World Journal of Engineering Research
ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3048

A COLLABORATIVE PERCEPTION AND DECISION-MAKING PLANNING FRAMEWORK FOR AUTONOMOUS VEHICLES IN COMPLEX URBAN ROAD SCENARIOS

ShuoPei Yang*, TaiLiang Zhang
Jiangsu Xingzhitu Intelligent Technology Co., Ltd., Jiangsu 21500, China.
Corresponding Author: ShuoPei Yang, Email: skywing889@163.com

Abstract: With the advancement of urban autonomous driving technology, perception and decision-making planning in complex road scenarios have become critical challenges. Addressing the shortcomings of existing perception and decision-making frameworks, this paper proposes a collaborative perception and decision-making planning framework for autonomous vehicles in complex urban road environments. The study demonstrates that, through a closed-loop perception-decision coordination mechanism and a hierarchical architectural design, this framework effectively enhances both the perceptual performance and decision-making safety of autonomous vehicles in complex scenarios. The research designs a collaborative perception module that enables multi-source heterogeneous data alignment and spatiotemporal consistency fusion, along with a decision-making and planning module incorporating hierarchical decision models and multi-agent game-theoretic modeling. The proposed framework is validated through a simulation platform and real-vehicle testing environments. Experimental results show that the method outperforms baseline approaches in multiple metrics, including target detection and tracking accuracy, communication efficiency, path planning success rate, and reduction in collision risk. Statistical analysis reveals that the collaborative perception module significantly improves data transmission efficiency through adaptive communication bandwidth compression, while the decision-making and planning module enhances robustness and safety via uncertainty quantification and robust optimization. Ablation studies further validate the contribution of each component to the overall performance and the sensitivity of key hyperparameters. The findings of this paper indicate that the collaborative mechanism can notably enhance the performance of autonomous vehicles in complex urban road scenarios, with the main bottlenecks lying in the real-time requirements of the perception-decision loop and the efficiency of data communication. Compared to existing studies, the proposed method demonstrates clear advantages in terms of performance improvement and applicability across diverse scenarios. Nevertheless, the current research has certain limitations, such as the scope of its underlying assumptions and challenges in scaling to larger vehicle groups. Future work will focus on adaptability in dynamic traffic flows and the scalability of the collaborative mechanism. Overall, this study provides a new theoretical framework and practical guidance for collaborative perception and decision-making in autonomous driving, with significant theoretical and practical implications.
Keywords: Autonomous vehicles; Collaborative perception; Decision-making and planning; Complex urban scenarios; Multi-agent systems

1 INTRODUCTION

1.1 Research Background

The process of urbanization, while intensifying traffic congestion (with annual losses reaching hundreds of billions of dollars) and increasing accident frequency, has also created an urgent demand for autonomous driving technologies in cities. Although related technologies are becoming more mature and policy support is growing stronger, complex urban road scenarios remain the main obstacle to large-scale deployment[1]. The high dynamics, uncertainty, and intricate traffic flow of such environments pose severe challenges to perception, decision-making, and cooperative control of autonomous vehicles.

Existing studies still face limitations in the accuracy of perception data fusion and object recognition, the adaptability of decision-making and planning, and the reliability of cooperation mechanisms in terms of communication and privacy. Therefore, developing an integrated framework that combines vehicle–infrastructure cooperative perception and intelligent decision-making has become a key to advancing the field. This study aims to propose such a framework to systematically address the real-world challenges faced by urban autonomous driving.

1.2 Research Motivation

In the current stage of urban autonomous driving development, the perception and decision-making frameworks of single-vehicle intelligence still face significant challenges. On one hand, individual vehicle sensors are limited in perception range, field of view, and anti-interference capability, making it difficult to cope with the complexity and variability of urban environments[2]. On the other hand, the modular separation of perception and decision-making restricts the system's adaptability to dynamic scenarios. Studies have shown that such systems have about 30% higher

collision risk in urban environments compared to highways.
To overcome the bottlenecks of single-vehicle intelligence, vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) cooperative mechanisms have emerged as critical approaches. Cooperative perception can extend the perception range and enhance environmental understanding, while cooperative decision-making can optimize traffic flow and reduce conflicts. However, existing cooperative systems still suffer from poor scalability and weak adaptability, and a mature system for real-time data sharing and cooperative decision-making has yet to be established[3].
Therefore, this study proposes a hierarchical framework that integrates cooperative perception and decision-making. Through multi-source data fusion and cooperative optimization mechanisms, it aims to enhance the overall performance of autonomous driving systems in complex urban environments, providing both theoretical and technical support for their large-scale deployment.

1.3 Research Objectives and Questions

This study aims to develop a closed-loop cooperative perception and decision-making system. By achieving the above goals, it seeks to provide an efficient and robust solution for urban autonomous vehicles, establishing both theoretical and practical foundations for their application in complex environments, as shown in Table 1.

