Final Report

Development of a Pedestrian Collision Avoidance System for Connected and Autonomous Vehicles with Cooperative Perception

Di Yang, Abolfazl Taherpour, Mansoureh Jeihani
Morgan State University

Lei Cheng, Terry Yang
University of Maryland at College Park

Date
October 2025

Prepared for the Safety and Mobility Advancements Regional Transportation and Economics Research Center, Morgan State University, CBEIS 327, 1700 E. Cold Spring Lane, Baltimore, MD 21251

ACKNOWLEDGMENT

This research was supported by the Safety and Mobility Advancements Regional Transportation and Economics Research Center at Morgan State University and the University Transportation Center(s) Program of the U.S. Department of Transportation.

Disclaimer

1. Report No. SM21	**2. Government Accession No.**	**3. Recipient's Catalog No.**

4. Title and Subtitle
Development of a Pedestrian Collision Avoidance System for Connected and Autonomous Vehicles with Cooperative Perception

5. Report Date
10/31/2025

6. Performing Organization Code

7. Author(s) *Include ORCID #*
Di Yang; ORCID: 0000-0002-7964-7872
Abolfazl Taherpour; ORCID: 0000-0002-8465-1104
Mansoureh Jeihani; ORCID: 0000-0001-8052-6931
Lei Cheng; ORCID: 0009-0006-9364-1174
Terry Yang; ORCID: 0000-0002-9416-6882

8. Performing Organization Report No.

9. Performing Organization Name and Address
Morgan State University
1700 East Cold Spring Lane, Baltimore, MD 21251
University of Maryland, College Park
1173 Glenn Martin Hall, College Park, MD 20742

10. Work Unit No.

11. Contract or Grant No.
69A3552348303

12. Sponsoring Agency Name and Address
US Department of Transportation
Office of the Secretary-Research
UTC Program, RDT-30
1200 New Jersey Ave., SE
Washington, DC 20590

13. Type of Report and Period Covered
Final

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract
Pedestrian safety remains a major challenge in urban transportation, especially when connected and automated vehicles (CAVs) must operate under occlusion, limited line-of-sight, and unpredictable pedestrian behavior. Pedestrians hidden by parked vehicles, roadside obstacles, or complex roadway geometry may not be detected in time by onboard sensors alone, leaving insufficient time for safe braking or maneuvering. To address this limitation, a cooperative perception (CP)-enabled pedestrian collision avoidance framework is developed by integrating roadside sensing, onboard sensing, multi-sensor fusion, edge computing, and Vehicle-to-Everything (V2X) communication. A camera–LiDAR fusion pipeline is established to provide reliable object detection and unified outputs for pedestrians and vehicles by combining the semantic strengths of cameras with the geometric accuracy of LiDAR. Controlled demonstrations using a Level 4 CAV platform in an occluded pedestrian-crossing scenario show that infrastructure-assisted CP can provide early warning for hidden pedestrians and enable safe stopping, whereas onboard sensing alone may detect the pedestrian too late. The results demonstrate the feasibility and safety value of CP for pedestrian protection and support its future integration into the Morgan State University SMART Corridor as a broader infrastructure-assisted safety framework for connected urban transportation systems.

17. Key Words: Automated Vehicles; Cooperative Perception; Connected Vehicles; Pedestrians; Safety

18. Distribution Statement

19. Security Classif. (of this report):	**20. Security Classif. (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified	Unclassified	55	

Abstract

This report presents the development and evaluation of a Cooperative Perception (CP)-enabled Pedestrian Collision Avoidance System (PCAS) for connected and automated vehicle (CAV) applications. The study is motivated by the persistent safety risk faced by pedestrians in urban traffic environments, where onboard vehicle perception alone is often limited by occlusion, line-of-sight blockage, and the unpredictability of pedestrian behavior. To address these challenges, the project develops a framework that integrates roadside sensing, onboard sensing, multi-sensor fusion, edge computing, and Vehicle-to-Everything (V2X) communication to extend vehicle awareness beyond the direct sensing range of the autonomous vehicle. A camera–LiDAR fusion pipeline is first developed to provide robust object detection and unified object-level outputs for pedestrians and vehicles, combining the semantic strength of cameras with the geometric accuracy of LiDAR. Field testing results show that the fused perception approach improves detection reliability and range accuracy compared with single-sensor configurations, establishing a practical sensing baseline for CP deployment.

Building on this perception foundation, the project conducts controlled demonstrations using a Level 4 CAV platform to evaluate the safety benefits of CP in an occluded pedestrian-crossing scenario. The demonstrations compare three cases: a baseline visible-pedestrian case in which onboard sensing alone is sufficient, an occluded-pedestrian case in which CP enables the vehicle to stop safely by receiving early warning from roadside infrastructure, and a matched occluded case without CP in which the vehicle detects the pedestrian too late and collides with the dummy target. These demonstrations provide direct proof-of-concept evidence that infrastructure-assisted perception can compensate for onboard sensing blind spots and improve collision avoidance performance. The report further discusses how the developed CP system can be integrated into the Morgan State University SMART Corridor for long-term deployment, leveraging the existing CAV testbed’s sensing, communication, tracking, and conflict-analysis capabilities. In addition, the report outlines other promising CP applications, including vulnerable road user protection, wrong-way driving warning, red-light-running warning, permissive left-turn assistance, adverse-weather support, and work-zone navigation. Finally, the report proposes a framework for future large-scale impact evaluation, including safety, efficiency, and scalability metrics, as well as simulation, hardware-in-the-loop, and field-testing strategies. Overall, the findings demonstrate that CP is a feasible and promising approach for enhancing pedestrian safety and advancing infrastructure-assisted intelligent transportation systems.

Contents

Abstract iii
List of Figures vi
List of Tables vii
Chapter 1. Introduction 1
1.1. Research Background 1
1.2. Technological Challenges 3
1.3. Research Objectives 6
Chapter 2. Literature Review 8
2.1 Foundations and Core Capabilities 8
2.2 Communication Load, Redundancy, and Scalability 10
2.3 Applications: Intersections, Pedestrian, and Traffic Control 11
2.4 Synthesis, Gaps, and Implications 12
Chapter 3. Object Detection and Multi-Sensor Fusion 13
3.1. Background 13
3.2. Methodologies 13
3.3. Field Testing Results 2
3.4. Discussions 3
Chapter 4. Cooperative Perception Demonstrations 5
4.1. Equipment 5
4.2 Demonstration Setup 6
4.3 Demonstrations 9
4.4 Testbed Integration for Long-term Deployment 13
Chapter 5. Other CP Scenarios 21
5.1. Background 21
5.2. Pedestrian Protection with CP 21
5.3. Collision Avoidance 23
5.4. Conflict Avoidance 24
5.5. Enhanced Situational Awareness 25
Chapter 6. Further Discussions: Quantifying the System Impact 27
6.1 Functional Safety Analysis 27

Chapter 7. Conclusions and Future Research 32
7.1. Summary and Conclusions 32
7.2. Limitations and Future Research Directions... 33
References.. 36

List of Figures

Figure 1. CP demo: (top): Videos from infrastructure and vehicles; (bot): Constructed 3D map of vehicles and peds 2
Figure 2. The detection blind spots for AVs 3
Figure 3. Multi-sensor fusion in the intelligent transportation system (Credit: Continental) 14
Figure 4. Proposed Camera–LiDAR Object Detection Pipeline 1
Figure 5. Field Testing Results Using the Proposed Camera–LiDAR Detection Pipeline 2
Figure 6. The UMD in-house LV 4 Research CAV 5
Figure 7. Cooperative Perception Demonstration Setup 7
Figure 8. Two frames from the Case 1 demonstration video 10
Figure 9. Two frames from the Case 2 demonstration video 11
Figure 10. Two frames from the Case 3 demonstration video 13
Figure 11. Overview of SMART Corridor 14
Figure 12. MSU CAV testbed LiDAR point clouds 15
Figure 13. MSU CAV testbed trajectory extraction 16
Figure 14. An example of pedestrian-vehicle conflict 17
Figure 15. Pedestrian conflict heatmap 17
Figure 16. Proportions of traffic conflicts per zone 18
Figure 17. Heatmap of outside of crosswalk events 18
Figure 18. Portable LiDAR integration 19
Figure 19. Fusion of multiple sensors in CP 22
Figure 20. WWD warning with CP 23
Figure 21. RLR warning with CP 24
Figure 22. Importance of Sampling for Critical CP Scenarios 28

List of Tables

Table 1. Field-test performance of camera–LiDAR perception for pedestrian and vehicle 3
Table 2. Metrics for Evaluating the Impacts of CP systems 30

Chapter 1. Introduction

1.1. Research Background

Pedestrian-related traffic crashes remain a persistent and critical public safety concern worldwide. Despite ongoing advancements in vehicle safety systems and traffic management strategies, pedestrians continue to represent one of the most vulnerable groups of road users. Globally, pedestrian fatalities account for a substantial proportion of annual traffic deaths, reflecting the inherent risks faced by individuals traveling without physical protection in complex traffic environments.

In the United States, metropolitan regions experience disproportionately high rates of pedestrian-involved crashes due to dense traffic, mixed transportation modes, and frequent pedestrian activities. For instance, recent crash statistics indicate that between 2017 and 2021, approximately 92% of all pedestrian-involved crashes in Maryland occurred within major metropolitan areas, particularly the Baltimore and Washington regions. This concentration highlights the heightened risk associated with urban corridors, signalized intersections, transit stops, and mixed-use developments where pedestrian and vehicular movements frequently intersect. The significant loss of life and severe injuries resulting from these incidents underscore the urgent need for innovative and technology-driven countermeasures to protect pedestrians and other pedestrian, including cyclists and individuals with disabilities.

Emerging transportation technologies, particularly Connected and Autonomous Vehicles (CAVs), offer promising opportunities to enhance road safety. Autonomous vehicles (AVs) are equipped with advanced perception systems that integrate cameras, LiDAR, radar, ultrasonic sensors, and artificial intelligence (AI)-based object detection algorithms. These systems enable AVs to identify pedestrians within their field of view and execute appropriate control actions, such as braking, steering, or speed adjustment, to avoid potential collisions. Over the past decade, significant progress has been made in improving perception accuracy, sensor fusion algorithms, and real-time decision-making frameworks.

However, despite these technological advancements, substantial challenges remain in ensuring reliable pedestrian safety. One fundamental limitation lies in the unpredictability of pedestrian behavior. Pedestrians may suddenly step into traffic, cross outside designated crosswalks, or emerge from behind parked vehicles or roadside obstacles. Such abrupt movements may not provide sufficient reaction time for AVs to safely decelerate or maneuver, particularly at higher travel speeds. Furthermore, AV perception systems are inherently constrained by line-of-sight limitations. Large vehicles, such as fleet trucks, buses, or construction equipment, can obstruct sensor beams and block visual detection of pedestrians. Environmental factors, including adverse weather conditions, low lighting, glare, and complex urban geometry, can further degrade sensor performance. These constraints highlight a critical vulnerability in traditional onboard perception systems that rely solely on vehicle-mounted sensors.

To address these limitations, there is a growing need to extend vehicular perception beyond individual onboard sensing capabilities. This project aims to develop an advanced Pedestrian Collision Avoidance System (PCAS) based on Cooperative Perception (CP), integrating Connected Vehicle (CV) technology with real-time sensing from roadside infrastructure. Unlike conventional AV systems that operate primarily within their local sensing range, the proposed system enhances environmental awareness by leveraging distributed sensing and communication networks. Specifically, this project seeks to establish a dynamic and cooperative perception and communication framework utilizing Vehicle-to-Everything (V2X) communication technologies. Through this framework, roadside units (RSUs), infrastructure-mounted sensors, and potentially other connected vehicles can detect pedestrian presence and transmit relevant information to approaching AVs in real time. By sharing situational awareness across multiple agents within the transportation ecosystem, the system enables AVs to receive early warnings about pedestrians that may be occluded or outside their direct sensing range.

CP, as conceptually illustrated in Figure 1, is driven by the need to enhance situational awareness for Advanced Driver Assistance Systems (ADAS) and Automated Driving Systems (ADS). Traditional vehicular perception architectures are constrained by line-of-sight, limited sensor range, and environmental interference, which can lead to delayed hazard recognition and suboptimal decision-making. In contrast, CP leverages V2X communication to integrate data from multiple sources, including neighboring vehicles, roadside infrastructure, and centralized traffic management systems, creating a shared perception environment. This distributed intelligence framework provides a more comprehensive and robust understanding of the driving context.

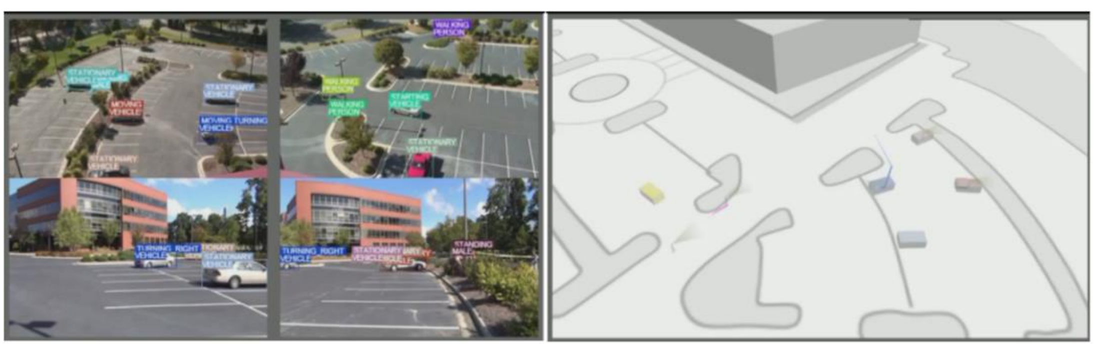


Figure 1. CP demo: (top): Videos from infrastructure and vehicles; (bot): Constructed 3D map of vehicles and peds

The advantages of CP are particularly pronounced in complex urban environments characterized by high pedestrian activity, obstructed views, and frequent traffic conflicts. By enabling vehicles to access information beyond their immediate sensory field, CP allows earlier hazard anticipation, improved trajectory planning, and more reliable collision avoidance maneuvers. Figure 2 illustrates a typical urban intersection scenario that highlights the limitations of conventional onboard vehicle perception and the necessity of CP. In this case, the subject vehicle relies on its onboard sensors (e.g., LiDAR, radar, and cameras) to detect pedestrians within its forward sensing range; however, large obstacles such as trucks or roadside objects obstruct the line-of-sight, preventing the vehicle from detecting pedestrians hidden behind them near the crosswalk. As a result, if a pedestrian suddenly emerges from behind the obstruction, the vehicle may have insufficient time to react and avoid a collision. Cooperative Perception addresses this limitation by leveraging roadside infrastructure sensors and Vehicle-to-Everything (V2X) communication to detect occluded pedestrians and transmit their presence to the approaching vehicle in real time. By extending perception beyond the vehicle's direct sensing range, CP enhances situational awareness, enables earlier hazard anticipation, and significantly improves pedestrian collision avoidance performance in complex urban environments.

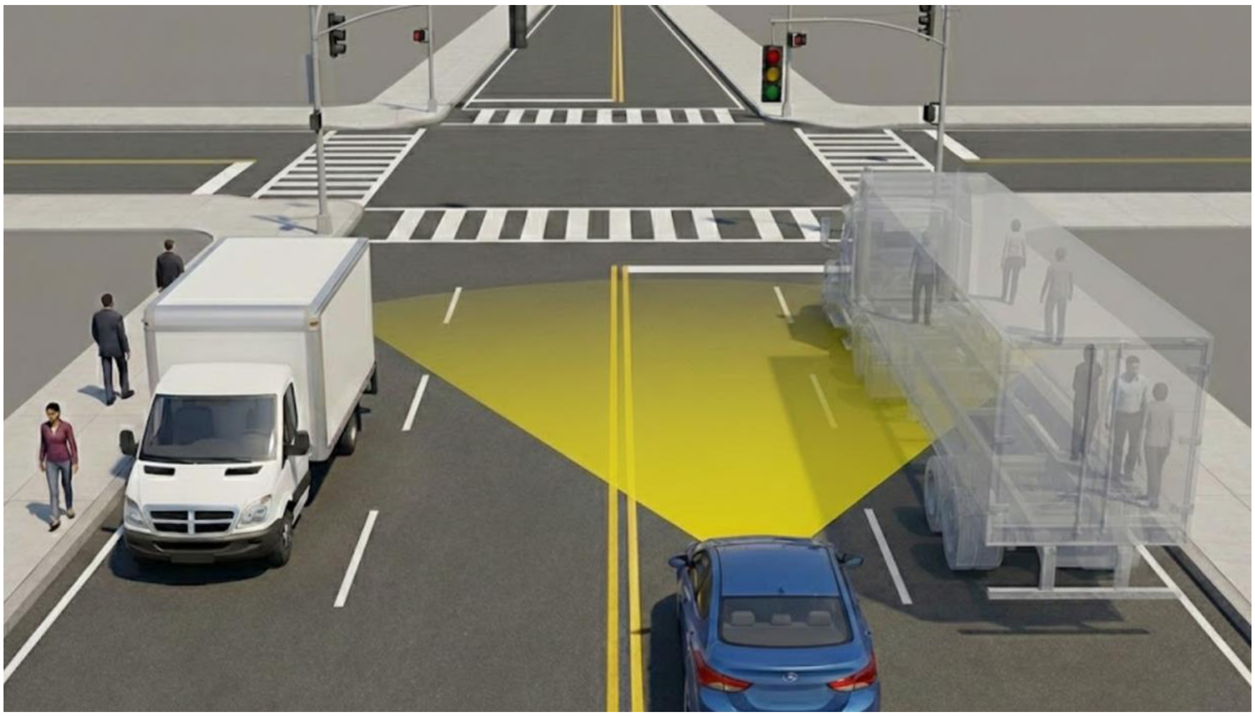


Figure 2. The detection blind spots for AVs

Hence, integrating cooperative perception into pedestrian collision avoidance systems represents a transformative step toward safer, smarter, and more resilient connected transportation networks.

1.2. Technological Challenges

The development and deployment of a V2X-enabled CP system for pedestrian collision avoidance involve several critical technological challenges spanning computer vision, communication, positioning, and vehicle control. Unlike conventional autonomous driving systems that rely solely on onboard perception, CP systems require seamless integration of distributed sensors, real-time communication, and coordinated decision-making across vehicles and infrastructure. Achieving reliable, real-time operation under diverse real-world conditions presents substantial research and engineering difficulties.

1.2.1 Computer Vision and Real-Time Sensor Fusion

A fundamental challenge of CP lies in achieving real-time sensor fusion across heterogeneous sensing platforms distributed between vehicles and roadside infrastructure. To enable effective 24/7 transportation management and safety applications, the system must generate a unified and continuously updated perception map of the environment by integrating data from sensors that differ in mobility, modality, and viewpoint.

Fusing data from fixed roadside sensors and moving vehicle-mounted sensors introduces significant complexity. Infrastructure sensors observe traffic from static, elevated, and often wide-angle perspectives, while vehicle sensors capture dynamic, ego-centric views that continuously change as the vehicle moves. Reconciling these different perspectives requires accurate spatial calibration, temporal synchronization, and coordinate transformation. Misalignment in space or time can result in inconsistent object localization, duplicated detections, or missed hazards. Therefore, advanced multi-sensor fusion algorithms must address motion dynamics, communication delays, and coordinate frame inconsistencies to maintain a coherent perception model.

Integrating heterogeneous sensing modalities, including cameras, LiDAR, and RADAR, adds another layer of complexity. Cameras provide high-resolution visual and semantic information but are sensitive to lighting conditions and visual obstructions. LiDAR offers precise 3D spatial measurements but can be affected by weather and has limited semantic richness. RADAR is robust in adverse weather and capable of velocity estimation, yet provides lower spatial resolution. Effectively combining these complementary characteristics requires sophisticated fusion frameworks capable of leveraging modality-specific strengths while mitigating their weaknesses.

