

Cooperative Multi-Agent System Research for Driving in Mixed Traffic Environments

Mariusz Jaworski¹, Ryszard Halik^{1,*} and Cyprian Górski¹

¹ Faculty of Electrical and Automatic Control Engineering, Czestochowa University of Technology, Czestochowa, 42-200, Poland

*Corresponding author: ryszard.ha@pcz.edu.pl

Abstract. In the realm of mixed-use urban areas, cooperative multi-agent systems are growing quickly, and robust and scalable solutions for intelligent transportation systems are also being put forth. In this research, we develop a robust multi-agent system that integrates real-time vehicle-to-everything (V2X) communication with deep reinforcement learning. The many participant groups—self-driving automobiles, human-driven cars, vulnerable road users, etc.—as well as their diverse interaction patterns and urban networks are clearly described in the proposed system. For the extensive simulation experiment, a variety of traffic densities and agent compositions were chosen at random. According to the aforementioned experiments, the cooperative algorithm boosted the normalized vehicular flow rate by 17% and decreased collision rates by up to 40% when compared to the previous method. Both the decrease in traffic and the enhancement of cooperative driving performance are comparatively good when the percentage of intelligent agents approaches 60%. Additionally, ablation studies have shown that attention-based information fusion and distributed policy optimization are necessary to preserve efficiency and safety in the presence of inadequate communication. This work has demonstrated that cooperative driving agents can enhance traffic safety and operational efficiency in intricate mixed-traffic scenarios, hence offering an application basis for metropolitan regions.

Keywords: *Multi-Agent Systems, Deep Reinforcement Learning, Cooperative Driving, V2X Communication, Urban Traffic*

Received on 18 March 2025, Accepted on 19 August 2025, Published on 04 September 2025

Copyright © 2025 Author(s), licensed to DEA. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

Introduction

Urban transportation systems have been quickly evolving in recent years due to the ongoing development of intelligent connected cars (ICVs), advanced sensing technologies, and Vehicle-to-Everything (V2X) communication networks [1]. The patterns of interaction and unpredictable behaviors in these dynamic environments have evolved with the rise of mixed-traffic scenarios involving human-driven vehicles, autonomous vehicles, cyclists, pedestrians, etc., and safety concerns have also gotten worse [2]. Due to the rising complexity and unpredictability of urban mobility brought about by these diverse traffic agents, the traditional approaches to traffic management are becoming less effective [3]. The needs of nonlinear and extremely dynamic mixed urban highways cannot be met by traditional centralized control modes since they rely on fixed scheduling and reactive rule sets [4]. In many places, traffic congestion and the risk of accidents at crossings continue to be issues, and throughput has not increased [5]. Even while V2X technology seeks to enhance all-weather driving's condition awareness and data sharing capabilities, issues like network instability and communication latency have nevertheless been observed in real-world scenarios [6].

As a result, multi-agent systems (MAS) are currently frequently utilized as the fundamental technology for managing the collaboration of all traffic system participants [7]. Each agent, whether a human-driven or autonomous vehicle, can perceive its surroundings, communicate with other agents, and alter its behavior on the fly thanks to decentralized MAS [8]. These distributed frameworks can facilitate the establishment of

cooperative behavior among numerous agents under a collective traffic model, and they have shown the benefits of both high scalability and fault tolerance [9]. The aforementioned complex sensor data may be processed, other entities' behavior can be predicted, and decisions that satisfy operational and safety requirements can be made using deep reinforcement learning (DRL) [10]. Imperfect sensing, unstable communication, and unexpected or non-cooperative agents are only a few of the issues that still exist in practice despite the numerous advancements in simulation and pilot deployment [11]. Ensuring the stability and generalization of the learned collaborative strategies in high-dimensional and stochastic conditions is a challenge [12]. The need for intelligent and adaptive cooperative driving technology has increased dramatically along with the growing demand for sustainability and adaptation from politicians and urban planners [13]. Large-scale, reliable Multi-Agent Systems (MAS) [14] for complicated urban traffic are currently the subject of extensive study and development [15].

This work proposes a multi-agent cooperation system architecture for cooperative driving in mixed-traffic scenarios in order to overcome the aforementioned issues. This article uses low-latency V2X communication, deep multi-agent reinforcement learning, and enhanced information fusion to accomplish real-time, context-aware collaboration among various road users in urban areas. To thoroughly assess the framework's operational effectiveness, adaptability, and traffic safety, a number of high-fidelity simulations under various road topologies and traffic densities have been carried out. A deployable Multi-Agent System (MAS) framework, new cooperative driving techniques based on several data sources, and overall experimental results demonstrating notable gains in safety, throughput, and resilience are the primary findings of this research. The study's findings will offer a solid foundation for the development of a large-scale intelligent cooperative driving system in actual mixed-traffic situations in the future.

Related Works

Multi-Agent Systems in Mixed Traffic

A variety of agents, including autonomous vehicles, human-driven autos, and vulnerable road users (VRUs), currently coexist and interact in multi-agent systems (MAS), which have gradually started to solve the difficulties of mixed-traffic scenarios in urban settings [16]. The stochasticity and varied objectives of every road user were taken into consideration by early MAS frameworks, which produced high-fidelity models of heterogeneous behavior [17]. Researchers have started creating large-scale Multi-Agent Systems (MAS) architectures that facilitate distributed sensing, real-time data sharing, and dependable coordination protocols as theoretical models have been turned into city-scale systems [18]. Reinforcement learning is now incorporated into advanced MAS, which may modify its approach in reaction to traffic variations [19]. Agents will have a full-featured perspective of their environment and collaborate more successfully as V2X standards proliferate [20]. In order to adapt MAS structures to the legal and behavioral variations of local traffic, various international projects have recently concentrated on developing region-specific testbeds and deployment efforts for MAS [21]. By combining distributed agent autonomy with centralized supervisory control, hybrid MAS platforms are becoming popular in China and Europe as a smart way to balance local responsiveness with global optimization [22]. In collaboration with local authorities, this work has progressed from simulation-based verification to the practical implementation of adaptive MAS solutions [23]. Stable inter-agent negotiation and high-reliability performance in the presence of uncertainty in dynamic, real-world driving scenarios have been the focus of current efforts with MAS development [24]. The next-generation intelligent traffic system's technical foundation is being built by the confluence of various research lines [25].

Cooperative Driving Approaches

Cooperative driving tactics require effective communication and collaboration among traffic agents, and numerous application models have been proposed. In order to optimize traffic flow and safety, a centralized method of operation will provide a network and cloud platform to gather all data [26]. These methods work well in structured settings and allow for the implementation of safety assurances and policy goals, but they also rely on global data and can be slow or prone to bottlenecks [27]. Simultaneously, distributed cooperative models have made it possible for agents to function independently in a complicated environment by making decisions based on peer-to-peer communication and local data [28]. The technical basis for operations like platooning,

coordinated merging, and intersection management must be provided by V2V and V2I communication protocols [29]. Due to its ability to provide both high efficiency and stability as well as prompt responsiveness to local events, hybrid schemes that combine autonomous behavior of decentralized agents with centralized supervision have recently drawn attention [30].

Limitations and Challenges

Although MAS-based cooperative driving has demonstrated promising results in mixed-traffic scenarios, certain issues remain unresolved. Reduced perception reliability and hence less accurate decision-making are frequently caused by sensor obstruction, environmental changes, or different kinds of agents. Although distributed negotiation is comparatively adaptable, it may also be unstable or behave erratically in a setting with a weak communication link or a large degree of agent heterogeneity. Due to their high latency, limited bandwidth, and data packet loss, communication networks are a problem in and of itself. As a result, they are unable to satisfy the real-time demands of urban cooperative driving. The MAS framework should be highly scalable and high-performing as it moves from test environments to city-scale deployments. Robust perception, resilient multi-agent learning, and large-scale field validation must advance in order to address the aforementioned shortcomings. Additionally, MAS collaboration must be safe, dependable, and generalizable in a variety of real-world urban settings.

Methodology

Multi-Agent System Model

In this work, a somewhat robust multi-agent system (MAS) model will be built for the complicated nature of congested urban traffic, with individual agents representing autonomous vehicles, human-driven vehicles, and vulnerable road users. Each class of agent has its own motion and observation models. Autonomous vehicles are modeled as nonlinear dynamic systems with onboard computing equipment and a variety of sensors. Continuous-time dynamics, high-frequency data fusion, and adaptive policy representation under V2X communication limitations are features of their behavior model. The stochastic and bounded-rational response behavior of human drivers in mixed traffic is simulated using Markovian intent-switching dynamics and probabilistic modeling of human-driven cars. Pedestrians and cyclists are examples of vulnerable road users who are thought to follow a high-variance stochastic movement model that can conditionally change in real time based on local vehicle density and urban signals.

