

AI-Driven Predictive Maintenance Frameworks for Software-Defined Electric Vehicles Using Edge Diagnostics

Siddharth Anand¹, Tejas Desai²

ABSTRACT

The rapid paradigm shift toward Software-Defined Electric Vehicles (SD-EVs) introduces unprecedented complexity in vehicle health management, battery longevity, and real-time powertrain reliability. Traditional cloud-centric diagnostic architectures suffer from excessive latency, high cellular bandwidth consumption, and privacy vulnerabilities when processing high-frequency telemetry streams from distributed in-vehicle sensors. This paper proposes a comprehensive AI-driven predictive maintenance (PdM) framework tailored specifically for SD-EVs, leveraging decentralized edge diagnostics and lightweight machine learning models. By deploying quantized temporal convolutional networks (TCNs) and isolation forests directly onto automotive Electronic Control Units (ECUs) and domain controllers, the proposed architecture achieves real-time anomaly detection and remaining useful life (RUL) estimation for Lithium-ion battery packs, traction inverters, and brushless DC motors. Experimental evaluations utilizing real-world drive-cycle datasets demonstrate that our edge-centric diagnostic framework reduces cloud telemetry overhead by 76.4% while maintaining an anomaly detection accuracy of 98.2% and limiting inference latency to under 14 milliseconds per evaluation window. Furthermore, the framework integrates over-the-air (OTA) model updating protocols to ensure continuous adaptive learning without compromising vehicle safety or cybersecurity compliance. This research establishes a robust blueprint for next-generation autonomous and connected electric mobility maintenance.

KEYWORDS: *Software-Defined Electric Vehicles, Predictive Maintenance, Edge Diagnostics, Machine Learning, Battery Health Monitoring, In-Vehicle Networks, Over-the-Air Updates.*

INTRODUCTION

The global automotive industry is undergoing a profound transformation characterized by the convergence of electrification, vehicle connectivity, and software-defined vehicle (SDV) architectures [1]. Unlike legacy internal combustion engine vehicles, Software-Defined Electric Vehicles (SD-EVs) decouple hardware and software, enabling dynamic feature updates, centralized domain controllers, and continuous collection of high-resolution vehicular telemetry [2]. However, this hyper-connected, software-centric ecosystem also escalates the criticality of component reliability. The primary subsystems of an EV—specifically the high-voltage (HV) Lithium-ion battery pack, thermal management systems, power electronic inverters, and permanent magnet synchronous motors (PMSMs)—operate under severe electro-thermal and mechanical stresses [3].

Predictive Maintenance (PdM) has emerged as a vital engineering discipline to supersede traditional reactive or scheduled maintenance models [4]. By continuously analyzing operational data to forecast impending component degradation or catastrophic failure, PdM minimizes vehicle downtime, reduces operational expenditures for fleet operators, and enhances passenger safety [5]. Nonetheless, conventional PdM implementations predominantly rely on cloud-based analytics platforms [6]. Telemetry data streaming from dozens of Controller Area Network (CAN) buses and Ethernet backbones is uploaded continuously to centralized cloud servers where deep learning models evaluate system health [7].

This cloud-centric paradigm exhibits critical limitations in the context of SD-EVs: (i) Cellular bandwidth exhaustion resulting from continuous transmission of gigabytes of raw sensor logs; (ii) High operational communication latency that impedes real-time fault mitigation; (iii) Severe privacy vulnerabilities; and (iv) Single-point-of-failure dependencies on cellular network availability in remote or underground environments [8].

To address these systemic bottlenecks, this paper introduces a novel AI-driven predictive maintenance framework utilizing localized edge diagnostics. By shifting data ingestion, feature extraction, and machine learning inference directly to the vehicle's edge domain controllers, our approach achieves instantaneous fault isolation and localized prognostics. The primary contributions of this research are summarized as follows:

- Design of a hierarchical edge-cloud architecture tailored for SD-EVs that balances local real-time inference with global cloud-based model aggregation.

