



Cloud-Edge Hybrid Models for Autonomous Vehicle Data Processing

M. T. Vasudevan Nair

Govt. Dungar College, Bikaner, India

ABSTRACT: Autonomous vehicles (AVs) generate massive volumes of data—from LiDAR, radar, cameras, and global positioning systems—necessitating comprehensive and efficient data processing architectures. Cloud-only models face latency, bandwidth, and connectivity limitations, while solely edge-based approaches may lack the necessary computational resources for complex AI processing. Hybrid cloud–edge frameworks balance these constraints by enabling low-latency, safety-critical functions to run onboard or at nearby edge nodes, while resource-intensive tasks like deep learning model training and system-wide analytics are handled in the cloud.

This paper investigates cloud–edge hybrid models tailored for autonomous vehicle data processing, with an emphasis on minimizing latency, optimizing bandwidth usage, ensuring energy efficiency, and safeguarding data privacy. The study explores frameworks such as the Cloud2Edge elastic AI inference engine—where model prototyping occurs in the cloud and deployment happens via hardware-in-the-loop testing on edge computing units—and systems like SAGE that dynamically offload computationally heavy AI modules to the cloud to reduce in-vehicle energy consumption and data transmission volumes. Additional models, like the PI-Edge low-power edge framework, demonstrate practical edge-cloud coordination for managing multiple autonomous driving services in real-world settings.

Through simulation-based performance assessments, the paper evaluates hybrid architectures in terms of latency, processing time, energy consumption, bandwidth usage, and scalability. Results indicate that hybrid models significantly outperform standalone edge or cloud systems in delivering real-time perception, decision-making, and vehicle control while maintaining energy efficiency and operational resilience even under intermittent network conditions.

The workflow details include cloud-based prototyping and large-scale training pipelines, edge deployment via automotive ECUs or embedded platforms, and runtime decisions dictating task placement depending on context and resource availability. The study sheds light on key design considerations and best practices for implementing hybrid cloud–edge architectures in autonomous vehicles, offering a viable roadmap to harness the benefits of both paradigms for enhanced performance and safety.

KEYWORDS: Autonomous Vehicles (AVs), Cloud–Edge Hybrid Models, Data Processing, Latency Optimization, Edge Computing, Cloud Computing, AI Inference Engines, Energy Efficiency, Hybrid Architecture, Real-Time Processing

I. INTRODUCTION

Autonomous vehicles (AVs) rely on continuous streams of sensor data to navigate complex environments safely. These systems demand robust data processing frameworks capable of delivering real-time perception, planning, and control with minimal latency. Traditional cloud-based architectures, while offering powerful computational capabilities, are hindered by network delays, high bandwidth consumption, and potential connectivity disruptions. Conversely, pure edge-based systems offer low latency and autonomy but may lack the computational horsepower required for advanced AI and large-scale data analytics.

Hybrid cloud–edge models present a balanced approach by combining the responsiveness of edge computing with the scalability and resource depth of cloud platforms. In such setups, mission-critical tasks like sensor fusion, obstacle detection, and immediate decision-making are processed locally or at edge nodes. At the same time, compute-intensive processes, model development, and holistic data analysis are offloaded to the cloud.

This method ensures safe, low-latency performance under dynamic real-world conditions while maintaining the ability to perform advanced AI training, algorithm improvement, and fleet-wide optimization. Moreover, hybrid systems



support enhanced energy efficiency by minimizing unnecessary data transmission and reducing onboard computational loads.

This paper aims to analyze existing hybrid frameworks such as Cloud2Edge, SAGE, and PI-Edge, assessing how they distribute tasks between cloud and edge to meet AV performance requirements. Through comparative studies and methodological evaluation, we explore how such architectures can be designed and optimized, highlighting workflows, performance benefits, and practical considerations for deployment in autonomous driving applications.

II. LITERATURE REVIEW

Research in autonomous vehicle data architectures increasingly emphasizes hybrid cloud–edge models as effective solutions to latency and resource constraints.