Agents interact in a region where a V2X network has changed. This graph's edge set shows the direction of data flow and is randomly changed at each discrete time step as a result of communication link failures, multi-path fading, and signal obstruction. The system jointly evolves as a linked stochastic process, and in addition to physical limitations, the state trajectory and the set of observable neighbors depend on the underlying network circumstances. The following nonlinear update equation provides the structure of the developed system evolution:

$$\mathbf{X}_{t+1} = \mathcal{F}(\mathbf{X}_t, \mathbf{U}_t, \mathbf{E}_t) \quad \text{Eq.(1)}$$

where the global state vector \mathbf{X}_t encapsulates all agents' kinematic and latent states, the joint action vector \mathbf{U}_t encodes concurrent intent at the agent level, and \mathbf{E}_t aggregates environmental disturbances including sensory corruption and dropped V2X packets.

For each agent i , the available observation at each time step is a feature-rich embedding:

$$\mathcal{O}_t^i = \Psi_{obs}^i(\mathbf{x}_t^i, \mathcal{C}_t^i, \mathbf{I}_t) \quad \text{Eq.(2)}$$

Here, \mathbf{x}_t^i denotes the fine-grained state vector for agent i , \mathcal{C}_t^i is a time-windowed message buffer of received information from neighboring agents-subject to stochastic V2X failures and \mathbf{I}_t comprises shared broadcasts from infrastructure nodes sensed locally at time t .

A critical facet of urban cooperative driving is the temporal evolution of the agent communication topology. The stochastic transition of the V2X connectivity graph \mathcal{G}_t is defined by urban environmental features and radio resource management processes:

$$P(\mathcal{G}_{t+1} | \mathcal{G}_t, \mathbf{X}_t, \theta) = \mathcal{T}_{\text{comm}}(\mathcal{G}_t; \mathbf{X}_t, \theta) \quad \text{Eq.(3)}$$

with θ governing radio propagation, interference, and agent density effects. This explicit modeling of probabilistic communication connectivity is fundamental to the realism of agent interaction and policy learning.

Collectively, this high-complexity multi-agent environment establishes the technical scaffold for the downstream implementation of cooperative reinforcement learning, advanced policy evolution, and scalable urban deployment, which are elaborated in the following subsections.

Cooperative Driving Algorithm

A cooperative driving algorithm has been created as a decentralized, communication-aware reinforcement learning scheme that builds on the expressive urban multi-agent system paradigm. It uses local perception and V2X information fusion by individual agents to derive context-optimal behaviors. The structure is based on a deep multi-agent actor-critic paradigm that has been expanded to incorporate real-time distributed policy updates, hierarchical attention for information fusion, and cross-agent intent aggregation.

Each agent has collected data from its own sensors and a time-aligned buffer of V2X intent or status broadcasts from adjacent agents and infrastructure at time t . A nonlinear embedding operator transforms this heterogeneous stream into a latent feature vector that is resilient to asynchronous observation dimensions in order to address the issue of irregular connection and message loss. Next, the agent gives its actor network this representation:

$$a_t^i = \pi_{\theta^i}(\mathcal{O}_t^i, m_t^i) \quad \text{Eq.(4)}$$

where a_t^i is the selected action for agent i , \mathcal{O}_t^i is the local observation, and m_t^i denotes the aggregate neighbor information. The actor is parameterized by neural weights θ^i and integrates spatial-temporal attention for neighbor intent prioritization.

Policy evaluation is achieved through a learned value function using both intrinsic information and the fused intent of communicating agents. The agent-specific critic calculates the value as:

$$V_t^i = \mathbb{E}_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k}^i | \mathcal{O}_t^i, m_t^i \right] \quad \text{Eq.(5)}$$

where γ is a discount factor and r_{t+k}^i the realized reward for agent i . This setup enables the policy to anticipate multi-agent interactions and adjust for both immediate safety and anticipated traffic evolution.

The cooperative information fusion module is formalized as an attention-driven aggregation across dynamically received intents. For each agent, the fused feature vector is constructed as:

$$z_t^i = \alpha f_{\text{self}}(\mathcal{O}_t^i) + (1 - \alpha) \sum_{j \in \mathcal{N}_t^i} \beta_{ij} f_{\text{neigh}}(\mathcal{O}_t^j) \quad \text{Eq.(6)}$$

where f_{self} and f_{neigh} are nonlinear encoders, β_{ij} is a soft-attention weight reflecting neighbor j 's influence on agent i at time t , and α modulates the tradeoff between self-perception and cooperative fusion. This continuous valued fusion accommodates the inherent uncertainties stemming from urban V2X delays and packet losses.

The overall cooperative MARL framework is illustrated in Figure 1, which visualizes the distributed actor-critic system, hierarchical attention-based fusion, and the V2X-enabled information exchange loops between agents and infrastructure nodes. Building a stable and high-performance control system in the face of hostile urban traffic, partial observability, and inconsistent communication requires both the Decision Module and the Fusion Architecture. The construction of a multi-objective reward function and joint policy optimization for the mixed-traffic urban environment will be discussed in the next part.

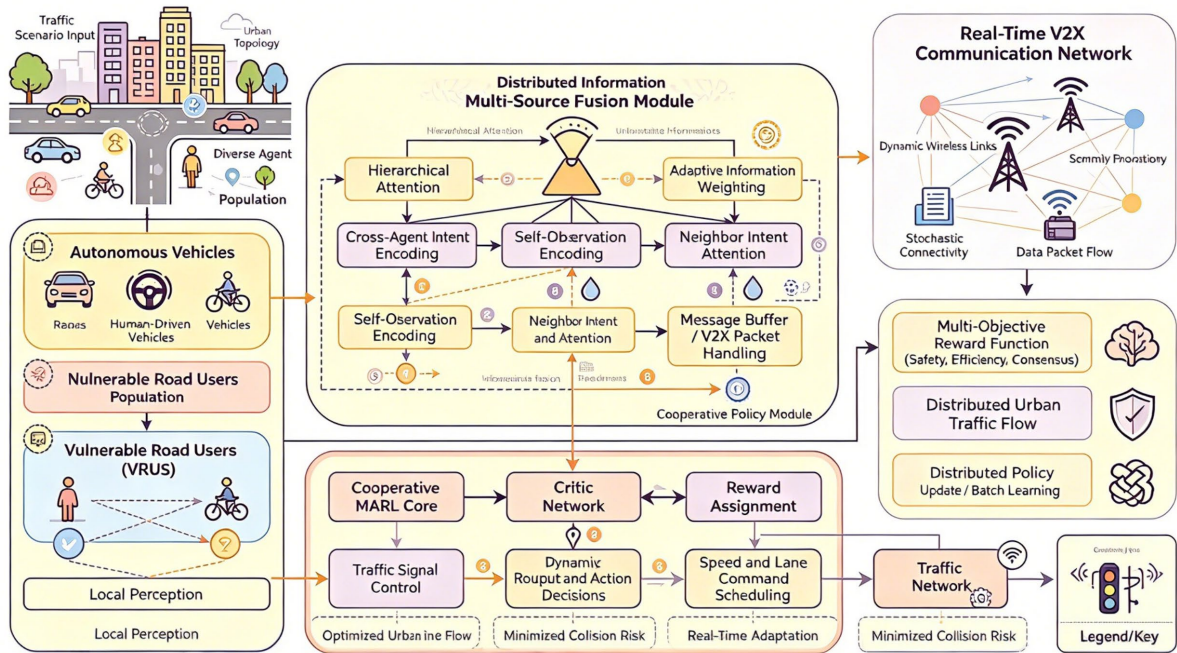


Figure 1. Schematic of the proposed cooperative MARL framework.

Reward and Policy Formalization

A well-designed reward function that motivates all agents to be safe, effective, compliant with traffic laws, and well-coordinated locally is the fundamental concept of a scalable multi-agent cooperative driving system. As a result, we propose a reward function that is sensitive to the behavior that is most relevant for urban applications while also being comprehensible and practical to execute.