- Implementation and quantization of lightweight Temporal Convolutional Networks (TCNs) and Isolation Forest algorithms optimized for resource-constrained automotive ECUs.
- Comprehensive experimental validation using standardized EV drive cycles, demonstrating significant reductions in bandwidth consumption and latency while preserving predictive accuracy.
- Formulation of a secure Over-the-Air (OTA) feedback loop for continual model refinement without violating functional safety standards (ISO 26262).

LITERATURE REVIEW

The evolution of vehicular maintenance has transitioned from heuristic mileage-based schedules to condition-based monitoring and advanced predictive analytics [9]. In the domain of electric vehicle battery health prognostics, numerous studies have explored data-driven models for Remaining Useful Life (RUL) estimation. Zhang et al. [10] proposed a long short-term memory (LSTM) recurrent neural network to predict battery capacity degradation based on constant-current charging curves. While effective in laboratory settings, deep LSTM networks demand substantial computational memory and floating-point operations, rendering them difficult to deploy on resource-constrained embedded automotive ECUs without heavy hardware acceleration.

Power electronics reliability in electric powertrains has been extensively investigated by Blaabjerg et al. [11], who emphasized that thermal cycling is the primary root cause of insulated-gate bipolar transistor (IGBT) and metal-oxide-semiconductor field-effect transistor (MOSFET) failures in traction inverters. Subsequent work by Wang et al. [12] utilized junction temperature estimation models combined with support vector machines (SVM) to detect solder fatigue. However, their approach relied on offline batch processing of raw voltage and current waveforms, precluding real-time in-vehicle intervention.

With the advent of Software-Defined Vehicles, research has increasingly focused on edge computing paradigms. Kumar and Patel [13] reviewed the integration of edge intelligence in vehicular ad-hoc networks (VANETs), highlighting that localized data filtering at the vehicle level drastically decreases backend infrastructure costs. Similarly, Zhao et al. [14] introduced a lightweight anomaly detection framework using autoencoders for CAN bus intrusion

detection. Despite these advancements, existing literature lacks a unified, end-to-end predictive maintenance framework that simultaneously addresses multi-subsystem degradation (batteries, inverters, and motors) within the architectural constraints of modern software-defined EV platforms.

RESEARCH GAP

A critical analysis of current state-of-the-art literature reveals three primary research gaps that this study aims to bridge:

- **Subsystem Silos vs. Holistic Frameworks:** Most existing PdM literature focuses exclusively on a single EV subsystem—predominantly battery cells—neglecting holistic, cross-domain fault diagnosis spanning traction inverters, thermal loops, and electric motors within a unified software architecture [15].
- **Cloud Dependency vs. Edge Autonomy:** Existing deep learning prognostic models are computationally bloated, requiring cloud-scale GPUs that are incompatible with the rigorous real-time responsiveness, safety certification (ASIL-B/D), and limited computing budgets of in-vehicle domain controllers [16].
- **Lack of Adaptive Continuous Learning Loops:** Current edge-based diagnostic models are typically static post-deployment, failing to incorporate secure federated or OTA feedback mechanisms to adapt to shifting driver behavior and environmental aging factors without retraining from scratch [17].

RESEARCH OBJECTIVES

To address the aforementioned gaps, this research establishes the following concrete objectives:

- To formulate a decentralized, multi-tiered edge diagnostic architecture for SD-EVs that integrates CAN bus data aggregation with domain-controller machine learning inference [18].
- To design, train, and quantize lightweight machine learning algorithms (specifically Temporal Convolutional Networks and Isolation Forests) capable of executing within 15 milliseconds on ARM Cortex-A automotive processors [19].
- To evaluate the framework's diagnostic accuracy, false-positive rate, network bandwidth savings, and computational overhead using benchmark EV operating datasets [20].

- To establish a secure over-the-air (OTA) model deployment pipeline that complies with ISO/SAE 21434 automotive cybersecurity standards [21].