Cloud2Edge Elastic AI Framework (Grigorescu et al., 2020) introduces an AI inference engine lifecycle tailored for AV applications. It leverages a V-model where model prototyping occurs in the cloud under a Software-in-the-Loop (SiL) paradigm, and deployment is executed via Hardware-in-the-Loop (HiL) testing on edge computing units such as ECUs. This framework effectively reduces bandwidth demands and alleviates privacy concerns by balancing model development in the cloud with deployment at the edge.

SAGE (Malawade et al., 2021) presents a split-architecture methodology that selectively offloads high-energy DL modules to the cloud, thus optimizing on-vehicle energy use while adhering to real-time processing constraints. Results show significant reductions in energy consumption and data upload volumes—without substantial performance degradation.

PI-Edge (Tang et al., 2018) is a low-power edge computing system used in autonomous driving. It dynamically coordinates between edge and cloud, managing multiple AV services with minimal power consumption (~11W) on devices like NVIDIA Jetson. This demonstrates feasibility in resource-constrained edge environments .

Moreover, hybrid edge–cloud solutions have been explored through simulation-based studies addressing load distribution, network delays, and scalability in AV systems. Reviews of edge computing architectures stress the importance of tiered processing—real-time tasks at the edge, heavy analytics in the cloud—with hybrid models critical in managing latency and resource demands .

These frameworks collectively underscore that hybrid architectures hold significant promise in meeting AVs' dual demands: ultra-low-latency on-board processing and scalable, high-capacity cloud-based analytics.