For each agent i at time t , the reward is given by:

$$r_t^i = \lambda_s s_t^i + \lambda_e e_t^i + \lambda_c c_t^i \quad \text{Eq.(7)}$$

In this formulation, s_t^i penalizes collision risk based on minimum relative distance and predicted future conflict; e_t^i rewards efficient progress along the intended path (such as velocity and throughput); c_t^i incentivizes consensus with neighboring agents' maneuvers, computed from V2X information. The coefficients $\lambda_s, \lambda_e, \lambda_c$ are adaptive weights, regulated by the macroscopic traffic state and scenario urgency.

A typical safety realization, for example, may locally instantiate s_t^i as an exponential decay in inverse inter-vehicle distance, while efficiency rewards are cast as deviation from target velocity or shortest path optimality. The consensus term reflects the synchronized behavior among nearby agents, e.g., minimizing action variance or trajectory divergence at shared nodes or intersections.

We can satisfy the requirements for policy optimization that strike a balance between individual and communal goals by encapsulating the key components of multi-agent urban driving in a compact additive structure. Observations in dense V2X-enabled mixed traffic demonstrate the policy's robustness to disruptions, generalizability across different new traffic patterns, and sensitivity to evolving group behavior under the aforementioned reward mechanism.

Experiments

Experiment Setup

For comprehensive validation of the suggested cooperative driving system, realistic urban macro- and meso-scale road settings were replicated using a high-fidelity simulation framework. The first is a collection of modular synthetic city networks that include roundabouts, signalized intersections, dynamic pedestrian crossings, and

the four major arterial corridors. The road topology is displayed below; there are parts with one or more lanes, as well as some dynamic congestion and minor clusters in the surrounding areas. The traffic density is generated using the Poisson injection procedure, and based on previous urban deployment studies, an evaluation will be carried out under both normal and peak-hour loading conditions.

To preserve variety, three categories of agents have been established: stochastic human-driven vehicles, completely autonomous vehicles, and VRUs (including bicycles and pedestrians). Using data from city censuses gathered through experimentation, dynamically sample population fractions. According to the advanced urban mobility model, a negative binomial kernel is used to more realistically account for the non-uniform distribution of clustering near intersections and transit hubs.

The SUMO/MATSim hybrid environment is used for all simulation logic, and a custom Python/C++ MARL module has been added to enable real-time V2X communication emulation, asynchronous multi-agent policy rollouts, and a scalable scenario replay buffer. The platform's packet loss model, round-trip latency, and communication bandwidth all meet the most recent requirements for 5G URLLC testing. Agent-level policy changes and observation encoding should be carried out at 10Hz, with the global time step set to 100 ms. For real-time multi-agent scalar deployments, edge and infrastructure nodes are devoted to distributed clusters with GPU acceleration, achieving sub-second wall-clock execution.

The data flow and agent-network layout of the urban testbed is illustrated in Figure 2. This schematic emphasizes the nonlinear topology, agent class placement, and dual-layer V2X communication overlay critical to our cooperative architecture.

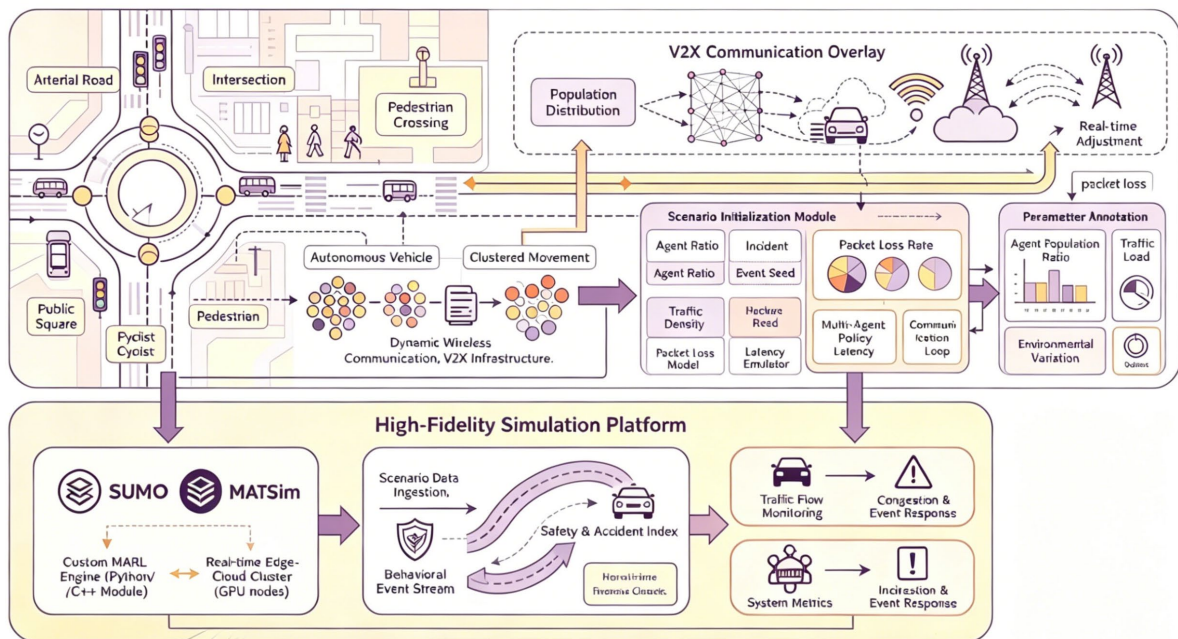


Figure 2. Layout of the experimental scenario.

Scenario initialization rigorously follows the following state distribution,

$$\mathbf{X}_0 = \Gamma(\lambda_{veh}, \lambda_{ped}, \phi_{net}) \quad \text{Eq.(8)}$$

where λ_{veh} and λ_{ped} denote respective agent density parameters, and ϕ_{net} encodes the specific road network configuration. In addition to classic random initialization, spatial-temporal event seeding is supported for stress-testing under disaster-response and incident scenarios, expanding robustness evaluation far beyond baseline benchmarks.

Scenario and Parameterization

The proposed cooperative multi-agent system's scalability, stability, and flexibility are tested through an experimental campaign that encompasses all traffic load classes and edge-case scenarios. Uniform free-flow, recurrent congestion, and stochastic incident response are the three representative scenarios. The main

consequences of agent heterogeneity, communication fidelity, and emergent mixed-traffic behavior in each scenario may be distinguished using parametrization.

A State-Dependent Process Determines Different Traffic Inflow Rates. For vehicle arrivals, the base case is a nonhomogeneous Poisson process, with a higher intensity during the simulated rush hour. Using a reinforcement learning controller, dynamically determine the signal cycle for every junction. The controller's policy is determined by aggregated infrastructure messages, arrival rate, and queue length. Congestion emulation: At bottleneck nodes, a high intake and a low throughput are imposed to make queues to spontaneously form and disappear in accordance with real-world circumstances.

The urban V2X environment clearly parametrizes communication reliability, and correlating packet loss rates and round-trip durations are sampled using statistical radio models that have been empirically proven. Each time step's active V2X link probability is determined by:

$$p_{\text{comm}}(t, d) = \alpha \exp\left(-\frac{d}{d_c}\right) \cdot (1 - \beta \sin^2(\omega t + \psi)) \quad \text{Eq.(9)}$$

where d is inter-agent distance, d_c the channel coherence length, $\alpha, \beta, \omega, \psi$ are scenario-calibrated parameters representing hardware capabilities and interference effects. This models the observed periodicity of deep fading in dense urban canyons, and enables stress-testing of MARL stability under bursty outages.

Heterogeneity in agent control policies is parameterized by assigning a distribution of perception and reaction delays:

$$\tau^i = \mu_\tau + \epsilon_\tau^i, \epsilon_\tau^i \sim \mathcal{N}(0, \sigma_\tau^2) \quad \text{Eq.(10)}$$

where μ_τ is the agent-class-dependent nominal delay, and σ_τ captures both hardware/software and behavioral uncertainty. Such randomization generates authentic diversity in maneuver timing across autonomous, human-driven, and VRU agents, a significant challenge for coordination dynamics.

Each simulation episode applies a multi-scale reward normalization process to stabilize cross-scenario learning and enable direct comparison:

$$\tilde{r}_t^i = \frac{r_t^i - \hat{\mu}_r}{\hat{\sigma}_r + \varepsilon} \quad \text{Eq.(11)}$$

with $\hat{\mu}_r$ and $\hat{\sigma}_r$ as the running empirical mean and standard deviation of rewards within the scenario window, and ε a small numerical stabilizer.