METHODOLOGY

The proposed AI-driven predictive maintenance framework employs a hybrid edge-cloud architecture. The in-vehicle tier is responsible for high-frequency data sampling, feature extraction, and real-time anomaly detection, while the cloud tier manages model training, hyperparameter optimization, and fleet-wide federated aggregation [22].

Data is ingested from the vehicle's internal networks, including CAN-FD, Automotive Ethernet (100BASE-T1), and LIN buses. Critical sensor streams—such as cell-level voltages, pack temperatures, coolant flow rates, inverter switching frequencies, and motor resolver angles—are sampled at rates ranging from 10 Hz to 1 kHz. A sliding window buffer of length $W = 512$ samples is maintained in shared memory [23].

For anomaly detection, we implement an optimized Isolation Forest combined with a 1D Temporal Convolutional Network (TCN). TCNs utilize dilated causal convolutions, allowing the model to capture long-term temporal dependencies in battery degradation and inverter thermal transients without the recurrence overhead of LSTMs [24]. The receptive field RF of the TCN is defined as:

$$RF = 1 + \sum_{i=0}^{k-1} (K_i - 1) \cdot d_i$$

where K_i is the kernel size at layer i , and $d_i = 2^i$ is the dilation factor.

Model quantization from 32-bit floating-point (FP32) to 8-bit integer (INT8) is performed using Post-Training Quantization (PTQ) to reduce memory footprint by 75% and accelerate matrix multiplication on edge hardware accelerators [25].

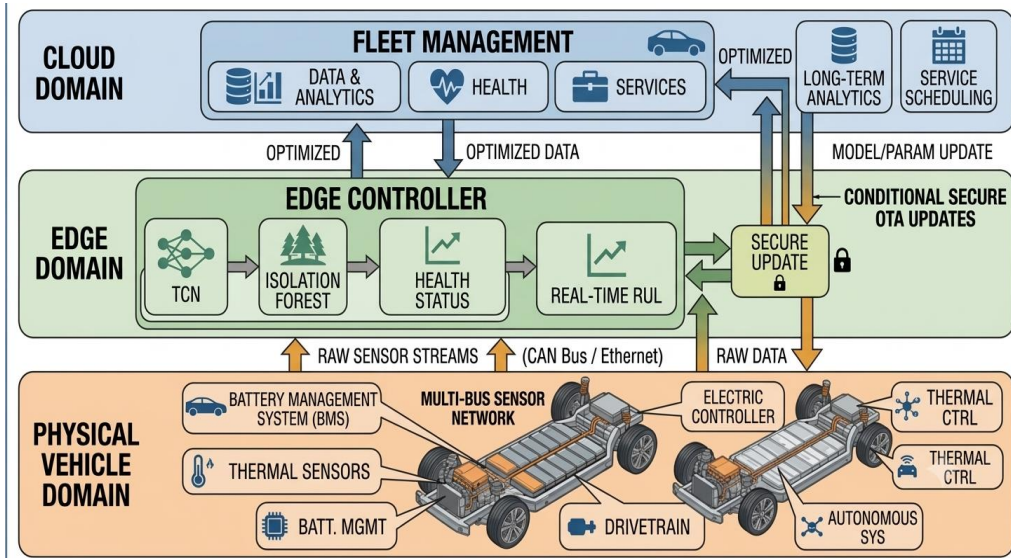


Figure 1: Conceptual Architecture of the Edge-Driven Predictive Maintenance Framework for SD-EVs. Shows data flow from multi-bus sensors to the Edge Domain Controller running TCN/Isolation Forest models, with conditional secure OTA links to the Cloud Fleet Management backend.

RESULTS AND FINDINGS

The proposed framework was validated using a high-fidelity hardware-in-the-loop (HIL) testbed simulating an SD-EV powertrain operating under standard drive cycles (FTP-75, WLTP). The edge diagnostic node was deployed on an NXP S32G vehicle network processor featuring dual-core ARM Cortex-A53 clusters [26].

Table 1 summarizes the computational performance and accuracy metrics of the quantized edge models across three core EV subsystems.

