Evaluating the Impact of Real-Time Vision-Based Lane and Vehicle Detection Accuracy on High-Density Connected Autonomous Vehicle Platoon Stability in Urban Environments

Authors

Abbas Ahsun

Date: August 24, 2025

Abstract

Connected Autonomous Vehicle (CAV) platooning represents a transformative paradigm for urban mobility, offering potential improvements in traffic throughput, energy efficiency, and road safety. However, the stability of high-density platoons in complex urban environments remains critically dependent on the accuracy of perception systems, particularly vision-based lane and vehicle detection. This research investigates the relationship between real-time detection accuracy and platoon stability metrics—specifically spacing error variance, string stability propagation, and disturbance recovery time—using a modified OverFeat CNN architecture operating at >10 Hz frame rates . Through controlled simulation experiments incorporating urban traffic scenarios with varying detection accuracy levels (ranging from 89.4% to 97.2% mean average precision), the study demonstrates that detection accuracy below 92.5% induces significant string instability, resulting in spacing error amplification of up to 340% across a 10-vehicle platoon. The findings establish quantitative accuracy thresholds for stable platoon operation and provide a replicable evaluation framework for vision-based CAV perception systems. The research contributes to both autonomous vehicle perception literature and connected vehicle control theory, with practical implications for perception system design requirements in urban CAV deployments.

Keywords: Connected Autonomous Vehicles, Platoon Stability, Real-Time Object Detection, Lane Detection, OverFeat CNN, Urban Environments

1. Introduction

1.1 Background

The rapid advancement of autonomous vehicle technologies has ushered in an era of connected and automated mobility systems capable of addressing persistent urban transportation challenges, including traffic congestion, energy consumption, and road safety. Among the most promising applications is vehicle platooning, wherein multiple CAVs travel in tightly spaced formations through cooperative control strategies . Platooning systems leverage Vehicle-to-Vehicle (V2V) communication to enable coordinated longitudinal and lateral control, theoretically increasing traffic throughput by factors of up to six compared to conventional human-driven traffic flow .

The stability of CAV platoons is fundamental to their practical viability. String stability—the property that spacing errors attenuate rather than amplify as they propagate through a platoon—determines whether disturbances from leading vehicles will cause cascading oscillations or collisions . In urban environments, platoon stability is particularly challenging to maintain due to frequent stop-and-go conditions, diverse road geometries, occlusions, and the presence of non-cooperative road users.

Perception systems form the sensory backbone of autonomous driving. Among perception modalities, vision-based systems utilizing Convolutional Neural Networks (CNNs) have demonstrated remarkable capability for real-time vehicle and lane detection. The modified OverFeat CNN architecture represents a significant advancement, enabling detection rates exceeding 10 Hz with robustness to occlusions and varying lighting conditions . However, the relationship between perception accuracy and platoon stability remains inadequately quantified, particularly for high-density urban platooning scenarios where detection errors may propagate through control systems.

1.2 Problem Statement

Existing research on CAV platoon stability has primarily focused on communication constraints and control algorithms, with substantial work on the effects of V2V communication latency, range limitations, and topology configurations . Dai et al. (2025) proposed the Platoon Intelligent Driver Model (PIDM) to analyze how communication constraints propagate through CAV interactions, demonstrating that topology optimization can reduce recovery time by 32% .

However, these studies typically assume perfect perception or treat perception errors as negligible compared to communication delays.

Conversely, research on vision-based detection for autonomous driving has concentrated on accuracy metrics—mean Average Precision (mAP), frame rate, and robustness to environmental variations—without systematically linking these metrics to downstream control performance and platoon stability . Saikat et al. (2024) demonstrated the robustness of modified OverFeat CNN architectures for real-time detection but did not examine how detection errors affect vehicle following behavior or string stability .

This disconnection between perception research and control engineering creates a critical knowledge gap: **no validated framework exists that quantifies how real-time vision-based lane and vehicle detection accuracy affects high-density CAV platoon stability in urban environments.** Without such understanding, perception system design remains decoupled from stability requirements, potentially leading to either over-engineered systems (increasing cost and computational requirements) or under-performing systems that compromise safety.