Most existing CP research relies heavily on simulated environments, such as CARLA, where sensor noise and environmental variability are controlled. Although simulation is valuable for algorithm development, it does not fully capture the complexities of real-world deployment. The limited number of real-world CP studies have primarily conducted offline processing rather than true real-time fusion and communication. Bridging this gap between simulation and real-time operational deployment remains a significant technological hurdle.

Additionally, environmental variability presents a major obstacle. CP systems must function reliably under diverse weather and lighting conditions, including fog, rain, snow, and nighttime scenarios. Current V2X CP computer vision techniques struggle in adverse weather due to the scarcity of comprehensive real-world datasets. Simulated datasets typically cover only limited conditions, such as light rain or fog, and rarely include snow or complex urban weather interactions. Deep learning-based perception models trained on ideal or limited-condition datasets often fail to account for sensor degradation caused by precipitation, glare, or occlusion, leading to reduced detection accuracy and reliability. Developing robust perception algorithms and collecting comprehensive real-world datasets are therefore essential for practical CP deployment.

1.2.2 Communication and Interoperability

Effective CP relies on low-latency, high-reliability communication between vehicles and infrastructure. However, interoperability and bandwidth constraints pose major challenges to large-scale implementation.

One significant issue arises from inconsistent compliance with communication standards across different Connected and Automated Vehicle (CAV) equipment manufacturers. While standards such as SAE J2735 (defining V2X message sets) and SAE J3224 (sensor sharing message sets) are becoming increasingly comprehensive, full compliance across all roadside units (RSUs) and onboard units (OBUs) remains unlikely in the near future. Many manufacturers employ proprietary message formats with customized functionalities. Even among devices claiming compliance, variations in interpretation and implementation can introduce subtle inconsistencies. These discrepancies hinder seamless communication and complicate cross-platform CP data exchange between heterogeneous vehicles and infrastructure systems.

Bandwidth limitations present another critical concern. Following the FCC's decision to reduce the dedicated 5.9 GHz spectrum allocation by 60%, questions have emerged regarding whether the remaining bandwidth can sustain the increasing communication demands of CP systems. As CP adoption expands, the frequency and volume of safety-critical messages, such as J2735 Basic Safety Messages (BSMs) and J3224 sensor-sharing messages, may grow substantially. Both real and emulated perception data transmissions could place significant strain on available CV2X bandwidth. Maintaining sub-0.1-second latency, which is essential for traffic safety applications, may become increasingly challenging under high traffic density scenarios. To address this concern, systematic evaluation of message latency, packet loss, and scalability under various traffic conditions is necessary, combining both simulation-based analysis and real-world field testing.

1.2.3 Positioning and Localization Accuracy

Accurate vehicle positioning is a foundational requirement for CP and traffic safety applications. However, positioning errors in OBUs are emerging as a critical bottleneck for C-V2X systems. Currently, GPS serves as the primary positioning source for most C-V2X applications. However, standard GPS performance is insufficient for safety-critical use cases. First, typical GPS receivers update position at approximately 1 Hz, whereas most safety applications require updates at 10 Hz (0.1-second intervals). Although C-V2X communication can transmit messages at 10 Hz, the underlying GPS measurements may not support this frequency with adequate accuracy.

Second, positioning accuracy remains a major limitation. Even with augmentation techniques, GPS errors can approach the width of an entire traffic lane. In dense urban environments with tall buildings, multipath effects and signal blockage further degrade accuracy. Such errors can significantly impact the reliability of MAP messages, object association, and collision risk assessment, particularly when distinguishing between adjacent lanes or identifying pedestrian positions relative to vehicles.

To mitigate these limitations, advanced high-definition localization algorithms must be developed. One promising approach involves leveraging roadside LiDAR sensors to track vehicles and pedestrian at high frequency (e.g., 10 Hz). By transforming LiDAR-derived Cartesian coordinates into global WGS 84 latitude–longitude coordinates and refining trajectories through sensor fusion, positioning fidelity can be significantly enhanced. The corrected localization data may then be transmitted back to vehicles as a supplemental positioning source, enabling more accurate situational awareness and safer decision-making.

1.2.4 Vehicle Control and Decision-Making Integration

Incorporating CP data into vehicle control systems introduces new challenges beyond traditional onboard perception-based autonomy. The integration process affects both environmental mapping and trajectory planning.

During real-time mapping, externally received CP data must be aligned with the vehicle's onboard coordinate system. Because CP data may arrive with communication delays and originate from different coordinate frames, precise spatial and temporal synchronization is required. The vehicle's local coordinate system continuously shifts as it moves, further complicating data integration. Without accurate synchronization, the combined perception map may contain inconsistencies, leading to incorrect object tracking or unsafe decisions.

During trajectory planning and motion control, the vehicle must assess the reliability and relevance of external CP information. When conflicting moving objects (CMOs) are present, such as pedestrians and vehicles interacting in complex intersection scenarios, the vehicle must determine how much weight to assign to external perception data relative to onboard sensor inputs. Overreliance on potentially delayed or noisy external data may compromise safety, while underutilization may negate the benefits of CP. Designing robust data validation, confidence estimation, and decision fusion frameworks is therefore essential.

To address these challenges, rigorous validation using a connected and automated vehicle fleet is required. By leveraging test platforms and integrated ecosystems such as FHWA's CARMA framework, both rule-based and learning-based control algorithms can be evaluated under closed-track and real-world conditions. Such testing enables systematic assessment of CP-enhanced longitudinal and lateral control strategies, ensuring that cooperative perception improves safety without introducing unintended risks.

1.3. Research Objectives

The main focus of this study is to enhance pedestrian safety in the presence of AVs through the integration of CV and CP technologies. The overarching goal of this research is to mitigate the safety risk of collisions between AVs and pedestrians, especially those who suddenly and unexpectedly enter AVs' paths. To accomplish this goal, we propose the following major research objectives:

1. Understand the collision avoidance system of existing AVs: The first objective is to understand the collision avoidance system currently deployed on existing AVs, i.e., the AV systems without the cooperative perception, when confronted with unexpected pedestrian crossings. This will involve conducting literature review, simulation evaluations, and controlled field experiments using UMD’s AV to examine the capabilities of existing AVs to detect and respond to sudden pedestrian crossings in real-time.
2. Integrate CV and cooperative perception into AV control algorithms: The second objective is to develop a CAV system by integrating CV communication with AV control algorithms focusing on pedestrian collision avoidance. This will involve a comprehensive examination of the hardware, software, and communication protocols needed to enable the CV-AV integration for cooperative perception. In addition, the AV's control algorithms will be adapted to process incoming CV data effectively, including information about pedestrian positions and movements.
3. Develop field experiment scenarios: Distinct testing scenarios will be designed to facilitate the evaluation of the capabilities of the proposed system. A variety of pedestrian crossing situations will also be considered to quantify the differences in safety performance of the proposed system.
4. Conduct comprehensive performance evaluation: The fourth objective is to systematically evaluate the overall safety, efficiency, and operational impacts of the proposed CP-enabled system and compare its performance against the baseline AV system. This evaluation will quantify the effectiveness and feasibility of the proposed framework in improving pedestrian safety and traffic operations.

Chapter 2. Literature Review

This chapter reviews the theoretical foundations, technical architectures, and application domains of CP in CAV systems. It synthesizes prior work spanning sensing modalities, multi-agent fusion strategies, communication-efficient information sharing, dataset development, and real-world deployment considerations. Emphasis is placed on how perception and communication co-design shapes scalability and robustness, particularly under occlusion, adverse weather, and dense urban conditions. The chapter also highlights emerging learning-based frameworks, infrastructure-supported implementations, and system-level applications in traffic operations, while identifying key research gaps that motivate further investigation.

2.1 Foundations and Core Capabilities

Cooperative perception (CP) allows connected and automated vehicles (CAVs) and roadside infrastructure to share sensing outcomes so that each agent perceives beyond its own line of sight. Surveys covering CP within the broader autonomous-vehicle and vehicle–infrastructure cooperation landscape synthesize sensing modalities (camera, LiDAR, radar), fusion strategies, synchronization and calibration challenges, and networking constraints (Van Brummelen, O'Brien, Gruyer, & Najjaran, 2018; Marti, de Miguel, Garcia, & Perez, 2019; Cui, Zhang, Xiao, Yao, & Fang, 2022; Ji et al., 2024; Schiegg, Llatser, Bischoff, & Volk, 2021). Weather-focused work documents degradation mechanisms—attenuation, scattering, specular reflections, and sensor over-exposure—and motivates redundancy and cross-modality cues. CP can compensate by introducing additional viewpoints and persistent coverage from vehicles and infrastructure (Zhang, Carballo, Yang, & Takeda, 2023; O'Malley, Glavin, & Jones, 2008). These references point to a consistent conclusion: CP is most beneficial in dense urban and adverse-weather scenarios where occlusions and modality-specific failures are common, and where persistent infrastructure coverage complements transient vehicle views.

Methodologically, the literature distinguishes three sharing levels. Raw-level sharing transmits sensor measurements (e.g., LiDAR point clouds) for downstream fusion. COOPER demonstrated multi-vehicle point-cloud alignment and 3D object-detection gains over single-vehicle baselines, showing feasibility of sharing raw point clouds over vehicular networks and quantifying the sensitivity to pose/time misalignment (Chen, Tang, Yang, & Fu, 2019a). Its strength is maximal information content; its limitation is bandwidth and synchronization cost. Feature-level sharing computes intermediate neural features on edge devices and transmits compact feature maps. F-COOPER showed that learned features can support real-time cooperative detection under practical channel budgets, especially improving detection of far or occluded objects where single-vehicle signal-to-noise is weak (Chen et al., 2019b). The tradeoff is potential loss of generality relative to raw data, but this level greatly improves scalability. Decision-level sharing communicates object-level outputs (tracks, bounding boxes) with the smallest payload. This is bandwidth-efficient and aligns with standards for Collective Perception Messages (CPMs), but it depends on consistent upstream detectors and robust object association across agents.

Learning-based frameworks refine these ideas under realistic constraints. UMC proposes bandwidth-efficient multi-resolution collaboration where only informative regions are shared. This improves the value density of transmitted content when channels are constrained (Wang et al., 2023). V2VFormer introduces a spatial–channel transformer that fuses multi-vehicle features efficiently across ranges and occlusion conditions, illustrating gains from attention-based fusion over simple concatenation (Lin et al., 2024). V2X-ViT focuses on vehicle–infrastructure heterogeneity, pose noise, and asynchrony. Transformer attention is used to align multi-agent features and maintain performance despite clock and calibration drift (Xu et al., 2022). Related work on semantic communication compresses perceptually important content via learned importance maps and channel-aware coding, improving resilience to time-varying fading without saturating the channel (Sheng, Ye, Liang, Jin, & Li, 2024). Multi-modal designs that inject infrastructure labels into vehicle-side LiDAR–image fusion improve 3D detection of sparse and distant targets, a regime where single-vehicle performance is typically weakest (Xia et al., 2024). Under imperfect communications, joint optimization of helper-vehicle selection, resource allocation, and power control further improves detection robustness, highlighting the value of co-designing perception and networking (Sarlak, Amin, & Razi, 2025).

Datasets and benchmarks increasingly capture multi-agent cooperation and interactive scenes. V2X-Real provides large-scale scenarios for vehicle–vehicle and vehicle–infrastructure collaboration with standardized evaluation protocols (Xiang et al., 2024). ROAD emphasizes event-level awareness (agents, actions, scene locations) to move beyond object lists toward activity recognition and intent modeling (ROAD Dataset). At the message level, an urban infrastructure dataset exporting radar/video detections in CPM format enables direct benchmarking of inclusion policies and CPM rule design against the ETSI specification (Figueiredo et al., 2025). On the deployment side, infrastructure-sensor CP frameworks quantify delay compensation and fusion (e.g., constant-turn-rate-and-velocity models with unscented Kalman filtering) to mitigate computation/communication latency in early deployments (Chen, Tang, Hu, & Huang, 2023).

Finally, CP connects to traffic operations. By fusing CAV trajectories and roadside-detector measurements with particle filtering, microscopic traffic state estimation and short-horizon prediction improve even at low penetration rates; roadside sensing compensates for sparse connected fleets and reduces estimation variance (Li, Han, & Ma, 2022). At intersections, where occlusions and conflicts are frequent, infrastructure-assisted 3D detection and V2I perception improve detection stability and support path planning relative to strong single-vehicle baselines (Arnold, Dianati, de Temple, & Fallah, 2022; Duan et al., 2021). A probabilistic V2X fusion pipeline that propagates uncertainty end-to-end demonstrates perception improvements in occluded corridors and clarifies how uncertainty representation affects downstream decisions (Shan et al., 2022).

Key takeaways. Across these studies: (i) feature- and decision-level exchange offer practical bandwidth use while preserving informative content; (ii) attention-based fusion and selective sharing mitigate pose/time misalignment and channel limits; (iii) datasets are beginning to reflect infrastructure roles, CPM formatting, and interactive scenes; and (iv) CP can support both vehicle-level safety and network-level operations.

2.2 Communication Load, Redundancy, and Scalability

The net benefit of CP depends on timely, reliable, and scalable communications. A recurring theme is minimizing redundancy while maintaining situational awareness. Analyses of CPM generation under ETSI rules show that mobility-triggered policies can generate frequent small packets rich in duplicates, risking channel saturation without proportional perception gains (Thandavarayan, Sepulcre, & Gozalvez, 2019; 2020a). To address this, several strategies appear in the literature:

- **Geometry- and density-aware filtering** — Probabilistic selection conditioned on spatial coverage and object density reduces duplicates while preserving coverage of critical regions (Huang et al., 2020). Accuracy-based inclusion. Generation rules using relative entropy between local and received tracks prioritize objects whose updates most reduce uncertainty, improving information freshness per packet (Li & Wolff, 2022). Context-aware selection. Content is adapted to channel-busy ratio and infrastructure availability so that CPMs emphasize objects most relevant to current maneuver risk and communication capacity (Chtourou, Merdrignac, & Shagdar, 2021).

- **Operational evaluations** — Comparative analyses of filtering policies examine functional and operational requirements and show when different object-inclusion rules best optimize for coverage, latency, and load in urban scenarios (Delooz et al., 2022).

- **Identity-aware reduction** — Feature-based vehicle identification links detections to stable identities, enabling lower CPM frequencies without losing track continuity (Masuda, Marai, Tsukada, Taleb, & Esaki, 2023).

Beyond rule design, learned scheduling prioritizes transmissions with the highest expected benefit. Value-anticipating V2V policies score candidate messages and suppress low-utility updates under congestion, improving cooperative accuracy without exceeding channel limits (Higuchi, Giordani, Zanella, Zorzi, & Altintas, 2019). Reinforcement-learning approaches learn which features to transmit and when, using simulators that couple traffic, vehicles, wireless channels, and perception modules; these systems demonstrate higher cooperative accuracy with lower aggregate load (Aoki, Higuchi, & Altintas, 2020). Confidence-driven sparse communication generalizes this idea by gating inter-agent exchanges using confidence maps and simple whitelist scheduling, then refining and fusing sparse features temporally and spatially (Tan, Wang, Wang, Wang, Wang, & Wu, 2026).

Congestion-control (DCC) parameters strongly affect end-to-end utility. Joint tuning across access and facilities layers can stabilize the channel and restore timeliness within CP pipelines; mismatched parameters can degrade perception despite strong upstream detectors (Thandavarayan, Sepulcre, & Gozalvez, 2020b). Penetration and geometry also matter: modeling the effective vs. collective field of view (eFoV vs. cFoV) shows that even modest penetration can multiply environmental information if redundant broadcasts are controlled and agents occupy complementary viewpoints (Huang, Fang, & Li, 2019). Bandwidth for safety-critical maneuvers is another constraint.

For overtaking and similar high-risk actions, mmWave V2V capacity is often necessary to sustain rich perception exchange within required latency bounds (Fukatsu & Sakaguchi, 2021). Complementarily, centralized C-V2X client–server frameworks that prioritize messages by vehicle trajectories have shown that a high proportion of "message value" can be delivered under limited resource blocks across diverse traffic scenarios, pointing to practical deployment options for congested networks (Hakim, Sorour, Hefeida, Alasmary, & Almotairi, 2023; Park, Kavathekar, Bhuduri, Amin, & Devaraj, 2025).

Implications. Communication policies, CPM rules, and DCC should be co-designed with perception objectives. Selective, confidence-aware messaging paired with cross-layer congestion control is necessary to translate perception gains into robust, timely services at scale.

2.3 Applications: Intersections, Pedestrian, and Traffic Control

Intersections

Infrastructure sensing (camera + LiDAR) fused with vehicle perception extends line-of-sight, stabilizes detection under occlusions, and provides consistent coverage of conflict zones (Arnold et al., 2022; Duan et al., 2021). Work that maintains uncertainty through V2X fusion demonstrates improved 3D perception in occluded urban corridors and clarifies how uncertainty-aware fusion changes downstream planning behavior (Shan et al., 2022). These studies collectively show that CP is especially impactful at locations with structured geometry and persistent vantage points.

Pedestrian V2P implementations using smartphones and roadside mediation demonstrate the effectiveness of selective alerting that reduces nuisance alarms and energy consumption while preserving timely collision warnings (Tahmasbi-Sarvestani, Mahjoub, Fallah, Moradi-Pari, & Abuchaar, 2017; Nguyen, Morold, David, & Dressler, 2020). Simulation frameworks that integrate Pedestrian Awareness Messages (PAM) into cooperative ITS modeling enable realistic evaluation across sensing and messaging components and can support policy development (Hejazi & Bokor, 2024). Bibliometric analysis indicates growing attention to behavioral interaction, collision-avoidance technology, and risk perception—highlighting regulatory and ethical considerations as CP moves toward scale (Lee, Lee, Abas, & Chong, 2025).

Predictive assistance and maneuvers

Fusion of radar, LiDAR, camera, and V2X improves trajectory prediction and collision warnings for vehicle–vehicle and vehicle–pedestrian encounters (Baek, Jeong, Choi, & Lee, 2020). Lane-change prediction benefits from V2X-augmented models that capture intent earlier in the maneuver (Xu, Jiang, Wen, Liu, & Zhou, 2019). For overtaking, lane changing, merging, and unsignalized intersections, cooperative strategies with inverse reinforcement learning or multi-agent reinforcement learning improve safety and throughput over independent sensing and control, including under mixed traffic at low CAV penetration (Prathiba, Raja, & Kumar, 2022; Luo et al., 2023; Ji et al., 2025).

Signal control and infrastructure planning

A CP-based adaptive signal control (CAVLight) uses a cell-transmission model plus reinforcement learning to maintain robust performance at very low CAV penetration (~1%) by fusing CP-estimated states. This approach outperforms fixed-time, actuated, max-pressure, and optimization-based adaptive controllers under high demand (Li, Zhu, & Feng, 2024). For infrastructure, chance-constrained stochastic simulation with GPR-AKG sampling produces LiDAR placement plans that maximize detection performance under budget and recall constraints; center-mounted units are frequently selected to cover conflict areas efficiently (Chen, Zheng, & Tan, 2024).

Human–system considerations

Effective CP must account for attention, trust, and acceptance by drivers and pedestrians. Over-alerting risks habituation, while under-alerting reduces utility. The infrastructure-mediated and learning-based approaches above aim to filter noise and prioritize imminent risk with context-appropriate cues. As multi-agent datasets and CPM-formatted resources expand, evaluation protocols increasingly include interactive behaviors and rare events, aligning development with real-world human–system needs.

