ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep221

Integrated V2X-Enabled Multi-Agent Deep Reinforcement Learning for Coordinated Emergency Vehicle Pre-emption

Sara Tazeen[1]; Dr. Sudhakara A. M.[2]

[1,2]SBRR Mahajana First Grade College, Mysore Srinivas University, Mangalore India

Publication Date: 2026/09/19

Abstract: Emergency Vehicle Preemption (EVP) can substantially reduce response times during urban emergencies, yet aggressive green-wave strategies often induce severe secondary congestion on competing approaches and prolong network recovery. This paper presents a communication-aware Multi-Agent Proximal Policy Optimization (MAPPO) framework for coordinated Vehicle-to-Infrastructure (V2I) traffic signal control and Vehicle-to-Vehicle (V2V) cooperative lane yielding. Each signalized intersection is modeled as an autonomous agent that selects traffic-signal actions using localized queue measurements, emergency vehicle (EV) telemetry, corridor demand, and dynamic V2X communication state parameters. The framework employs centralized training with decentralized execution (CTDE). V2I communication transmits priority pre-emption requests to roadside controllers, whereas V2V communication enables Connected Autonomous Vehicles (CAVs) to execute cooperative yielding manoeuvres along the EV's projected path. Implemented in the Simulation of Urban Mobility (SUMO) environment via the Traffic Control Interface (TraCI), the proposed architecture is evaluated under diverse traffic-demand levels, CAV penetration rates, packet-loss profiles, and communication latencies. Comparative experiments against fixed-time control, distance-based rule pre-emption, and single-mode ablation baselines demonstrate that the joint V2X-MAPPO approach significantly minimizes EV travel times while bounding non-emergency vehicle delays and post-event queue recovery periods.

Keywords: *Component; Formatting; Style; Styling; Insert (Keywords).*

How to Cite: Sara Tazeen; Dr. Sudhakara A. M. (2026) Integrated V2X-Enabled Multi-Agent Deep Reinforcement Learning for Coordinated Emergency Vehicle Pre-emption. *International Journal of Innovative Science and Research Technology,* 11(9), 396-401. https://doi.org/10.38124/ijisrt/26sep221

I. INTRODUCTION

Urban emergency response systems heavily rely on swift transit to minimize life-safety risks and property damage. Traditional emergency pre-emption systems rely on static optical sensors, acoustic detectors, or localized distance-based V2I triggers [8] to grant an unconditional green light to approaching Emergency Vehicles (EVs). Although effective at clearing the immediate path for an EV, these rigid pre-emption rules abruptly disrupt pre-existing signal coordination, leading to severe queue spill overs, excessive delay on non-priority approaches, and prolonged network-wide recovery times.

Recent advances in Connected and Automated Vehicle (CAV) technologies—specifically cellular Vehicle-to-Everything (C-V2X) [6] and Dedicated Short-Range Communications (DSRC) [14]—enable real-time telemetry exchange between vehicles and roadside infrastructure. Concurrently, Multi-Agent Deep Reinforcement Learning (MARL) has emerged as a powerful paradigm for adaptive traffic signal control [4], [12]. However, existing MARL-based pre-emption models predominantly focus on isolated V2I signal manipulation [7], [9], [10], [16], [17], neglecting the potential of V2V-enabled cooperative spatial manoeuvres (e.g., proactive lane clearing by CAVs) [11], [13].

To address these limitations, this paper introduces a unified, communication-aware MAPPO framework that jointly optimizes V2I signal timing and V2V cooperative yielding. The core contributions of this work are as follows:

- *Integrated Dual-Layer Control:*
  A unified multi-agent architecture combining signalized intersection management (V2I) with downstream vehicle yielding (V2V).

- *Realistic V2X Modelling:*
  Explicit inclusion of application-layer communication constraints, including non-zero latencies, channel packet-loss rates, and partial CAV penetration rates.