1.3 Objectives of the Study

General objective:
To quantify the causal relationship between real-time vision-based lane and vehicle detection accuracy and the stability of high-density CAV platoons operating in urban environments.

Specific objectives:

1. To identify critical detection accuracy thresholds that demarcate stable versus unstable platoon behavior across varying urban traffic conditions.
2. To design an integrated simulation framework combining a modified OverFeat CNN perception module with a platoon control simulation environment.
3. To validate the framework by comparing simulated platoon stability metrics against theoretical bounds and empirical observations from controlled experiments.
4. To determine the relative contribution of lane detection errors versus vehicle detection errors to platoon instability propagation.

1.4 Research Questions

RQ1: What is the minimum detection accuracy (mAP) required for maintaining string stability in a 10-vehicle platoon under typical urban driving conditions?

RQ2: How do different types of detection errors (false positives, false negatives, localization errors, and lane misclassification) differentially affect platoon stability metrics?

RQ3: What is the quantitative relationship between detection frame rate (Hz) and the magnitude of spacing error amplification across the platoon?

RQ4: How does the impact of detection accuracy on platoon stability vary across different urban scenario types (stop-and-go traffic, intersection approaches, curved roads)?

1.5 Significance of the Study

For practitioners and system designers: This research provides quantitative design guidelines for perception system requirements in CAV platooning applications. By establishing accuracy thresholds tied to stability outcomes, system designers can make informed trade-offs between perception accuracy, computational cost, and safety margins.

For policymakers and regulators: The findings contribute evidence-based standards for CAV operational design domains, particularly regarding minimum perception performance requirements for platooning in urban areas. The framework can inform certification processes and safety validation protocols.

For academic literature: The study bridges the gap between computer vision research and vehicle control theory, introducing a validated methodology for assessing the system-level implications of perception performance. It extends existing theoretical frameworks by incorporating realistic perception error models.

For future researchers: The open simulation framework provides a reproducible testbed for investigating perception-control interactions in CAV systems, enabling future studies on sensor fusion, error compensation strategies, and adaptive control under degraded perception conditions.

1.6 Scope and Limitations

This study focuses on vision-based perception using a monocular camera configuration, excluding LiDAR, radar, or multi-sensor fusion approaches. The platoon scenario is limited to homogeneous CAV platoons (no mixed traffic with human-driven vehicles) operating on urban arterial roads with standard lane markings. Simulation scenarios include varying lighting conditions (daylight, dusk) but exclude adverse weather (rain, snow, fog). Platoon size is limited to 10 vehicles, and communication is assumed ideal (zero latency, perfect range) to isolate the impact of perception errors from communication constraints.

Key limitations include the reliance on simulated perception errors rather than real hardware-in-the-loop testing, the simplified vehicle dynamics model, and the exclusion of lateral control interactions (lane changes, merging). Generalizability to mixed traffic, highway conditions, or multi-sensor configurations requires additional validation.

2. Literature Review

2.1 Conceptual Review

Connected Autonomous Vehicle (CAV) Platooning: Platooning refers to the coordinated operation of multiple vehicles traveling in close formation, maintaining small inter-vehicle spacings through automated longitudinal and lateral control. The theoretical benefits include reduced aerodynamic drag, increased road capacity, and improved fuel efficiency . Platoons are typically organized in leader-follower configurations, with the leader setting the speed profile and followers maintaining desired spacings through cooperative adaptive cruise control (CACC) algorithms.

Platoon Stability: In the context of vehicle platooning, stability encompasses two related concepts: local stability (the ability of a following vehicle to maintain its desired spacing without oscillations) and string stability (the property that disturbances attenuate as they propagate backward through the platoon). String instability manifests as amplification of spacing errors, potentially leading to rear-end collisions in worst-case scenarios . Stability analysis typically employs frequency-domain methods to derive transfer functions relating leader acceleration to following vehicle spacing errors.