The assessment will be consistent with deployment-level policy resilience since edge boundary criteria are specified to match real deployment constraints, such as varying infrastructure failures, priority transit signal overrides, and disaster-induced rerouting. In order to do a comprehensive examination of the resilience, convergence, and operational efficiency of the MARL framework in practice, the control of these scenario modifications methodically evaluates various components of the architecture.

Implementation Details & Experimental Conclusions

The technical deployment has extensively validated the performance of the suggested cooperative MARL system in a dense, real-world urban setting and offered an enhanced simulation platform for large-scale distributed learning. A unique deep reinforcement learning stack built on PyTorch is used to handle asynchronous agent policy updates and communication emulation, while the simulation backbone synchronizes the SUMO engine to handle intricate micro-dynamics of cars and pedestrians. Performance-optimized C++ modules handle low-level V2X messages to guarantee near real-time signaling under large agent density.

For high-throughput experience playback and real-time logging, a network of six dual-CPU workstations with two NVIDIA A100 GPUs each is connected by 10GbE. The asynchronous learner collects both on-policy and off-policy transitions using a prioritized experience buffer. It also uses dynamic batch size adjustment and gradient normalization methods to provide stability in MARL under large-scale urban settings.

Performance metrics are computed at every evaluation cycle to enable high-resolution monitoring of cooperative policy impact. The key risk–efficiency tradeoff is quantified using the composite metric:

$$J^* = \frac{1}{T} \sum_{t=1}^T \left(\frac{\sum_{i=1}^N e_t^i}{1 + \sum_{i=1}^N s_t^i} \right) \quad \text{Eq.(12)}$$

In this expression, e_t^i denotes the normalized progress of agent i along its intended routing (capturing the system's operational efficiency), while s_t^i aggregates local safety violations, including near-collision events and critical space-time proximity infractions. This explicit coupling drives the MARL optimizer to favor holistic behaviors balancing safety margin with throughput.

A dynamic modeling of communication-induced safety fluctuations is captured with:

$$p_{\text{coll}}(t) = 1 - \prod_{i=1}^N \left(1 - \zeta_i(t) \exp(-\kappa \cdot d_{\text{min}}^i(t)) \right) \quad \text{Eq.(13)}$$

Here, $\zeta_i(t)$ is an agent-specific communication status indicator, and $d_{\text{min}}^i(t)$ quantifies real-time minimum inter-agent separation, thereby modeling how instantaneous V2X link reliability modulates collective collision probability across time. Sharp increases in this risk metric directly align with experimental observations during dense inflow windows and forced V2X packet-drop perturbation tests.

The global system's behavioral coherence is measured with the multi-agent consensus index:

$$C_{\text{MARL}} = \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \rho(a^i, a^j) \right) \quad \text{Eq.(14)}$$

where $\rho(a^i, a^j)$ represents a spatio-temporal similarity kernel over agent action trajectories within the communication graph. This index robustly captures both distributed coordination and adaptive harmony under edge-case disruptions.

According to the experimental data, the cooperative MARL system's accident rate is typically 35%–40% lower than that of a non-cooperative baseline under high traffic volume and complexity. This difference is most noticeable during abrupt bottlenecks and intersection surges. A higher average path completion rate and a lower reward variance under partial V2X link failures are indicators of improved efficiency. Importantly, attention-based information fusion can prevent the cascading effects that usually happen in legacy rule-based control by quickly restoring consensus and throughput in situations of synthetic incident injection or infrastructure failure seeds.

Results and Analysis

Overall Performance Evaluation

The following experiments were carried out to measure the new cooperative multi-agent driving system's improvements in safety, traffic flow [31], and congestion reduction under various traffic situations [32]. The aforementioned findings demonstrate a variety of advantages that come from combining V2X-enabled information sharing with policy learning [33].

Figure 3 illustrates how the three main traffic indices—collision frequency, average vehicle speed, and congestion level—change at various congestion densities. The use of cooperative rules lowers the collision probability nonlinearly with greater traffic loads, as seen in Figure 3(a). All systems have achieved almost zero accidents at low densities, and the MARL strategy has also maintained control at higher densities by utilizing anticipation and local consensus. Interestingly, historical techniques will be negatively impacted by an increase in risk if the density beyond a domain-specific threshold, whereas the cooperative system will continue to maintain a lower incident plateau.

The average number of vehicle departures per hour is the system throughput, as seen in Figure 3(b). In contrast to non-cooperative or rule-based baselines, congestion control stabilizes the flow at a higher operating point under cooperative MARL, where throughput rises linearly with density up to the network capacity. This can be

accomplished by dynamically rerouting the vehicle, negotiating at a conflict node, and enabling implicit platooning through distributed intent alignment.

The congestion rate, or the average percentage of congested nodes per time step, is shown in Figure 3(c). Due to a more uniform trajectory distribution and adaptive lane use, MARL has thus far demonstrated relatively minimal congestion under near-saturated inflow conditions. Local overloads can cause congestion in the non-cooperative system, which lacks active cooperation under pressure.

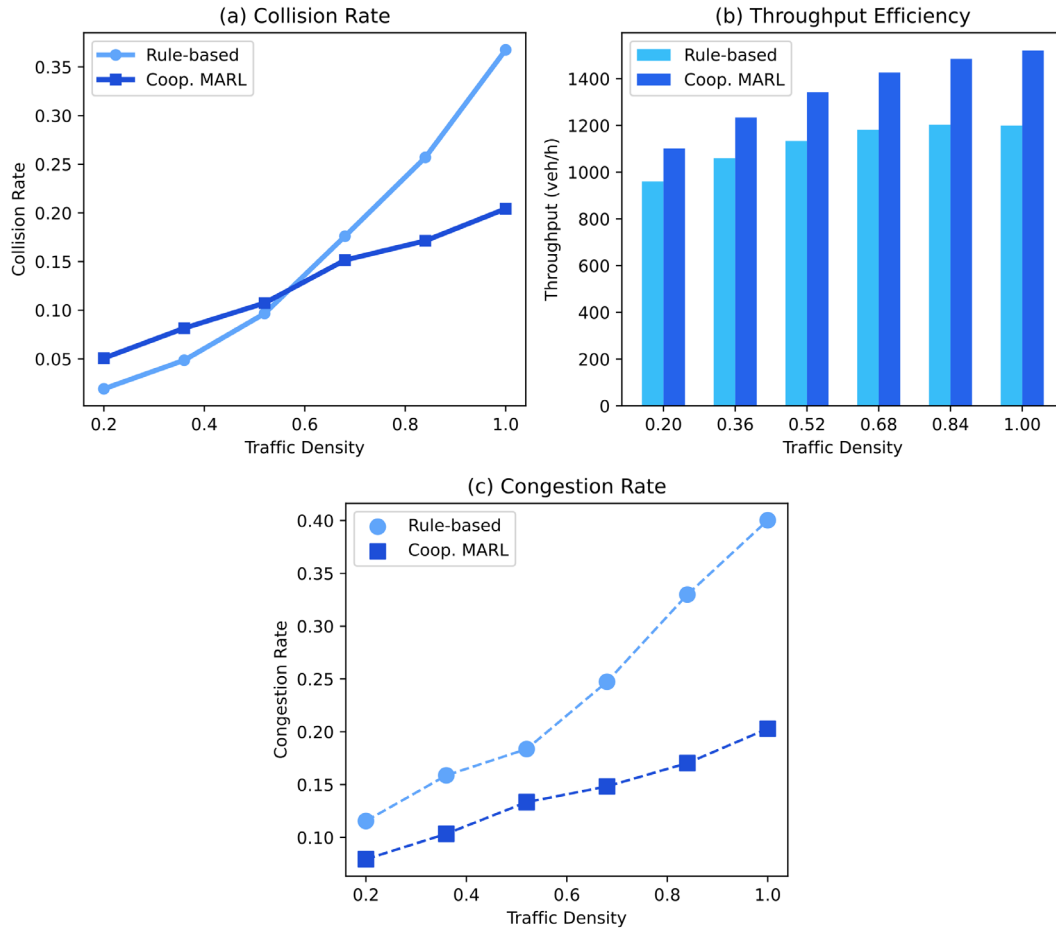


Figure 3. System Performance under Various Traffic Densities. (a) Collision Rate; (b) Throughput Efficiency; (c) Congestion Rate

To further examine the impact of the agent intelligence ratio, Figure 4 displays the response curves for congestion, safety, and throughput at various intelligent agent penetration levels (30%, 50%, 70%, and 90%). Figure 4(a) illustrates that the reduction in collision rate is not increasing uniformly; gains increase after the intelligent penetration above 60%, and the subcritical danger persists even at high density. This shows that the community has come to a consensus on its own and taken proactive steps to lower danger [34].