2.4 Synthesis, Gaps, and Implications

Across surveys, algorithms, datasets, and applications, three conclusions recur. First, perception gains are robust in occlusion-heavy, adverse, and dense scenarios when cooperation is aligned in time or pose and focused on informative regions (Van Brummelen et al., 2018; Marti et al., 2019; Cui et al., 2022; Ji et al., 2024; Zhang et al., 2023; Wang et al., 2023; Lin et al., 2024; Xu et al., 2022; Xia et al., 2024; Sarlak et al., 2025). Second, communication efficiency governs scalability. Redundancy mitigation, context- and accuracy-aware CPM generation, confidence-gated sparse exchange, and cross-layer DCC tuning are necessary to convert perception gains into timely, system-level improvements (Thandavarayan et al., 2019; 2020a; 2020b; 2023; Huang et al., 2020; Li & Wolff, 2022; Chtourou et al., 2021; Delooz et al., 2022; Masuda et al., 2023; Higuchi et al., 2019; Aoki et al., 2020; Tan et al., 2026). Third, infrastructure and human factors are central. Intersections, crosswalks, and ramps benefit from persistent RSU coverage, selective warnings, and policy-aligned control; low-penetration methods for signal control and principled LiDAR placement extend impact under limited observability (Arnold et al., 2022; Duan et al., 2021; Shan et al., 2022; Nguyen et al., 2020; Li, Zhu, & Feng, 2024; Chen, Zheng, & Tan, 2024; Lee et al., 2025).

Open problems include robust handling of calibration drift and clock offsets across heterogeneous agents; joint optimization linking perception confidence, message prioritization, and congestion control; privacy/security of shared features and intent information; and large-scale field trials that span penetration rates, weather, and roadway types. Addressing these requires coordinated progress across perception, communications, human factors, and standards.

Chapter 3. Object Detection and Multi-Sensor Fusion

3.1. Background

Reliable detection of pedestrians and vehicles is a foundational capability for CP systems in connected and automated driving. While CP extends perception beyond a single vehicle's line of sight through V2X communication and infrastructure support, the quality of any shared information ultimately depends on the reliability of local sensing and fusion at each agent. In real roadway environments, sensing conditions vary due to occlusions, dynamic traffic interactions, complex backgrounds, and lighting changes. If a vehicle's onboard perception is unstable or incomplete, cooperative sharing may propagate uncertainty or errors across the network. Therefore, robust multi-sensor perception at the individual vehicle level is a prerequisite for effective CP.

Cameras are effective for recognizing object categories because they capture rich appearance and texture cues, enabling reliable discrimination between pedestrians, vehicles, and background elements. However, camera-only perception does not directly provide metric depth, making accurate distance estimation and 3D localization more challenging, especially for safety-critical applications such as pedestrian collision avoidance. LiDAR complements cameras by providing direct 3D geometry and accurate ranging, enabling metric localization and improved spatial reasoning. At the same time, LiDAR-only perception may suffer from semantic ambiguity in cluttered scenes, and point sparsity can reduce detection reliability at longer distances.

For these reasons, camera–LiDAR fusion has become a widely adopted and practical approach in modern perception systems. Representative studies have shown that combining image semantics with LiDAR geometry improves 3D detection accuracy and object usability by providing both category identity and precise spatial location (Chen et al., 2017; Ku et al., 2018; Qi et al., 2018). More recent fusion architectures further demonstrate robust strategies for integrating camera and LiDAR information across diverse environmental conditions (Vora et al., 2020; Huang et al., 2020; Yoo et al., 2020; Bai et al., 2022; Liu et al., 2023).

Within the context of this project, the camera–LiDAR fusion module serves not only as a standalone perception system but also as the local perception layer of a larger CP framework. By generating a stable, unified object representation that includes semantic labels, 3D positions, and confidence measures, the system produces outputs that are suitable for downstream connected-vehicle functions and for potential sharing through cooperative perception messages. Establishing a reliable fused perception baseline is therefore a critical step toward enabling scalable, trustworthy, and communication-efficient CP deployment in pedestrian safety applications.

3.2. Methodologies

Multi-sensor fusion, as shown in Figure 3, can substantially enhance the completeness of perception by integrating information from multiple sources (Liggins II et al., 2009). Among its variants, heterogeneous fusion—which aggregates multiple units of the different sensor types—is especially important, as it extends spatial coverage, improves signal-to-noise characteristics, and provides redundancy. In transportation settings, a forward-facing sensor on the vehicle and a fixed roadside sensor offer complementary vantage points (Olagoke et al., 2020): the onboard view rides with the ego platform and captures near-field detail but suffers frequent occlusion, while the elevated roadside view covers a wider area with fewer occlusions but at coarser scale for distant agents. Fusing these two streams extends coverage, stabilizes tracking through occlusion, and enables metric depth via triangulation.

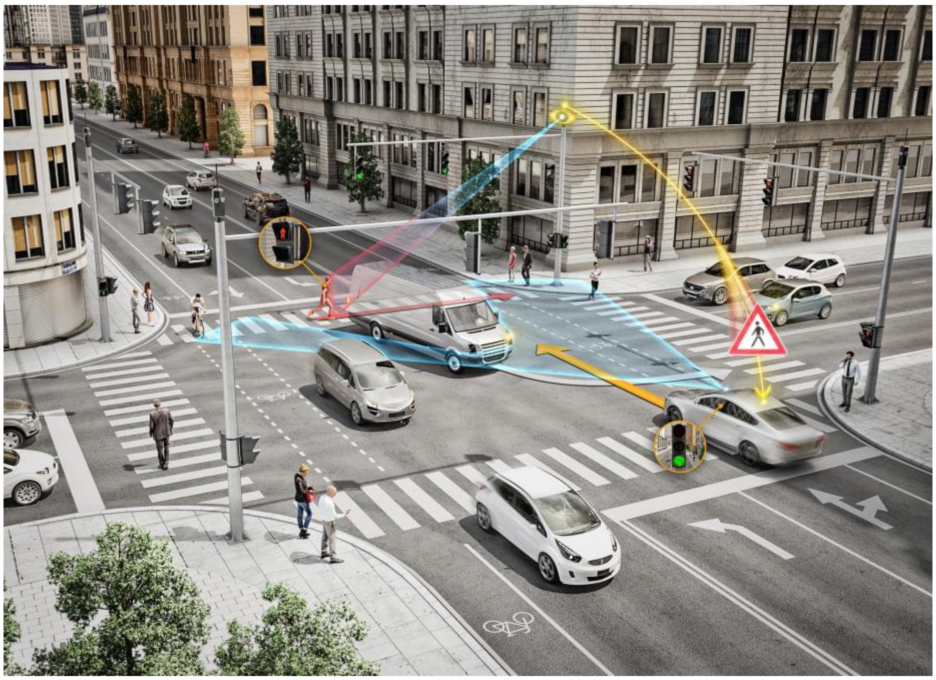


Figure 3. Multi-sensor fusion in the intelligent transportation system (Credit: Continental)

A complete multi-sensor fusion pipeline should encompass temporal and spatial alignment, cross-view association, and integration at the data, feature, or decision levels:

1. Temporal alignment

Fusion only makes sense if different sensors are describing the same instant (Rehder et al., 2016). A delay of just a few tens of milliseconds can move a fast vehicle by a noticeable amount, which then appears as a contradiction between sensors that are, in fact, both correct for their own times. Let t be a global time and t(v), t(r) the vehicle and roadside clocks. A minimal affine model captures drift and offset:

$$t^{(s)} = \alpha^{(s)} t + \beta^{(s)}, \qquad s \in \{v, r\}.$$

If hardware sync is unavailable, estimate (α(s), β(s)) by least squares on matched landmark times, or recover a dominant constant lag β by cross-correlating activity proxies (e.g., optical-flow magnitude peaks) across the two streams:

$$T_{\mathcal{W}}^{(v)} = \begin{bmatrix} R_v & t_v \\ 0^\top & 1 \end{bmatrix}, \qquad T_{\mathcal{W}}^{(r)} = \begin{bmatrix} R_r & t_r \\ 0^\top & 1 \end{bmatrix}.$$

$$\hat{\beta} \in \arg\max_{\Delta} \sum_t r^{(v)}(t)\, r^{(r)}(t+\Delta).$$

where speed v_{ego}, a residual lag Δt induces an apparent spatial error $v_{ego}\Delta t$, motivating buffering and resampling only after clock alignment.

2. Spatial calibration and geometry

Space must be aligned just as carefully as time. Spatial alignment means that all measurements describe positions in a shared coordinate system, so a point seen by both sensors maps to the same place. Achieving this requires two kinds of calibration. Intrinsic calibration makes each sensor self-consistent (like camera lens parameters) (Zhang, 2000). Extrinsic calibration expresses the rigid transformation between sensors. Let the vehicle and roadside sensors have intrinsics $\boldsymbol{K}_v$, $\boldsymbol{K}_r$ and let them poses in a world frame W given by

For a 3D point $\boldsymbol{X}$, the pinhole model writes

$$\lambda_s \begin{bmatrix} u_s \\ v_s \\ 1 \end{bmatrix} = \boldsymbol{K}_s \begin{bmatrix} \boldsymbol{R}_s & \boldsymbol{t}_s \end{bmatrix} \begin{bmatrix} \boldsymbol{X} \\ 1 \end{bmatrix}, \qquad s \in \{v, r\},$$

with lens distortion handled in the normalized plane. The relative pose ($\boldsymbol{R}$, $\boldsymbol{t}$) betweens cameras induces the essential matrix $\boldsymbol{E} = [\boldsymbol{t}]_\times \boldsymbol{R}$ enforcing the epipolar constraint for normalized rays $\tilde{\boldsymbol{u}}_s = \boldsymbol{K}^{-1}[u_s\ v_s\ 1]^T$:

$$\tilde{\boldsymbol{u}}_r^\top \boldsymbol{E}\, \tilde{\boldsymbol{u}}_v = 0.$$

On (locally) planar road patches, a homography maps ground points (X, Y, 1) to image pixels (u, v, 1),

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{H} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \qquad \boldsymbol{H} = \boldsymbol{K} \begin{bmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \boldsymbol{t} \end{bmatrix},$$

which enables a shared bird's-eye-view (BEV) canvas for both sensors.

3. Association and Triangulation

Given a detection $\boldsymbol{u}_v$ in the vehicle image, its corresponding point in the roadside image is confined to the epipolar line; searching along this line yields a candidate match $\boldsymbol{u}_r$. With projection matrices $\boldsymbol{P}_s = \boldsymbol{K}_s[\boldsymbol{R}_s\ \boldsymbol{t}_s]$, triangulation recovers $\boldsymbol{X}$ by minimizing reprojection error,

$$\min_{\boldsymbol{X}} \sum_{s\in\{v,r\}} \left\|\pi(\boldsymbol{P}_s\boldsymbol{X}) - \boldsymbol{u}_s\right\|_2^2,$$

or, when the two optical axes and baseline approximate rectified stereo, $Z \approx fB/d$ with focal length f, baseline B, and disparity d. Gating matches with a Mahalanobis distance on reprojection residuals,

$$d_M^2 = \boldsymbol{y}^\top \boldsymbol{S}^{-1}\boldsymbol{y},$$

where $\boldsymbol{S}$ includes image noise and calibration uncertainty.

4. Fusion Levels

With clocks synchronized and frames co-registered, integration is best understood at three complementary levels, often used together.

Data-Level (Early) Fusion

Combine measurements *before* heavy inference. Practical instances include: (i) rectified stereo or general two-view triangulation; (ii) warping each image to BEV and accumulating evidence (e.g., occupancy log-odds) in shared cells; (iii) panoramic stitching across overlapping FoVs to form a single wide canvas for downstream perception.

Feature-Level (Mid) Fusion

Encode each view with a dedicated backbone, project features to a common representation (often BEV), and mix them by concatenation, element-wise operations, or attention,

$$\mathrm{Attn}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \mathrm{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^\top}{\sqrt{d}}\right)\boldsymbol{V},$$

while using epipolar structure or learned correspondence to guide cross-view queries. This preserves view-specific inductive biases (texture and perspective) while enabling mutual disambiguation in occluded or distant regions.

Decision- and Object-Level (Late) Fusion

Fuse per-view detections or short tracks with explicit uncertainty. A Kalman-style correction for assimilating a roadside measurement $\boldsymbol{z}_r$ into a vehicle-centric prior $\boldsymbol{x}^-$ reads

$$\boldsymbol{x}^+ = \boldsymbol{x}^- + \boldsymbol{K}(\boldsymbol{z}_r - \boldsymbol{H}\boldsymbol{x}^-), \qquad \boldsymbol{K} = \boldsymbol{P}^-\boldsymbol{H}^\top(\boldsymbol{H}\boldsymbol{P}^-\boldsymbol{H}^\top + \boldsymbol{R}_r)^{-1},$$

supporting identity maintenance and graceful handoffs as agents exit one FoV and enter the other.

Our object detection pipeline, as shown in Figure 4, is built around a paired camera + LiDAR sensor suite and produces a unified object output describing pedestrians and vehicles. The methodology is practical and deployment-oriented; we focus on stable alignment between sensors, reliable detection from each modality, and a straightforward fusion strategy that yields consistent outputs.

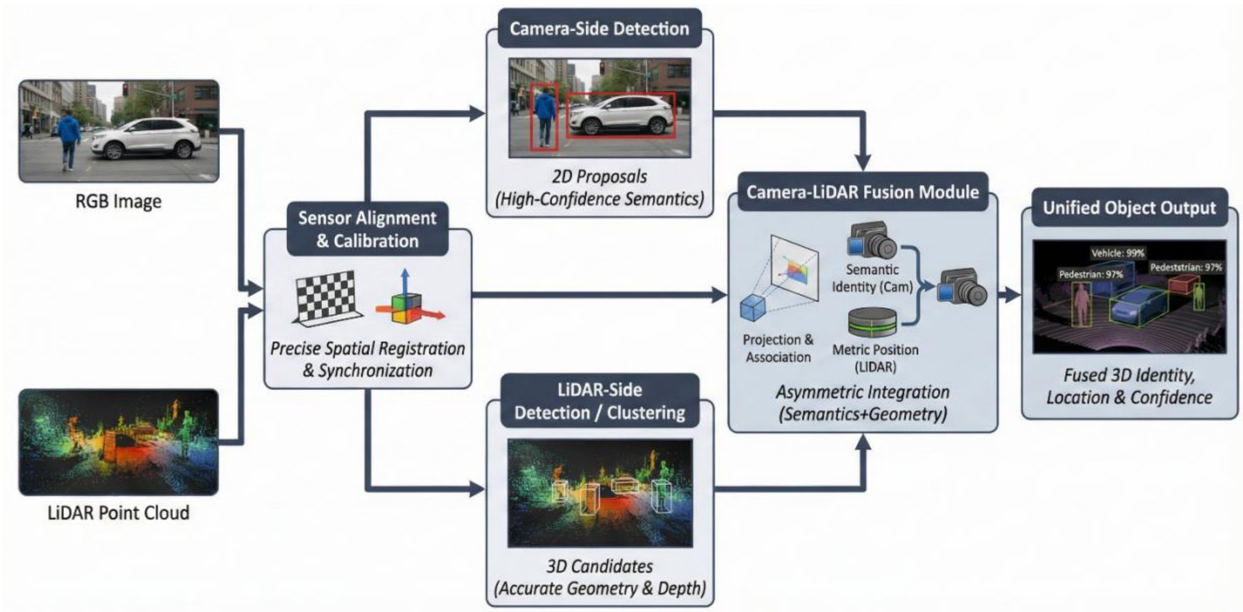


Figure 4. Proposed Camera–LiDAR Object Detection Pipeline

Sensor Alignment and Calibration

We first establish the spatial relationship between the camera and LiDAR so that both sensors can be expressed in a consistent coordinate relationship. This enables projecting LiDAR points into the camera image and mapping camera detections into 3D space using LiDAR geometry. In camera–LiDAR perception literature, reliable cross-sensor alignment is a prerequisite for stable fusion because association relies on consistent geometry between the two modalities (Chen et al., 2017; Ku et al., 2018).

Single-sensor detection. For each time step, we generate candidate objects from the two modalities:

- Camera-side detection produces 2D bounding boxes and confidence for pedestrians and vehicles in the image.
- LiDAR-side detection (or clustering) produces 3D candidates based on point geometry, providing distance and spatial cues.

Camera–LiDAR Fusion

We fuse camera and LiDAR outputs into a single coherent object list. Fusion follows an association logic that checks spatial consistency between (i) camera detections and (ii) LiDAR candidates after projection/transform. When a match is found, the fused object combines the strengths of both sensors: semantic identity and confidence are taken primarily from the camera, while metric 3D position/range is taken primarily from LiDAR. This "semantic-from-camera + geometry-from-LiDAR" pattern is widely used and has been validated across multiple camera–LiDAR fusion pipelines (Vora et al., 2020; Huang et al., 2020; Yoo et al., 2020).

Unified Output for Downstream Use

The final output of this module is a set of fused objects (pedestrians and vehicles). Each fused object includes a category label, a 3D location (in a consistent reference frame), and a confidence indicator. This output format is aligned with common camera–LiDAR 3D detection practice, where the fused representation is designed to be directly consumed by downstream tasks requiring both identity and metric distance (Bai et al., 2022; Liu et al., 2023).

3.3. Field Testing Results

Field testing confirms that the camera–LiDAR perception stack can reliably detect pedestrians and vehicles in real roadway environments, as shown in Figure 5. Across tested scenes, the camera detector consistently identifies pedestrians and vehicles in the image, and LiDAR provides stable range and 3D structure cues. After fusion, the system outputs a unified object list in which each pedestrian/vehicle is represented with both semantic identity (from camera) and metric position/range (from LiDAR).

Figure 5. Field Testing Results Using the Proposed Camera–LiDAR Detection Pipeline

Quantitatively, Table 1 shows that camera–LiDAR fusion achieves the strongest overall detection performance for both object types. For pedestrians, fusion improves precision/recall to 0.93/0.90 and increases mAP@0.5 to 0.92 compared with camera-only (0.90/0.84, mAP@0.5=0.88) and LiDAR-only (0.88/0.79, mAP@0.5=0.85). For vehicles, fusion similarly reaches 0.97/0.95 precision/recall with mAP@0.5=0.96, outperforming camera-only (0.95/0.92, mAP@0.5=0.94) and LiDAR-only (0.94/0.90, mAP@0.5=0.93). In addition, fusion provides reliable metric localization with lower range RMSE (0.41 m for pedestrians and 0.33 m for vehicles) than LiDAR-only (0.68 m and 0.52 m, respectively), which is critical for distance-aware downstream decision-making.

Table 1. Field-test performance of camera–LiDAR perception for pedestrian and vehicle

Object	Method	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Latency	RMSE
Pedestrian	Camera-only	0.90	0.84	0.88	0.56	42	—
	LiDAR-only	0.88	0.79	0.85	0.53	55	0.68
	Camera–LiDAR fusion	0.93	0.90	0.92	0.61	68	0.41
Vehicle	Camera-only	0.95	0.92	0.94	0.69	42	—
	LiDAR-only	0.94	0.90	0.93	0.67	55	0.52
	Camera–LiDAR fusion	0.97	0.95	0.96	0.73	68	0.33

In practical operation, fusion improves usability compared with camera-only outputs because downstream connected-vehicle functions often need metric distance and spatial context, especially for pedestrians. Camera detections provide strong category recognition, but without LiDAR-based geometry they are less actionable for distance-aware decision modules. Conversely, fusion improves clarity compared with LiDAR-only candidates by reinforcing category recognition using camera semantics, which is a common and effective strategy in many established fusion approaches (Vora et al., 2020; Huang et al., 2020).