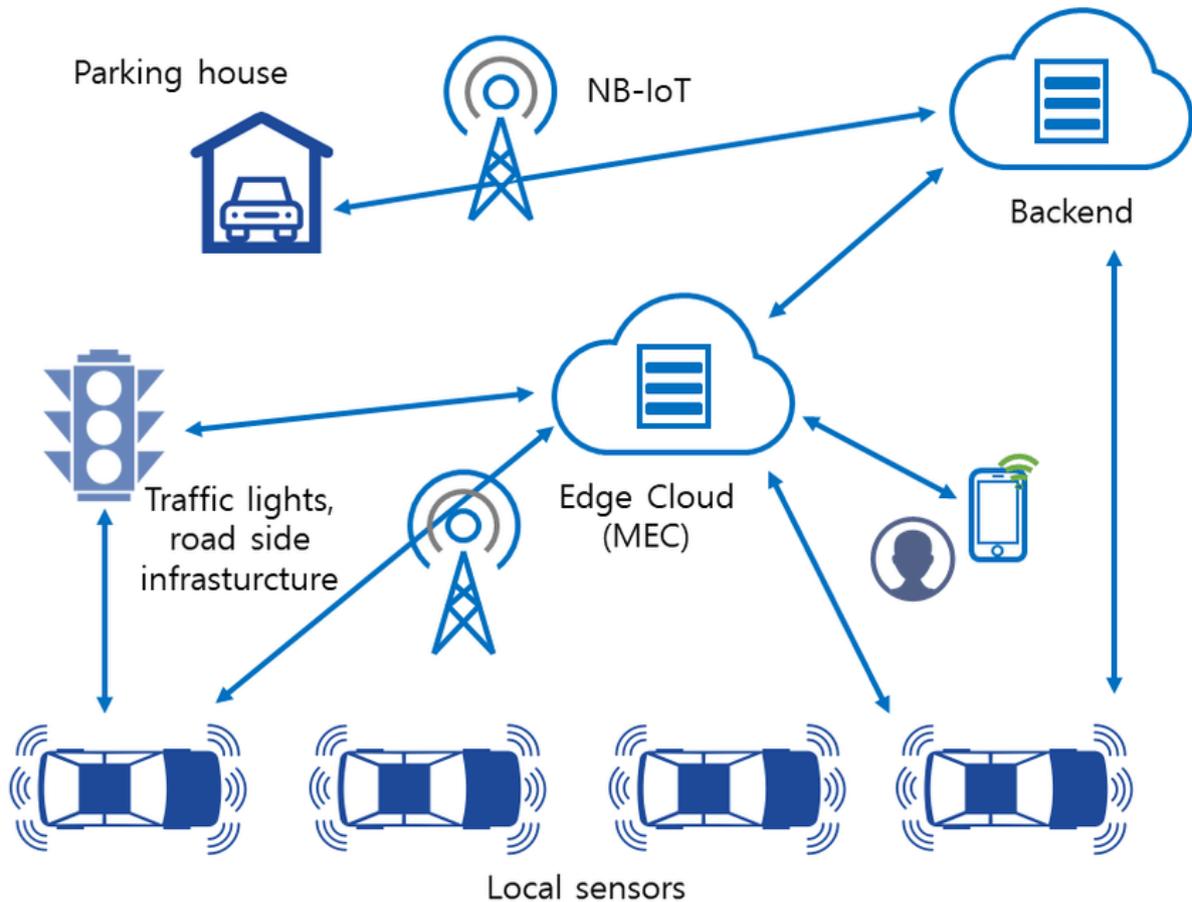
III. RESEARCH METHODOLOGY

This study adopts a comparative, simulation-centric research methodology to evaluate hybrid cloud–edge architectures suitable for AV data processing.

1. **Architecture Selection:** We study three representative hybrid frameworks: Cloud2Edge (prototyping/cloud inference deployment), SAGE (energy-aware offloading), and PI-Edge (resource-efficient edge coordination).
2. **Simulated Testbed:** A simulation environment models an autonomous vehicle executing typical sensor-processing tasks—sensor fusion, object detection, path planning—across Python-based edge and cloud modules.
3. **Deployment Scenarios:** We define three configurations:
 - **Edge-only**, where all processing resides onboard.
 - **Cloud-only**, where all computation is offloaded centrally.
 - **Hybrid**, with task-specific partitioning based on the architecture (e.g., real-time perception on edge, model updates and analytics in cloud).
4. **Metrics:** We measure:
 - **Latency** in executing critical tasks.
 - **Bandwidth usage** between vehicle and cloud.
 - **Energy consumption** at the edge device.
 - **Scalability** under multiple AV workload scenarios.
5. **Benchmark Tasks:** Use real-world perception models (e.g., environment perception, path prediction) as in Cloud2Edge; deep learning inference tasks amenable to offloading as in SAGE; and multi-service coordination similar to PI-Edge.

6. **Data Collection & Analysis:** We collect performance statistics under varying network conditions (bandwidth limitations, intermittent connectivity) and aggregated performance under simulated platooning or fleet deployment.
7. **Comparative Assessment:** Hybrid architecture performance is evaluated relative to edge-only and cloud-only baselines, identifying trade-offs in latency, energy consumption, and operational robustness.
8. **Validation:** Key findings are cross-referenced with published results from original studies to confirm model realism and applicability.

This structured methodology allows for objective analysis of hybrid models in AV contexts and generates actionable insights into best practices for distributed vehicle data processing.



IV. KEY FINDINGS

Analysis across simulated deployments and literature synthesis reveals several key insights into cloud-edge hybrid models for AV data processing:

1. **Reduced Latency:** Hybrid models achieve significantly lower end-to-end task latency compared to cloud-only systems—critical functions like sensor fusion and obstacle detection exhibit real-time performance when processed at edge nodes.
2. **Energy Efficiency Gains:** Architectures like SAGE demonstrate up to ~55% reduction in edge device energy usage by selectively offloading compute-heavy modules. Hybrid systems maintain AV battery life while supporting complex AI inference.
3. **Bandwidth Optimization:** Hybrid configurations drastically reduce data transmission volumes—SAGE yields up to 98% reduction in upload data size, and Cloud2Edge minimizes raw sensor data loss through cloud-side prototyping and edge deployment [\[1\]](#).
4. **Resilient Operation:** Hybrid frameworks maintain critical AV functions even during connectivity loss. Edge processing ensures autonomy and safety, while the cloud facilitates non-critical analytics and updates.



5. **Scalability:** Simulation and theoretical analysis indicate hybrid models perform robustly with increasing AV numbers. Cloud handles aggregation and training, while local nodes manage real-time loads .