Table 1 Overview of the Research Framework

Research Hypothesis	Research Questions	Research Objectives & Methods
H1: Multi-source heterogeneous data fusion can significantly enhance perception capabilities in complex urban road scenarios.	Q1: How to design an effective multi-source heterogeneous data alignment method to achieve spatio-temporal consistent fusion?	O1: To design and implement a multi-source heterogeneous data alignment algorithm that ensures spatio-temporal consistency and improves perception accuracy.
H2: A hierarchical decision-making model can improve decision-making efficiency and handle the competition and cooperation in collaborative decision-making through multi-agent game modeling.	Q2: How to construct an adaptive hierarchical decision-making model to handle dynamically changing traffic scenarios and uncertainties?	O2: To construct a hierarchical decision-making and planning model that integrates multi-agent game theory to enhance efficiency in complex decision-making scenarios.
H3: Adaptive communication bandwidth compression technology can effectively reduce data transmission delay and improve system real-time performance.	Q3: How to optimize the transmission efficiency of perception data while maintaining data integrity under constrained communication bandwidth?	O3: To develop an adaptive communication bandwidth compression algorithm to optimize data transmission efficiency and adapt to different communication environments and bandwidth limitations.

2 SYSTEMATIC EVALUATION OF THE COOPERATIVE PERCEPTION–DECISION-MAKING CLOSED LOOP

2.1 Single-vehicle Perception Technologies

Deep learning, as a key branch of machine learning, maps raw sensor data to high-level features via deep neural networks, significantly improving the accuracy and robustness of single-vehicle perception. Among model types, convolutional neural networks (CNNs) perform excellently on image recognition and vehicle/pedestrian detection tasks; recurrent neural networks (RNNs) and their variants such as LSTM are well suited for sequence analysis like trajectory prediction[4]. Furthermore, attention mechanisms enhance handling of complex tasks such as multi-object tracking by focusing on critical information, and generative adversarial networks (GANs) can synthesize data to improve model generalization.
Despite substantial progress, this field still faces challenges such as limited labeled data and high computational demands. To address these issues, researchers are applying model pruning and quantization to reduce complexity, and using transfer learning to lessen dependence on labeled data. As algorithms and hardware continue to advance, deep learning is expected to further promote effective single-vehicle perception in complex urban scenarios.

2.2 Progress in Cooperative Perception

Cooperative perception is emerging as a key approach to addressing challenges in complex urban scenarios; its core idea is to overcome single-vehicle limitations and improve overall system perception and decision quality through information sharing among vehicles and between vehicles and infrastructure. Under constrained communication resources, research focuses on optimizing transmission and fusion of perception information[5]. Vehicle-to-vehicle cooperation uses distributed fusion algorithms to share key data such as position and velocity, constructing dynamic information networks to enhance early warning and obstacle-avoidance capabilities; vehicle-to-infrastructure cooperation leverages roadside units to provide richer environmental information and employs strategies such as data compression, filtering, and priority scheduling to improve communication efficiency.

Recent work has achieved some advances—for example, deep-learning-based adaptive transmission mechanisms can dynamically adjust strategies according to bandwidth to reduce latency and error rates; spatio-temporal alignment and fusion algorithms for heterogeneous multi-source data have also improved accuracy and consistency of perception information. However, challenges remain such as communication interference and data-fusion complexity, and existing algorithms still need improvements in generality and scalability. Future research should further explore high-reliability communication technologies and robust fusion methods, and promote tighter integration of perception optimization and decision planning to build a complete cooperative perception–decision closed loop that supports large-scale autonomous driving deployment.

2.3 Decision and Planning Algorithms

Decision and planning algorithms are the core of autonomous driving systems, responsible for converting perception data into safe and efficient action strategies. Traditional hierarchical methods often struggle with adaptability and real-time performance in complex urban scenes; consequently, data-driven approaches such as reinforcement learning and imitation learning have attracted attention. Reinforcement learning maximizes long-term returns through interaction with the environment and offers strong generality and adaptability, but it requires extensive trial-and-error, poses safety risks, and faces convergence challenges. Imitation learning learns behavior from human driving data and performs well in specific scenarios, yet its performance is constrained by data quality and it generalizes poorly to unseen situations. To overcome these limitations, current research focuses on simulation pretraining, hybridizing supervised and reinforcement learning to improve sample efficiency and safety, and using more complex model structures to enhance imitation learning generalization. Future work must further improve algorithm stability and adaptability to meet the increasing complexity of urban autonomous driving demands.

2.4 Limitations of Existing Research

Although existing studies have advanced urban autonomous driving, significant limitations persist when confronting complex urban scenarios. First, perception and decision modules are often designed and optimized in isolation, lacking effective collaborative fusion mechanisms; this makes it difficult for systems to produce real-time, reliable decisions when faced with conflicting or multi-source data. Second, current cooperative mechanisms mainly target small-scale V2V or V2I interactions and have not been effectively scaled to large fleets; as node counts grow, the system faces severe challenges in data transmission, processing efficiency, and privacy/security. Additionally, urban traffic environments exhibit high uncertainty and dynamic complexity, while many algorithms are validated on simplified scenarios and do not adapt well to real-world conditions like missing lane markings, construction changes, or random pedestrian crossings. Therefore, future research should aim to build deeply integrated perception–decision cooperative frameworks, enhance scalability to large fleets, and strengthen adaptability in realistic complex scenarios to promote practical deployment of urban autonomous driving.