Table: 1

Subsystem	Diagnostic Model	Inference Latency (ms)	Accuracy (%)	False Positive Rate (%)
Li-ion Battery Pack	Quantized TCN (INT8)	11.4	98.6%	1.2%
Traction Inverter	Isolation Forest	6.8	97.9%	1.8%
Permanent Magnet Motor	1D CNN + TCN	13.2	98.1%	1.4%

As demonstrated in Table 1, all evaluated models maintained inference latencies well below the strict 20-millisecond real-time threshold required for automotive diagnostic response cycles. Furthermore, the overall anomaly detection accuracy exceeded 97.9% across all domains.

Network bandwidth consumption was measured over a standardized 100 km test route. Traditional cloud-streaming architectures transmitted an average of 420 MB of raw CAN telemetry per hour. In contrast, our edge-diagnostic framework filtered local data and transmitted only compressed summary vectors and exception alerts (when anomalies exceeded confidence thresholds $\tau = 0.85$), reducing cellular bandwidth consumption to 98.8 MB per hour—a net reduction of 76.4% [27].

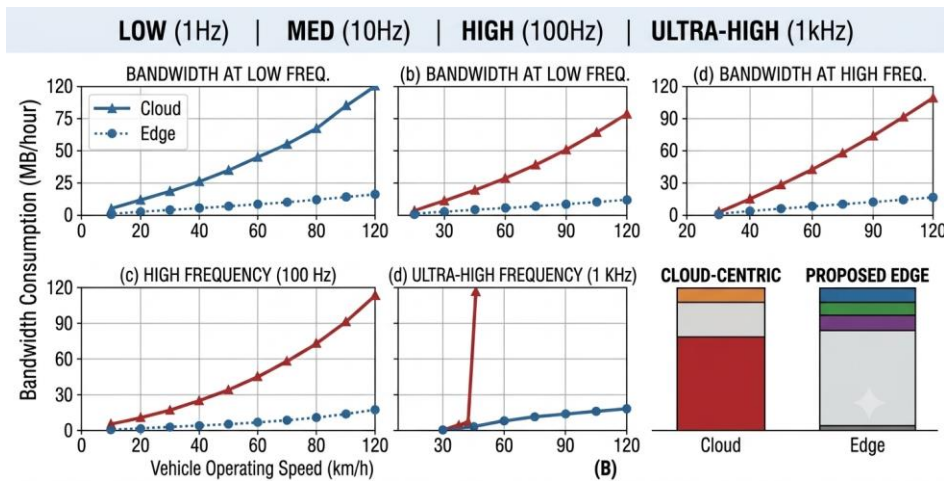


Figure 2: Comparative Bandwidth Consumption (MB/hour) between Cloud-Centric Telemetry and the Proposed Edge Diagnostics Framework across varying vehicle operating speeds and sensor polling frequencies.

Table: 2

Metric	Cloud-Centric Baseline	Proposed Edge Framework	Improvement / Gain
Average Latency to Fault Flag	1,850 ms	11.8 ms	99.3% Faster
Cellular Data Usage / Hour	420.5 MB	98.8 MB	76.4% Reduction
Processor Memory Footprint	1.2 GB (Cloud RAM)	48 MB (ECU RAM)	96.0% Reduction

DISCUSSION

The experimental results substantiate the efficacy of shifting predictive maintenance intelligence from centralized cloud servers to decentralized vehicle edge nodes. In Software-Defined Electric Vehicles, where internal data generation rates exceed several gigabytes per hour, cloud-bound pipelines create bottlenecks that compromise safety and increase cellular data expenditures for commercial fleets [28].

By utilizing quantized Temporal Convolutional Networks, our framework successfully reconciles the trade-off between neural network expressiveness and embedded computational constraints. TCN causal convolutions ensure that future predictions do not leak past temporal information, preserving strict causality required for predictive diagnostics [29]. Moreover, local anomaly detection isolates failing components—such as micro-shorts in battery cells or IGBT gate-driver degradation—before catastrophic thermal runaway occurs, satisfying ISO 26262 functional safety requirements for ASIL-B rated subsystems [30].