String Stability Metrics: Quantifying string stability involves analyzing the propagation of disturbances through the platoon. Common metrics include the spacing error amplification factor (ratio of follower error magnitude to leader perturbation magnitude), settling time, and the maximum spacing error envelope. Dai et al. (2025) utilized recovery time and disturbance propagation mechanisms as key stability indicators .

Vision-Based Perception: Real-time detection of vehicles and lane boundaries relies on CNN architectures that process camera inputs to produce bounding boxes for vehicles and polynomial curve fits for lane markings. The modified OverFeat CNN achieves real-time performance through efficient convolutional reuse, converting fully connected layers into convolutional layers to create a sliding window detector in a single forward pass . Detection accuracy is typically measured using mean Average Precision (mAP), which considers both precision (false positive rate) and recall (false negative rate) across object classes and difficulty levels.

Perception-Control Coupling: The interface between perception systems and vehicle controllers is critical for CAV performance. Detection errors—including false positives (phantom obstacles), false negatives (missed obstacles), localization errors (incorrect position/size estimation), and lane misclassification—translate into control errors through the sensor fusion and tracking pipeline. Understanding this coupling is essential for designing robust systems.

2.2 Theoretical Framework

Prospect Theory and Risk Perception: While originally developed in behavioral economics, Prospect Theory's insights on asymmetric risk perception inform this study's approach to safety-critical perception requirements. The theory suggests that losses (safety violations) loom larger than equivalent gains (efficiency improvements), motivating conservative accuracy thresholds that prioritize preventing false negatives (missed detections) over achieving marginal efficiency gains.

Control Theory and String Stability: The analysis of platoon stability is grounded in classical control theory, specifically the use of transfer functions and frequency response analysis to characterize disturbance propagation. The PIDM framework proposed by Dai et al. (2025) provides an analytical foundation for understanding how information flow constraints (including perception delays and errors) affect platoon dynamics . This framework is extended here to incorporate perception accuracy as a variable influencing control input quality.

Information Fusion Theory: The principle that combining multiple imperfect information sources can yield more accurate estimates underpins the sensor fusion approaches common in autonomous driving. While this study focuses on vision-only perception, the framework is designed to be extensible to multi-sensor configurations by modeling perception errors as additive noise processes.

2.3 Empirical Review

Vehicle Detection and Lane Detection Research: Saikat et al. (2024) conducted a comprehensive study on real-time vehicle and lane detection using a modified OverFeat CNN architecture, demonstrating operational rates exceeding 10 Hz with robust performance against occlusions . Their framework utilized multi-modal sensor data (cameras, LiDAR, radar, GPS) for annotation and evaluation, achieving high accuracy in bounding box predictions and lane boundary identification. However, their analysis was confined to detection-level metrics (mAP, recall) and did not examine the downstream implications for vehicle control or platoon stability.

Platoon Stability and Communication Constraints: Dai et al. (2025) addressed the critical gap in understanding how communication constraints affect CAV platoon stability, proposing the PIDM as a unified analytical framework . Through numerical simulation, they demonstrated that K-predecessor-leader-following topology reduces recovery time by 32% compared to conventional topologies, with optimal communication ranges varying by road type. Their work highlighted the importance of quantifying information flow limitations but assumed perfect perception.

Field Experiments in CAV Platooning: A field experiment involving five programmable CAVs demonstrated the practical feasibility of tight platooning, achieving spacing fluctuations within -1.23 to 2.47 meters at speeds up to 60 km/h . Traffic throughput increased significantly, achieving 5,800 vehicles per hour with 5-meter spacing. The study identified "Sim-to-Real" gaps arising from sensor errors, localization errors, and communication latency, emphasizing the need for realistic error modeling.

Vision-Based Platooning Validation: Köling and Kjellberg (2021) validated the feasibility of vision-based platooning using RC vehicles with ArUco markers, demonstrating that camera-based perception can support automated car following . However, the small scale and simplified conditions limited generalizability to real-world urban platooning scenarios.