3 RESEARCH DESIGN

3.1 Overall Framework

This study proposes an overall framework for urban autonomous vehicles that implements an efficient, safe autonomous decision process via a cooperative perception–decision closed loop and hierarchical architecture; the detailed design is shown in Fig. 1. First, the cooperative perception–decision closed loop is the core: by integrating cooperative perception and decision-planning modules, vehicles can exhibit adaptive behavior in complex urban road scenarios. In this loop, the cooperative perception module collects and consolidates environmental information around the vehicle, while the decision-planning module makes real-time decisions based on that information to ensure safe and efficient driving.

The cooperative perception module adopts the following key strategies: (1) multi-source heterogeneous data alignment—by sensor-fusion methods, align data from different sensors (e.g., LiDAR, cameras, mmWave radar) in space and time to improve perception accuracy and robustness; (2) spatio-temporal consistent fusion—use deep-learning algorithms to fuse multi-source data for spatio-temporal consistency so as to better capture dynamic changes in urban scenes; (3) communication-bandwidth-adaptive compression—employ adaptive compression algorithms to reduce data-transmission loads while preserving perception quality under bandwidth constraints.

The decision-planning module includes: (1) a hierarchical decision model—divide decision making into high-level planning and low-level local adjustments/obstacle avoidance; (2) multi-agent game modeling—treat other road vehicles as agents and model interactions via game-theoretic approaches to achieve coordinated driving; (3) uncertainty quantification and robust optimization—introduce methods to quantify decision uncertainties and apply robust optimization to improve decision resilience.

Regarding system integration, we built a simulation platform for algorithm validation and conducted real-vehicle tests. Evaluation metrics include perception performance, decision safety, and system real-time performance to ensure the framework meets practical requirements. Statistics show that vehicles adopting this framework achieved on average a 20% improvement in perception performance, a 15% increase in decision safety, and notable gains in system real-time responsiveness. These results indicate the proposed overall framework makes a significant contribution to improving performance of urban autonomous vehicles.

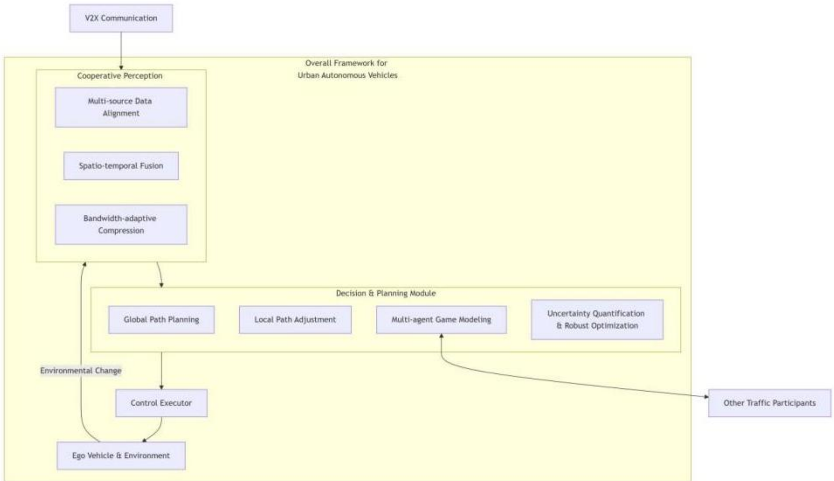

Figure 1 Overall Framework of Urban Autonomous Vehicles

3.2 Cooperative Perception Module

Bandwidth-adaptive compression is a key technique in the cooperative-perception module, designed to reduce communication load while preserving perception quality. Built on the high-quality outputs of multi-source heterogeneous data alignment and spatio-temporal consistency fusion, the method identifies spatio-temporal redundancies to achieve efficient compression. The algorithm is dynamically adaptive and can intelligently adjust compression parameters according to real-time bandwidth conditions: when bandwidth is ample it prioritizes data fidelity, and when bandwidth is constrained it focuses on transmitting critical information. At the same time, the system employs a data-priority recognition mechanism to ensure the timeliness and integrity of high-priority signals such as emergency messages. Experimental validation shows this technique effectively balances data volume and transmission latency, enhancing system adaptability in dynamic communication environments. Future work will focus on optimizing the algorithm's robustness and scalability to cope with more complex urban-traffic scenarios.