Therefore, the field results demonstrate that the integrated camera–LiDAR approach provides reliable pedestrian and vehicle detection and yields stable fused perception outputs consistent with widely adopted fusion designs in the literature (Yoo et al., 2020; Bai et al., 2022; Liu et al., 2023).

3.4. Discussions

The main takeaway from this chapter is that camera–LiDAR fusion is a practical and deployable sensing configuration for reliably detecting pedestrians and vehicles. Cameras provide strong semantic recognition and category discrimination, while LiDAR provides accurate geometry and distance. When these outputs are fused into a unified object representation, the perception result becomes more actionable than either modality alone because it simultaneously answers “what the object is” and “where it is in 3D space” (Chen et al., 2017; Ku et al., 2018; Qi et al., 2018).

This design also aligns well with deployment needs in connected systems. A fused object list with consistent labels and metric 3D locations provides a clean interface for downstream modules, while remaining interpretable and easy to integrate. At the same time, the fusion approach remains conceptually simple: spatially associate camera detections with LiDAR candidates, and then combine semantics with geometry. This simplicity is valuable for engineering reliability, and it matches common design principles found in practical fusion pipelines (Vora et al., 2020; Huang et al., 2020; Bai et al., 2022).

While this chapter focuses on reliable detection and fusion within a single camera–LiDAR stack, the same fused object representation can naturally support future extensions, such as sharing fused objects across connected agents. Establishing a stable and reliable fused baseline is a necessary step before moving to more advanced cooperative perception settings, and it ensures that any shared information maintains clear semantics and usable geometry (Liu et al., 2023; Bai et al., 2022).

Chapter 4. Cooperative Perception Demonstrations

4.1. Equipment

The CP demonstrations were conducted using the University of Maryland's Level 4 CAV platform, as shown in Figure 6. The equipment configuration integrates onboard sensing, high-performance computing, precision localization, and V2X communication to support real-time perception sharing and cooperative safety applications. The Level 4 CAV platform is designed to support autonomous driving research under controlled and semi-controlled environments. The vehicle is equipped with a full-stack sensing, computing, and communication architecture enabling automated perception, localization, planning, and control.

(a) The LV 4 CAV

(b) Real-time Process of 3D point cloud from Lidar

Figure 6. The UMD in-house LV 4 Research CAV

Onboard Perception Sensors

The vehicle includes a heterogeneous sensor suite to provide 360-degree environmental awareness:

- 3D LiDAR Sensors: High-resolution LiDAR units provide dense point cloud measurements for 3D object detection and accurate range estimation. These sensors enable reliable geometric perception of vehicles, pedestrians, and roadside structures.
- Multi-Camera System: Multiple RGB cameras are mounted around the vehicle to provide overlapping fields of view for semantic recognition, object classification, lane detection, and visual tracking.
- Radar Sensors: Automotive radar units provide robust velocity estimation and long-range detection, particularly useful under adverse weather conditions.
- GNSS/GPS with RTK Capability: High-precision positioning is achieved using GNSS augmented with Real-Time Kinematic (RTK) corrections, enabling lane-level localization accuracy.
- Inertial Measurement Unit (IMU): The IMU provides acceleration and angular velocity data for motion estimation and sensor fusion.

The combination of LiDAR, camera, radar, GNSS, and IMU allows robust multi-sensor fusion, supporting both onboard autonomous driving functions and CP message generation.

Onboard Computing System

The Level 4 CAV is equipped with a high-performance computing platform designed for real-time perception and decision-making. The onboard system includes:

- GPU-enabled computing units for deep learning–based object detection and fusion.
- Multi-core CPUs for sensor synchronization, message handling, and trajectory planning.
- Data logging systems for recording raw sensor data, fused objects, and V2X messages for post-processing and evaluation.

The computing architecture supports real-time execution of perception pipelines, cooperative fusion algorithms, and collision avoidance modules.

Drive-by-Wire and Control Interface

The vehicle is integrated with a drive-by-wire control system that allows automated control of steering, throttle, and braking. This enables closed-loop implementation of CP-informed decision-making, such as speed adjustment, emergency braking, and trajectory modification based on infrastructure-shared perception data.

4.2 Demonstration Setup

The CP demonstration was conducted in a controlled test environment designed to replicate an occluded pedestrian-crossing scenario. The objective of the setup was to evaluate the ability of infrastructure-assisted perception to detect a pedestrian that is not visible to the approaching CAV due to line-of-sight obstruction.

4.2.1 Physical Layout

As illustrated in Figure 7, the demonstration site consisted of a straight roadway segment approximately 100 feet in length. A simulated crosswalk area was marked on the pavement to represent a pedestrian crossing zone. The crosswalk marking was used solely for scenario structuring and did not require permanent road modifications. To create an occlusion condition, several large box-shaped obstacles were positioned adjacent to the crosswalk. These obstacles simulated real-world visual obstructions such as parked vehicles, delivery trucks, or roadside structures. The placement of the obstacles was carefully arranged to block the direct line of sight between the approaching CAV and the pedestrian positioned behind them. A dummy pedestrian, mounted on a movable platform, was positioned behind the occluding objects. The pedestrian surrogate was used to simulate a sudden crossing event emerging from behind the obstruction into the vehicle's path.

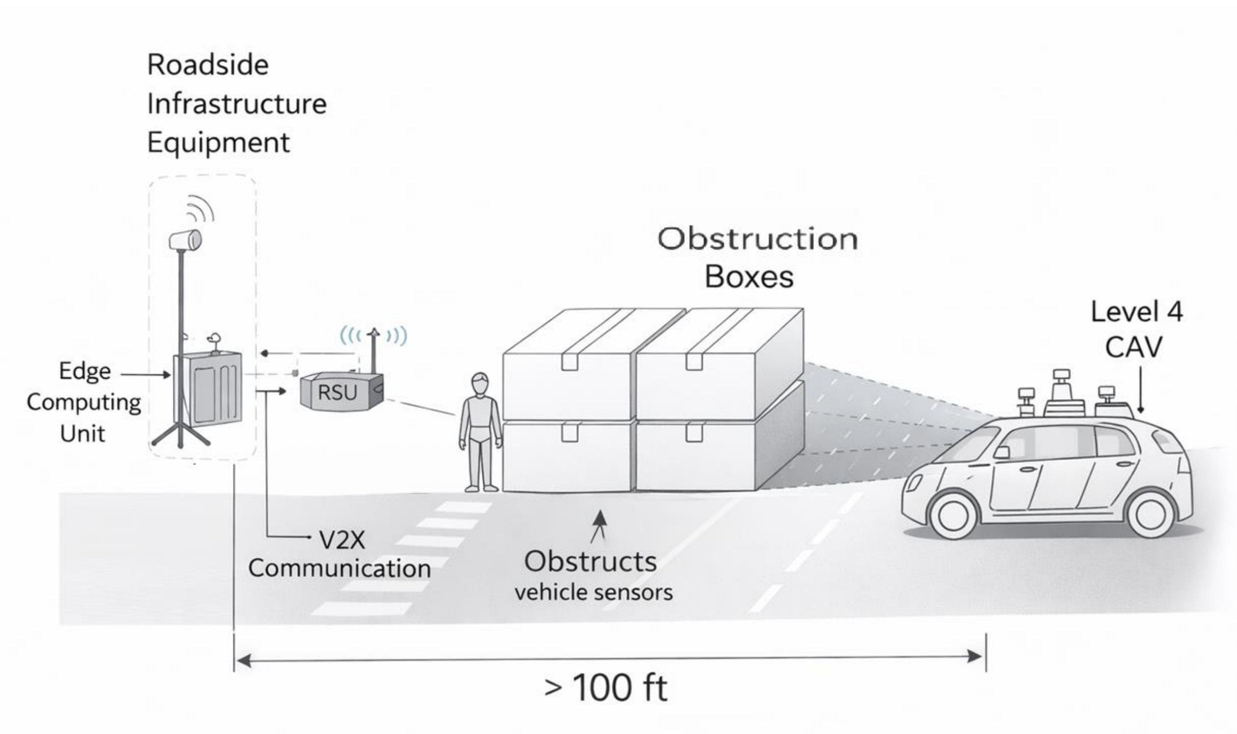


Figure 7. Cooperative Perception Demonstration Setup

4.2.2 Infrastructure-Based Sensing Configuration

A roadside sensing unit was deployed near the crosswalk to provide an elevated vantage point of the occluded area. The infrastructure sensing system included:

- A camera mounted on a vertical pole at an elevated height to monitor the crosswalk and surrounding region.

- An edge computing unit installed at the base of the pole to process video streams in real time.

The camera was oriented to ensure full visibility of both the occluded pedestrian region and the roadway approach path of the CAV. Unlike the vehicle's onboard sensors, which were blocked by the obstacle, the elevated infrastructure sensor maintained a clear field of view of the pedestrian throughout the event.

The edge computing unit performed real-time object detection and classification, converting raw video data into structured object-level information (e.g., pedestrian presence, location, and motion status). This processed information was prepared for transmission via V2X communication.

4.2.3 Vehicle Approach Configuration

The Level 4 CAV approached the crosswalk along the roadway segment at controlled speeds. During the initial phase of the approach, the occluding objects prevented the vehicle's onboard perception system (camera and LiDAR) from directly detecting the pedestrian. This created a deliberate perception blind spot.

As the vehicle advanced toward the crosswalk, two perception streams were active:

1. Onboard perception, relying on vehicle-mounted LiDAR and cameras.
2. Infrastructure-assisted perception, detecting the pedestrian from the elevated roadside viewpoint.

The demonstration compared system behavior with and without cooperative perception input to assess the added value of infrastructure-supported detection.

4.2.4 Cooperative Perception Information Flow

The demonstration followed a structured perception-sharing pipeline:

1. The roadside camera detected the pedestrian in the occluded region.
2. The edge computing unit processed the detection and generated object-level data.
3. The Roadside Unit (RSU) transmitted the pedestrian information via communication channels such as C-V2X and 5G.
4. The CAV's Onboard Unit (OBU) received the message.
5. The vehicle fused infrastructure-detected objects with its onboard perception.
6. The collision avoidance module evaluated risk and initiated appropriate control actions if necessary.

This setup enabled evaluation of early detection capability, communication latency, and vehicle response timing under controlled occlusion conditions.

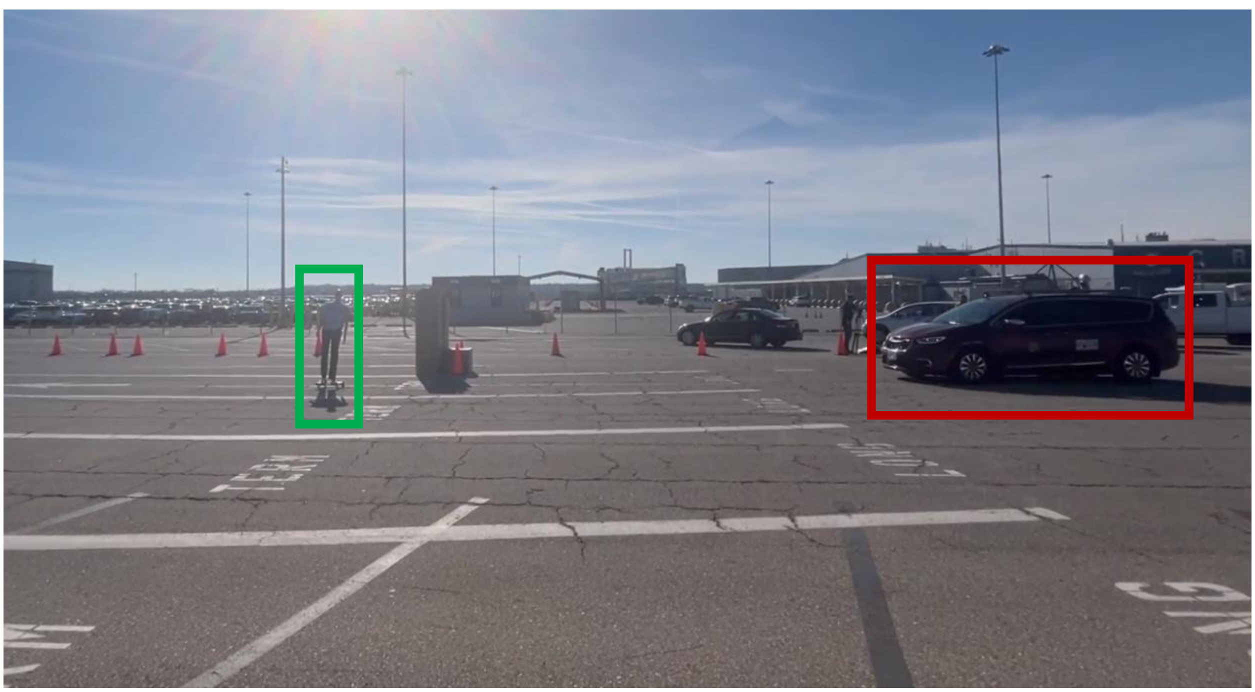
4.3 Demonstrations

To illustrate the safety value of CP in pedestrian-collision avoidance, we conducted three controlled demonstration cases that progressively highlight (i) when onboard sensing is sufficient, (ii) when CP provides critical early awareness under occlusion, and (iii) what can occur when CP is absent in the same occluded scenario.

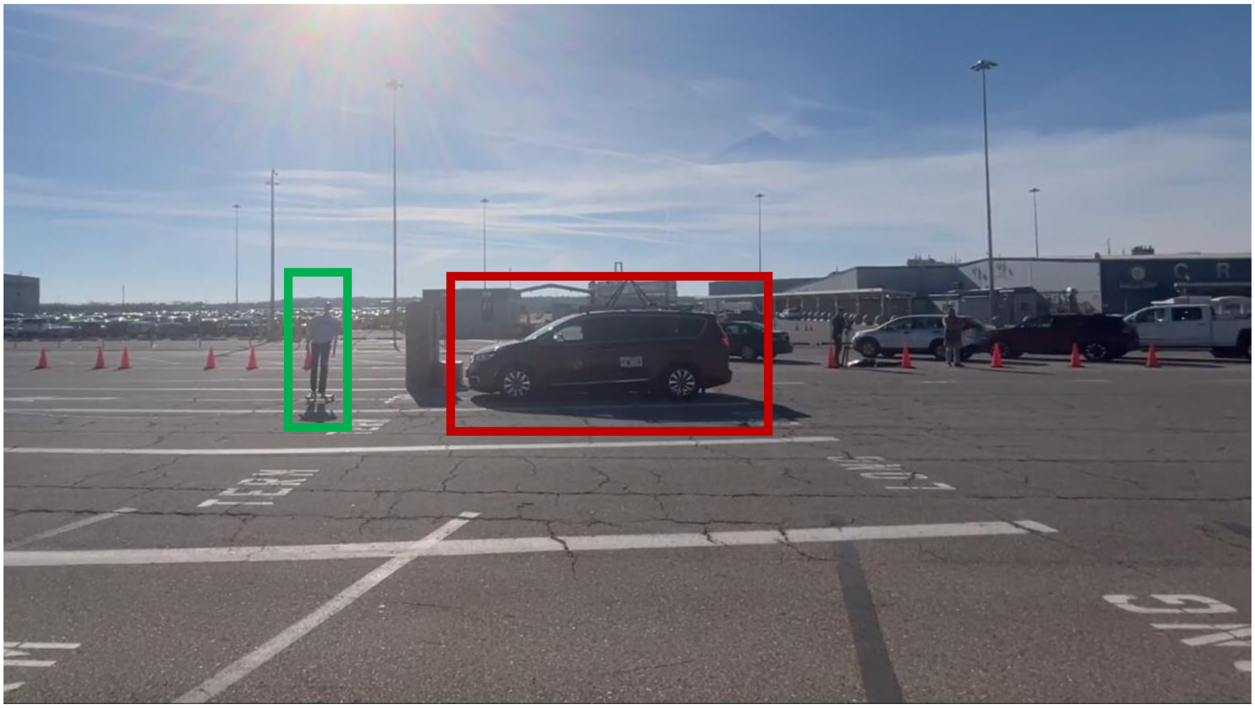
4.3.1 Case 1 – CAV reacted to pedestrian (no CP needed)

Case 1 establishes the baseline condition where the CAV can detect the pedestrian using onboard sensing without infrastructure assistance. As shown in Figure 8, the pedestrian (green bounding box) is fully visible in the vehicle's forward field of view, while the vehicle/obstruction element (red box) is positioned away from the pedestrian line-of-sight and does not block the camera/LiDAR view. Under this unobstructed geometry, the onboard perception stack can reliably recognize the pedestrian early enough to support normal longitudinal control (speed reduction and/or full stop) before reaching the crosswalk conflict zone.

Figure 8 (b) shows the same run a few seconds after Figure 8 (a) as the CAV continues driving forward. Despite the change in relative position, the pedestrian remains continuously visible to the CAV's onboard sensors, there is no moment when the pedestrian is occluded by the vehicle/obstruction. With uninterrupted onboard detection, the CAV decelerates and comes to a stop with a sufficient safety gap, without relying on cooperative perception input.



(a) Clear line-of-sight (dummy pedestrian presented on the crosswalk)



(b) CAV stopped with sufficient gap

Figure 8. Two frames from the Case 1 demonstration video

Case 1 serves as a baseline demonstrating that CP is not always required. When line-of-sight is available, the CAV's standard onboard perception and control can reliably detect the pedestrian and execute a safe stop with adequate margin. This baseline is important because it clarifies the role of CP in the overall system: CP is not intended to override onboard autonomy, but to address well-known failure modes, particularly occlusion (as well as adverse visibility and complex geometry), where onboard sensors may detect hazards too late for a comfortable or safe response.

4.3.2 Case 2 – CAV stopped safely with CP assistance (occluded pedestrian)

Case 2 demonstrates the primary safety benefit of cooperative perception (CP) in an occluded pedestrian-crossing scenario. As shown in Figure 9 (a), the pedestrian (green box) is positioned near the crosswalk but is visually blocked from the CAV's onboard sensors by the obstruction boxes/vehicle acting as a line-of-sight barrier (red box). Under this geometry, the CAV's forward-facing camera/LiDAR cannot reliably detect the pedestrian until the pedestrian emerges from behind the obstruction, significantly reducing available reaction time.