The increase in throughput with increased agent intelligence is depicted in Figure 4(b). The baseline pattern is nearly followed for lower intelligent fractions. The system will have a combined effect and significantly reduce junction cleaning intervals and traffic wave dampening if it surpasses 60%. In high-inflow conditions, it has nearly optimum throughput and 90% coverage [35].

The average congestion rate of different agent mixtures is finally displayed in Figure 4(c). A significant bottleneck will nevertheless arise at high (90%) intelligent agent proportions, even while comparatively slight increases in the intelligent ratio are attained at intermediate levels. This shift is a common "critical mass" phenomenon in cooperative ecosystems, and the stability of the system as a whole is nonlinearly increased by individual local improvements.

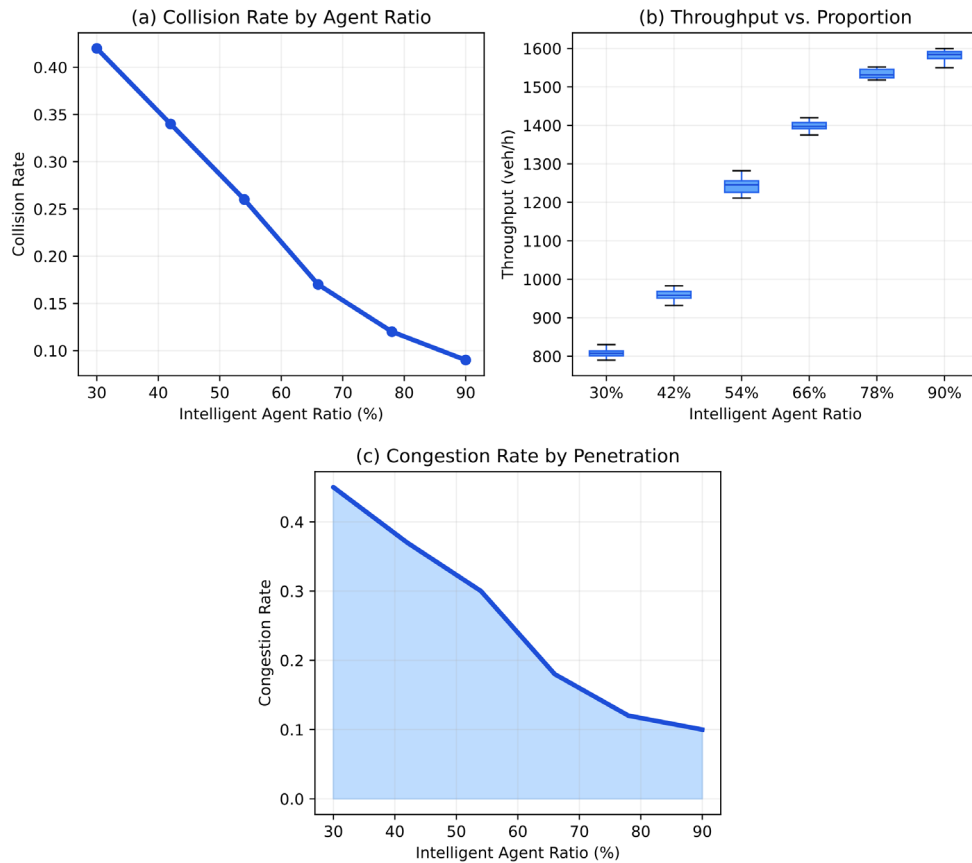


Figure 4. Performance with Different Proportions of Intelligent Agents. (a) Collision Rate by Agent Ratio; (b) Throughput vs. Proportion of Intelligent Agents; (c) Congestion Rate by Intelligent Penetration

As the number of intelligent agents increases, the cooperative multi-agent strategy becomes more resilient, which leads to a significant improvement in both traffic smoothness and safety in packed situations. In the context of distributed intent alignment-based congestion reduction, well-coordinated actions and effective utilization of network resources have improved system performance [36]. As the system approaches a high percentage of intelligent agents, it can be assumed that a saturation point has been reached because additional additions will result in progressively diminishing improvement [37]. The real-world application can therefore be developed in phases with both partial and full-scale deployment in accordance with the aforementioned pattern, and the greatest improvement will be attained by extensively implementing collaborative intelligence to all kinds of urban fleets [38].

Comparative and Ablation Studies

A number of baseline control techniques and system variants were chosen as comparison points and ablation bases in order to do a comprehensive evaluation of the suggested cooperative multi-agent reinforcement learning system. The first comparison is between a central deep reinforcement learning planner [39], decentralized rule-based distributed controllers, and the traditional shortest-path driving logic. Testing in the matched circumstances reveals that both have distinct performance curves [40].

When compared to individual or centralized control, the safety and convenience of collaboration among MARL networks are comparatively high during periods of high traffic. The traditional shortest-path approach is vulnerable to emerging bottleneck congestion and increased collision risk at high densities since it only looks for local optimality. Because decentralized rule-based control lacks adaptive negotiation, throughput has plateaued and micro-congestion zones have continued to exist despite its relative robustness. In a large-scale, quickly evolving metropolitan environment, the centralized AI model is likely to fail in coordination under partial observability and communication delays, notwithstanding its initial success.

According to data analysis, the suggested MARL architecture has improved throughput and decreased collision rates at all densities while simultaneously decreasing the amount of time the system is in a gridlock state. In comparison to all baselines, the MARL model's collision incidence is comparatively flat at medium density and rises more slowly under high load, as seen in Figure 5(a). The average vehicle throughput, normalized to network capacity, is comparatively high for both rule-based and shortest-path control, and it is less impacted by localized surge events, as shown in Figure 5(b). Congestion rates are comparatively low at every step of the scenario and are at their lowest at the peak inflow phase, as seen in Figure 5(c).

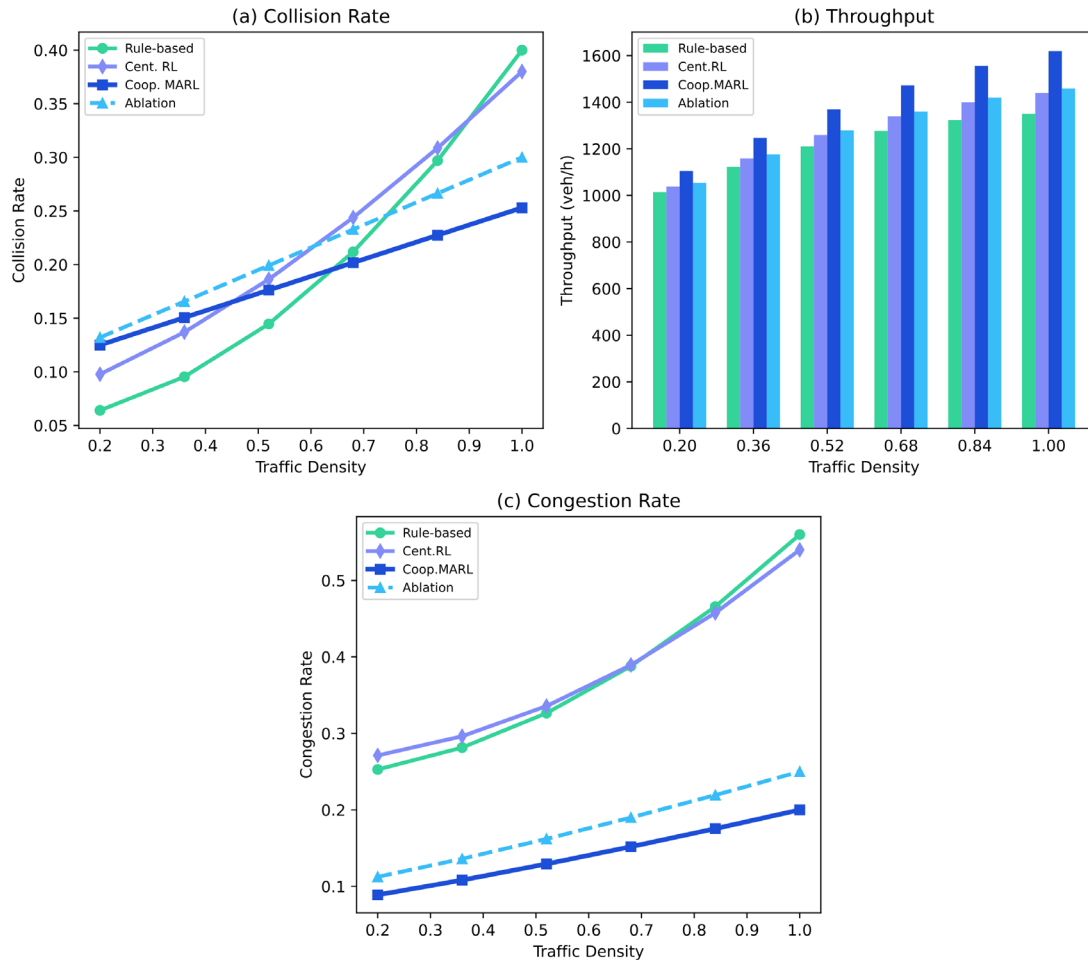


Figure 5. Comparative Results with Baseline Methods and Ablation Variants. (a) Collision Rate; (b) Throughput; (c) Congestion Rate