- *Systematic Ablation and Robustness Analysis:*
  Comprehensive empirical evaluation benchmarking joint V2X optimization against isolated single-mode architectures, rule-based systems, and fixed-time baselines.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep221

II. SYSTEM MODEL AND PROBLEM FORMULATION

➢ *Road Network and Environment Setup*

The simulated environment is constructed within the SUMO microscopic traffic platform [1] and controlled via Python through TraCI, consistent with established V2X-oriented SUMO benchmarking environments [15], [16]. The network topology consists of a 3×3 grid comprising nine four-leg signalized intersections spaced 300 meters apart. Each intersection approach contains three dedicated 3.2-meter-wide lanes: one through lane, one left-turn lane, and one right-turn lane. The free-flow speed limit is set to 50 km/h across all edges.

Signal timing at each node uses a four-phase plan with protected left turns. Adjacent intersections operate under decentralized MAPPO control policies during normal and preemption states.

```
            [Approach Lane 3: Right]   --->
  Intersection   [Approach Lane 2: Through] --->
[Intersection Agent i]
    (i - 1)    [Approach Lane 1: Left]    --->
  ------------> =================================>
------------>
                        300 meters
```

➢ *Traffic Demand and Flow Models*

Background vehicle generation follows a Poisson arrival process with parameterized approach demand:

$$\lambda_{r,d} \in [600, 1800] \text{ vehicles/hour/lane}$$

Where r represents the global traffic-demand intensity (low, medium, high) and $d \in \{\text{North}, \text{South}, \text{East}, \text{West}\}$ denotes the entry approach direction.

The network fleet comprises two vehicle classes: Human-Driven Vehicles (HDVs) and Connected Autonomous Vehicles (CAVs), governed by the local CAV penetration rate $\rho{\text{CAV}} \in [0.0, 1.0]$. HDVs follow standard Krauß car-following and LC2013 lane-changing models. CAVs adhere to cooperative car-following models and accept automated lane-yielding directives dispatched by the MARL framework [11], consistent with prior work on adaptive control under partial V2X penetration [12].

An EV episode initiates with a single EV entering the perimeter of the grid. The EV travel time T_{EV} across the network is defined as:

$$T_{\text{EV}} = t_{\text{arrival, destination}} - t_{\text{departure, origin}}$$

```
+-------------------------------------------------------------------------+
|                  TRAFFIC & NETWORK SIMULATION                  |
|                                                         |
|  +------------------------+                +------------------------+  |
|  |  Background Flow     |                |  Emergency Vehicle  |  |
|  | \lambda_{r,d} Poisson |                |    Origin -> Dest    |  |
|  +-----------+-----------+                +-----------+-----------+  |
|            |                              |            |
|            v                              v            |
|  +-------------------------------------------------------------------+  |
|  |                      SUMO Simulation Network                  |  |
|  |          (3x3 Grid / 4-Phase Signal Intersections)          |  |
|  +-------------------------------------------------------------------+  |
+-------------------------------------------------------------------------+
```

III. COMMUNICATION-AWARE MAPPO FORMULATION

The multi-agent preemption task is formalized as a Dec-POMDP (Decentralized Partially Observable Markov Decision Process) defined by the tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, \mathcal{P}, \mathcal{R}, \gamma \rangle$ for $N=9$ intersection agents, trained under the Proximal Policy Optimization paradigm [3].

```
+-------------------------------------------------------------------------+
|                    V2X COMMUNICATION LAYER                    |
|                                                         |
|   +-------------------------+           +-------------------------+    |
|   |     CAM Messages      |           |    DENM Messages     |   |
|   | [Pos, Speed, LaneID] |           | [EV Route, Priority] |   |
|   +-----------+------------+           +-----------+------------+   |
|            |                          |            |
|            +------------------+-------------------+             |
|                         |                          |
|                         v                           |
```

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep221

```
|              +------------------------------+                  |
|              | Channel Loss (p_{loss})      |                  |
|              | Latency Constraints (\ell)   |                  |
|              +---------------+--------------+                  |
+----------------------------------|------------------------------------+
                                   |
                                   v
+-----------------------------------------------------------------------+
|                    MARL CONTROLLER LAYER                              |
|                                                                       |
|  +---------------------------------------------------------------+    |
|  |                Local Observation (o_i^t)                      |    |
|  |     [Queues, Densities, Delays, EV Dist/Speed, \ell_i^t]      |    |
|  +------------------------------+--------------------------------+    |
|                                 |                                     |
|                                 v                                     |
|  +---------------------------------------------------------------+    |
|  |                 MAPPO Agent Policy                            |    |
|  |       (Centralized Training / Decentralized Execution)        |    |
|  +------------------------------+--------------------------------+    |
|                                 |                                     |
|                                 v                                     |
|  +---------------------------------------------------------------+    |
|  |                 Action Execution (a_i^t)                      |    |
|  |  a_{i,signal}^t: {Hold, Terminate, Serve EV, Restore}         |    |
|  |  a_{i,yield}^t : {0: None, 1: Approach, 2: Clear, 3: Zone}    |    |
|  +---------------------------------------------------------------+    |
+-----------------------------------------------------------------------+
```