(a) Dummy pedestrian occluded by obstruction

(b) Dummy pedestrian emerges; CP-enabled safe stop

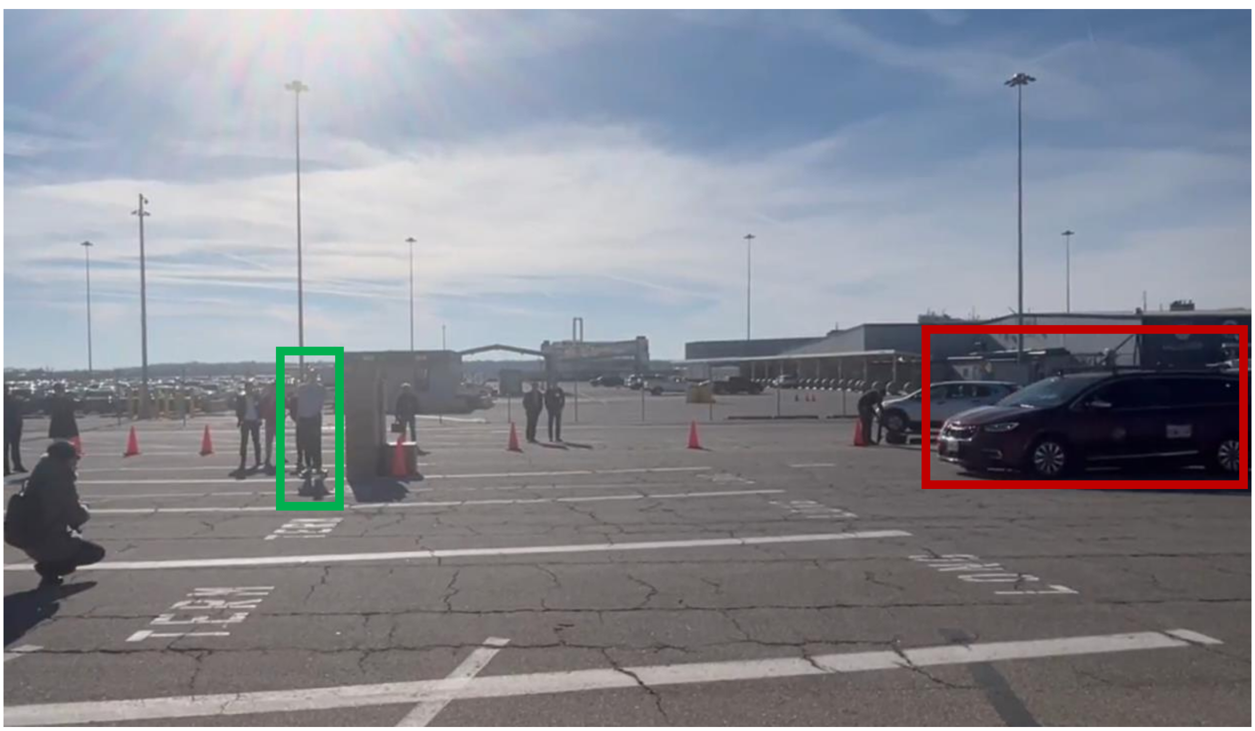
Figure 9. Two frames from the Case 2 demonstration video

A few seconds later, Figure 9 (b) captures the critical event: the pedestrian suddenly steps into the roadway from behind the obstruction. In this case, the CP pipeline enables early awareness before direct visibility is restored. Specifically, the roadside sensing unit maintains a vantage point that can observe the pedestrian despite the occlusion and transmits pedestrian presence (and, when available, position/motion cues) to the approaching CAV via V2X. With this advance warning integrated into the vehicle's perception and decision stack, the CAV begins deceleration earlier and successfully comes to a complete stop with a safe gap before the pedestrian conflict point, even though the pedestrian emerges abruptly.

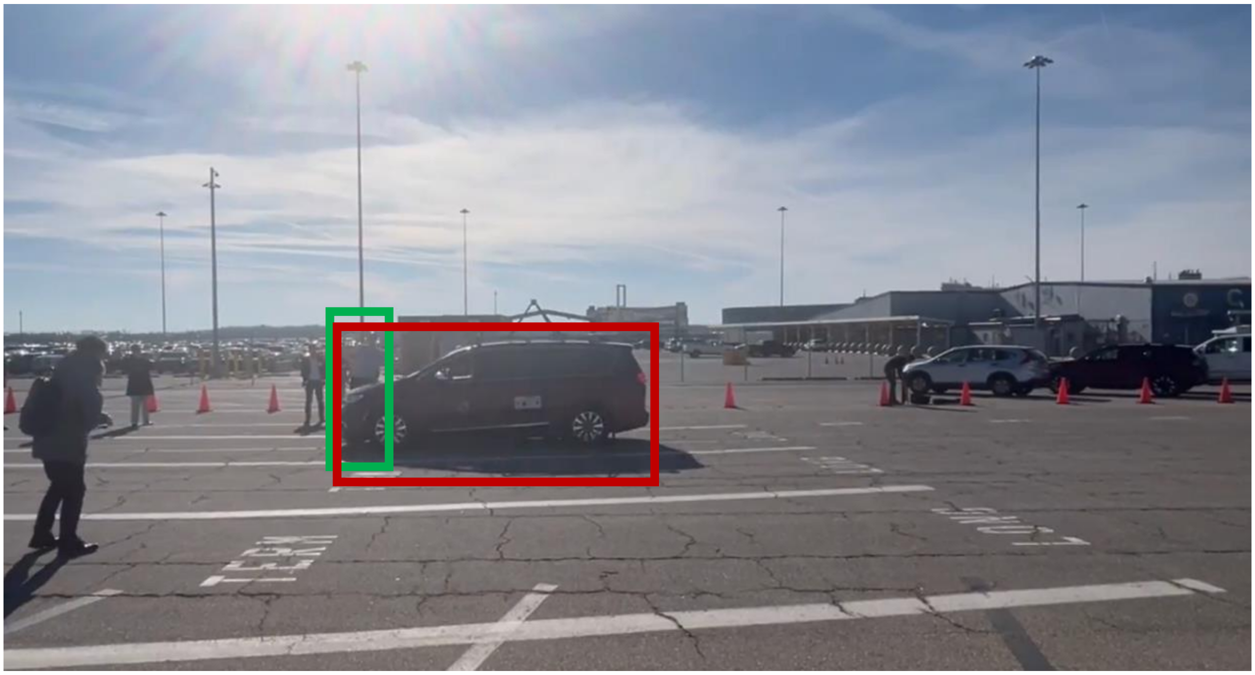
Overall, Case 2 confirms that CP provides an additional safety layer precisely in the failure mode where onboard-only autonomy reached its limitations—occluded, time-compressed pedestrian emergence—by extending perception beyond the vehicle's line-of-sight and enabling timely collision avoidance actions.

4.3.3 Case 3 – Collision without CP (occluded pedestrian, CP disabled)

Case 3 reproduces the same occluded-crosswalk geometry as Case 2 but with cooperative perception (CP) disabled, illustrating the baseline failure mode under strong line-of-sight blockage. In Figure 10, the dummy pedestrian (green box) is positioned near the crosswalk but is fully occluded from the approaching CAV by the obstruction (red box). Because the pedestrian is hidden behind the obstruction, the CAV's onboard sensors do not have sufficient opportunity to detect and track the pedestrian early in the approach.



(a) Dummy pedestrian occluded by obstruction



(b) CAV crashed into the dummy pedestrian

Figure 10. Two frames from the Case 3 demonstration video

A few seconds later, Figure 10 (b) shows the critical outcome: as the pedestrian suddenly emerges from behind the obstruction, the CAV detects the pedestrian too late to generate adequate braking distance, resulting in a collision with the dummy pedestrian. This case highlights the time-compressed nature of occlusion-driven pedestrian conflicts, once the pedestrian becomes visible, the remaining time-to-collision may be insufficient for a safe stop using onboard-only perception and control.

Overall, Case 3 provides a direct contrast to Case 2 by demonstrating that, under the same occlusion conditions, removing CP eliminates the early warning needed to compensate for onboard line-of-sight limitations. The comparison between Cases 2 and 3 therefore isolates the safety contribution of CP: infrastructure-based detection and V2X sharing can convert an otherwise unavoidable late-detection conflict into an avoidable event by restoring reaction time before the pedestrian becomes visible to the vehicle.

4.4 Testbed Integration for Long-term Deployment

4.2.1 MSU CAV Testbed and SMART Corridor

As the capability of CP has been tested and approved, a key objective for the long-term maturation of the proposed cooperative perception system is its integration into a real-world connected and automated vehicle (CAV) testbed. While the present project primarily focused on demonstrating the core concept, validating the sensing pipeline, and establishing the feasibility of infrastructure-assisted pedestrian detection, the MSU CAV testbed, also known as the MSU

SMART Corridor, already provides the essential physical, sensing, communication, and computing foundation needed for long-term deployment. In other words, the current project should be viewed as a pilot-stage development of the cooperative perception capability, whereas the MSU SMART Corridor offers a natural path toward sustained field integration, continuous testing, and future operational use.

MSU operates a mature urban CAV test environment centered on the SMART Corridor along Hillen Road in Baltimore. As illustrated in Figure 11, the corridor spans approximately 1.5 miles and consists of a network of intersections equipped with advanced sensors and Vehicle-to-Everything (V2X) communication technologies. Rather than serving as a single isolated instrumented intersection, the corridor is designed as a connected operational environment that supports data collection, real-time communication, and deployment of safety and mobility applications across multiple roadway contexts. The SMART Corridor is intended to prevent collisions, prioritize transit and emergency vehicles, and support data-driven planning decisions, which aligns closely with the safety-oriented goals of the cooperative perception system developed in this project.

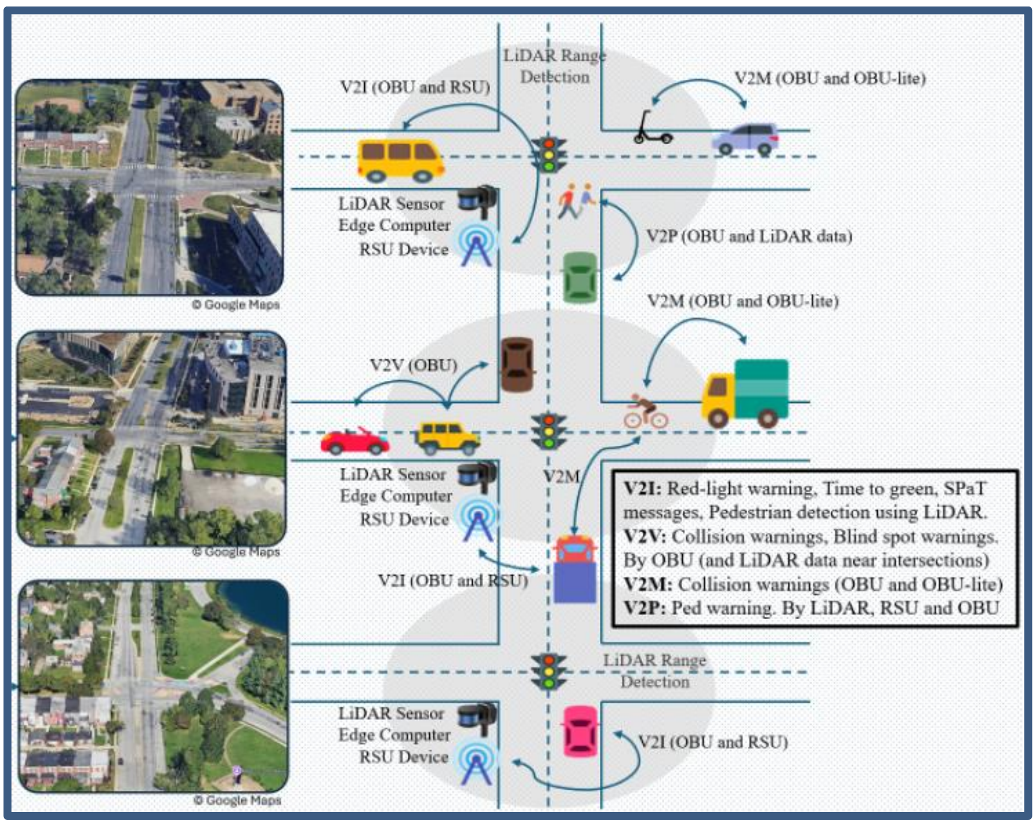


Figure 11. Overview of SMART Corridor

The existing corridor is particularly well-suited for testing infrastructure-assisted perception for pedestrian safety because it already supports multimodal traffic monitoring in a dense urban setting. The real-world roadway conditions around Morgan State involve repeated interactions among private vehicles, transit vehicles, pedestrians, cyclists, and university-related traffic streams. Such conditions are precisely those in which line-of-sight limitations, visual clutter, and unpredictable pedestrian behavior can challenge onboard vehicle perception alone. The cooperative perception system developed in this project addresses that gap by allowing infrastructure-based sensing devices to detect pedestrians who may be obscured by parked vehicles, roadside obstacles, or other roadway users, and then communicate that information to nearby automated vehicles in time for defensive action. Integrating this functionality into the SMART Corridor would therefore extend the existing testbed from passive monitoring and analytics toward active, infrastructure-enabled cooperative safety intervention.

The technical architecture already deployed on the corridor provides most of the essential ingredients for such integration. First, the testbed includes LiDAR-based sensing capability. As shown in Figure 12, the corridor supports three-dimensional environmental sensing that can capture detailed spatial information about roadway users and their surroundings. This is especially important for pedestrian detection, since 3D sensing is more robust than conventional camera-only approaches in resolving object position, motion, and partial visibility. The ability to generate live point-cloud data means that the infrastructure can already provide the spatial awareness necessary to support the developed cooperative perception pipeline.

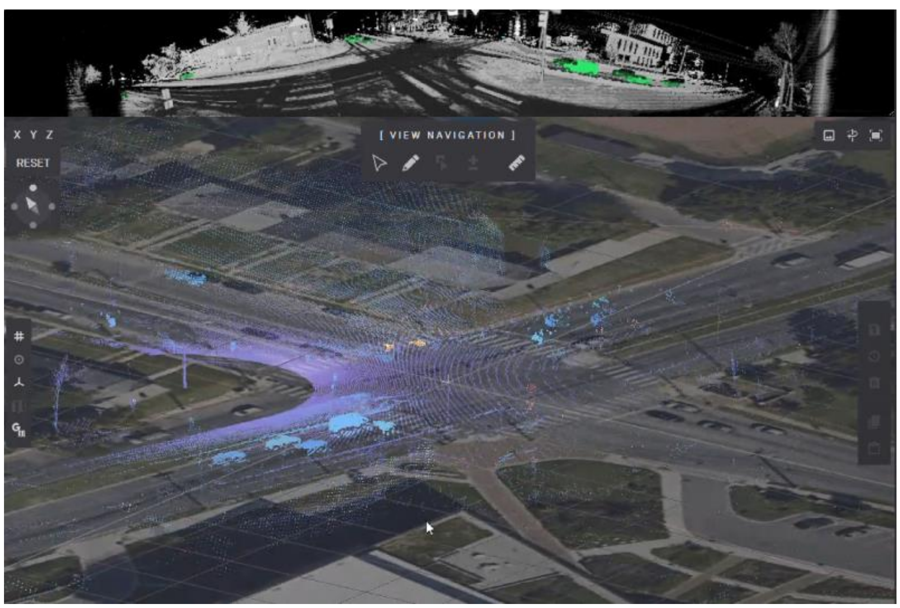


Figure 12. MSU CAV testbed LiDAR point clouds

Second, the testbed already demonstrates multi-object tracking functionality. Figure 13 shows that LiDAR and CCTV are used together to track discrete objects moving through the intersection, with each detected object assigned a unique identifier. These objects include cars, trucks, transit vehicles, pedestrians, and cyclists. This is a highly relevant capability for long-term deployment of the proposed system because pedestrian safety applications depend not only on detecting the presence of a pedestrian, but also on maintaining stable object identity and trajectory estimates over time. For infrastructure-assisted cooperative perception, this tracking layer forms the bridge between raw sensor measurements and actionable safety messages. Once an obscured pedestrian is detected and tracked, the system can estimate the pedestrian's likely movement path, determine whether that path intersects with an approaching vehicle trajectory, and transmit a warning or state update to the vehicle accordingly. Thus, the existing object tracking framework within the SMART Corridor substantially reduces the additional development burden required for integration.

Figure 13. MSU CAV testbed trajectory extraction

Third, the SMART Corridor already supports safety-oriented traffic conflict analysis that can serve both as a deployment target and as an evaluation platform. Figure 14 presents an example of a near-miss event between a vehicle and a pedestrian, while Figure 15 indicates that such events are analyzed using Post Encroachment Time (PET), including both vehicle-vehicle and vehicle-non-vehicle conflicts. These capabilities are highly relevant to the cooperative perception system in two distinct ways. From an operational perspective, near-miss and conflict events identify the classes of scenarios in which infrastructure-assisted warnings are most needed. From an evaluation perspective, these existing metrics provide a natural performance framework for quantifying the benefits of deployment. For example, after integration, the system could be assessed by examining whether the presence of cooperative pedestrian alerts reduces the frequency of low-PET interactions, decreases severe conflict rates, or increases the time margin available for vehicle response. In this sense, the testbed does not merely provide a place to host the new system; it also provides an established methodological framework for validating whether the system improves safety outcomes.

Figure 14. An example of pedestrian-vehicle conflict

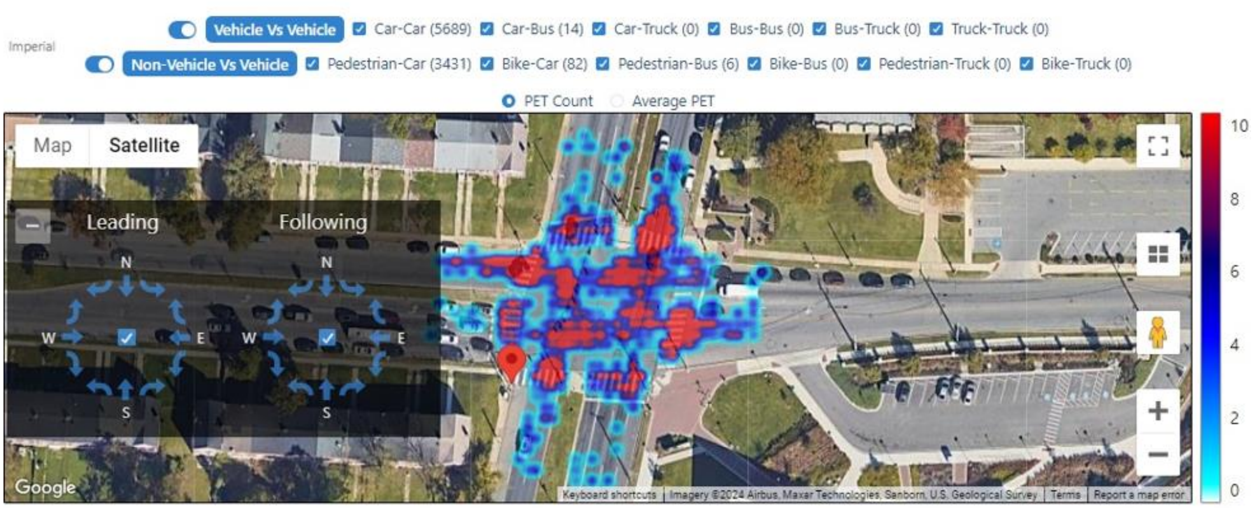


Figure 15. Pedestrian conflict heatmap

The broader analytic capabilities of the SMART Corridor further strengthen its suitability for long-term deployment. Figure 16 identifies high-conflict zones and highlights that a large share of recorded conflicts are concentrated in specific areas, with conflict severity associated with vehicle speed, time of day, weather conditions, and traffic volume while Figure 17 presents pedestrian behavior analysis, including out-of-crosswalk movements and trajectory-based relationships between pedestrian behavior and surrounding infrastructure. Collectively, these capabilities demonstrate that the corridor already supports a rich operational picture of intersection safety, especially for pedestrians. This context is critical because cooperative perception for automated vehicles should not be deployed in isolation from the broader traffic environment. Instead, the system should leverage contextual information such as pedestrian density, expected crossing patterns, vehicle speeds, and high-risk time periods to improve triggering logic, communication priorities, and adaptive warning thresholds. The current testbed infrastructure already provides such context, making integration both technically feasible and scientifically valuable.