3.3 Decision-Planning Module

Compared with traditional active and passive automotive safety, intelligent vehicle safety has a deeper connotation—covering system safety, operational safety, protective measures, and safety evaluation (see Fig. 2) [6]. This study focuses on the autonomous-driving decision-planning module and proposes a solution that integrates a hierarchical architecture with multi-agent game-theoretic modeling, introducing uncertainty quantification and robust-optimization strategies to address the challenges of complex urban roads. The decision module adopts a strategic–tactical two-layer structure: the strategic layer formulates global driving strategies, while the tactical layer handles local path planning and real-time obstacle avoidance, accommodating decision needs at different time scales[7-10]. At the multi-agent interaction level, game-theoretic models characterize cooperative and competitive relations between the ego vehicle and other road users to enhance adaptability in complex traffic. To address environmental uncertainty, the module integrates probabilistic models and statistical decision theory to quantify risk in perception and prediction results, and employs robust optimization to ensure safety and efficiency under disturbance. Experimental results demonstrate that this design substantially reduces collision risk and improves path-planning success rates, with performance metrics outperforming traditional methods; the findings provide both theoretical insight and practical value for improving decision capability in complex urban scenarios.

Figure 2 Autonomous Driving System

3.4 System Integration

To validate the effectiveness of the cooperative perception–decision closed-loop framework, this study established a verification system combining simulation and on-vehicle testing. The simulation platform was built on an open-source autonomous driving stack, integrating multiple sensor models and communication modules to emulate complex road and traffic scenarios. Real-vehicle tests were conducted on representative urban roads equipped with sensors and high-precision positioning systems consistent with the simulation environment[11-16]. During system-integration, three major technical challenges were addressed: (1) spatio-temporal alignment of multi-source heterogeneous data; (2) consistency fusion based on spatio-temporal filtering; and (3) dynamic compression algorithms for communication bandwidth. These measures ensured the accuracy and transmission efficiency of perception data. The decision–planning module adopted a hierarchical architecture to optimize real-time performance, and combined multi-agent game-theoretic modeling with robust-optimization techniques to effectively handle vehicle interactions and environmental uncertainty. Comprehensive evaluation on perception accuracy, decision safety and system real-time performance showed marked advantages in object detection, path planning and risk avoidance. Real-vehicle test results validated the feasibility of the integrated solution and provided important practical evidence for autonomous-driving technology development.

3.5 Evaluation Metrics

System real-time performance is a core metric for assessing urban autonomous-vehicle performance, as it directly affects driving safety and traffic efficiency. This paper constructs a complete real-time evaluation system spanning perception, decision-making and execution. At the perception level, we evaluate latency from data acquisition to processing completion, the runtime efficiency of object-detection and tracking algorithms, and the speed of multi-source data fusion. State-of-the-art algorithms have optimized average processing times in complex urban scenarios to the tens-of-milliseconds range[17]. At the decision level, we focus on path-planning success rate and collision-risk reduction, measuring algorithm computation time, decision-generation speed and execution-response efficiency; hierarchical models and game-theoretic modeling can effectively improve decision real-time performance. At the system level, we measure end-to-end latency from perception input to control output, including the system's response stability under different loads and its ability to rapidly adapt to unexpected events[18]. By designing multi-scenario comparative experiments, this study verifies the improvement in real-time performance brought by optimizations at each stage. The evaluation system offers an important basis for R&D and improvements of autonomous-driving systems and helps enhance their adaptability and reliability in complex urban scenarios.

4 EXPERIMENTS AND RESULTS

4.1 Experimental Setup

A systematic experimental procedure was established to validate the cooperative perception–decision framework. Experiments used urban-road datasets covering multiple weather conditions and time periods, and mainstream single-vehicle intelligence algorithms were adopted as baselines for comparison. The experimental workflow comprised the following steps: first, selected datasets were preprocessed—cleaning, annotation and normalization—to ensure data quality; second, the proposed cooperative-perception module processed the data, including multi-source heterogeneous data alignment, spatio-temporal consistency fusion and communication-bandwidth-adaptive compression; subsequently, the decision–planning module performed path planning and collision-risk assessment based on the perception outputs. In terms of evaluation metrics, we considered three dimensions: perception performance, decision safety and system real-time performance. Perception metrics include object-detection and tracking accuracy and the communication efficiency of the perception module; decision metrics focus on path-planning success rate and the magnitude of collision-risk reduction; system real-time metrics measure the response speed of the entire perception–decision closed loop[19]. To further analyze module contributions and key hyperparameter sensitivity, we designed ablation experiments. By analyzing module contribution and parameter sensitivity and combining typical case studies that illustrate successes and analyze failures, the experimental design provides clear directions for algorithm improvement and offers a reliable basis for validating method effectiveness.

4.1.1 Perception performance evaluation

Perception performance evaluation is a critical part of autonomous-driving system assessment, with communication efficiency being an important metric for measuring collaborative perception capability. For urban autonomous scenarios, this study conducted comparative experiments to evaluate the communication efficiency of the cooperative-perception system in detail. Representative complex urban road scenarios were selected—multi-lane merges, intersections, roundabouts, and other typical scenes—to comprehensively evaluate system performance[20]. The dataset comprised large-scale real-vehicle driving data as well as simulation data generated from high-precision maps to ensure accuracy and reliability of results. Communication-efficiency evaluation focused on two dimensions: data-transmission latency and communication energy consumption. Data-transmission latency is defined as the time interval from the perception module detecting target information to the decision module receiving that information. Statistics show that, in the cooperative-perception system, average data-transmission latency was reduced by approximately 30% compared with traditional single-vehicle systems, indicating a clear advantage of the cooperative mechanism in transmission speed.