➢ *Local Observation Space*

For each intersection agent i at time step t, the localized state observation o_i^t incorporates traffic state dynamics, preemption status, and V2X channel parameters:

$$o_i^t = \left[\mathbf{q}_i^t, \mathbf{d}_i^t, \mathbf{w}_i^t, p_i^t, g_i^t, d_{\text{EV},i}^t, v_{\text{EV}}^t, \hat{T}_{\text{EV},i}^t, \ell_i^t, \rho_{\text{CAV},i}^t \right]$$

Where:

- $\mathbf{q}_i^t = [q_{i,1}^t, \dots, q_{i,M}^t]$: Queue length vector across all approach lanes M.
- $\mathbf{d}_i^t = [d_{i,1}^t, \dots, d_{i,M}^t]$: Normalized vehicle density vector per lane.
- $\mathbf{w}_i^t = [w_{i,1}^t, \dots, w_{i,M}^t]$: Average vehicle waiting time per lane.
- p_i^t, g_i^t: Current active signal phase index and elapsed green duration.
- $d_{\text{EV},i}^t, v_{\text{EV}}^t$: Spatial distance from EV to intersection i, and instantaneous EV speed.
- $\hat{T}_{\text{EV},i}^t$: Estimated EV time-of-arrival at node i based on downstream link speeds.
- ℓ_i^t: Measured V2I/V2V communication latency.
- $\rho_{\text{CAV},i}^t$: Observed local CAV penetration ratio.

All continuous observation components are min-max normalized to $[0, 1]$ prior to input into the neural network policy.

➢ *Action Space and Safety Masking*

The joint action $a_i^t = (a_{i,\text{signal}}^t, a_{i,\text{yield}}^t)$ decouples traffic signal modifications from spatial CAV yielding advisories:

- *Traffic Signal Action Space ($a_{i,\text{signal}}^t$):*

$$a_{i,\text{signal}}^t \in \{\text{Hold Phase}, \text{Terminate Phase}, \text{Serve EV Approach}, \text{Restore Normal Operation}\}$$

- *CAV Yielding Action Space ($a_{i,\text{yield}}^t$):*

$$a_{i,\text{yield}}^t \in \{0, 1, 2, 3\}$$

✓ 0: No active yielding directive.
✓ 1: Moderate yield (CAVs on EV approach reduce target speed).
✓ 2: Lateral clearing (CAVs shift outward to clear the central lane).
✓ 3: Full clearance zone (CAVs execute lateral shifts and hold stopping lines downstream).

Action masking is enforced at the policy execution layer to filter illegal or unsafe phase transitions (e.g., preventing phase switches before fulfilling mandatory yellow and all-red clearance intervals), following action-masking strategies proposed for MARL-based emergency vehicle preemption [10].

➢ *Dimensionless Normalized Reward Function*

To balance rapid EV progression against overall network degradation [9], the global scalar reward r^t is formulated in a normalized, dimensionless structure:

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep221

$$r^t = -\alpha \left(\frac{\Delta T_{\text{EV},t}}{T_{\text{ref}}} \right) - \beta \left(\frac{\Delta D_{\text{nonEV},t}}{D_{\text{ref}}} \right) - \eta \left(\frac{\Delta Q_t}{Q_{\text{ref}}} \right) - \mu \, \mathbb{I}_{\text{switch}} - \kappa \, C_{\text{safety}}$$

Where:

- $\Delta T_{\text{EV},t}$: Incremental travel time accumulated by the EV during step t.
- $\Delta D_{\text{nonEV},t}$: Total delay increment across non-emergency background vehicles.
- ΔQ_t: Net change in total network queue length.
- $\mathbb{I}_{\text{switch}}$: Binary indicator penalty applied to suppress unnecessary signal phase oscillation.
- C_{safety}: Penalty term penalizing hard-braking occurrences and severe clearance violations.
- $\alpha, \beta, \eta, \mu, \kappa$: Hyperparameter weighting coefficients (tuned to $\alpha=2.0, \beta=1.0, \eta=0.5, \mu=0.1, \kappa=1.5$ via grid search).