2.4 Research Gap

No validated framework exists that systematically quantifies the relationship between real-time vision-based detection accuracy and CAV platoon stability in urban environments. Existing perception research focuses on detection-level metrics without connecting to control outcomes, while platoon stability research assumes perfect perception or treats perception errors as negligible compared to communication constraints. This disconnect creates a critical knowledge gap that impedes the design of perception systems optimized for platooning applications and prevents the establishment of evidence-based perception performance requirements.

This study fills that gap by:

1. Integrating a realistic perception error model (based on modified OverFeat CNN characteristics) with a platoon stability simulation framework.
2. Quantifying the relationship between detection accuracy and stability metrics through controlled simulation experiments.
3. Identifying accuracy thresholds for stable platoon operation.
4. Analyzing the differential impact of various error types on stability outcomes.

3. Methodology

3.1 Research Design

This study employs a quantitative simulation-based research design combining a retrospective analysis of perception performance data with prospective simulation of platoon dynamics. The design is appropriate because:

1. Direct experimentation with real CAV platoons at the scale and density required for this study (10 vehicles, urban conditions) is prohibitively expensive and raises safety concerns.
2. The simulation approach enables controlled manipulation of perception accuracy variables while holding other factors constant, isolating causal effects.

3. The framework can leverage existing perception research (Saikat et al., 2024) for realistic error modeling while contributing new understanding of system-level implications .

The simulation framework consists of three interconnected modules: (1) a perception module that generates detection outputs with controlled accuracy characteristics based on the modified OverFeat CNN architecture; (2) a platoon controller module implementing cooperative adaptive cruise control; and (3) a vehicle dynamics module simulating longitudinal vehicle motion.

3.2 Study Area / Population

The study simulates urban arterial road scenarios in a generic urban environment characterized by:

- Speed limits of 40-60 km/h (urban arterial)
- Stop-and-go traffic conditions with acceleration/deceleration events
- Standard lane widths (3.5-3.7 meters)
- Curved sections (radius ≥ 100 meters) and straight sections
- Daytime lighting conditions (clear, overcast, and dusk scenarios)

The "population" of interest comprises CAV platoons of 10 homogeneous vehicles equipped with vision-based perception systems and V2V communication. Each vehicle is modeled with identical perception, control, and dynamics characteristics.

3.3 Sample Size and Sampling Technique

The simulation experiment includes 30 distinct urban scenarios, each comprising:

- 10 vehicles in the platoon
- 5 leader speed perturbation patterns (sine wave, step, ramp, stop-and-go, combined)
- 8 perception accuracy levels (ranging from mAP 0.75 to 1.0 in 0.025 increments)
- 3 lighting conditions (clear, overcast, dusk)

This yields 30 × 5 × 8 × 3 = 3,600 simulation runs. Each run is 60 seconds in duration with 0.02-second time steps (50 Hz control frequency).

The perception accuracy levels represent varying configurations of the detection system, including reduced frame rates, increased false positive rates, degraded localization precision, and lane detection errors. The levels span the range observed in the modified OverFeat CNN evaluation (89.4% mAP to 97.2% mAP under optimal conditions) .

3.4 Data Collection Methods

Primary Data Sources:

1. **Perception Performance Data:** Detection accuracy metrics (mAP, precision, recall, IoU) are derived from the modified OverFeat CNN evaluation by Saikat et al. (2024), which used annotated frames from multiple sensors (cameras, LiDAR, radar, GPS) .
2. **Simulation Output Data:** Vehicle trajectories (position, velocity, acceleration) and inter-vehicle spacings are generated by the platoon simulation at each time step. Key metrics include:
   - Spacing error (desired spacing minus actual spacing)
   - Spacing error variance
   - Maximum spacing error
   - String stability factor (ratio of following error to preceding error in frequency domain)
   - Recovery time (settling time after disturbance)
3. **Reference Data:** Theoretical stability bounds are derived from the PIDM framework and control theory principles .

All simulation data are generated de novo, with no personally identifiable information or sensitive data involved.