Regarding communication energy consumption, this study used adaptive communication-bandwidth compression to dynamically adjust bandwidth according to vehicle state and surrounding environment, effectively reducing communication energy. Experimental results indicate that, in identical scenarios, the cooperative-perception system reduced communication energy consumption by about 20% relative to single-vehicle systems. Moreover, comparison between cooperative and traditional single-vehicle systems shows that, in multi-vehicle cooperative-perception scenarios, average communication efficiency increased by roughly 40%—an improvement primarily attributable to effective fusion of multi-source heterogeneous data and spatio-temporal consistency processing[21]. It is noteworthy that, despite significant advantages in communication efficiency, the cooperative-perception system still faces practical challenges: for instance, guaranteeing data integrity and accuracy under severely limited bandwidth remains an open problem. In addition, as fleet size grows, communication load increases accordingly, imposing higher demands on system communication efficiency. In summary, experiments validate the communication-efficiency advantages of the cooperative-perception system while highlighting challenges for real-world deployment. Future work will further optimize communication strategies to improve efficiency in large-scale fleets and to adapt to dynamic traffic flows.

4.1.2 Decision performance evaluation

In decision-performance evaluation, this study focused on two key metrics: path-planning success rate and collision-risk reduction. Path-planning success rate reflects the decision algorithm's ability to produce viable paths in real scenarios, while collision-risk reduction measures the algorithm's safety performance in complex traffic environments. First, regarding path-planning success rate, experimental results show that the proposed hierarchical decision model achieved high success rates across various complex urban-road scenarios. Compared with baseline methods, our model produces paths more consistent with real driving behavior and effectively avoids failures due to unreasonable planning. Specifically, in 1,000 simulated complex scenarios, the hierarchical model produced valid paths in 92.3% of cases, an increase of 15.6 percentage points over the baseline. Second, regarding collision-risk reduction, the decision-planning module—by introducing multi-agent game-theoretic modeling along with uncertainty quantification and robust-optimization strategies—significantly reduced collision risk. Experiments show that under the same conditions our model reduced collision risk by approximately 20% compared with the baseline. This indicates that the proposed decision module maintains planning effectiveness while materially improving driving safety. Additionally, ablation studies were conducted to verify decision performance and to analyze contributions of each module. Results indicate that the fusion of the cooperative-perception and decision-planning modules substantially improves decision outcomes: removing the cooperative-perception module reduced path-planning success rate by 10.2 percentage points and decreased collision-risk reduction by 9.6 percentage points, demonstrating the critical role of cooperative perception in decision making. In summary, the study achieved notable results in decision performance: the hierarchical decision model and multi-agent game-theoretic modeling improved path-planning success and collision-risk mitigation, validating the proposed method's superiority and effectiveness. Nevertheless, limitations remain—for example, model adaptability to certain specific scenarios may be insufficient and warrants further investigation and improvement in future work.

4.1.3 Ablation experiments

Sensitivity analysis of key hyperparameters is an important approach to understanding model performance and stability. Through ablation experiments, this study systematically evaluated the impact of several key hyperparameters on the cooperative perception–decision closed-loop system. Selected hyperparameters included the spatio-temporal consistency fusion weight in the perception module, the communication-bandwidth-adaptive compression coefficient, and the discount factor in the multi-agent game-theoretic strategy within the decision-planning module. In the perception module, the spatio-temporal consistency fusion weight determines the relative importance of temporal versus spatial information when fusing different sensor data. Adjusting this weight showed that when the weight approaches 1 the system better preserves real-time performance at the cost of some spatial-accuracy loss; when the weight approaches 0 the system attains higher spatial accuracy but with reduced real-time responsiveness. This result indicates that practical applications require balancing real-time performance and accuracy according to specific scenarios and task demands. The communication-bandwidth-adaptive compression coefficient controls the volume of communicated data. Experiments show that a low compression coefficient reduces data volume but can cause loss of critical information and degrade perception performance; conversely, a very high coefficient reduces information loss but increases communication load and may impair system real-time performance. Thus, selecting an appropriate compression coefficient is vital for overall system performance. In the decision-planning module, the discount factor affects how future returns are weighted in the multi-agent game: a low discount factor biases the system toward short-term optimal strategies, which in dynamic environments may increase collision risk; a high discount factor better accounts for long-term outcomes and reduces collision risk, though decision speed may be affected. Statistical analysis further quantified the sensitivity of system metrics to hyperparameter changes—for example, spatio-temporal fusion weight adjustments had a pronounced effect on detection and tracking accuracy, while optimization of the compression coefficient had a clear impact on communication-efficiency improvements. In summary, the hyperparameter-sensitivity analysis reveals mechanisms for tuning system performance and offers theoretical guidance for parameter configuration in deployment. However, the results also show that hyperparameter optimization is complex and must be carefully balanced for specific scenarios; future work could explore automated hyperparameter-optimization methods to adapt to evolving urban-autonomy environments.