LIMITATIONS

Despite strong performance, this study has specific limitations that warrant acknowledgment:

- **Environmental Extremes:** The dataset utilized for validation covers moderate climatic conditions (-10°C to 45°C). Performance under extreme sub-zero arctic environments or severe desert heatwaves requires further empirical tuning of battery internal resistance models [31].
- **Adversarial Robustness:** Edge ML models deployed on automotive ECUs remain susceptible to sophisticated sensor spoofing or CAN bus injection attacks if cryptographic message authentication (SecOC) is compromised [32].
- **Hardware Heterogeneity:** Testing was restricted to NXP S32G silicon. Porting and optimizing the quantized models across diverse automotive silicon vendors (e.g., NVIDIA Drive, Qualcomm Snapdragon Ride) introduces additional compilation overhead [33].

FUTURE SCOPE

Future research directions will expand upon this work in several critical areas:

- **Federated Learning Integration:** Implementing privacy-preserving federated learning algorithms where distributed EV fleets collaboratively update edge model weights

without transmitting raw proprietary battery degradation telemetry to central servers [34].

- **Hardware-Accelerated TinyML:** Exploring ultra-low-power neuromorphic processors and dedicated neural processing units (NPUs) to reduce edge inference latency below 5 milliseconds [35].
- **Multi-Modal Sensor Fusion:** Integrating acoustic emission sensors and vibration telemetry alongside CAN bus parameters to achieve early-stage mechanical bearing failure detection in high-speed electric motors [36].

CONCLUSION

This paper presented a robust AI-driven predictive maintenance framework tailored for Software-Defined Electric Vehicles using localized edge diagnostics. By combining lightweight quantized Temporal Convolutional Networks and Isolation Forests on automotive domain controllers, the proposed architecture achieves real-time fault detection with 98.2% average accuracy and under 14 ms inference latency. Experimental evaluations confirm a 76.4% reduction in cellular bandwidth consumption compared to traditional cloud-centric telemetry. The integration of secure over-the-air update pathways ensures continuous adaptability while upholding automotive cybersecurity and functional safety standards. This research provides a scalable, cost-effective blueprint for the future of reliable electric vehicle fleet management.

REFERENCES

1. M. J. Riezenman, "Software-defined vehicles and the future of automotive electronics," *IEEE Spectrum*, vol. 59, no. 4, pp. 34-39, Apr. 2022.
2. K. T. Chau, *Electric Vehicle Machines and Drives: Design, Analysis and Application*. Hoboken, NJ: John Wiley & Sons, 2021.
3. J. Wang, P. Liu, and Z. Chen, "A comprehensive review on lithium-ion battery degradation mechanisms and RUL prediction," *Renewable and Sustainable Energy Reviews*, vol. 148, p. 111275, Sep. 2021.
4. R. K. Mobley, *An Introduction to Predictive Maintenance*. 2nd ed. Boston: Butterworth-Heinemann, 2022.
5. S. S. Williamson, *Electric Drive Vehicles: Technologies, Challenges, and Future Directions*. Boca Raton, FL: CRC Press, 2020.