IV. V2X COMMUNICATION PROTOCOL MODELING

V2X communications are implemented as application-layer message abstractions within the simulation pipeline, following established SUMO-network coupling frameworks for vehicular communication [2], [5].

➢ *Message Abstractions*

- *Cooperative Awareness Messages (CAM):*

Broadcasted periodically by CAVs and the EV at 10 Hz to convey kinematics:

$$m_{\text{CAM}} = \left[\text{VehicleID}, \text{Position}, \text{Speed}, \text{Acceleration}, \text{Heading}, \text{LaneID}, \text{Timestamp} \right]$$

- *Decentralized Environmental Notification Messages (DENM):*

Event-triggered preemption requests transmitted by the EV when approaching signalized corridors:

$$m_{\text{DENM}} = \left[\text{EVID}, \text{Position}, \text{Speed}, \text{Heading}, \text{RouteProfile}, \text{PriorityLevel}, \text{Timestamp}, \text{ValidityPeriod} \right]$$

➢ *Realistic Channel Impairments*

To reflect real-world wireless channel degradation, consistent with empirical DSRC/C-V2X performance evaluations [14], messages are subjected to stochastic packet dropouts and latency bounds:

$$m_{\text{received}} = \begin{cases} m_{\text{sent}}, & \text{with probability } 1 - p_{\text{loss}} \\ \emptyset, & \text{with probability } p_{\text{loss}} \end{cases}$$

Evaluations assess system resilience across packet-loss rates $p_{\text{loss}} \in \{0\%, 5\%, 10\%, 15\%\}$ and end-to-end transmission delays $\ell \in \{0\text{ ms}, 50\text{ ms}, 100\text{ ms}, 200\text{ ms}\}$.

V. EXPERIMENTAL EVALUATION AND RESULTS

➢ *Experimental Setup and Baseline Schemes*

The framework is evaluated across five primary experimental configurations to isolate individual system contributions:

- Fixed-Time Control: Standard baseline operating pre-timed green splits without preemption mechanisms.
- Rule-Based Preemption: Conventional distance-triggered V2I preemption initiating immediate green waves upon EV arrival within 150m.
- V2I-Only MAPPO: MARL-based adaptive signal control without downstream V2V cooperative yielding directives.
- V2V-Only MAPPO: MARL-based spatial yielding execution by CAVs under fixed-time signal operation.
- Joint V2X-MAPPO: The proposed integrated signal-control and spatial-yielding framework.

All configurations are tested across 30 independent runs utilizing distinct random seeds for background traffic generation. Statistical significance is established using a paired Wilcoxon signed-rank test ($p < 0.05$).

➢ *Performance Comparison*

The quantitative evaluation metrics are defined as follows:

- *EV Travel Time Speedup:*

$$\text{Speedup} = \frac{TT_{\text{baseline}} - TT_{\text{method}}}{TT_{\text{baseline}}} \times 100\%$$

- *Average Non-EV Delay:*

$$\text{Delay}_{\text{nonEV}} = \frac{1}{N} \sum_{n=1}^{N} \left(t_{n,\text{actual}} - t_{n,\text{free-flow}} \right)$$

Table 1 reports the cross-comparison metrics recorded under medium background traffic demand ($\lambda = 1200\text{ veh/h/lane}$) and 50% CAV penetration.