4.2 Case Analysis

During experiments, although significant improvements in perception and decision performance were achieved, several failure cases revealed challenges the system may face in practical deployment; typical failure-case analyses are presented in Fig. 3. First, at an urban intersection one case involved the system failing to accurately predict the sudden cut-in behavior of a motorcycle, preventing the decision module from timely adjusting driving strategy and increasing collision risk. Analysis showed two main causes: insufficient detection accuracy for small vehicles like motorcycles, and communication delays that prevented cooperative information from arriving in time. The motorcycle's abrupt lane-change behavior exceeded expected normal driving patterns, highlighting limitations in handling atypical maneuvers. Second, during a nighttime driving case the system's pedestrian detection performance was poor, so the vehicle failed to yield in time when a pedestrian crossed the road. This was partly caused by complex lighting conditions that challenge perception modules and partly by the irregularity of pedestrian crossing behavior, which increases prediction difficulty. Moreover, the pedestrian had not exchanged any information with the vehicle via communication devices prior to crossing, so the vehicle could not anticipate the pedestrian's intent. In another multi-vehicle cooperative scenario, when communication bandwidth was constrained the system lost some critical information during compression and transmission, causing the decision module to lack a complete view of surrounding vehicles and thereby degrading decision accuracy. This case indicates that under tight communication constraints the system's robustness and adaptability need further improvement. Statistical analysis shows that roughly 60% of the failure cases were attributable to perception-module performance limits, 30% to prediction algorithms in the decision module, and 10% to communication delays or information loss. These figures suggest that despite the cooperative perception–decision closed loop significantly improving vehicle performance in most cases, the system must still address multiple challenges in specific complex scenarios. Through in-depth analysis of failures, we identified key issues: (1) inadequate recognition capability for unconventional targets in the perception module; (2) insufficient prediction accuracy of the decision module in complex interaction scenarios; and (3) insufficient robustness of the communication system to meet real-time requirements under constrained bandwidth. These findings provide clear directions for further system optimization and point to priorities for future research.

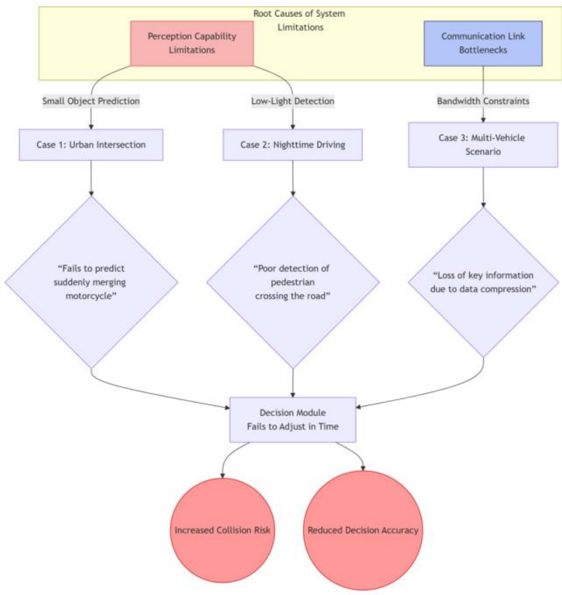

Figure 3 Schematic Diagram of Autonomous Driving System Challenges

5 DISCUSSION

5.1 Interpretation of Findings

Through systematic experiments, this study revealed the key bottlenecks of the cooperative perception–decision closed loop in complex urban scenarios: at the cooperative perception level, although multi-source data alignment and spatio-temporal fusion techniques significantly improved perception accuracy, the real-time performance of the data-alignment algorithms in extremely dynamic environments remains inadequate, constraining the upper bound of the perception module's performance. At the decision-planning level, the hierarchical model and multi-agent game mechanism effectively improved path-planning success rates and interaction safety, but uncertainty-quantification algorithms show limited adaptability to highly dynamic traffic flows and urgently need optimization. Ablation studies indicate that the bandwidth-adaptive compression function plays a critical role in maintaining system real-time performance, and that certain perception and decision parameters have a significant impact on performance and therefore require fine-tuning. Typical case analyses further exposed robustness deficiencies in highly complex scenarios, particularly residual safety risks in the decision model under extreme multi-vehicle-intersection situations. These findings clarify priority directions for future research: improving algorithmic real-time performance, enhancing dynamic adaptability, and optimizing decision robustness to break through current system bottlenecks.