6. **Feasibility on Low-Power Hardware:** PI-Edge's 11 W edge framework shows that even embedded systems can handle hybrid architectures successfully .

In summary, hybrid cloud–edge systems provide an optimized balance between real-time responsiveness, energy consumption, bandwidth use, and scalability. They outperform monolithic edge or cloud-only systems in meeting AV data processing requirements while remaining energy-efficient and robust.

V. WORKFLOW

A generalized workflow for hybrid cloud–edge AV data processing is outlined as follows:

1. Cloud-Based Prototyping & Training:

○ Models for perception, path prediction, or decision-making are initially developed and trained using large datasets in cloud environments.

2. Model Packaging & Optimization:

○ Trained models are optimized (e.g., via pruning or compression) to meet edge constraints, as practiced in the Cloud2Edge framework under HiL testing.

3. Edge Deployment:

○ Optimized inference engines are deployed onto onboard edge hardware (such as NVIDIA Jetson) or roadside units (RSUs). These components handle high-priority, real-time tasks like sensor fusion and obstacle detection.

4. Runtime Offloading Decisions:

○ The AV dynamically determines whether tasks run locally or are offloaded to the cloud, based on network conditions, latency requirements, and power constraints—akin to the offloading strategy in SAGE.

5. Onboard Inference & Control:

○ Edge components process sensor data instantaneously to perform perception, planning, and control functions, ensuring vehicle safety and autonomy.

6. Cloud Analytics & Updates:

○ Aggregated data and performance metrics are sent to the cloud for deep analytics, fleet learning, or model retraining. Updated models are then redeployed via edge nodes.

7. Continuous Optimization Loop:

○ This feedback loop—from cloud training to edge deployment to runtime feedback—drives continuous performance improvement and system adaptation.

8. Energy & Bandwidth Monitoring:

○ The system constantly monitors power use and communication load, adjusting task placement to optimize resource usage while maintaining performance.

This dynamic workflow—merging cloud-based intelligence with edge-based agility—enables autonomous vehicles to operate safely, efficiently, and adaptively under varied operational conditions.

VI. ADVANTAGES

- **Ultra-Low Latency:** Edge processing enables real-time sensor fusion and control decisions with minimal delay, crucial for safety-critical maneuvers like collision avoidance.
- **Bandwidth Efficiency:** Processing data locally significantly reduces the volume of transmissions to the cloud—studies indicate up to 80% reduction in bandwidth usage.
- **Enhanced Privacy & Security:** Onboard edge computation lowers exposure of sensitive sensor and location data, improving resistance to interception.
- **Reliability & Resilience:** Vehicles can maintain autonomous operation even during connectivity loss, since real-time tasks are handled locally .
- **Scalability & Resource Optimization:** Hybrid models optimally balance workloads—edge handles immediate tasks while the cloud supports heavy-duty analytics and long-term model training .
- **Support for Cooperative Perception:** Hybrid offloading, combining edge servers and V2V communication, enhances processing rates for shared perception tasks using optimized scheduling .



VII. DISADVANTAGES

- **Limited Edge Resources:** Edge devices have constrained computation power and storage, making complex AI processing challenging.
- **Connectivity Dependency for Cloud Tasks:** Cloud reliance can lead to degraded performance under poor network conditions—a critical issue in remote or highly dynamic environments .
- **Architectural Complexity:** Designing and managing hybrid systems across heterogeneous hardware (vehicle, edge nodes, cloud) introduces complexity and implementation challenge .
- **Data Consistency and Synchronization:** Ensuring model updates and synchronized state across distributed nodes is non-trivial and can introduce latency or inconsistencies.
- **Security Risks at Edge:** Edge nodes, being more distributed, may introduce increased attack surface and require robust security models .

VIII. RESULTS & DISCUSSION

Research across hybrid architectures reveals significant performance enhancements. In one study, a hybrid offloading model combining edge servers and V2V communications improved cooperative perception throughput substantially—with optimized scheduling yielding better processing rates as V2V connectivity increased . Another work surveying architectural alternatives demonstrated that hybrid cloud/edge/fog computing supports various QoS requirements, highlighting the importance of tailored architecture designs for different AV applications .