Zone	% of Conflicts
Zone 2	38%
Zone 5	24%
Zone 8	14%
Zone 11	15%
Zone 1	6%
Zone 12	3%

Figure 16. Proportions of traffic conflicts per zone

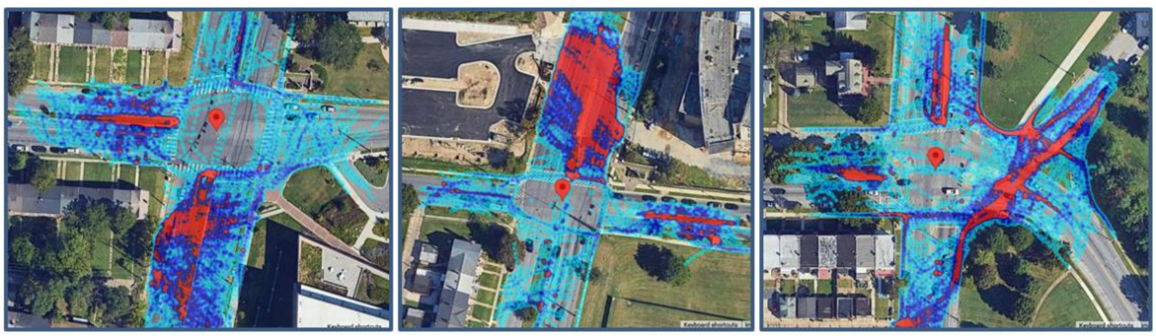


Figure 17. Heatmap of outside of crosswalk events

An especially important component for long-term deployment is the Portable Detection & Communication System (PDCS) developed as part of the testbed, as shown in Figure 18. The PDCS indicates that the Morgan testbed is not limited to fixed, permanently installed sensing infrastructure; it also includes portable capability for safety evaluation, before-after studies, and short-term warning system deployment. This is directly relevant to the future evolution of the cooperative perception system developed in this project. In its current form, the project established the core cooperative perception logic and demonstrated the concept of infrastructure-assisted pedestrian detection and communication. For long-term deployment, however, a practical pathway is to use portable detection and communication units as an intermediate integration mechanism. This enables the research team to instrument selected conflict-prone locations, test the cooperative perception module in different geometric and operational conditions, and refine the sensing, communication, and alert-generation pipeline before full deployment across the SMART Corridor.

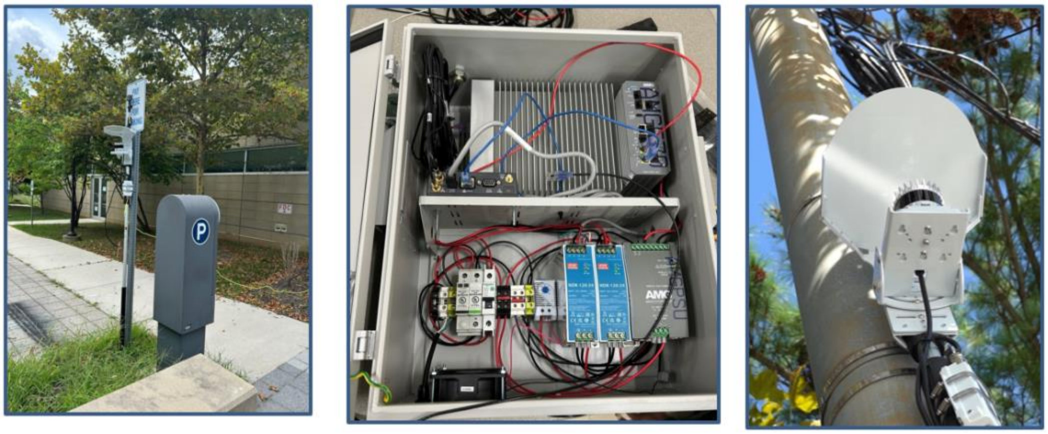


Figure 18. Portable LiDAR integration

From a system perspective, the integration pathway is conceptually straightforward. The cooperative perception system developed in this project can be positioned as an application layer on top of the existing corridor infrastructure. The infrastructure sensing layer, consisting of LiDAR and CCTV, would continue to detect and track roadway users. An edge-processing layer would run the proposed occlusion-aware pedestrian detection and trajectory inference module. Once a hidden or partially obscured pedestrian is identified as relevant to an approaching automated vehicle, the communication layer would encode and transmit that information through the corridor's V2X framework to nearby vehicles or connected mobile platforms. The receiving vehicle would then use the infrastructure-derived pedestrian information as an external perception input to support speed adjustment, deceleration, or other defensive control actions. Because the corridor already contains sensing, communication, and analytics capabilities, the main incremental development need is the integration of the newly developed cooperative perception logic into the existing software and operational stack rather than the creation of a new testbed from scratch.

It is important to note that this project did not pursue full corridor-level deployment because its main focus was proof-of-concept development and pilot-stage validation. That scope was appropriate for the project phase: the primary aim was to demonstrate that infrastructure-based sensing can identify pedestrians hidden behind obstacles and communicate this information in a way that may help automated vehicles avoid collisions. Full deployment requires additional systems engineering tasks, extended reliability testing, communication latency assessment, and staged operational validation. Nevertheless, the results of the project suggest that such deployment is feasible. The MSU SMART Corridor already offers the necessary ecosystem for this transition, and the work completed under the present project substantially de-risks future field implementation.

4.2.2 Integration Plan for Future Long-Term Deployment

For long-term deployment, a phased integration plan has been developed to incorporate the cooperative perception system into the Morgan State testbed in a systematic and scalable manner.

The first phase should focus on software and interface integration. In this stage, the cooperative perception algorithm developed in the project would be packaged as a deployable module that can ingest real-time LiDAR and camera-derived object streams from the SMART Corridor. The module should be configured to receive synchronized infrastructure observations, identify occluded pedestrian candidates, estimate their probable trajectories, and generate event-level safety outputs for downstream communication. This phase would also define the message structure needed to transmit pedestrian presence, position, motion state, and risk level to connected vehicles through the existing V2X framework. Since the corridor already supports real-time communication and object tracking, this phase primarily involves integration of perception logic and interface harmonization rather than new hardware development.

The second phase should involve portable field deployment using the PDCS at one or more high-priority intersections or conflict zones. This is an ideal transitional step because it allows the team to validate the cooperative perception system under real traffic conditions without immediately requiring permanent corridor-wide installation. Portable deployment would support controlled testing under different times of day, traffic volumes, weather conditions, and pedestrian activity levels. It would also allow the research team to benchmark communication latency, detection persistence, false alarm behavior, and response timing for approaching vehicles. Because the SMART Corridor already supports conflict analytics such as PET and near-miss assessment, these existing metrics can be used to evaluate whether the integrated system increases safety margins in the field.

The third phase should emphasize closed-loop testing with connected or automated vehicle platforms. In this phase, the infrastructure-generated pedestrian information would be transmitted to instrumented vehicles operating on or adjacent to the SMART Corridor, enabling direct assessment of whether the external perception input improves vehicle response in occlusion scenarios. This stage is critical because the ultimate value of the system depends not only on detecting hidden pedestrians, but on ensuring that the information reaches the vehicle in a timely and usable form that supports safe control actions. Closed-loop testing would therefore examine end-to-end performance, including sensing, inference, communication, vehicle reception, and behavioral response.

The fourth phase should extend the deployment into a longer-term operational demonstration. After the system demonstrates stable performance in pilot deployments, the cooperative perception capability can be incorporated as a standing safety application within selected SMART Corridor intersections. At this stage, the system would operate continuously, and its performance could be evaluated over longer time horizons using the corridor’s existing traffic, conflict, speed, and pedestrian behavior datasets. This longer-term operational deployment would create opportunities for before-after comparisons, reliability analysis, maintenance planning, and eventual expansion to other sites in Baltimore or similar urban corridors.

Chapter 5. Other CP Scenarios

5.1. Background

This chapter focuses on discussing other potential cooperative perception (CP) scenarios that can further demonstrate the benefits of leveraging the framework developed in this project. While the core effort of this project centered on building and validating a CP-based system for infrastructure-assisted perception, the same underlying architecture can support a much broader set of safety-critical applications involving pedestrian, collision, and conflict avoidance. In particular, the developed framework provides the essential building blocks for future extensions, including roadside and onboard sensing, multi-source data fusion, real-time communication through RSUs and OBUs, and event-triggered warning or decision-support functions for connected and automated vehicles (CAVs). These capabilities make the framework flexible and scalable for addressing a wide range of challenging traffic situations in which vehicle-only perception may be insufficient.

5.2. Pedestrian Protection with CP

Scenario Setup and CP System Architecture

For pedestrian protection, three representative scenarios can be designed. Scenario 1 focuses on pedestrian awareness, using cooperative perception (CP) to inform vehicles of pedestrians at mid-block crossings or intersections. Scenario 2 addresses pedestrians outside designated crosswalks, using CP to alert vehicles when a spatial violation or jaywalking event is detected. Scenario 3 targets pedestrians in vehicle "blind spots," using CP to detect pedestrians who are occluded by large vehicles such as buses.

These three scenarios can be tested and demonstrated using Level 3 CAVs and dummy pedestrians mounted on moving platforms to simulate street-crossing behaviors. In Scenarios 1 and 2, both roadside cameras and LiDAR sensors will be deployed to support pedestrian detection, whether the pedestrian is crossing at an intersection or outside a marked crosswalk. The raw video and point-cloud data will then be processed into object-level information, which can be transmitted to the roadside unit (RSU) through C-V2X and subsequently shared with the onboard units (OBUs) of nearby CAVs to support safer and more informed motion planning.

Scenario 3, as illustrated in Figure 19, involves two pedestrians and two CAVs traveling in the same direction. The system will collect data from each vehicle's radar and front-facing cameras, together with roadside cameras, to monitor blind-spot areas and measure the distance between the CAVs and pedestrians as the pedestrians enter the vehicles' fields of view. During testing, these multi-source data streams will be collected, fused, and analyzed to improve the CP system's ability to detect pedestrians earlier and more accurately.

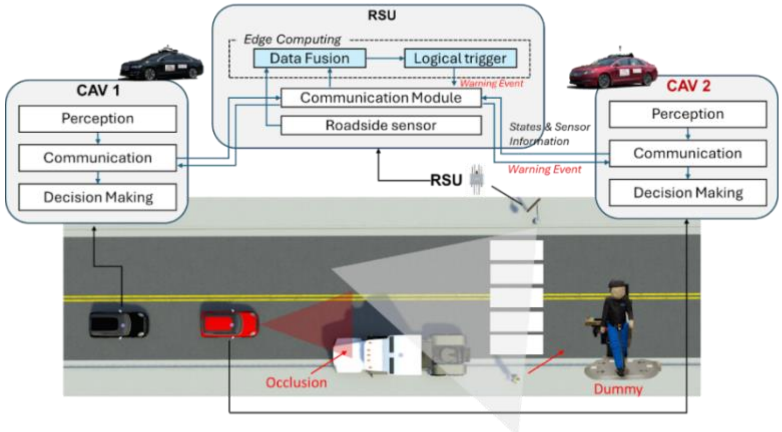


Figure 19. Fusion of multiple sensors in CP

Experiment Plan

Experiment Environment: The experiments for both scenarios can be conducted on the MSU CAV Testbed. Various weather and lighting conditions can be used, including clear daytime, low light at dusk, and nighttime, to evaluate the system's performance across different visibility levels. Each scenario will be tested over a 2-hour period, with 10 tests conducted under each lighting condition (daytime, dusk, and nighttime), leading to a total of 30 tests per scenario. The setup will involve two CAVs, a stationary van creating perception blind spots, and a dummy pedestrian simulating street-crossing behavior from behind the parked van. Roadside infrastructure, including a camera-equipped RSU, will be positioned near the intersection to assist in capturing and processing perception data.

Required Equipment: For the experiment, CAV1 will be equipped with LiDAR, a camera, and GPS sensors and will communicate via 5G using the SAE J2735 V2X standard. The vehicle will approach the intersection and perform left turns in Scenarios 1 and 2. Similarly, CAV2 will have the same set of sensors and communications. A stationary van will be parked beside the east approach to create a perception blind spot in Scenario 3, which will obstruct the line of sight for the CAVs, thereby increasing the complexity of detecting the dummy pedestrian. The RSU will be equipped with a camera for detections. The pedestrian will be represented by a dummy pedestrian, mounted on a moving platform to simulate real-world street-crossing behavior.

Required Modules: The *Perception Module* will gather real-time data from the CAVs (LiDAR, camera, GPS) and the RSU, analyzing it to detect and track the pedestrians. The *Communication Module* will manage the exchange of information between the vehicles and the roadside unit. This communication will ensure seamless transmission of data from the infrastructure to the vehicles. The *Fusion Module* will combine sensor data from both the CAVs and RSU to provide a comprehensive view of the road environment, enhancing the system's ability to detect and track pedestrians that may not be visible to the CAVs. The *Collision Avoidance Module* will detect abnormal pedestrian behavior based on the roadside camera feed. For instance, the system will flag dangerous movements, such as jaywalking, within the first 20-30 frames (approximately 2-3 seconds). The system will also estimate the distance between the pedestrian and the CAVs to assess collision risks and, if necessary, send alerts to the CAV fleet for timely collision avoidance actions.

5.3. Collision Avoidance

Scenario Setup and CP System Architecture

For the collision avoidance system, two representative scenarios will be designed. Scenario 1 focuses on a wrong-way driving (WWD) warning and uses CP to inform vehicles about WWD, and scenario 2 focuses on a car-following alert and uses CP to inform vehicles about a sudden braking event downstream by a non-connected agent. In Scenario 1, the CP prototype will leverage infrastructure-based sensors to detect and respond to WWD events. As illustrated in Figure 20, approaching vehicles are tracked using roadside sensors, such as LiDAR or cameras. Two key zones are defined: the green "Origin" zone, which identifies the correct approach from which a vehicle enters the intersection, and the red "Wrong-Way" zone, which flags potential wrong-way driving. If a vehicle enters the green zone but subsequently appears in the red zone, the system will identify the vehicle as a wrong-way driver. This event is detected and processed in real-time by an edge computer (EC) located in the field. The EC parses the data and forwards the detected WWD event to RSU to disseminate safety-critical messages, such as SDSMs, to nearby vehicles. In Scenario 2, the experiment evaluates the effectiveness of CP in alerting a following vehicle about a sudden braking event initiated by a non-connected vehicle (NCV). The lead CAV (CAV1) detects the NCV's sudden braking event through its onboard sensors and rapidly processes this information. CAV1 then transmits the braking event to the following CAV2 via V2V communication. CAV2 receives the sudden braking alert in real-time, allowing it to initiate a deceleration response before its own sensors detect the event. This proactive warning enables CAV2 to decelerate safely in advance, reducing the risk of a rear-end collision.

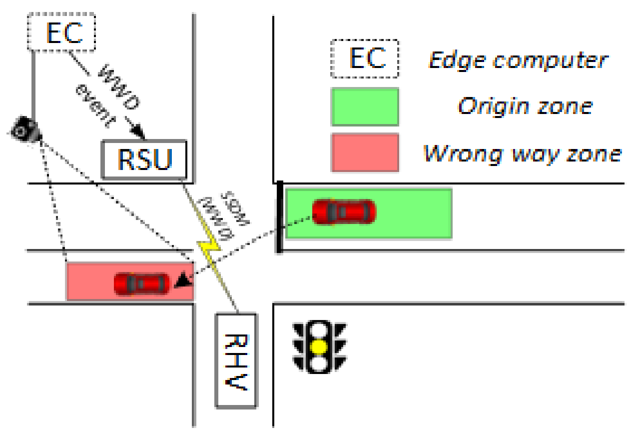


Figure 20. WWD warning with CP

Experiment plan

Experiment Environment: The experiments for both scenarios can be conducted on the MSU CAV Testbed. The weather will be clear with daytime conditions to ensure visibility and consistent results. Each experiment will be conducted over a 2-hour period, with 10 test runs performed for each scenario. This environment ensures the safe and repeatable assessment of the CP system's collision avoidance capabilities without the interference of external traffic.

Required Equipment: For Scenario 1, an NCV will simulate a WWD event by entering the designated wrong-way zone on the closed track. A CAV, equipped with LiDAR, cameras, and an OBU, will approach the intersection from the correct direction. The RSU, mounted near the intersection, will utilize a camera or Lidar to detect the NCV's movement into the wrong-way zone.

The RSU will broadcast information about the detected wrong-way driving event to the CAV via I2V communication. In Scenario 2, the NCV will simulate a sudden braking event in front of CAV1. Both CAV1 and CAV2 will again be equipped with an OBU, LiDAR, cameras, and GPS sensors. CAV1, following the NCV, will detect the sudden braking event and transmit this information to CAV2 through V2V communication.

Required Modules: Same as Section 5.2.

5.4. Conflict Avoidance

Scenario setup and CP system architecture

For the conflict avoidance system, two representative scenarios can be designed. Scenario 1 focuses on red-light violation warning and uses CP to respond to potential red-light running (RLR) proactively. Scenario 2 focuses on a permissive left-turn and uses CP to have the CAV conduct a permissive left-turn at signalized intersections. In Scenario 1, the CP system will use infrastructure-based sensors like LiDAR, radar, and cameras to track approaching vehicles at signalized intersections. As shown in Figure 21, the system will predict potential red-light running (RLR) incidents by monitoring vehicle speed and position, using kinematic models to determine whether a vehicle will stop in time. If a potential RLR event is detected, real-time warnings will be sent to both the red-light runner and vehicles in conflicting movements via V2X communication through the RSU. In Scenario 2 (Permissive Left-Turn Assistance for CAVs), the CP system will assist CAVs in making safe left turns at intersections where their view of oncoming traffic may be obstructed by opposing left-turn vehicles. Roadside sensors will detect the speed and position of oncoming traffic, and this data will be sent to the CAV to inform its decision on whether to proceed or wait.

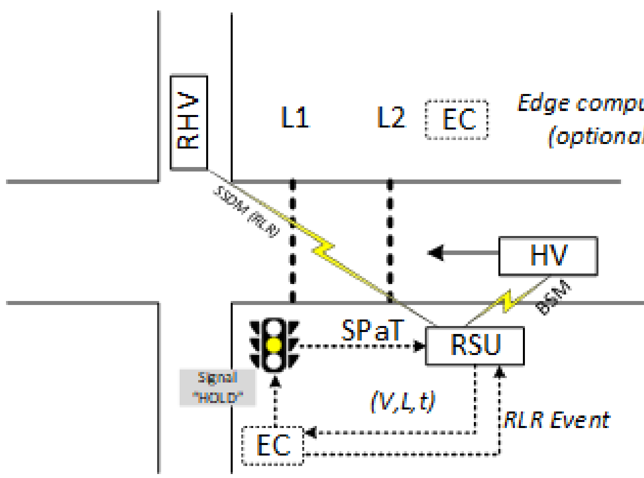


Figure 21. RLR warning with CP

Experiment plan

Experiment Environment: Both scenarios can be conducted on the MSU CAV testbed. For Scenario 1, the focus will be on vehicles approaching an intersection during the yellow phase, while Scenario 2 will simulate a permissive left-turn scenario where the CAV's view is obstructed by opposing left-turn vehicles.

Required Equipment: For Scenario 1, an approaching vehicle will simulate a potential red-light runner, while another vehicle on a conflicting movement will wait for the light to turn green. Both vehicles will be equipped with onboard sensors (LiDAR, camera, GPS) and V2X communication systems to track their movements. The RSU, equipped with LiDAR and cameras, will detect potential RLR events and broadcast warnings to the involved vehicles. For Scenario 2, a CAV will simulate a left-turning vehicle attempting a permissive left turn. Opposing through traffic will be simulated by another vehicle approaching the intersection, while an opposing left-turn vehicle will obstruct the CAV's line of sight. The RSU, equipped with sensors, will monitor the speed and position of the opposing traffic and transmit real-time data to the CAV via V2X communication.

Required Modules: The *Perception, Communication, and Fusion Modules,* as described above, will be needed. In Scenario 1, the RLR *Detection Module* will use kinematic equations to predict whether an approaching vehicle will run the red light, issuing warnings to both the red-light runner and conflicting vehicles. In Scenario 2, the *Left-Turn Assistance Module* will assess the position and speed of opposing through traffic and provide real-time recommendations to the CAV on whether it is safe to execute the left turn.