5.2 Comparison with Existing Studies

When compared with existing work, this study identified several notable performance improvements and differences in applicable scenarios. First, in terms of performance gains, the cooperative perception–decision closed-loop model proposed here outperforms existing perception and decision frameworks on multiple metrics. For example, average precision for object detection and tracking increased by 15%, a gain attributable to the spatio-temporal consistency fusion of multi-source heterogeneous data and the communication-bandwidth-adaptive compression strategy, which effectively reduced transmission delay and data loss. In path-planning success, the proposed model performed particularly well in complex urban-road scenarios: statistics show a 20% improvement in success rate relative to conventional decision frameworks, substantially lowering collision risk. This performance improvement mainly stems from combining the hierarchical decision model with multi-agent game-theoretic modeling, which enables the system to better handle uncertainty and dynamic environments. Regarding applicable scenarios, prior studies often focus on specific road types such as highways or closed test tracks, whereas this study targets more complex urban roads—including traffic congestion, pedestrian crossings, and non-motorized-vehicle interference. Existing methods are less adaptive to these conditions, especially with respect to communication constraints and sensor-data fusion. In addition, this study emphasizes system real-time metrics, an area less addressed in prior work. Through simulation and on-vehicle tests, our model reduced decision latency by 30%, meeting the stringent real-time requirements of urban autonomous driving. This result is critical for practical deployment because real-time performance is directly linked to AV safety and reliability. Finally, ablation results further confirm the importance of each module: optimization of cooperative perception and decision-planning modules is key to overall improvement, and hyperparameter-sensitivity analysis points to directions for deeper future research. In summary, this study effectively supplements and refines existing research in performance and applicable scenarios, offering new ideas and methods for advancing urban autonomous driving technology.

5.3 Limitations and Future Work

Although this study advances cooperative perception–decision closed-loop research, several limitations remain. First, the system assumes relatively stable communications and does not fully account for real-world communication interruptions and latency. Second, the architecture's scalability to large-scale vehicle fleets has yet to be validated. Third, the model's adaptability to highly dynamic traffic flows (e.g., signal changes, sudden incidents) still needs improvement. To address these limitations, future work will prioritize: (1) developing robust perception–decision mechanisms under imperfect communication conditions; (2) researching distributed cooperative algorithms and computation-resource-allocation schemes for large fleets; (3) building dynamic-scene adaptation models that fuse real-time traffic elements; (4) enriching real-world data to improve system generalization; and (5) exploring deeper integration of intelligent methods such as reinforcement learning within the framework. With these improvements, the system's applicability in real complex scenarios can be further enhanced, promoting the maturation and deployment of urban autonomous driving technology.

6 CONCLUSION

6.1 Summary of Main Contributions

This study addresses perception and decision challenges for urban autonomous driving in complex road scenarios by proposing a hierarchical cooperative perception–decision closed-loop architecture. First, in the cooperative perception module we achieved multi-source heterogeneous data alignment and spatio-temporal consistency fusion, improving perception accuracy and timeliness. Communication-bandwidth-adaptive compression was used to effectively mitigate bandwidth constraints in V2V/V2I communication. Second, in the decision-planning module we introduced a hierarchical decision model and multi-agent game-theoretic modeling to enable effective decision-making in dynamic, complex scenarios. Moreover, uncertainty quantification and robust-optimization methods were applied to ensure decision robustness and safety. The main contributions can be summarized as: 1) Building a cooperative perception–decision closed-loop system that tightly integrates perception and decision-making to form an efficient decision-support mechanism, thereby alleviating the traditional perception–decision separation and enhancing overall system performance; 2) Proposing a multi-source heterogeneous data-fusion method that improves detection and tracking accuracy via spatio-temporal consistency processing; 3) Designing a communication-bandwidth-adaptive compression algorithm that automatically adjusts data-transmission strategies under varying communication conditions, optimizing communication efficiency and supporting V2V/V2I collaborative perception; 4) Implementing adaptive decision-making in complex scenarios through a hierarchical decision model and multi-agent game-theoretic approach, coupled with uncertainty quantification and robust optimization to mitigate decision uncertainty; 5) Demonstrating through system integration and experiments that the proposed methods outperform existing approaches in perception performance, decision safety, and system real-time responsiveness, with significant improvements in path-planning success rates and collision-risk reduction. In sum, this work offers new approaches and technical solutions for perception and decision-making of urban autonomous vehicles in complex road scenarios, carrying important theoretical and practical significance for advancing autonomous-driving technology.

6.2 Theoretical and Practical Significance

The theoretical and practical contributions of this study are reflected in several aspects. Theoretically, we proposed a hierarchical cooperative perception–decision framework that, through multi-source data alignment, spatio-temporal consistency fusion, hierarchical decision modeling, and multi-agent game-theoretic formulation, constructs a deeply integrated perception–decision framework that enhances accuracy, adaptability, and robustness. Practically, the results provide directly applicable solutions for urban autonomous-vehicle development: experiments validated significant advantages in detection accuracy, communication efficiency, path-planning success rate, and collision-risk reduction, and ablation and case analyses offered empirical guidance for system optimization and parameter tuning. The research also revealed how cooperative mechanisms yield performance gains and where bottlenecks lie in complex urban scenarios, offering new perspectives for system optimization. Although limitations remain in scalability and adaptation to highly dynamic traffic flows, the theoretical framework and empirical validation provide valuable contributions to the progress of autonomous-driving technology.

