l=0}^{L} \beta^l r^{(k-l)}, \tag{8}$$

where β weighs the importance of the previous risk levels.

V. Performance Analysis

According to [18], the computational complexity of RLCP mainly relies on formulating the modified Boltzmann distribution for choosing the collaborative perception policy. More specifically, the hierarchical RL structure repeatedly chooses the transmit power with M_1 levels and one out of the M_2 sub-frames within K time slots based on the distribution $\pi_{\mathrm{T/L}}$, which is formulated based on a weighted sum of the Q and E values under the given state that is calculated by traversing the sub-policy space $\boldsymbol{A_1}$ and $\boldsymbol{A_2}$, yielding the computational complexity given by $\mathcal{O}(K(M_1 + M_2))$.

The upper performance bound of the collaborative perception in terms of the accuracy, latency and utility is provided based on the Nash equilibrium of the cooperative game, in which collaborative CAV $i \in \mathcal{N}$ chooses the transmit power $p_i \in \{jP/M_1 | 0 \leq j \leq M_1\}$ and one out of the M_2 sub-frames $m_i \in \{1, 2, ..., M_2\}$ to send the feature maps to the ego CAV to enhance the accuracy with less latency and PLR given by Eq. (6). According to [1], [19], the perception accuracy decreases with the latency $\boldsymbol{\tau} = [\tau_i]_{1\leq i\leq N}$ and PLR $\boldsymbol{b} = [b_i]_{1\leq i\leq N}$, which can be denoted by $\rho = \mathbb{F}(\boldsymbol{\tau}, \boldsymbol{b})$ with the regression parameters $\boldsymbol{\varpi} = [\varpi_i]_{1\leq i\leq N}$ and $\boldsymbol{\varphi} = [\varphi_i]_{1\leq i\leq N}$ for simplicity given by

$$\mathbb{F}(\boldsymbol{\tau}, \boldsymbol{b}) = \rho_0 - \sum_{i=1}^{N}(\varpi_i\tau_i + \varphi_i b_i), \tag{9}$$

where ρ_0 is the benchmark accuracy in the ideal transmission conditions.

The latency and packet loss for sharing a D-bit feature map depend on the SINR given by

$$\gamma_i = \frac{p_i h_i}{\sum\limits_{l=1, l\neq i}^{N} p_l h_l \mathcal{I}(m_i = m_l) + N_0} \tag{10}$$

Without loss generality, the packet timestamp of the feature map generated by CAVs is assumed to follow $t_1 < \cdots < t_i < t_j < \cdots < t_N < \delta_1 + \delta_2$. Thus, the latency of feature map sharing is given by

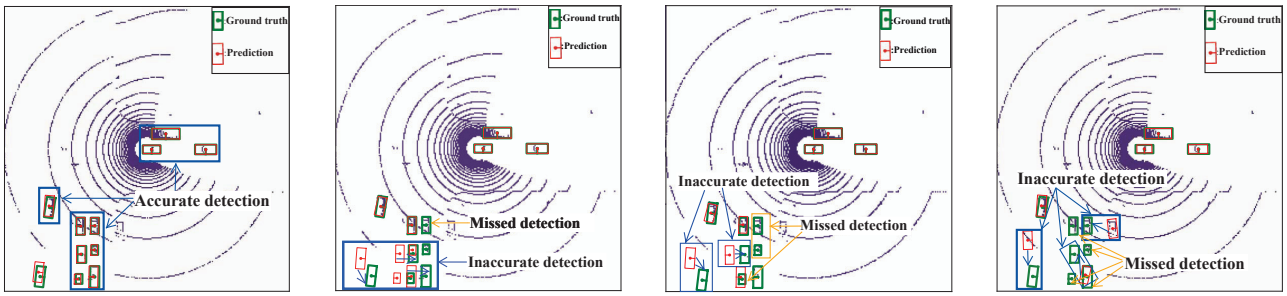
$$\tau_i = m_i - t_i + T\mathcal{I}(m_i < t_i) + \frac{\iota D}{B\log_2(1+\gamma_i)}, \tag{11}$$

where ι is the modulation and coding rate. The packet loss at each time slot is denoted by $b_i = \mathcal{I}(\gamma_i < \gamma_0)$.