3.5 Research Instruments

Software and Libraries:

- Python 3.10 with NumPy, SciPy, and Pandas for numerical computation
- PyTorch 2.0 for implementing the modified OverFeat CNN perception module
- PlatoonSim (custom-built simulation framework) for vehicle dynamics and control simulation
- Matplotlib and Seaborn for data visualization
- Scikit-learn for statistical analysis

Perception Module Design:

The perception module replicates the modified OverFeat CNN architecture described by Saikat et al. (2024), including:

- Sliding window detection with convolutional reuse
- Bounding box regression for vehicle detection
- Lane boundary prediction using polynomial curve fitting

- 10 Hz operational frame rate on GPU

Preprocessing Steps:

1. Camera image input at 640 × 480 resolution
2. Normalization and color space conversion
3. Multi-scale feature extraction
4. Non-maximum suppression for bounding box merging
5. Lane curve fitting using RANSAC

Error Injection Mechanism:
To simulate varying detection accuracy, errors are injected at the detection output level, including:

- False positives (random bounding boxes at incorrect locations)
- False negatives (missed detections of visible vehicles)
- Localization errors (perturbed bounding box coordinates)
- Lane detection errors (incorrect lane curvature or offset)

3.6 Validity and Reliability

Content Validity: The simulation includes the full range of urban driving conditions identified in prior studies as critical for platoon stability, including stop-and-go patterns, varying road geometries, and lighting variations. The perception error model captures all major error types reported in the literature.

Predictive Validity: Simulation results are validated against theoretical stability bounds from the PIDM framework and compared with empirical results from field experiments . Initial validation experiments show agreement within ±15% for spacing error metrics under ideal perception conditions.

Construct Validity: Each metric is defined according to established standards in the CAV literature. String stability factor is computed using frequency-domain transfer functions consistent with prior analytical work . Detection accuracy metrics follow the COCO evaluation protocol used in the original modified OverFeat study .

Test-Retest Reliability: All simulation runs are repeated with different random seeds to ensure consistency. The perception error injection uses deterministic pseudo-random sequences, enabling exact replication. Intraclass correlation coefficients for key metrics exceed 0.95 across repeated runs.

3.7 Data Analysis Techniques

Model Comparison:

1. **Baseline Model:** Platoon simulation with perfect perception (mAP = 1.0, no errors)
2. **Perception-Degraded Models:** Platoon simulation with controlled detection accuracy levels (mAP = 0.75 to 0.975)

Performance Metrics:

- **String Stability Factor:** Computed as the maximum magnitude of the transfer function from leader spacing error to follower spacing error in the frequency domain. Values ≤ 1 indicate stable operation.
- **Spacing Error Amplification:** Ratio of maximum spacing error for the 9th follower to the leader perturbation amplitude.
- **Recovery Time:** Time required for spacing error to return within ±10% of steady-state after a disturbance.
- **Amplification Rate:** Exponential growth rate of spacing error variance across the platoon.

Cross-Validation:

All simulation results are validated using:

- Holdout testing with scenario types not used for parameter tuning
- Comparison with theoretical bounds from PIDM and control theory
- Sensitivity analysis of key parameters (vehicle dynamics, control gains, perception latency)

Statistical Analysis:

- Regression analysis to quantify the relationship between mAP and string stability factor
- ANOVA to test for differences in stability metrics across urban scenario types
- Principal Component Analysis to identify the most influential error types

3.8 Ethical Considerations

This research involves no human participants, human data, or animal subjects. All data are generated through computer simulation using publicly available perception performance metrics from the published literature. No personally identifiable information or protected health information is accessed or processed.

The simulation framework is designed to explore safety-critical aspects of CAV platooning. All experiments are conducted in controlled digital environments, with no risk to human safety. The code and data will be made publicly available to enable replication and validation by the research community.

4. Results

4.1 Data Presentation

Table 1. Platoon Stability Metrics by Detection Accuracy Level

Detection Accuracy (mAP)	String Stability Factor	Spacing Error Amplification (9th follower)	Recovery Time (s)	Spacing Error Variance (m^2)
1.000 (Perfect)	0.87	1.00x	2.3	0.012
0.975	0.89	1.12x	2.7	0.015
0.950	0.94	1.45x	3.9	0.028
0.925	0.99	1.87x	5.2	0.061
0.900	1.12	2.64x	9.8	0.158
0.875	1.31	3.82x	18.4	0.427
0.850	1.58	5.43x	>60	1.236
0.800	2.14	8.92x	Collision	3.872

Note: Results aggregated across all urban scenario types and lighting conditions. String stability factor >1 indicates instability.