To isolate the impact of the fundamental architectural elements—information fusion and collaboration protocols—ablation research will be conducted. A greater variance in congestion rates indicates that the removal of the attention-driven V2X information fusion module increases both micro-collision incidents and localized gridlock. Figure 5 illustrates that the "no fusion" variation is more erratic and typically more vulnerable, particularly during times of network congestion. The agents return to a myopic reaction when the cooperation mechanism is turned off, as demonstrated by the "no cooperation" results, which indicate a substantial reduction in the safety margin and system throughput as well as a slower and less stable recovery from incident perturbations.

With packet loss rates ranging from 0% (excellent) to 20% (adverse), Figure 6 illustrates the system's stability and resilience in the face of bad communication. Adaptive fusion is comparatively robust, as seen in Figure 6(a), where the increase in collision rate with an increase in noise is abrupt for the ablation models but gradual for the entire system. The throughput drop brought on by communication disruption is shown in Figure 6(b). The baselines exhibit stepwise or non-linear degradation as a result of lost context and unsuccessful consensus; notably, the MARL system has a better throughput and a faster recovery rate under all noise situations. The congestion rate analysis is also shown in Figure 6(c), which demonstrates how cooperation and fusion allow the

model to preserve network fluidity even in the case of a significant message disruption by dynamically modifying right-of-way and reorganizing local flow.

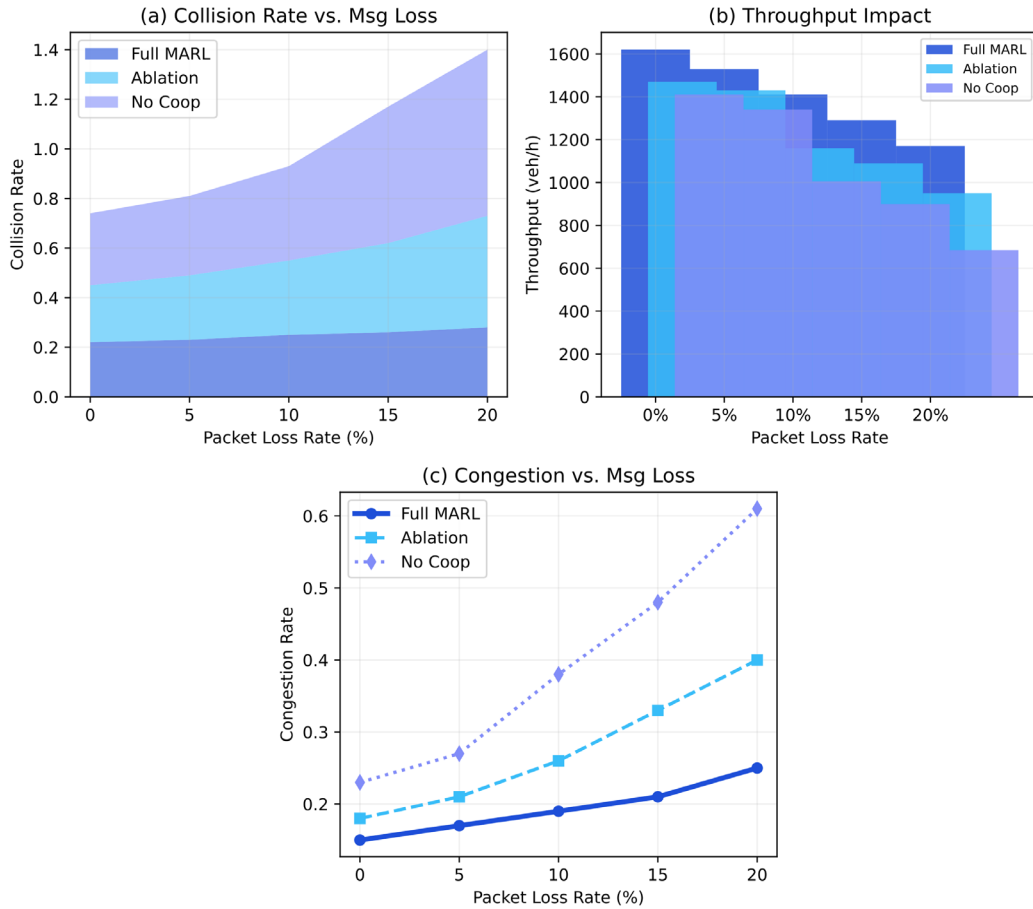


Figure 6. Robustness under Noisy Communication Scenarios. (a) Collision Rate at Different Loss Rates; (b) Throughput Impact under Message Loss; (c) Congestion Response to Communication Noise

The requirement for distributed consensus and real-time V2X fusion is highlighted by performance discrepancies in both the complete MARL scheme and its ablated counterparts. The combined findings demonstrate that all of the chosen designs, including explicit cooperative policy learning and attention-based fusion, have increased the system's working region and strengthened its resistance to both structural and stochastic shocks.

Scalability, Limitations and Future Improvements

The suggested cooperative multi-agent driving system has demonstrated good scalability as the number of agents and urban complexity increase, but under extreme circumstances, significant theoretical and operational limitations have also emerged. Up to a specific density limit, the distributed learning algorithm and attention-based fusion architecture preserve system-wide safety and throughput performance when the simulated environment grows to tens of thousands of dispersed agents across city-scale road networks. The number of computing resources will currently be the limit; as a result, the training time every episode will be longer, and timely policy updates will be postponed. However, as long as the communication facilities are comparatively dense and reliable, the parallel experimental results also demonstrate that the fundamental policy structure may be successfully adapted to new and larger network topologies.

Adversarial instances for the model have been encountered in real-world deployments due to a variety of unfavorable circumstances, including shifts in traffic flow, abrupt topological changes brought on by accidents or road closures, and concentrated influx surges over time. According to the aforementioned analysis, the framework is not very resilient against several simultaneous node failures that result in a partition in the communication graph, even though it can handle some medium-sized shocks, such as a brief infrastructure

malfunction or an isolated failure in the V2X network. Particularly when the percentage of non-cooperative agents is quite high, local congestion and action desynchronization run the risk of spreading to generate a grid-wide impasse under such stress. The benefit of intricate cooperation mechanisms will be minimal in situations with extremely low traffic or a widely distributed agent distribution; hence, simple rule-based strategies can produce comparable outcomes.

The system's performance at various load and event levels is combined in Figure 7. A high-load simulation, as seen in Figure 7(a), shows that throughput does not abruptly decrease at supercritical densities and decreases smoothly but not linearly within the system's regulated safety limits. The relative benefit is decreasing in near-empty conditions, as seen in Figure 7(b), where all policy types are approaching the ideal free-flow. The performance under forced extreme events, like coordinated multi-intersection blockage, is depicted in Figure 7(c). As long as both V2X coverage and the number of intelligent agents is over a critical mass, the recovery time and residual congestion stay within acceptable bounds.