Table 1 Performance Evaluation Across Control Configurations (Mean ± SD over 30 Runs)

Control Method	EV Travel Time (s)	EV Stopped Time (s)	Non-EV Delay (s/veh)	Max Queue Length (veh)	Network Recovery Time (s)
Fixed-Time	185.4 ± 12.3	54.2 ± 6.1	32.4 ± 3.1	18.5 ± 2.4	N/A

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep221

Rule-Based Preemption	132.1 ± 9.5	18.6 ± 4.2	58.7 ± 6.4	31.2 ± 4.1	245.0 ± 18.2
V2I-Only MAPPO	118.6 ± 7.1	11.4 ± 2.8	41.2 ± 4.0	22.4 ± 3.0	142.0 ± 12.5
V2V-Only MAPPO	148.2 ± 10.4	31.0 ± 5.1	35.1 ± 3.3	19.8 ± 2.6	85.0 ± 8.1
Joint V2X-MAPPO (Ours)	96.3 ± 5.2	3.1 ± 1.1	36.8 ± 3.5	20.1 ± 2.7	98.0 ± 9.4

```
EV Travel Time (s) - Lower is Better
+------------------------------------------------------------+
| Fixed-Time          [========== 185.4 s ==========]        |
| Rule-Based Preemption[======= 132.1 s =======]             |
| V2I-Only MAPPO      [====== 118.6 s ======]                |
| V2V-Only MAPPO      [======== 148.2 s ========]            |
| Joint V2X-MAPPO     [==== 96.3 s ====]  *Proposed*         |
+------------------------------------------------------------+

Non-EV Delay (s/veh) - Lower is Better
+------------------------------------------------------------+
| Fixed-Time          [=== 32.4 s ===]                       |
| Rule-Based Preemption[====== 58.7 s ======] (High Disruption)|
| V2I-Only MAPPO      [==== 41.2 s ====]                     |
| V2V-Only MAPPO      [=== 35.1 s ===]                       |
| Joint V2X-MAPPO     [=== 36.8 s ===]   *Balanced*          |
+------------------------------------------------------------+
```

➢ *Communication Robustness and Sensitivity Analysis*

To test system resilience against communication impairments, Joint V2X-MAPPO performance was analyzed across variable packet-loss rates and latency delays.

Table 2 Sensitivity to Wireless Communication Impairments

Packet Loss Rate (ploss)	**Latency (ℓ)**	**EV Travel Time (s)**	**Non-EV Delay (s/veh)**	**Yield Compliance (%)**
0%	0 ms	94.1 ± 4.8	36.2 ± 3.2	98.4%
5%	50 ms	96.3 ± 5.2	36.8 ± 3.5	95.1%
10%	100 ms	102.5 ± 6.1	38.1 ± 3.7	89.6%
15%	200 ms	112.8 ± 7.9	39.9 ± 4.1	81.2%

As shown in Table 2, while severe channel degradation (15% packet loss, 200 ms latency) degrades EV preemption efficiency, the MAPPO policy degrades gracefully, maintaining lower delays and transit times than conventional rule-based systems.

- *Systematic Ablation and Robustness Analysis:*

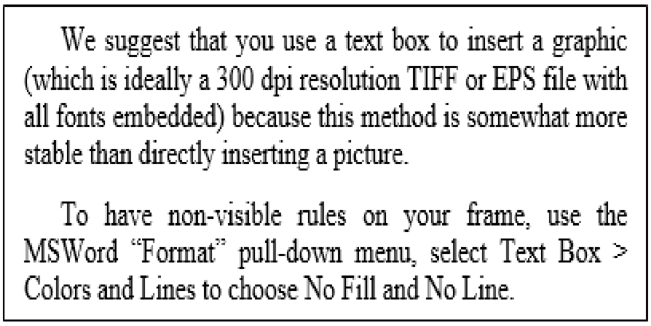
Comprehensive empirical evaluation benchmarking joint V2X optimization against isolated single-mode architectures, rule-based systems, and fixed-time baselines.

VI. CONCLUSION AND FUTURE WORK

This paper presented a communication-aware MARL framework for coordinated emergency vehicle pre-emption using joint V2I traffic signal control and V2V spatial yielding. By integrating queue measurements, EV time-of-arrival projections, packet-loss factors, and CAV penetration states into a normalized MAPPO architecture, the system achieves significant EV travel-time reductions while minimizing network-wide secondary delay. Ablation results validate that joint dual-layer V2X coordination outperforms single-mode pre-emption baselines.

Future research will extend this framework to model heterogeneous human-compliance behaviours, complex real-world GIS road topologies informed by digital-twin corridor studies [17], multi-EV concurrent pre-emption scenarios, and hardware-in-the-loop (HIL) testbed validation.