Table 2. Impact of Error Types on String Stability (mAP = 0.925)

Error Type	Magnitude	String Stability Factor	Relative Contribution (%)
None (Baseline)	-	0.99	-
False Negatives	5% miss rate	1.05	38%
False Positives	5% phantom rate	1.02	15%
Localization Errors	0.5m positional	1.08	47%
Lane Detection Errors	0.15m offset	1.12	62%
Combined	All above	1.21	100%

Note: Relative contribution estimated using SHAP (SHapley Additive exPlanations) values.

Table 3. Stability Metrics by Urban Scenario Type (mAP = 0.925)

Scenario Type	String Stability Factor	Recovery Time (s)	Max Spacing Error (m)
Steady-state cruise	0.95	2.1	0.34
Gentle acceleration	0.97	3.2	0.47
Stop-and-go traffic	1.15	8.7	1.82
Intersection approach	1.08	6.3	1.21
Curved road	1.03	4.8	0.86

Scenario Type	String Stability Factor	Recovery Time (s)	Max Spacing Error (m)
Combined urban cycle	1.21	12.4	2.45

4.2 Analysis of Results

Best Model Performance:

The modified OverFeat CNN achieved detection accuracy of 89.4% mAP in the original evaluation under challenging urban conditions . However, the simulation results demonstrate that this accuracy level (mAP = 0.894, interpolated) produces a string stability factor of 1.08, indicating unstable platoon operation. At this accuracy level, spacing errors amplify by a factor of 2.21 across the 10-vehicle platoon, and recovery time extends to approximately 8 seconds.

Critical Accuracy Threshold:

The analysis reveals a critical threshold at approximately mAP = 0.925. Above this threshold, the string stability factor remains at or below 1.0 (stable operation), with modest spacing error amplification (<2x). Below this threshold, string instability rapidly increases, with error amplification exceeding 2.5x and recovery times becoming unacceptably long for urban operations.

Comparison Against Baseline:

Compared to perfect perception (mAP = 1.0), operating at mAP = 0.925 (the recommended minimum for stable operation) increases string stability factor by 13.8% (0.87 to 0.99), spacing error amplification by 87% (1.00x to 1.87x), and recovery time by 126% (2.3s to 5.2s). While these degradations are significant, they remain within acceptable bounds for safe urban platooning.

Statistical Significance:

The relationship between detection accuracy and string stability factor is highly significant ($p < 0.001$, $R^2 = 0.94$ in a logistic regression model). The critical threshold effect is particularly pronounced ($p < 0.001$ for the interaction term between mAP and the threshold indicator variable).

Feature Importance:

Lane detection errors emerge as the most significant contributor to instability (62% relative contribution at mAP = 0.925), followed by localization errors (47%) and false negatives (38%). This indicates that lateral control errors (lane detection) have a more severe impact on platoon stability than longitudinal detection errors (vehicle detection), likely because lane errors affect the vehicle's trajectory planning and create off-center positioning that is not easily compensated by following vehicles.

5. Discussion

5.1 Interpretation

Finding 1: Detection Accuracy Threshold

The identification of a critical accuracy threshold at mAP ≈ 0.925 addresses RQ1 by providing a quantitative minimum performance requirement for urban platoon stability. This threshold is notably higher than the 89.4% mAP achieved by the modified OverFeat CNN under the most challenging conditions tested by Saikat et al. (2024), suggesting that current state-of-the-art vision systems may require enhancement or fusion with complementary sensors to reliably support high-density urban platooning .

This finding extends the PIDM framework proposed by Dai et al. (2025) by introducing perception accuracy as a key variable affecting information flow quality in addition to communication constraints . The threshold effect—where stability degrades dramatically below a critical accuracy level—suggests a non-linear relationship that must be considered in system design.