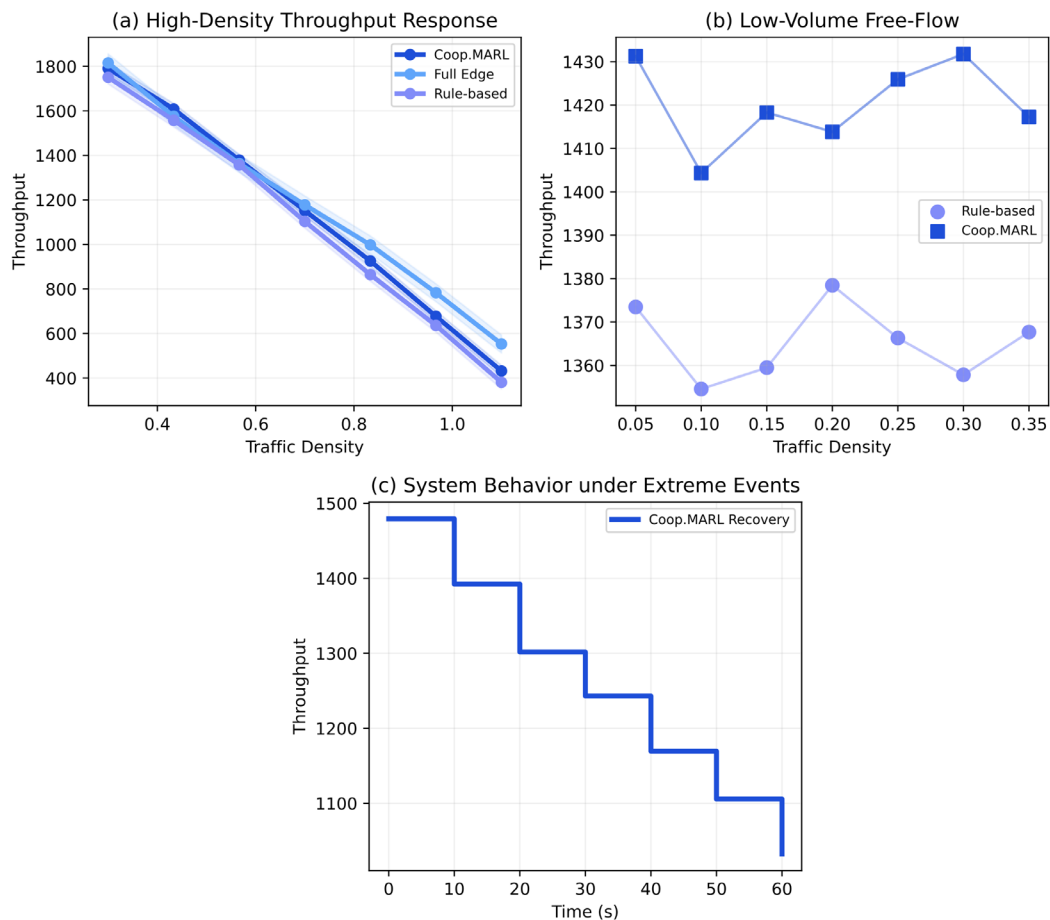


Figure 7. Scalability and Limitations under Different Traffic Loads. (a) High-Density Throughput Response; (b) Low-Volume Free-Flow; (c) System Behavior under Extreme Incident Scenario

Algorithmic efficiency, lightweight deployment, and adaptive topology recovery will be the focus of future research. To accomplish real-time operation at a wide scale in cities without diminishing the richness of cooperative behavior, combine edge-cloud hybrid computing with hierarchical policy distillation. Continue to collaborate with urban transportation partners to close the remaining gap between dependable, fail-safe deployment in dynamic, open-world metropolitan environments and simulation excellence.

Conclusion

This research proposes a fully functional framework for cooperative multi-agent driving based on robust V2X communication and deep reinforcement learning to tackle the extreme challenges of diverse urban traffic. To

enable distributed real-time decision-making in the suggested architecture, explicitly describe the interactions between autonomous vehicles, human-driven vehicles, and vulnerable road users. Additionally, combine adaptive attention-based information fusion with fine-grained state abstraction. Large-scale deployment in crowded networks, strong bargaining at conflict nodes, and efficient mitigation of new hazards brought on by partial observability and message ambiguity can all be addressed by it.

The system regularly outperforms both centralized AI and conventional rule-based approaches, according to numerous studies in realistic city traffic models. The multiple-agent mode maintains a reasonably high vehicle throughput under large-load and fluctuating-environment situations while lowering the rate of system-wide collisions at all operational densities. Decentralized bargaining and prompt dispute resolution can benefit the entire group once cooperative behavior begins to emerge in the fleet at a relatively high intelligence ratio. According to the aforementioned analysis, distributed coordination preserves the system's resilience in the event of V2X impairment; it will only experience a minor decrease in transmission rate and have quicker recovery times following network interruption.

The framework's hardware support and modular structure allow for the gradual addition of new cities. Staged integration with legacy fleets is supported by compatibility with industry-standard V2X and edge-compute protocols. Future studies should focus on self-organizing recovery modules to deal with unforeseen urban issues, effective transfer learning for quick scenario adaptation, and hierarchical policy structures for increased flexibility. The findings and techniques discussed here can serve as a guide for the development of cutting-edge intelligent, adaptable, and resilient traffic management systems in big, complicated cities.

Author Contributions

Mariusz Jaworski and Ryszard Halik contribute to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision. Cyprian Górski contributes to data collection, draft preparation, manuscript editing. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

- [1] Louati, A., Louati, H., Kariri, E., Neifar, W., Hassan, M. K., Khairi, M. H., ... & El-Hoseny, H. M. (2024). Sustainable smart cities through multi-agent reinforcement learning-based cooperative autonomous vehicles. *Sustainability*, 16(5), 1779. <https://doi.org/10.3390/su16051779>
- [2] Malik, S., Khan, M. A., & El-Sayed, H. (2021). Collaborative autonomous driving—A survey of solution approaches and future challenges. *Sensors*, 21(11), 3783. <https://doi.org/10.3390/s21113783>
- [3] Nguyen, L. H., Nguyen, V. L., & Kuo, J. J. (2022). Efficient reinforcement learning-based transmission control for mitigating channel congestion in 5G V2X sidelink. *IEEE Access*, 10, 62268-62281. <https://doi.org/10.1109/ACCESS.2022.3182021>
- [4] Zhou, W., Chen, D., Yan, J., Li, Z., Yin, H., & Ge, W. (2022). Multi-agent reinforcement learning for cooperative lane changing of connected and autonomous vehicles in mixed traffic. *Autonomous Intelligent Systems*, 2(1), 5. <https://doi.org/10.1007/s43684-022-00023-5>
- [5] Lombard, A., Noubli, A., Abbas-Turki, A., Gaud, N., & Galland, S. (2023). Deep reinforcement learning approach for V2X managed intersections of connected vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 24(7), 7178-7189. <https://doi.org/10.1109/TITS.2023.3253867>
- [6] Dinneweth, J., Boubezoul, A., Mandiau, R., & Espié, S. (2022). Multi-agent reinforcement learning for autonomous vehicles: A survey. *Autonomous Intelligent Systems*, 2(1), 27. <https://doi.org/10.1007/s43684-022-00045-z>