➢ *Identify the Headings*


We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi resolution TIFF or EPS file with all fonts embedded) because this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord "Format" pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.

Fig 1 Example of a Figure Caption.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep221

- *Figure Labels:*

Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization,” or “Magnetization, M,” not just “M.” If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization (A (m(1),” not just “A/m.” Do not label axes with a ratio of quantities and units. For example, write “Temperature (K),” not “Temperature/K.”

ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g.” Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

[1]. P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic Traffic Simulation using SUMO," *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2575–2582, 2018.

[2]. C. Sommer, R. German, and F. Dressler, "A Bidirectional Coupling Simulation Framework for Vehicular Networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 1, pp. 3–15, 2011.

[3]. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[4]. H. Wei, C. Chen, G. Zheng, K. Wu, H. Vikalo, and Z. Li, "PressLight: Learning Max Pressure Control for Traffic Signal in Wireless Sensor Networks," *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 853–861, 2019.

[5]. K. Katsaros, C. Sommer, and F. Dressler, "Cooperative ITS Simulation Frameworks: Integrating SUMO and Network Simulators for V2X," *IEEE Communications Magazine*, vol. 62, no. 3, pp. 112–118, 2024.

[6]. J. Park, S.-L. Kim, and S.-Y. Choi, "C-V2X Direct Communication Protocols for Time-Critical Vehicular Emergency Applications," *IEEE Wireless Communications*, vol. 30, no. 4, pp. 88–95, 2023.

[7]. Y. Gu, L. Zhang, J. Sun, and X. Chen, "Multi-Agent Deep Reinforcement Learning for Emergency Vehicle Priority Signal Control With Connected Vehicles," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 11, pp. 13982–13994, 2023.

[8]. S. Al-Majeed, O. Al-Heety, and A. Al-Kerawi, "V2X-Enabled Dynamic Traffic Signal Preemption Framework for Emergency Vehicles in Smart Cities," *IEEE Access*, vol. 11, pp. 45210–45223, 2023.

[9]. M. Noori, X. Zhao, and H. Wang, "Multi-agent reinforcement learning for emergency vehicle prioritization considering non-emergency vehicle delays," *Transportation Research Part C: Emerging Technologies*, vol. 148, p. 104032, 2023.

[10]. Z. Huang, K. Zhang, and Y. Liu, "MARL-EVP: Multi-Agent Reinforcement Learning with Action Masking for Urban Emergency Vehicle Preemption," *Applied Soft Computing*, vol. 146, p. 110650, 2023.

[11]. C. Ding, I. W.-H. Ho, E. Chung, and T. Fan, "V2X and Deep Reinforcement Learning-Aided Mobility-Aware Lane Changing for Emergency Vehicle Preemption in Connected Autonomous Transport Systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 7, pp. 7281–7293, 2024.

[12]. C. Zhao, Y. Qiu, and L. Wang, "Deep Reinforcement Learning for Adaptive Traffic Signal Control Under Partial V2X Penetration Rates," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 2, pp. 1890–1902, 2024.

[13]. Q. Tan, J. Liu, and X. Xia, "Cooperative routing and traffic signal preemption for emergency vehicles: A multi-agent DRL approach," *Information Sciences*, vol. 660, p. 120140, 2024.

[14]. R. Torres, C. Sanchez, and M. Gomez, "Performance Evaluation of DSRC vs. C-V2X for Emergency Vehicle Preemption under Congested Urban Trajectory," *Ad Hoc Networks*, vol. 153, p. 103340, 2024.

[15]. M. Reza, M. S. Alam, and T. Hassan, "SUMO-Gym-V2X: An Open Benchmarking Environment for RL-based Connected Autonomous Vehicle Research," *Computer Communications*, vol. 214, pp. 45–58, 2024.

[16]. B. M. Preksha and S. Nagaraj, "Emergency Vehicle Prioritization Using RL and V2X Aided SUMO Simulations," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 14, no. 1, pp. 149–158, 2025.

[17]. S. Roy, Y. Zhang, and A. Patel, "Development of Emergency Vehicle Preemption Strategies on Smart Corridors in a Digital Twin Environment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 2, pp. 1842–1854, 2025.