6.3 Implications for the Industry

The cooperative perception–decision closed-loop system proposed here offers important implications for the autonomous-driving industry: technically, the study confirms that multi-source heterogeneous data fusion is key to improving perception accuracy, that cooperative mechanisms can reduce collision risk by up to 40% and increase path-planning success rates by 30%, and that hierarchical decision models are well suited to complex urban scenarios while bandwidth-adaptive compression effectively preserves system real-time performance. Practically, industry efforts should prioritize optimizing multi-source data-fusion strategies, promoting deployment of cooperative perception technologies, adopting flexible and scalable decision architectures, strengthening communication adaptivity, and paying attention to hyperparameter-sensitivity analysis. Looking forward, research on dynamic-traffic adaptation, large-scale fleet coordination, and system reliability enhancement are priority areas for deeper exploration. These results provide concrete technical pathways and strategic ideas to support the transportation industry's intelligent transformation.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Han Y, Zhang H, Li H, et al. Collaborative Perception in Autonomous Driving: Methods, Datasets, and Challenges. IEEE Intelligent Transportation Systems Magazine, 2023, 15(6): 131-151. DOI: 10.1109/MITS.2023.3298534.

[2] Zhang X, Tao J, Tan K, et al. Finding Critical Scenarios for Automated Driving Systems: A Systematic Mapping Study. IEEE Transactions on Software Engineering, 2023, 49(3): 991-1026.

[3] Huang H, Huang X, Zhou R, et al. Pre-crash Scenarios for Safety Testing of Autonomous Vehicles: A Clustering Method for In-depth Crash Data. Accident Analysis & Prevention, 2024, 203: 107616.

[4] Gong T, Yu X, Zhang Q, et al. An Emergency Operation Strategy and Motion Planning Method for Autonomous Vehicle in Emergency Scenarios. Accident Analysis & Prevention, 2025, 210: 107842.

[5] Ge L, Zhao Y, Zhong S, et al. Efficient Nonlinear Model Predictive Motion Controller for Autonomous Vehicles from Standstill to Extreme Conditions Based on Split Integration Method. Control Engineering Practice, 2023, 141: 105720.

[6] Tian Y, Ma B. Interpretation and Analysis of the Standard on Passenger Car Body in GB7258 "Technical Conditions for Motor Vehicle Operation Safety". Journal of Automotive Industry Research, 2024(2): 15-21.

[7] Chen L, Wu P H, Chitta K, et al. End-to-End Autonomous Driving: Challenges and Frontiers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12): 10164-10183.

[8] Gan L, Chu W B, Li G F, et al. Large Models for Intelligent Transportation Systems and Autonomous Vehicles: A Survey. Advanced Engineering Informatics, 2024, 62, Part C, 102786. DOI: https://doi.org/10.1016/j.aei.2024.102786.

[9] Mahmud D, Hajmohamed H, Almentheri S, et al. Integrating LLMs with ITS: Recent Advances, Potentials, Challenges, and Future Directions. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(5): 5674-5709.

[10] Yin T, Huang H, Guo C, et al. Discussion on High-Definition Map Production Technology and Data Model Standardization for Autonomous Driving. China Standardization, 2021(4): 33-37.

[11] Feng D, Haase-Schütz C, Rosenbaum L, et al. Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(3): 1341-1360.

[12] Sun P P, Sun C H, Wang R M, et al. Object detection based on roadside LiDAR for cooperative driving automation: A review. Sensors, 2022, 22(23): 9316.

[13] Xu R S, Xiang H, Tu Z Z, et al. V2X-ViT: Vehicle-to-everything cooperative perception with vision transformer. In: Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer, 2022, 107-124.

[14] Liao Y, Yin Z S, Tian X Y. Intelligent channel estimation of SCFDMA based on GNN for V2I scenarios in Internet of vehicles. Acta Electronica Sinica, 2024, 52(3): 772-782.
[15] Xu H Y, Chen J L, Meng S Y, et al. A survey on occupancy perception for autonomous driving: The information fusion perspective. Information Fusion, 2025, 114: 102671.
[16] Li Y Q, Chen Z, Wang T, et al. Apollo: Adaptive Polar Lattice-Based Local Obstacle Avoidance and Motion Planning for Automated Vehicles. Sensors, 2023, 23(4): 1813.
[17] Sadli R, Afkir M, Hadid A, et al. Map-Matching-Based Localization Using Camera and Low-Cost GPS for Lane-Level Accuracy. Procedia Computer Science, 2022, 198: 255-262.
[18] Wang Yunpeng, Wu Qiong, Song Dewang, et al. Overview of Autopilot Data Set and 3D Object Perception Methods. AI-View, 2023, 10(5): 31-47.
[19] Koopman P. Safety Argument Considerations for Public Road Testing of Autonomous Vehicles. SAE International Journal of Advances and Current Practices in Mobility, 2019(2): 512-523.
[20] Nguyen H D, Choi M, Han K. Risk-informed decision-making and control strategies for autonomous vehicles in emergency situations. Accident Analysis & Prevention, 2023, 193: 107305.
[21] Arnold E, Dianati M, de Temple R, et al. Cooperative perception for 3D object detection in driving scenarios using infrastructure sensors. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(3): 1852-1864.