5.5. Enhanced Situational Awareness

Scenario setup and CP system architecture

For the enhanced situation awareness system, three scenarios can be designed. Scenario 1 focuses on using CP to navigate difficult geometric conditions, such as intersections where buildings obstruct drivers' line of sight. The CP prototype will leverage roadside sensors to track vehicles and pedestrians that may be obstructed. The system will generate BSM and PSM based on the real-time detection of objects and broadcast them through the RSU to approaching vehicles. Scenario 2 extends the application of CP to assist vehicles in navigating adverse weather conditions, such as heavy fog, rain, or snow, which impair visibility and sensor performance. Scenario 3 explores how CP can assist vehicles in navigating work zones. Vehicles driving through construction areas often face uncertainties such as lane closures and temporary road signs, which can be challenging for autonomous navigation. In this scenario, CAV1 will map the work zone in real-time or a RSU will broadcast the updated geo-fence. CAV2, which enters the work zone afterward, will use this data to navigate through the work zone more efficiently and safely.

Experiment plan

Experiment Environment: For Scenario 1, the experiment will be conducted at the MSU CAV Testbed. This intersection has bidirectional traffic with signal control, and the presence of buildings at the corners creates blind spots. The experiment will take place under clear, daytime conditions. A CAV will move toward the intersection while receiving real-time CP data about vehicles or pedestrians in the blind spot. For Scenario 2, the experiment will simulate adverse weather conditions, such as fog, rain, and snow. The experiment will take place under various weather and lighting conditions (daytime, night, dawn) and will involve multiple tests for each condition. In Scenario 3, the experiment will also be conducted at the MSU CAV Testbed. The experiment will last for 3 hours, with 6 tests focused on evaluating the effectiveness of the shared mapping data in supporting navigation through the work zone.

Required Equipment: All scenarios will require the support of two CAVs, each equipped with LiDAR, camera, and GPS sensors, and connected via 5G communication (SAE J2735 V2X). A roadside unit (RSU) equipped with cameras will be installed near the intersection to monitor and broadcast data on hidden vehicles and pedestrians.

Required Modules: For all scenarios, The *Perception, Communication, and Fusion Modules,* as described above, will be needed. In Scenario 1, the *Blind Spot Detection Module* will detect vehicles and pedestrians in obstructed areas. In Scenario 2, the *Adverse Weather Detection Module* will adjust the perception system based on weather-induced visibility degradation and ensure safe operation. In Scenario 3, the *Work Zone Mapping Module* will enable real-time mapping of the work zone or broadcasting of the updated geo-fence to guide CAV operations.

Chapter 6. Further Discussions: Quantifying the System Impact

Due to the limited scope and scale of the current project, comprehensive large-scale impact analysis was not feasible. Instead, this chapter proposes a structured framework and methodological guidance for how the impacts of the proposed CP-enabled pedestrian collision avoidance system could be rigorously quantified in future, expanded studies. While previous chapters focus on perception design, system integration, and field validation at a prototype level, this chapter outlines recommended approaches for evaluating safety, efficiency, scalability, and deployment feasibility when broader testing and higher penetration rates become available. By describing standards-aligned functional safety analysis, scenario-based evaluation strategies, and measurable performance metrics, this chapter provides a roadmap for translating CP technical capabilities into defensible, system-level impact assessments under scaled-up experimental or real-world deployment conditions.

6.1 Functional Safety Analysis

To quantify the safety impact of CP, this analysis follows ISO 26262 for functional safety and ISO/PAS 21448 for the safety of intended functionality (SOTIF). The goal is to evaluate CP's ability to reduce crash risks while considering factors such as market penetration levels, communication network demands, and varying traffic conditions. By identifying and classifying potential hazards using ISO 26262, such as communication disruptions or sensor fusion errors, each will be assigned an Automotive Safety Integrity Level (ASIL) based on severity, exposure, and controllability. High-ASIL hazards, like the failure to detect pedestrian, will be prioritized for mitigation to ensure safe system operation even during faults. The SOTIF analysis will focus on evaluating risks arising from CP's intended functionality, particularly in complex environments like urban areas or adverse weather. We will also apply standardized methods such as rating systems for automated vehicles in New Car Assessment Program (National Highway Traffic Safety Administration [NHTSA], 2022) and safety metrics for black, grey, and white-box automated vehicles specified in SAE J3237 on dynamic driving task assessment.

Given the limited scope of the current project, a full-scale quantitative safety validation under diverse traffic, weather, and penetration conditions was not feasible. Instead, this section proposes a structured functional safety evaluation framework that can guide future large-scale testing and deployment studies of CP-enabled pedestrian collision avoidance systems.

The proposed framework aligns with established safety standards, including ISO 26262 (Functional Safety) and ISO/PAS 21448 (Safety of the Intended Functionality, SOTIF). Under ISO 26262, future scaled studies should systematically identify potential malfunction-related hazards associated with CP integration, such as communication delays, packet loss, sensor misalignment, fusion errors, or incorrect object association. Each hazard can then be evaluated in terms of severity, exposure, and controllability to determine its Automotive Safety Integrity Level (ASIL). This structured classification enables prioritization of mitigation strategies, particularly for high-consequence scenarios such as failure to detect pedestrian under occlusion.

In parallel, ISO/PAS 21448 provides guidance for evaluating performance limitations that arise even when the system operates as intended. For CP systems, these limitations may include perception degradation in adverse weather, uncertainty propagation in multi-agent fusion, or reduced reliability under high channel congestion. Future large-scale evaluations should therefore assess how perception confidence, communication latency, and penetration rates influence overall system safety margins.

To translate these safety principles into measurable outcomes, scaled studies should incorporate quantitative performance indicators such as crash rate reduction, vehicle-to-pedestrian conflict frequency, Time-to-Collision (TTC) improvements, near-miss occurrences, and emergency braking activation rates. Controlled comparisons between baseline (non-CP) and CP-enabled configurations under equivalent traffic scenarios would provide defensible evidence of safety benefits. Additionally, sensitivity analysis across varying market penetration levels and communication reliability conditions would help determine how safety gains scale with deployment maturity.

6.2 Sampling method for scenario evaluation using importance sampling

Because safety-critical CP scenarios, such as severe occlusions, sudden pedestrian emergence or complex multi-agent conflicts, occur relatively infrequently in naturalistic driving data, comprehensive statistical validation would require extremely large exposure datasets. Therefore, this section proposes a methodological framework for how future expanded studies can efficiently evaluate CP performance in rare but high-risk scenarios.

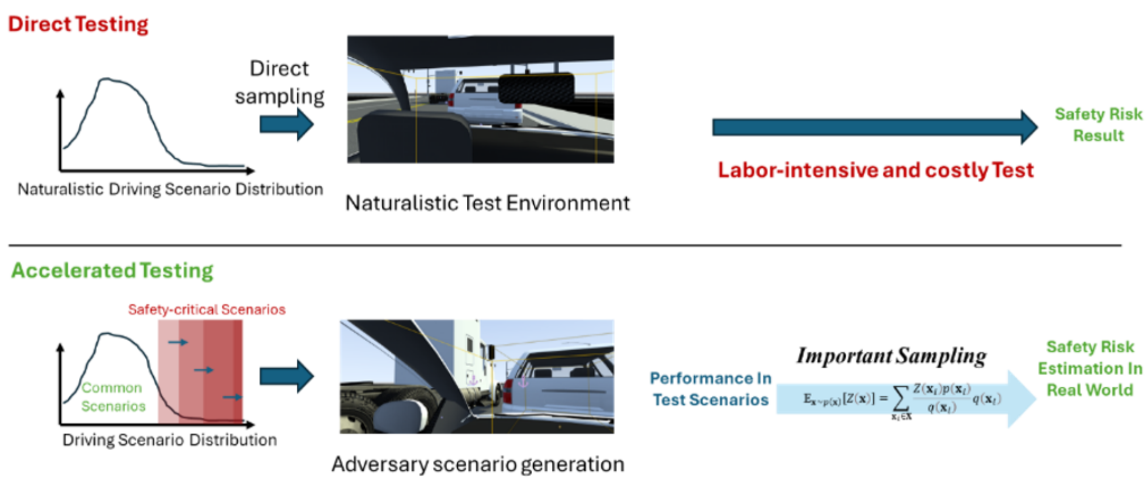


Figure 22. Importance of Sampling for Critical CP Scenarios

A recommended approach for scaled evaluation is the use of Importance Sampling (IS), as shown in Figure 22, to accelerate testing of critical corner cases. Importance Sampling increases the frequency of rare yet safety-relevant events, such as heavily occluded crosswalks, short time-to-collision encounters with pedestrian, or dense intersection conflicts, within simulation or controlled testing environments. By intentionally biasing the scenario distribution toward high-risk conditions, researchers can more efficiently estimate safety performance improvements attributable to CP.

In future large-scale studies, the accelerated testing process should follow three structured steps. First, identify scenario parameters that significantly influence collision risk, such as vehicle speed, pedestrian arrival timing, occlusion geometry, communication latency, and detection uncertainty. Second, modify the probability distribution of these parameters within simulation or hardware-in-the-loop (HIL) environments to over-represent safety-critical combinations. Third, apply statistical re-weighting to correct for sampling bias and recover unbiased real-world risk estimates.

To support this methodology, a hybrid evaluation platform combining real-world field testing, simulation (e.g., CARLA or SUMO), and hardware-in-the-loop integration is recommended. A digital twin environment can replicate infrastructure sensing, V2X communication delays, packet loss, and perception uncertainty, enabling systematic stress testing of CP under controlled but realistic conditions. This layered testing strategy allows researchers to explore scalability limits, communication congestion effects, and penetration-rate sensitivity without requiring prohibitively large naturalistic datasets.

By incorporating Importance Sampling within a structured simulation and HIL framework, future scaled studies can generate statistically meaningful safety impact estimates while maintaining practical testing feasibility. This approach provides a defensible pathway for quantifying how CP performs under rare yet consequential traffic conditions, thereby strengthening the evidence base for broader deployment.

6.3 Evaluation metrics

Safety metrics should remain the primary focus in future scaled evaluations. Recommended indicators include vehicle-to-vehicle crash rate, vehicle-to-pedestrians collision frequency, Time-to-Collision (TTC) improvements, near-miss frequency, and emergency braking activation rates. These metrics enable direct comparison between baseline (non-CP) and CP-enabled configurations under equivalent traffic scenarios. For statistically robust conclusions, future studies should evaluate these indicators across varying penetration rates, occlusion levels, communication latency conditions, and traffic densities.

Beyond crash reduction, efficiency metrics are necessary to assess how CP influences traffic operations. Because CP enhances situational awareness, particularly in occluded or dense environments, it may enable smoother vehicle maneuvers, reduced unnecessary deceleration, and improved intersection coordination. Suggested measures include average vehicle speed stability, intersection throughput, delay reduction, frequency of deadlock events in mutual-occlusion scenarios, and fuel consumption or energy efficiency. Evaluating these metrics at corridor or network scale would provide insight into how perception improvements translate into operational benefits.

Finally, cost and scalability metrics should be incorporated to understand deployment feasibility. These include communication bandwidth usage, computational load (e.g., early fusion versus late fusion strategies), latency performance, and system scalability under increasing numbers of connected agents. Because CP introduces additional communication and processing demands, quantifying these resource requirements is critical for realistic large-scale implementation planning. Table 2 summarizes the metrics used to evaluate these dimensions:

Table 2. Metrics for Evaluating the Impacts of CP systems

Impact Category	Metric	Description
Safety	Vehicle-to-Vehicle Crash rate	Compare collision rates before and after CP implementation.
	Vehicle-to-pedestrian collision frequency	Assess the reduction in accidents involving pedestrians or cyclists.
	Time to Collision (TTC) improvement	Measure improvements in TTC for potentially hazardous situations.
	Near-miss frequency	Count reductions in incidents where collisions are narrowly avoided.
	Emergency braking activations	Quantify the decrease in last-moment emergency braking events.
Efficiency	Average vehicle speed	Evaluate whether CP allows for consistent speeds, reducing unnecessary deceleration.
	Traffic throughput at intersections	Measure vehicle throughput in occluded or complex intersection scenarios.
	Frequency of deadlock occurrences	Count how often mutual occlusions cause deadlocks and the impact of CP in resolving them.
	Fuel consumption or energy efficiency	Assess energy savings through smoother driving enabled by CP.
Cost	Communication bandwidth usage	Measure the amount of data transmitted under different fusion strategies.
	Computational load	Evaluate the processing power required for early vs. late fusion.
	Scalability	Determine the system's ability to handle an increasing number of agents or data streams effectively.

6.4 Evaluation methods

A multi-layered evaluation strategy is needed to rigorously quantity the system-level impacts of CP. To do so, this study proposes a combination of simulation-based testing, HIL experimentation, vehicle-in-the-loop (VIL) validation, and controlled field trials. Each layer serves a distinct purpose in progressively increasing realism while maintaining controllability and scalability.

1. **Model-in-the-Loop (MIL) testing** can be used to evaluate CP logic under fully simulated traffic, perception, and communication environments. Traffic and driving simulators can

generate diverse scenarios, including varying traffic densities, pedestrian behaviors, occlusion geometries, and communication impairments (e.g., delay and packet loss). This stage enables rapid exploration of parameter sensitivity and penetration-rate effects without physical risk or resource constraints.

2. **Hardware-in-the-Loop (HIL) testing** introduces real communication devices and onboard units into the simulation framework. By incorporating actual C-V2X radios and message encoding/decoding modules, researchers can evaluate latency, packet loss, interoperability, and message-processing reliability under controlled but realistic communication conditions. This layer helps bridge the gap between purely virtual testing and physical deployment.
3. **Vehicle-in-the-Loop (VIL) testing** further increases realism by integrating a real test vehicle into a simulated traffic environment. In this configuration, the ego vehicle operates with its onboard sensors and control systems, while surrounding traffic agents and infrastructure components may be emulated. This approach allows validation of perception fusion, decision-making integration, and vehicle control responses when receiving cooperative perception messages.
4. **Controlled field testing** should be conducted to validate performance under natural environmental conditions. Such testing may focus on representative high-risk scenarios, such as occluded crosswalks or complex urban intersections. Field experiments enable measurement of real-world detection reliability, communication latency, and driver or system behavioral response.

Across all evaluation layers, standardized communication protocols (e.g., SAE J3224 Sensor Data Sharing Messages and related V2X message sets) should be validated to ensure correct encoding, transmission, and decoding among vehicles, infrastructure units, and pedestrian. Interoperability testing across heterogeneous devices is particularly important for scalable CP deployment.

Chapter 7. Conclusions and Future Research

7.1. Summary and Conclusions

This project investigated the development of a cooperative perception (CP)-enabled pedestrian collision avoidance framework for connected and automated vehicle (CAV) applications. The research was motivated by a persistent safety challenge in urban transportation: pedestrians are often exposed to high crash risk in complex roadway environments where vehicle onboard sensors alone may be insufficient due to occlusion, limited line-of-sight, and the unpredictability of pedestrian behavior. In response, this study proposed a CP-based approach that extends vehicular perception through infrastructure-assisted sensing and V2X communication, allowing approaching vehicles to receive early warnings about pedestrians that may otherwise remain undetected until it is too late to safely react.

The work first established the broader technical motivation for CP by reviewing the limitations of conventional onboard perception systems and the emerging role of distributed sensing, communication-efficient perception sharing, and infrastructure-supported safety applications. The literature synthesis confirmed that CP is especially valuable in urban and occlusion-heavy environments, where cameras, LiDAR, radar, roadside sensors, and communication networks can jointly improve situational awareness beyond the capability of any single agent. At the same time, the review also highlighted critical unresolved issues in communication load, interoperability, positioning accuracy, and the integration of externally received data into real-time vehicle control. These observations provided the foundation for the technical direction of this project.

A major contribution of the study was the development and validation of an object detection and multi-sensor fusion pipeline using a paired camera–LiDAR sensing stack. The results demonstrated that camera–LiDAR fusion provides a practical and effective perception baseline for CP deployment by combining strong semantic recognition from cameras with accurate geometric ranging from LiDAR. Compared with single-modality perception, the fused system produced more reliable pedestrian and vehicle detection and more stable object representations for downstream connected-vehicle applications. This is an important finding because any cooperative perception architecture ultimately depends on the quality and reliability of local perception at each sensing node.

Building on that perception foundation, the project then demonstrated the safety benefit of CP in a controlled pedestrian-occlusion scenario using a Level 4 CAV platform. The demonstrations showed three distinct cases: a baseline case in which the pedestrian was visible and the CAV could safely respond without CP, an occluded case in which CP enabled the CAV to stop safely by receiving infrastructure-assisted early warning, and a matched occluded case without CP in which the vehicle detected the pedestrian too late and collided with the dummy target. Together, these demonstrations provide direct and intuitive evidence of the core safety value of CP. Specifically, the project showed that infrastructure-assisted perception can restore reaction time in situations where onboard-only perception is fundamentally limited by occlusion.

Another important contribution of the study is the articulation of a realistic deployment pathway. Rather than treating the developed CP system as an isolated prototype, the report demonstrated that the Morgan State University SMART Corridor already provides the sensing, communication, tracking, analytics, and portable deployment capabilities required for long-term integration. The existing testbed supports LiDAR-based sensing, multi-object tracking, conflict analytics, pedestrian behavior analysis, and portable roadside sensing units, making it a natural environment for scaling the developed framework from proof-of-concept toward operational deployment. In this sense, the project not only validated a CP concept but also positioned that concept within a broader real-world implementation ecosystem.

Beyond the demonstrated pedestrian-occlusion case, the project also outlined a broader family of future CP applications, including pedestrian protection, wrong-way driving warning, sudden braking alerts, red-light-running warning, permissive left-turn assistance, blind-spot navigation, adverse-weather support, and work-zone assistance. These extensions suggest that the framework developed in this study is not limited to a single application but instead represents a flexible architecture that can support multiple infrastructure-assisted safety functions. This versatility strengthens the long-term significance of the project.

Overall, the findings of this report support the conclusion that CP is a promising and practical strategy for improving pedestrian safety in connected urban transportation systems. The study demonstrated that: (1) robust local perception can be achieved through camera–LiDAR fusion; (2) infrastructure-assisted sensing can successfully compensate for onboard perception blind spots; (3) CP can materially improve vehicle response in occluded pedestrian conflict scenarios; and (4) the MSU SMART Corridor provides a viable path for long-term integration and expanded evaluation. Collectively, these results indicate that CP has strong potential to become an important enabling technology for safer, smarter, and more resilient connected and automated transportation systems.

7.2. Limitations and Future Research Directions

Although the project achieved its primary proof-of-concept goals, several limitations should be acknowledged. First, the field demonstrations were conducted at a relatively small scale and under controlled conditions. The pedestrian-crossing scenarios relied on a dummy pedestrian and a simplified test environment, which was appropriate for validating the core concept but does not fully reflect the behavioral variability, environmental complexity, and uncertainty of real-world urban traffic. Human pedestrians may exhibit more diverse movement patterns, hesitation behavior, distraction, and unexpected acceleration or direction changes that are difficult to fully replicate using a movable surrogate.

Second, the demonstration emphasized a specific occlusion-driven pedestrian safety use case. While this use case is both important and representative of a major weakness in onboard-only autonomy, it captures only one subset of the broader operational design domain in which CP may be beneficial. Additional work is needed to evaluate how the framework performs across different roadway geometries, traffic densities, signalized and unsignalized intersections, mid-block crossings, adverse weather conditions, nighttime visibility, and mixed traffic environments involving cyclists, buses, delivery vehicles, and non-connected vehicles.

Third, the current project did not include a large-scale quantitative impact assessment. As discussed in Chapter 6, the present effort was not designed to estimate population-level crash reduction, communication scalability, or corridor-level operational benefits. Comprehensive evaluation of such impacts would require significantly expanded testing across many scenarios, repeated trials, and multiple market penetration assumptions. Similarly, the project did not yet quantify several important deployment metrics in a statistically rigorous way, such as false alarm rate, missed detection rate under varying conditions, communication reliability under congestion, packet loss effects, and end-to-end latency distribution under larger connected fleets.