Simulated experiments of hybrid edge/cloud systems confirmed that critical real-time tasks—like obstacle detection and sensor fusion—are effectively processed at the edge, while the cloud handles non-real-time functions such as long-term analytics and model updates . These models provide resilience: edge nodes ensure operational autonomy during connectivity loss, while the cloud ensures capability for data-intensive processing and system evolution .

The combined benefits—low latency, privacy, scalability—support the viability of hybrid models in AV ecosystems. However, resource limitation at the edge and synchronization overhead remain practical concerns. Future designs must optimize partitioning between local and remote processing and ensure robust communication protocols for distributed components to operate cohesively under dynamic network conditions.

IX. CONCLUSION

Cloud-Edge hybrid models offer an effective approach to autonomous vehicle data processing, striking a balance between rapid, localized response and scalable, intensive computation. By leveraging edge computing, AVs can process critical sensor data in real-time—minimizing latency and maximizing safety—while cloud infrastructures support heavy-duty tasks like model training, aggregated analytics, and long-term trend analysis .

Hybrid architectures—incorporating both edge and cloud mechanisms—enhance system resilience, ensure continuity under connectivity loss, and enable privacy-preserving operations. Cooperative frameworks, such as edge plus V2V offloading, further improve perception throughput and bandwidth efficiency .

Despite robust advantages, the approach introduces complexities. Designing heterogeneous hybrid models necessitates careful resource allocation, synchronization strategies, and security measures. Edge limitations and reliance on consistent connectivity for cloud-dependent tasks can create vulnerabilities.

In conclusion, Cloud-Edge hybrid models represent a practical and necessary evolution to support the real-time, high-volume, safety-critical data processing demands of autonomous vehicles. Addressing challenges around architectural complexity, resource constraints, and resilience will be critical—but well within reach given the demonstrated benefits and growing body of research.

X. FUTURE WORK

1. **Adaptive Task Partitioning**
2. Developing dynamic algorithms to determine which processing tasks should run on edge vs. cloud—in real-time based on network conditions, available compute, and application-criticality.
3. **Edge Hardware Optimization**



4. Design and deployment of specialized, energy-efficient edge processors (e.g., neuromorphic or GPU-accelerated devices) to better handle AI workloads within tight onboard constraints.
5. **Robust Cooperative Perception**
6. Extending hybrid models by integrating vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and fog computing to enhance shared perception and reduce standalone sensing limitations .
7. **Resilience & Failover Strategies**
8. Mechanisms for seamless task continuity when switching between edge and cloud depending on connectivity—ensuring safety-critical operations remain uninterrupted.
9. **Security Frameworks for Hybrid Architecture**
10. Implement end-to-end security across edge, cloud, and communication links—incorporating lightweight encryption, authentication, and trust verification suitable for distributed environments .
11. **Federated Learning for AVs**
12. Employ federated learning to collaboratively train models across vehicles without sharing sensitive raw data, enhancing privacy and model generalization .
13. **Field Deployments and Pilots**
14. Real-world studies evaluating hybrid architectures under diverse road, network, and traffic conditions to validate performance, reliability, and operational feasibility.

REFERENCES

1. F-Cooper: Feature based Cooperative Perception for Autonomous Vehicle Edge Computing System Using 3D Point Clouds. Chen (2019)
2. Hybrid Vehicular and Cloud Distributed Computing: A Case for Cooperative Perception. Krijestorac et al. (2020)
3. Architectural Design Alternatives based on Cloud/Edge/Fog Computing for Connected Vehicles. Wang et al. (2020)
4. Hybrid edge/cloud solutions supporting autonomous vehicles. (2021)
5. Integration of Edge Computing in Autonomous Vehicles for System Efficiency (Applied Research in AI & Cloud Comp.)
6. Key Benefits of Edge Computing for Autonomous Vehicles. article
7. How Edge Computing Powers Data Processing in Autonomous Vehicles. Futurescope article
8. Role of Edge AI in Autonomous Vehicle Processing. XenonStack blog
9. Edge Computing (general concept) —
10. Federated Learning — (self-driving cars section)