Theorem 1. *The upper performance bound of the RL-based collaborative perception is given by*

$$\rho^* = \rho_0 - \sum_{i=1}^{N}\varpi_i(\lceil t_i \rceil - t_i) - \frac{\iota D}{B}\sum_{i=1}^{N}\frac{\varpi_i}{\log_2(1+\frac{Ph_i}{N_0})} \tag{12}$$

$$\tau_i^* = \lceil t_i \rceil - t_i + \frac{\iota D}{B\log_2(1+\frac{Ph_i}{N_0})} \tag{13}$$



(a) RLCP (b) ABFRL [10] (c) DARP [6] (d) Single perception

Fig. 4. Example detection performance of LiDAR-based collaborative perception, in which the red and green bounding boxes represent the prediction and ground truth of vehicle locations, respectively.

$$u^* = \rho_0 - \frac{1}{N}\sum_{i=1}^{N}\left(c_1 + N\varpi_i\right)\left(\lceil t_i \rceil - t_i\right) - \frac{\iota D}{NB}\sum_{i=1}^{N}\frac{c_1 + N\varpi_i}{\log_2(1+\frac{Ph_i}{N_0})}, \tag{14}$$

if

$$h > \frac{\gamma_0 N_0 M_1}{P} \tag{15}$$

$$t_j > \delta_2/M_2 + t_i \tag{16}$$

Proof. See Appendix A. □

Remark: If the packet arrival time of the feature map between any two CAVs surpasses the time duration of a sub-frame given by (16), the feature map is transmitted in the nearest sub-frame with the maximum transmit power to minimize the latency given by (13). Feature maps are successfully decoded if the minimum channel gain between the ego and collaborative CAVs exceeds the threshold given by (15) that depends on the minimum SINR requirement γ_0, the maximum transmit power P and the power level M_1, yielding the performance bound of utility and perception accuracy given by (12) and (14).

VI. Simulation Results

Simulations based on the aggregation model in [12] and the autonomous driving dataset in [13] were performed to evaluate the perception accuracy and latency that involves 5 CAVs to share feature maps of point clouds in object detection. Each CAV selects one out of the 10 sub-frames and transmit power up to $P = 100$ mW with 10 levels to share the 250-KB feature map. Based on the vehicular channel model as presented in [20], the reference path-loss h_0 is 68.83 at distance $d_0 = 1$ m, the path-loss exponent λ is 2.75 and the standard deviation of shadowing σ is 5.5. The received packet is successfully decoded if the SINR exceeds the threshold $\gamma_0 = 15$.

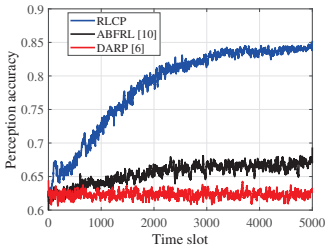
Each CAV optimizes the collaborative perception policy with the learning rate $\alpha = 0.4$ and discounted factor $\delta = 0.3$ to maximize the utility with coefficients $c_1 = 0.1$ and $c_2 = 2$. The policy distributions make a trade-off between the policy quality and exploration risk with the weights $\xi_{\mathrm{T}} = \xi_{\mathrm{L}} = 10$, and the risk function in Eq. (7) takes $c_{l,1} = c_{l,2} = 0.5$ to evaluate the risk level based on both the latency and PLR that are quantified into two levels with the threshold $\mu_{l,1} = 50$ and $\mu_{l,2} = 0.1$, respectively.

The LiDAR-based detection performance of proposed RLCP scheme outperforms the benchmarks ABFRL in [10], DARP in [6], and the single perception scheme as shown in Fig. 4, in which RLCP scheme accurately detects neighboring CAVs due to the lower latency and PLR for exchanging the feature maps of point cloud frames (i.e., the bounding boxes of vehicle prediction (red boxes) are well aligned with ground truths (green boxes)). In addition, the single perception scheme fails to detect the vehicles around the bottom left corner due to the occlusions and limitation of sensing range, and DARP fails to detect 5 vehicles due to time asynchronization and information loss of the point cloud frames by using the reservation-based sub-frames allocation.

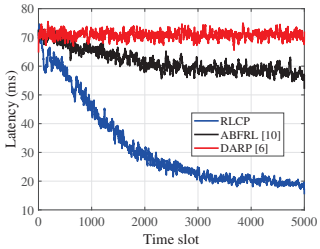
The performance of collaborative perception in Fig. 5 averaged by 100 runs each with 5000 time slots shows that our proposed scheme outperforms the benchmark in terms of perception accuracy, latency and utility, in which the perception accuracy is evaluated based on the Intersection-over-Union threshold of 0.5. The accuracy of RLCP increases from 62.2% to 84.7% over time slots due to the lower latency of the feature map sharing, which decreases from 70.9 to 19.5 ms. In addition, our proposed scheme improves the perception performance and communication efficiency compared with ABFRL in [10] against path-loss and interference, which improves 17.1% perception accuracy and reduces 67.3% latency after 4000 time slots.