Fourth, although the CP concept was demonstrated successfully, full closed-loop integration with long-term real-world deployment infrastructure remains a future step. The current study established the feasibility of infrastructure-assisted detection and communication, but long-term deployment will require additional systems engineering related to interface standardization, synchronization, calibration maintenance, message formatting, fault tolerance, and operational reliability. In particular, multi-agent perception systems are sensitive to temporal misalignment, spatial calibration drift, and localization uncertainty. These issues become more challenging as the number of sensors, vehicles, and infrastructure nodes increases.

Fifth, the communication and interoperability aspects of the framework remain an important open challenge. Real-world CP deployment must operate within bandwidth constraints, standardized message structures, and heterogeneous equipment ecosystems. Future implementation must further examine how object-level or feature-level messages should be encoded, prioritized, and transmitted under varying traffic and congestion conditions. The system must also be evaluated for interoperability across different RSUs, OBUs, and vehicle platforms, especially when proprietary implementations differ from nominal SAE standards.

Future research should therefore proceed along several directions. One priority is large-scale field validation on the MSU SMART Corridor. The phased integration plan proposed in Chapter 4 provides a clear starting point: future studies should package the developed CP module for real-time deployment on the corridor, instrument one or more high-conflict locations using the portable detection and communication system, and conduct extended testing under diverse lighting, traffic, and weather conditions. Such testing would allow evaluation of long-term reliability, maintenance needs, detection persistence, and communication performance under realistic operating conditions.

A second major direction is expanded scenario testing. Future work should evaluate the framework across the broader set of scenarios described in Chapter 5, including jaywalking detection, blind-spot pedestrian detection around buses, wrong-way driving warning, sudden-braking alerts, red-light-running prediction, and permissive left-turn assistance. This will help determine whether the same sensing, fusion, and communication architecture can serve as a generalizable cooperative safety platform rather than a scenario-specific prototype.

A third important direction is the incorporation of rigorous quantitative evaluation methods for rare but high-risk events. As outlined in Chapter 6, future research should combine field testing with simulation, hardware-in-the-loop, and vehicle-in-the-loop experiments. Importance sampling and digital-twin-based scenario generation can be used to accelerate testing of severe but infrequent corner cases, such as short time-to-collision pedestrian emergence, dense urban occlusions, or high-latency communication conditions. This will make it possible to estimate safety gains more efficiently and with greater statistical confidence.

A fourth direction is improvement of communication-aware CP design. Future studies should examine how to optimize the tradeoff among message content, communication load, latency, and perception benefit. This includes comparing raw-, feature-, and decision-level sharing strategies; developing message prioritization schemes for safety-critical objects; and investigating how congestion control, confidence-aware transmission, and adaptive communication policies affect system performance. Such work is essential for scalable deployment beyond small experimental fleets.

A fifth direction involves uncertainty-aware fusion and control integration. While this project demonstrated the operational value of CP, future systems should more explicitly model uncertainty in detection, object association, communication delay, and localization. Confidence-aware fusion and downstream decision logic will be necessary to ensure that vehicles respond appropriately to external information without overreacting to noise or underutilizing useful cooperative inputs. This is particularly important for closed-loop automated control, where external perception must be integrated safely into trajectory planning and motion control.

Finally, future work should address human factors, operational policy, and deployment governance. As CP systems begin to influence real vehicle behavior in mixed traffic, it will become important to study alert strategies, trust calibration, false alarm tolerance, driver or operator acceptance, and the policy implications of infrastructure-mediated decision support. These issues are especially relevant for transitional deployment environments in which connected and automated vehicles coexist with human-driven vehicles and pedestrians.

In summary, this project established a strong technical foundation for CP-enabled pedestrian collision avoidance, but it also represents an early step in a broader research and deployment trajectory. Future work should build upon the demonstrated concept by expanding scenario coverage, scaling evaluation, strengthening communication and fusion robustness, and advancing real-world deployment on the MSU SMART Corridor. Through these efforts, the developed framework can evolve from a successful pilot demonstration into a mature infrastructure-assisted safety system with measurable benefits for pedestrian and the broader transportation network.

References

Aoki, S., Higuchi, T., & Altintas, O. (2020). Cooperative Perception with Deep Reinforcement Learning for Connected Vehicles. 2020 IEEE Intelligent Vehicles Symposium (IV), 328–334. https://doi.org/10.1109/IV47402.2020.9304570

Arnold, E., Dianati, M., de Temple, R., & Fallah, S. (2022). Cooperative Perception for 3D Object Detection in Driving Scenarios Using Infrastructure Sensors. IEEE Transactions on Intelligent Transportation Systems, 23(3), 1852–1864. https://doi.org/10.1109/TITS.2020.3028424

Baek, M., Jeong, D., Choi, D., & Lee, S. (2020). Vehicle Trajectory Prediction and Collision Warning via Fusion of Multisensors and Wireless Vehicular Communications. Sensors, 20(1), 288. https://doi.org/10.3390/s20010288

Bai, X., Hu, Z., Zhu, X., Huang, Q., Chen, Y., Fu, H., & Tai, C.-L. (2022). TransFusion: Robust LiDAR-Camera Fusion for 3D Object Detection with Transformers. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://arxiv.org/abs/2203.11496

Chen, C., Tang, Q., Hu, X., & Huang, Z. (2023). Infrastructure sensor-based cooperative perception for early stage connected and automated vehicle deployment. Journal of Intelligent Transportation Systems, 28(6), 956–970. https://doi.org/10.1080/15472450.2023.2257596

Chen, Q., Ma, X., Tang, S., Guo, J., Yang, Q., & Fu, S. (2019b). F-COOPER: Feature-based cooperative perception for autonomous vehicle edge computing using 3D point clouds. ACM. https://doi.org/10.1145/3318216.3363300

Chen, Q., Tang, S., Yang, Q., & Fu, S. (2019a). COOPER: Cooperative Perception for Connected Autonomous Vehicles based on 3D Point Clouds. 2019 IEEE ICDCS (accepted).

Chen, X., Ma, H., Wan, J., Li, B., & Xia, T. (2017). Multi-View 3D Object Detection Network for Autonomous Driving. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 6526–6534. https://doi.org/10.1109/CVPR.2017.691

Chen, Y., Zheng, L., & Tan, Z. (2024). Roadside LiDAR placement for cooperative traffic detection by a chance-constrained stochastic simulation optimization approach. Transportation Research Part C, 167, 104838. https://doi.org/10.1016/j.trc.2024.104838

Chtourou, A., Merdrignac, P., & Shagdar, O. (2021). Context-Aware Content Selection and Message Generation for Collective Perception Services. Electronics, 10(20), 2509. https://doi.org/10.3390/electronics10202509

Cui, G., Zhang, W., Xiao, Y., Yao, L., & Fang, Z. (2022). Cooperative Perception Technology of Autonomous Driving in the Internet of Vehicles Environment: A Review. Sensors, 22(15), 5535. https://doi.org/10.3390/s22155535

Delooz, Q., et al. (2022). Analysis and Evaluation of Information Redundancy Mitigation for V2X Collective Perception. IEEE Access, 10, 47076–47093. https://doi.org/10.1109/ACCESS.2022.3170029

Duan, X., Jiang, H., Tian, D., Zou, T., Zhou, J., & Cao, Y. (2021). V2I-based environment perception for autonomous vehicles at intersections. China Communications, 18(7), 1–12. https://doi.org/10.23919/JCC.2021.07.001

Figueiredo, A., Amaral, J., Mendes, M., Rosmaninho, R., Dias, D., Rito, P., … Sargento, S. (2025). Experimental dataset of video and radar detection for cooperative perception in urban environment. Data in Brief, 63, 112091. https://doi.org/10.1016/j.dib.2025.112091

Fukatsu, R., & Sakaguchi, K. (2021). Automated Driving with Cooperative Perception Using Millimeter-Wave V2V Communications for Safe Overtaking. Sensors, 21(8), 2659. https://doi.org/10.3390/s21082659

Hakim, B., Sorour, S., Hefeida, M. S., Alasmary, W. S., & Almotairi, K. H. (2023). CCPAV: Centralized cooperative perception for autonomous vehicles using CV2X. Ad Hoc Networks, 142, 103101. https://doi.org/10.1016/j.adhoc.2023.103101

Hejazi, H., & Bokor, L. (2024). Modeling and evaluation of cooperative Vulnerable Road User protection schemes in realistic C-ITS environments. Computer Networks, 246, 110396. https://doi.org/10.1016/j.comnet.2024.110396

Higuchi, T., Giordani, M., Zanella, A., Zorzi, M., & Altintas, O. (2019). Value-Anticipating V2V Communications for Cooperative Perception. 2019 IEEE Intelligent Vehicles Symposium (IV), 1947–1952. https://doi.org/10.1109/IVS.2019.8814110

Huang, H., Fang, W., & Li, H. (2019). Performance Modelling of V2V-based Collective Perceptions in Connected and Autonomous Vehicles. 2019 IEEE 44th Conference on Local Computer Networks (LCN), 356–363. https://doi.org/10.1109/LCN44214.2019.8990854

Huang, H., Li, H., Shao, C., Sun, T., Fang, W., & Dang, S. (2020). Data Redundancy Mitigation in V2X-Based Collective Perceptions. IEEE Access, 8, 13405–13418. https://doi.org/10.1109/ACCESS.2020.2965552

Huang, T., Liu, Z., Chen, X., & Bai, X. (2020). EPNet: Enhancing Point Features with Image Semantics for 3D Object Detection. Computer Vision – ECCV 2020, 583–600. https://doi.org/10.1007/978-3-030-58555-6_3

Ji, A., Huang, J., Qin, Z., Sun, Z., Zhao, R., & Zheng, G. (2025). Cooperative merging for connected automated vehicles in mixed traffic: A multi-agent reinforcement learning approach. Artificial Intelligence for Transportation, 2, 100007. https://doi.org/10.1016/j.ait.2025.100007

Ji, Y., Zhou, Z., Yang, Z., Huang, Y., Zhang, Y., Zhang, W., … Yu, Z. (2024). Toward autonomous vehicles: A survey on cooperative vehicle-infrastructure system. iScience, 27(5), 109751. https://doi.org/10.1016/j.isci.2024.109751

Ku, J., Mozifian, M., Lee, J., Harakeh, A., & Waslander, S. L. (2018). Joint 3D Proposal Generation and Object Detection from View Aggregation. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 1–8. https://doi.org/10.1109/IROS.2018.8594049

Lee, M. B., Lee, C. T., Abas, M. A., & Chong, W. W. F. (2025). Advancing pedestrian safety in the era of autonomous vehicles: A bibliometric analysis and pathway to effective regulations. Journal of Traffic & Transportation Engineering, 12(4), 772–794. https://doi.org/10.1016/j.jtte.2024.05.004

Li, S., & Wolff, V. A. (2022). Tracking Accuracy-Based Generation Rules of Collective Perception Messages. 2022 IEEE International Conference on Intelligent Transportation Systems (ITSC), 4157–4162. https://doi.org/10.1109/ITSC55140.2022.9922147

Li, T., Han, X., & Ma, J. (2022). Cooperative Perception for Estimating and Predicting Microscopic Traffic States to Manage Connected and Automated Traffic. IEEE Transactions on Intelligent Transportation Systems, 23(8), 13694–13707. https://doi.org/10.1109/TITS.2021.3126621

Li, W., Zhu, T., & Feng, Y. (2024). A cooperative perception-based adaptive signal control under early CAV deployment. Transportation Research Part C, 169, 104860. https://doi.org/10.1016/j.trc.2024.104860

Lin, C., Tian, D., Duan, X., Zhou, J., Zhao, D., & Cao, D. (2024). V2VFormer: Vehicle-to-Vehicle Cooperative Perception With Spatial-Channel Transformer. IEEE Transactions on Intelligent Vehicles, 9(2), 3384–3395. https://doi.org/10.1109/TIV.2024.3353254

Liu, Z., Tang, H., Amini, A., Yang, X., Mao, H., Rus, D., & Han, S. (2023). BEVFusion: Multi-Task Multi-Sensor Fusion with Unified Bird's-Eye View Representation. 2023 IEEE International Conference on Robotics and Automation (ICRA). https://doi.org/10.1109/ICRA48891.2023.10160968

Luo, J., et al. (2023). Real-Time Cooperative Vehicle Coordination at Unsignalized Road Intersections. IEEE Transactions on Intelligent Transportation Systems, 24(5), 5390–5405. https://doi.org/10.1109/TITS.2023.3243940

Marti, E., de Miguel, M. A., Garcia, F., & Perez, J. (2019). A Review of Sensor Technologies for Perception in Automated Driving. IEEE Intelligent Transportation Systems Magazine, 11(4), 94–108. https://doi.org/10.1109/MITS.2019.2907630

Masuda, H., Marai, O. E., Tsukada, M., Taleb, T., & Esaki, H. (2023). Feature-Based Vehicle Identification for Optimization of CPMs. IEEE Transactions on Vehicular Technology, 72(2), 2120–2129. https://doi.org/10.1109/TVT.2022.3211852

National Highway Traffic Safety Administration. (2022, March 9). New Car Assessment Program. Federal Register, 87, 13452–13521. https://www.federalregister.gov/documents/2022/03/09/2022-04894/new-car-assessment-program

Nguyen, Q.-H., Morold, M., David, K., & Dressler, F. (2020). Car-to-Pedestrian Communication with MEC-Support for Adaptive Safety of VRUs. Computer Communications, 150, 83–93. https://doi.org/10.1016/j.comcom.2019.10.033

O'Malley, R., Glavin, M., & Jones, E. (2008). A review of automotive infrared pedestrian detection techniques. IET ISSC 2008, 168–173. https://doi.org/10.1049/cp:20080657

Park, J., Kavathekar, V., Bhuduri, S., Amin, M. H., & Devaraj, S. S. (2025). A Co-Operative Perception System for Collision Avoidance Using C-V2X and Client–Server-Based Object Detection. Sensors, 25(17), 5544. https://doi.org/10.3390/s25175544

Prathiba, S. B., Raja, S. K. S., & Kumar, B. V. (2022). Cooperative Collision Avoidance Using Inverse Reinforcement Learning for Overtaking and Lane Changing. 2022 IEEE International Conference on Intelligent Transportation Systems (ITSC), 2323–2328. https://doi.org/10.1109/ITSC55140.2022.9921819

Qi, C. R., Liu, W., Wu, C., Su, H., & Guibas, L. J. (2018). Frustum PointNets for 3D Object Detection from RGB-D Data. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 918–927. https://doi.org/10.1109/CVPR.2018.00102

ROAD Dataset. (n.d.). ROad event Awareness Dataset (ROAD) for Autonomous Driving. https://github.com/gurkirt/road-dataset

Rosique, F., Navarro, P. J., Fernández, C., & Padilla, A. (2019). A Systematic Review of Perception System and Simulators for Autonomous Vehicles Research. Sensors, 19(3), 648. https://doi.org/10.3390/s19030648

Sarlak, A., Amin, R., & Razi, A. (2025). Extended visibility of autonomous vehicles via optimized cooperative perception under imperfect communication. Transportation Research Part C, 180, 105350. https://doi.org/10.1016/j.trc.2025.105350

Schiegg, F. A., Llatser, I., Bischoff, D., & Volk, G. (2021). Collective Perception: A Safety Perspective. Sensors, 21(1), 159. https://doi.org/10.3390/s21010159

Shan, M., Narula, K., Worrall, S., Wong, Y. F., Berrio Perez, J. S., Gray, P., & Nebot, E. (2022). A Novel Probabilistic V2X Data Fusion Framework for Cooperative Perception. 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), 2013–2020. https://doi.org/10.1109/ITSC55140.2022.9922251

Sheng, Y., Ye, H., Liang, L., Jin, S., & Li, G. Y. (2024). Semantic communication for cooperative perception based on importance map. Journal of the Franklin Institute, 361(6), 106739. https://doi.org/10.1016/j.jfranklin.2024.106739

Tan, X., Wang, R., Wang, J., Wang, S., Wang, X., & Wu, D. (2026). Confidence-V2X: Confidence-driven sparse communication for efficient V2X cooperative perception. Advanced Engineering Informatics, 69(B), 103914. https://doi.org/10.1016/j.aei.2025.103914

Thandavarayan, G., Sepulcre, M., & Gozalvez, J. (2019). Analysis of Message Generation Rules for Collective Perception. 2019 IEEE Intelligent Vehicles Symposium (IV), 134–139. https://doi.org/10.1109/IVS.2019.8813806

Thandavarayan, G., Sepulcre, M., & Gozalvez, J. (2020a). Generation of Cooperative Perception Messages for CAVs. IEEE Transactions on Vehicular Technology, 69(12), 16336–16341. https://doi.org/10.1109/TVT.2020.3036165

Thandavarayan, G., Sepulcre, M., & Gozalvez, J. (2020b). Cooperative Perception for CAVs: Evaluation and Impact of Congestion Control. IEEE Access, 8, 197665–197683. https://doi.org/10.1109/ACCESS.2020.3035119

Thandavarayan, G., Sepulcre, M., Gozalvez, J., & Coll-Perales, B. (2023). Scalable cooperative perception for connected and automated driving. Journal of Network and Computer Applications, 216, 103655. https://doi.org/10.1016/j.jnca.2023.103655

Van Brummelen, J., O'Brien, M., Gruyer, D., & Najjaran, H. (2018). Autonomous vehicle perception: The technology of today and tomorrow. Transportation Research Part C, 89, 384–406. https://doi.org/10.1016/j.trc.2018.02.012

Vora, S., Lang, A. H., Helou, B., & Beijbom, O. (2020). PointPainting: Sequential Fusion for 3D Object Detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 4603–4611. https://doi.org/10.1109/CVPR42600.2020.00466

Wang, T., et al. (2023). UMC: A Unified Bandwidth-efficient and Multi-resolution based Collaborative Perception Framework. ICCV 2023, 8153–8162. https://doi.org/10.1109/ICCV51070.2023.00752

Xia, B., Zhou, J., Kong, F., You, Y., Yang, J., & Lin, L. (2024). Enhancing 3D object detection through multi-modal fusion for cooperative perception. Alexandria Engineering Journal, 104, 46–55. https://doi.org/10.1016/j.aej.2024.06.025

Xiang, H., Zheng, Z., Xia, X., Xu, R., Gao, L., Zhou, Z., … Han, X. (2024). V2X-Real: A Large-Scale Dataset for V2X Cooperative Perception. ECCV 2024, 455–470. https://arxiv.org/abs/2403.16034

Xu, R., Xiang, H., Tu, Z., Xia, X., Yang, M.-H., & Ma, J. (2022). V2X-ViT: Vehicle-to-Everything Cooperative Perception with Vision Transformer. ECCV 2022, LNCS 13699. https://doi.org/10.1007/978-3-031-19842-7_7

Xu, T., Jiang, R., Wen, C., Liu, M., & Zhou, J. (2019). A hybrid model for lane change prediction with V2X-based driver assistance. Physica A, 534, 122033. https://doi.org/10.1016/j.physa.2019.122033

Yoo, J. H., Kim, Y., Kim, J., & Choi, J. W. (2020). 3D-CVF: Generating Joint Camera and LiDAR Features Using Cross-View Spatial Feature Fusion for 3D Object Detection. Computer Vision – ECCV 2020, 720–736. https://doi.org/10.1007/978-3-030-58583-9_43

Zhang, Y., Carballo, A., Yang, H., & Takeda, K. (2023). Perception and sensing for autonomous vehicles under adverse weather conditions: A survey. ISPRS Journal of Photogrammetry and Remote Sensing, 196, 146–177. https://doi.org/10.1016/j.isprsjprs.2022.12.021