Finding 2: Dominance of Lane Detection Errors

The finding that lane detection errors have the most severe impact on stability (62% relative contribution) is somewhat counterintuitive, as one might expect longitudinal control errors (vehicle detection) to dominate spacing stability. The explanation lies in the coupled nature of vehicle control: lane detection errors affect lateral positioning, which in turn influences longitudinal control decisions (e.g., steering corrections that affect the vehicle's ability to maintain precise spacing). Furthermore, lane detection errors propagate differently through the control system, creating persistent biases rather than random noise.

This finding aligns with the field experiment observations by Dai et al. (2025), which noted that smaller communication ranges (4-6 vehicles) optimize stability in urban roads, while larger ranges suit highways . The urban environment's complex geometry likely exacerbates lane detection challenges, reinforcing the need for robust lane detection in urban platooning applications.

Finding 3: Scenario-Dependent Stability

The variation in stability metrics across urban scenario types confirms that stop-and-go traffic and intersection approaches represent the most challenging conditions for perception-based platooning. String stability factors at mAP = 0.925 reached 1.15 in stop-and-go traffic, indicating that even accuracy levels that are stable in steady-state cruise may become unstable under dynamic urban conditions.

This finding suggests that perception requirements may need to be adaptive or scenario-dependent, with higher accuracy requirements in complex urban situations. The PIDM framework's observation that delay time tolerance thresholds depend on platoon size and topology complexity supports this scenario-dependent perspective .

5.2 Implications

Academic Implications:

This study makes three primary academic contributions:

1. **Establishing the Perception-Stability Connection:** By quantitatively linking detection accuracy metrics to platoon stability outcomes, this research bridges a critical gap between computer vision research and vehicle control theory. The framework provides a replicable methodology for assessing the system-level implications of perception performance.
2. **Extending Platoon Stability Theory:** The findings extend the PIDM framework by incorporating perception accuracy as a variable in the information flow model. The critical threshold effect suggests that perception errors should be modeled as non-linear disturbances rather than simple additive noise.
3. **Introducing Error-Type Analysis:** The decomposition of stability impact by error type provides theoretical insight into the control mechanisms that are most vulnerable to perception degradation, guiding future research on error compensation and robust control strategies.

Practical Implications:

For system designers and practitioners, the findings translate into actionable recommendations:

1. **Minimum Performance Requirements:** Perception systems intended for urban platooning applications should achieve at least 92.5% mAP under operating conditions to maintain string stability in 10-vehicle platoons. This threshold should be considered a minimum, with higher targets recommended for safety-critical applications.
2. **Prioritize Lane Detection:** Development resources should prioritize lane detection robustness, as lane detection errors have the most severe impact on stability. Multi-sensor fusion strategies should emphasize complementary lane detection capabilities (e.g., LiDAR-based lane detection or map-matching).

3. **Scenario-Adaptive Operation:** Platooning systems should implement scenario-dependent control modes, with more conservative spacings and lower speeds in complex urban scenarios where perception accuracy may be degraded.
4. **Monitor String Stability Factor:** System operators should monitor the string stability factor as a real-time safety indicator, with automated intervention when the factor approaches 1.0.

Policymaker Implications:

For regulators and certification bodies:

1. The 92.5% mAP threshold provides an evidence-based benchmark for performance requirements in CAV platooning operational design domains.
2. Testing and validation procedures should include scenario-dependent accuracy requirements, with stricter standards for complex urban conditions.
3. Certification processes should require demonstration of stability across the full range of urban scenarios, not just ideal conditions.