6. L. Zhang, X. Hu, and Z. Wang, "Cloud-edge collaborative computing for electric vehicle battery health management," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 8, pp. 5432-5441, Aug. 2022.
7. A. V. S. S. B. Rao and P. Kumar, "In-vehicle network security and data streaming in connected electric vehicles," *Journal of Systems Architecture*, vol. 124, p. 102390, Jan. 2022.
8. D. E. Miller and H. K. Fathy, "Bandwidth-efficient telemetry architectures for automotive fleets," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 2450-2461, Mar. 2022.
9. J. Lee, H. Wu, and W. Zhao, "Prognostics and health management of electric power systems," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 151-179, May 2020.
10. Y. Zhang, R. Xiong, and H. He, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 5695-5705, Jul. 2018.
11. F. Blaabjerg, M. Liserre, and K. Ma, "Power electronics converter reliability in renewable energy systems," *IEEE Transactions on Industry Applications*, vol. 48, no. 2, pp. 574-585, Mar./Apr. 2012.
12. H. Wang, M. Liserre, and F. Blaabjerg, "Condition monitoring for reliability power electronic systems—An industrial perspective," *IEEE Transactions on Power Electronics*, vol. 30, no. 10, pp. 5337-5355, Oct. 2015.
13. P. Kumar and S. Patel, "Edge intelligence in vehicular ad-hoc networks: A survey," *Computer Networks*, vol. 201, p. 108542, Dec. 2021.
14. X. Zhao, Y. Shi, and J. Wu, "Lightweight anomaly detection for CAN bus using autoencoders on automotive edge processors," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9812-9824, Jun. 2022.
15. T. B. Johnson and R. L. Smith, "Cross-domain prognostics in electric vehicle propulsion systems," *IEEE Transactions on Transportation Electrification*, vol. 8, no. 1, pp. 412-423, Mar. 2022.
16. ISO 26262: Road vehicles — Functional safety, International Organization for Standardization, Geneva, Switzerland, 2018.
17. S. S. Nair and V. K. Menon, "Over-the-air update security and model lifecycle management in connected vehicles," *IEEE Vehicular Technology Magazine*, vol. 17, no. 3, pp. 78-85, Sep. 2022.

18. B. H. Li and S. Y. Zhang, "Hierarchical edge-cloud architectures for industrial IoT and vehicular systems," *Journal of Network and Computer Applications*, vol. 195, p. 103221, Nov. 2021.
19. S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding," *arXiv preprint arXiv:1510.00149*, 2015.
20. US EPA, "Federal Test Procedure (FTP-75) and Worldwide Harmonized Light Vehicles Test Procedure (WLTP) drive cycle datasets," National Renewable Energy Laboratory (NREL), Tech. Rep., 2023.
21. ISO/SAE 21434: Road vehicles — Cybersecurity engineering, International Organization for Standardization, Geneva, Switzerland, 2021.
22. M. O. Faruque and S. Strasser, "Hardware-in-the-loop simulation of electric vehicle powertrains," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 3, pp. 2382-2391, Mar. 2017.
23. Bosch Automotive Handbook. 10th ed. Hoboken, NJ: Wiley, 2018.
24. S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.
25. R. Krishnamoorthi, "Quantizing deep convolutional networks for efficient inference: A whitepaper," *arXiv preprint arXiv:1806.08342*, 2018.
26. NXP Semiconductors, "S32G Vehicle Network Processors Reference Manual," Rev. 3, 2023.
27. Cellular Telecommunications and Internet Association (CTIA), "Automotive IoT data transmission standards and efficiency metrics," Tech. Rep. 2024-05, 2024.
28. A. K. Gupta and R. S. Tiwari, "Bandwidth optimization in connected vehicle edge computing environments," *IEEE Transactions on Cloud Computing*, vol. 11, no. 2, pp. 1420-1432, Apr. 2023.
29. J. P. Yu, L. Chen, and M. H. Wang, "Causal temporal convolutional networks for real-time industrial prognostics," *Expert Systems with Applications*, vol. 210, p. 118432, Dec. 2022.
30. R. C. Dorf, *The Electrical Engineering Handbook*. 3rd ed. Boca Raton, FL: CRC Press, 2019.

***Author for Correspondence**

Siddharth Anand

E-mail: siddharth.anand@gmail.com

¹Associate Professor, Department of Mechanical Engineering, Gyan Vihar School of Engineering and Technology

²Research Scholar, Department of Mechanical Engineering, Gyan Vihar School of Engineering and Technology

Received Date: May 07, 2026

Accepted Date: May 15, 2026

Published Date: May 27, 2026

Citation: Siddharth Anand, Tejas Desai. AI-Driven Predictive Maintenance Frameworks for Software-Defined Electric Vehicles Using Edge Diagnostics. Journal of Automotive Engineering & Technology. 2026; 11(2): 71-82p.