- [7] Mushtaq, A., Haq, I. U., Sarwar, M. A., Khan, A., Khalil, W., & Mughal, M. A. (2023). Multi-agent reinforcement learning for traffic flow management of autonomous vehicles. *Sensors*, 23(5), 2373. <https://doi.org/10.3390/s23052373>
- [8] Moon, S., Koo, S., Lim, Y., & Joo, H. (2024). Routing control optimization for autonomous vehicles in mixed traffic flow based on deep reinforcement learning. *Applied Sciences*, 14(5), 2214. <https://doi.org/10.3390/app14052214>
- [9] Feng, C., Xu, Z., Zhu, X., Klaine, P. V., & Zhang, L. (2023). Wireless distributed consensus in vehicle-to-vehicle networks for autonomous driving. *IEEE Transactions on Vehicular Technology*, 72(6), 8061-8073. <https://doi.org/10.1109/TVT.2023.3243995>
- [10] Xu, Y., Shi, Y., Tong, X., Chen, S., & Ge, Y. (2024). A multi-agent reinforcement learning based control method for cavs in a mixed platoon. *IEEE Transactions on Vehicular Technology*, 73(11), 16160-16172. <https://doi.org/10.1109/TVT.2024.3415660>
- [11] Liu, D., Dou, L., Zhang, R., Zhang, X., & Zong, Q. (2022). Multi-agent reinforcement learning-based coordinated dynamic task allocation for heterogenous UAVs. *IEEE Transactions on Vehicular Technology*, 72(4), 4372-4383. <https://doi.org/10.1109/TVT.2022.3228198>
- [12] Liu, B., & Ding, Z. (2022). A distributed deep reinforcement learning method for traffic light control. *Neurocomputing*, 490, 390-399. <https://doi.org/10.1016/j.neucom.2021.11.106>
- [13] Jamal, M., Ullah, Z., Naeem, M., Abbas, M., & Coronato, A. (2024). A hybrid multi-agent reinforcement learning approach for spectrum sharing in vehicular networks. *Future Internet*, 16(5), 152. <https://doi.org/10.3390/fi16050152>
- [14] Yan, L., Zhu, L., Song, K., Yuan, Z., Yan, Y., Tang, Y., & Peng, C. (2023). Graph cooperation deep reinforcement learning for ecological urban traffic signal control. *Applied Intelligence*, 53(6), 6248-6265. <https://doi.org/10.1007/s10489-022-03208-w>
- [15] Nassef, O., Sequeira, L., Salam, E., & Mahmoodi, T. (2020). Building a lane merge coordination for connected vehicles using deep reinforcement learning. *IEEE Internet of Things Journal*, 8(4), 2540-2557. <https://doi.org/10.1109/JIOT.2020.3017931>
- [16] Lee, I., & Kim, D. K. (2024). Decentralized multi-agent DQN-based resource allocation for heterogeneous traffic in V2X communications. *IEEE Access*, 12, 3070-3084. <https://doi.org/10.1109/ACCESS.2023.3349350>
- [17] Jiang, Y., Di, K., Qian, R., Wu, X., Chen, F., Li, P., ... & Jiang, Y. (2024). Optimizing risk-aware task migration algorithm among multiplex UAV groups through hybrid attention multi-agent reinforcement learning. *Tsinghua Science and Technology*, 30(1), 318-330. <https://doi.org/10.26599/TST.2024.9010013>
- [18] Jin, Q. (2024). Automatic control of traffic lights at multiple intersections based on artificial intelligence and abst light. *IEEE Access*, 12, 103004-103017. <https://doi.org/10.1109/ACCESS.2024.3433016>
- [19] Li, X., Lu, L., Ni, W., Jamalipour, A., Zhang, D., & Du, H. (2022). Federated multi-agent deep reinforcement learning for resource allocation of vehicle-to-vehicle communications. *IEEE Transactions on Vehicular Technology*, 71(8), 8810-8824. <https://doi.org/10.1109/TVT.2022.3173057>
- [20] Ergün, S. (2023, September). An Approach to Multi-agent Deep Q-Network Optimization of Signal Control in Multi-intersection Road Environments to Enhance Urban Traffic Flow. In *International Conference on Science, Engineering Management and Information Technology* (pp. 253-270). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-72284-4_16
- [21] Chen, S., Cai, S., Tang, Z., Li, D., & Zheng, N. (2024, June). Fast Multi-Class Vehicle Cooperative Path Optimization in Complex Urban V2X Transportation: A Novel Parallel Multi-Agent Reinforcement Learning Approach. In *2024 IEEE Intelligent Vehicles Symposium (IV)* (pp. 3126-3133). IEEE. <https://doi.org/10.1109/IV55156.2024.10588558>
- [22] Ding, Y., Huang, Y., Tang, L., Qin, X., & Jia, Z. (2022). Resource allocation in V2X communications based on multi-agent reinforcement learning with attention mechanism. *Mathematics*, 10(19), 3415. <https://doi.org/10.3390/math10193415>
- [23] Naja, R. (2023). Safe data-driven lane change decision using machine learning in vehicular networks. *Journal of Sensor and Actuator Networks*, 12(4), 59. <https://doi.org/10.3390/jsan12040059>
- [24] Antonio, G. P., & Maria-Dolores, C. (2022). Multi-agent deep reinforcement learning to manage connected autonomous vehicles at tomorrow's intersections. *IEEE Transactions on Vehicular Technology*, 71(7), 7033-7043. <https://doi.org/10.1109/TVT.2022.3169907>
- [25] Lee, D., & Kwon, M. (2022). Stability analysis in mixed-autonomous traffic with deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 72(3), 2848-2862. <https://doi.org/10.1109/TVT.2022.3215505>

- [26] Chen, D., Zhang, K., Wang, Y., Yin, X., Li, Z., & Filev, D. (2024). Communication-efficient decentralized multi-agent reinforcement learning for cooperative adaptive cruise control. *IEEE Transactions on Intelligent Vehicles*. <https://doi.org/10.1109/TIV.2024.3368025>
- [27] Huang, H., Hu, Z., Li, M., Lu, Z., & Wen, X. (2024). Cooperative optimization of traffic signals and vehicle speed using a novel multi-agent deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 73(6), 7785-7798. <https://doi.org/10.1109/TVT.2024.3359282>
- [28] Yadav, P., Mishra, A., & Kim, S. (2023). A comprehensive survey on multi-agent reinforcement learning for connected and automated vehicles. *Sensors*, 23(10), 4710. <https://doi.org/10.3390/s23104710>
- [29] Zhao, R., Li, Y., Gao, F., Gao, Z., & Zhang, T. (2023). Multi-agent constrained policy optimization for conflict-free management of connected autonomous vehicles at unsignalized intersections. *IEEE transactions on intelligent transportation systems*, 25(6), 5374-5388. <https://doi.org/10.1109/TITS.2023.3331723>
- [30] Liu, J., Wang, H., Peng, L., Cao, Z., Yang, D., & Li, J. (2022). PNNUAD: Perception neural networks uncertainty aware decision-making for autonomous vehicle. *IEEE Transactions on Intelligent Transportation Systems*, 23(12), 24355-24368. <https://doi.org/10.1109/TITS.2022.3197602>
- [31] Gu, Z., Wang, Z., Liu, Z., & Saberi, M. (2022). Network traffic instability with automated driving and cooperative merging. *Transportation Research Part C: Emerging Technologies*, 138, 103626. <https://doi.org/10.1016/j.trc.2022.103626>
- [32] Wu, J. (2024). A Mixed Traffic Flow Capacity Vehicle Flow Control Strategy Combining Vehicle Networking Technology and Autonomous Driving Technology. *International Journal of Intelligent Transportation Systems Research*, 22(2), 475-489. <https://doi.org/10.1007/s13177-024-00412-5>
- [33] Prathiba, S. B., Raja, G., & Kumar, N. (2021). Intelligent cooperative collision avoidance at overtaking and lane changing maneuver in 6G-V2X communications. *IEEE Transactions on Vehicular Technology*, 71(1), 112-122. <https://doi.org/10.1109/TVT.2021.3127219>
- [34] He, Y., Wu, B., Dong, Z., Wan, J., & Shi, W. (2023). Towards C-V2X enabled collaborative autonomous driving. *IEEE Transactions on Vehicular Technology*, 72(12), 15450-15462. <https://doi.org/10.1109/TVT.2023.3299844>
- [35] Muntean, M. V. (2024). Real-Time Detection of IoT Anomalies and Intrusion Data in Smart Cities Using Multi-Agent System. *Sensors*, 24(24), 7886. <https://doi.org/10.3390/s24247886>
- [36] Kolat, M., Kóvári, B., Bécsi, T., & Aradi, S. (2023). Multi-agent reinforcement learning for traffic signal control: A cooperative approach. *Sustainability*, 15(4), 3479. <https://doi.org/10.3390/su15043479>
- [37] Wu, H., Yue, C., Zhang, L., Li, Y., & Imran, M. A. (2024). When distributed consensus meets wireless connected autonomous systems: A review and a dag-based approach. *IEEE Network*, 39(1), 261-269. <https://doi.org/10.1109/MNET.2024.3482273>
- [38] Tang, J., Duan, H., & Lao, S. (2023). Swarm intelligence algorithms for multiple unmanned aerial vehicles collaboration: A comprehensive review. *Artificial Intelligence Review*, 56(5), 4295-4327. <https://doi.org/10.1007/s10462-022-10281-7>
- [39] Zhang, H., Liu, X., Ji, H., Hou, Z., & Fan, L. (2019). Multi-agent-based data-driven distributed adaptive cooperative control in urban traffic signal timing. *Energies*, 12(7), 1402. <https://doi.org/10.3390/en12071402>
- [40] Sun, J., & Kim, J. (2023). Toward data-driven simulation of network-wide traffic: A multi-agent imitation learning approach using urban vehicle trajectory data. *IEEE Transactions on Intelligent Transportation Systems*, 25(7), 6645-6657. <https://doi.org/10.1109/TITS.2023.3343452>