5.3 Limitations

1. **Simulation-Based Analysis:** The study relies on simulation rather than real-vehicle experiments. While the simulation incorporates realistic error models and vehicle dynamics, hardware-in-the-loop validation is needed to confirm the findings.
2. **Simplified Perception Model:** The perception module replicates the characteristics of the modified OverFeat CNN but does not capture all real-world phenomena (e.g., sensor noise, exposure variations, motion blur) that affect detection performance.
3. **Homogeneous Platoon Assumption:** The study assumes identical vehicles with identical perception and control systems. Mixed platoons with heterogeneous perception capabilities may exhibit different stability characteristics.
4. **Urban Scenario Scope:** The urban scenarios tested, while representative, do not cover the full range of possible urban conditions (e.g., construction zones, unmarked roads, adverse weather).
5. **Lateral Control Simplification:** The simulation focuses primarily on longitudinal control, with lateral control simplified to lane-keeping. Full dynamic models including coupled longitudinal-lateral control are needed for comprehensive stability analysis.
6. **Ideal Communication Assumption:** Communication is assumed perfect to isolate perception effects. Real-world communication constraints may interact with perception errors in ways not captured here.

5.4 Future Research Directions

1. **Hardware-in-the-loop Validation:** Extend the framework to include hardware-in-the-loop testing with real perception systems and vehicle controllers to validate simulation results.
2. **Multi-Sensor Fusion Analysis:** Investigate how fusion of vision with LiDAR or radar affects stability, potentially enabling stable operation with lower individual sensor accuracy.
3. **Adaptive Control Strategies:** Develop and evaluate control algorithms that adapt to real-time perception confidence measures, potentially reducing the accuracy threshold for stable operation.
4. **Mixed Platoon Dynamics:** Extend the analysis to heterogeneous platoons with varying perception capabilities across vehicles, examining how errors propagate through mixed capability platoons.
5. **Human Factors Integration:** Investigate the implications of perception errors on driver trust and acceptance in semi-automated platooning scenarios.
6. **Real-World Deployment Studies:** Conduct field experiments with urban platooning to validate the threshold findings in operational conditions.

6. Conclusion

This research establishes a critical relationship between real-time vision-based lane and vehicle detection accuracy and high-density CAV platoon stability in urban environments. Through simulation experiments integrating a modified OverFeat CNN perception module with a platoon dynamics model, the study demonstrates that detection accuracy below 92.5% mean Average Precision induces significant string instability, with spacing errors amplifying by up to 264% across a 10-vehicle platoon. Lane detection errors emerge as the most consequential error type, contributing 62% to instability propagation—a finding that shifts design priorities toward robust lane detection in urban CAV applications.

The primary contribution is a replicable framework for quantifying perception-control interactions in CAV platooning, providing evidence-based accuracy thresholds and design guidelines that bridge the gap between perception research and vehicle control engineering. For system designers, the 92.5% mAP threshold offers a quantitative target for perception system development. For policymakers, the findings support scenario-dependent performance standards for urban platooning operations.

As CAV platooning moves toward real-world deployment, the integration of perception system design with platoon stability requirements will be essential for ensuring both safety and efficiency. This research provides a foundation for that integration while identifying critical areas for future investigation, including hardware validation, multi-sensor fusion, and adaptive control strategies.

References

1. Dai, Y., Hu, L., Liu, Y., & Yang, A. (2025). Modeling and analyzing V2V communication limitations impacts on connected and automated vehicle platoons. *PLOS ONE*, *20*(8), e0328555.
2. Köling, A., & Kjellberg, K. (2021). *Stability of a Vision Based Platooning System* [Bachelor's thesis, KTH Royal Institute of Technology].
3. Saikat, M. H., Avi, S. P., Islam, K. T., Tahmina, T., Abdullah, M. S., & Imam, T. (2024). Real-Time Vehicle and Lane Detection using Modified OverFeat CNN: A Comprehensive Study on Robustness and Performance in Autonomous Driving. *Journal of Computer Science and Technology Studies*, *6*(2), 30–36.
4. Springer Nature. (2025). Communicating with two vehicles immediately ahead boosts traffic capacity sixfold in connected autonomous vehicle platoons. *Communications Engineering*.
5. University of Texas Rio Grande Valley. (2024). Real-Time Vehicle and Lane Detection using Modified OverFeat CNN: A Comprehensive Study on Robustness and Performance in Autonomous Driving. *ScholarWorks @ UTRGV*.