VII. Conclusion

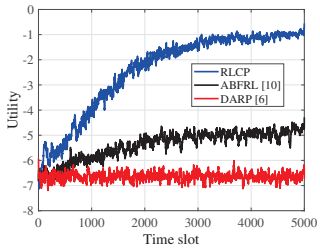
In this paper, we proposed a RL-based collaborative perception scheme for each CAV to choose the transmit power and sub-frame for the feature map exchange to enhance the perception performance. The upper performance bound of the latency and perception accuracy is provided, showing that the perception accuracy grows linearly with the bandwidth and logarithmically with the maximum transmit power. Simulation results based on five CAVs equipped with LiDAR sensors



(a) Perception accuracy



(b) Latency



(c) Utility

Fig. 5. Performance of the LiDAR-based collaborative perception over 100 runs, each with 5000 time slots.

to detect vehicles in town areas show the performance gain over ABFRL [10] with 17.1% higher perception accuracy and 67.3% less latency.

Appendix A
Proof of Theorem 1

Proof: By (6), (9)-(11), if (15) and (16) holds, $\forall p_i \in \{jP/M_1 | 0 \le j \le M_1\}$ and $\forall m_i \in \{1, 2, ..., M_2\}$, we have

$$\begin{aligned} u(p_i, m_i) &= \rho_0 - \frac{1}{N}\sum_{i=1}^{N}\big(c_2 + N\varphi_i\big)\mathcal{I}(\gamma_i < \gamma_0) \\ &\quad - \frac{1}{N}\sum_{i=1}^{N}\big(c_1 + N\varpi_i\big)\big(m_i - t_i + T\mathcal{I}(m_i < t_i)\big) \\ &\quad - \frac{\iota D}{NB}\sum_{i=1}^{N}\frac{c_1 + N\varpi_i}{\log_2(1+\gamma_i)} \\ &\le \rho_0 - \frac{1}{N}\sum_{i=1}^{N}\big(c_1 + N\varpi_i\big)\big(\lceil t_i\rceil - t_i\big) \\ &\quad - \frac{\iota D}{NB}\sum_{i=1}^{N}\frac{c_1 + N\varpi_i}{\log_2(1+\frac{Ph_i}{N_0})} = u\big(P, \lceil t_i\rceil\big) \end{aligned} \tag{17}$$

Therefore, we have the Nash equilibrium of the game that the collaborative perception policy is maximized at $\big[P, \lceil t_i\rceil\big]$, yielding the performance bound given by (12)-(14).

References

[1] 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, 2023.

[2] K. Qu, W. Zhuang, Q. Ye *et al.*, "Model-assisted learning for adaptive cooperative perception of connected autonomous vehicles," *IEEE Trans. Wireless Commun.*, 2024.

[3] X. Ye, K. Qu, W. Zhuang *et al.*, "Accuracy-aware cooperative sensing and computing for connected autonomous vehicles," *IEEE Trans. Mobile Comput.*, 2023.

[4] Y. Hu, S. Fang, Z. Lei *et al.*, "Where2comm: Communication-efficient collaborative perception via spatial confidence maps," in *Proc. Advances in Neural Inf. Process. Syst. (NeurIPS)*, New Orleans, LA, 2022.

[5] R. Xu, H. Xiang, X. Xia *et al.*, "OPV2V: An open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Philadelphia, PA, 2022, pp. 2583–2589.

[6] H. Mosavat-Jahromi, Y. Li, Y. Ni *et al.*, "Distributed and adaptive reservation MAC protocol for beaconing in vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 20, no. 10, pp. 2936–2948, 2021.

[7] F. Mason, M. Giordani, F. Chiariotti *et al.*, "An adaptive broadcasting strategy for efficient dynamic mapping in vehicular networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5605–5620, 2020.

[8] F. Lyu, H. Zhu, H. Zhou *et al.*, "SS-MAC: A novel time slot-sharing MAC for safety messages broadcasting in VANETs," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 3586–3597, 2018.

[9] H. Ye, G. Y. Li, Juang *et al.*, "Deep reinforcement learning based resource allocation for V2V communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3163–3173, 2019.

[10] 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, 2022.

[11] Y. Ji, Y. Wang, H. Zhao *et al.*, "Multi-agent reinforcement learning resources allocation method using dueling double deep Q-network in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13 447–13 460, 2023.

[12] 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)*, vol. 34, Virtual, 2021, pp. 29 541–29 552.

[13] Y. Li, D. Ma, Z. An *et al.*, "V2X-Sim: Multi-agent collaborative perception dataset and benchmark for autonomous driving," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 10 914–10 921, 2022.

[14] 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.

[15] 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.

[16] 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, 2022.

[17] 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.

[18] 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, 2022.

[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] 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, 2007.