Article

Research on a One-Stop After-sales Service Platform for Vehicles Based on Intelligent Connected Vehicles

Xinghai Yang[1,*]

[1] ShiJiHengTong Technology Co., Ltd., Guiyang, China
* Correspondence: Xinghai Yang, ShiJiHengTong Technology Co., Ltd., Guiyang, China

Abstract: The rapid evolution of intelligent connected vehicle (ICV) technologies has intensified demand for integrated, responsive, and predictive after-sales service ecosystems. This study presents the design, implementation, and empirical validation of a one-stop after-sales service platform tailored specifically for ICVs. The platform unifies telematics data ingestion, real-time fault diagnostics, automated service scheduling, dynamic parts logistics coordination, and personalized customer interaction through a unified cloud-edge architecture. A multi-layered service orchestration framework enables adaptive response to vehicle health signals, reducing mean time to resolution by 42% in field trials across 12,850 vehicles over six months. System performance was evaluated using latency benchmarks, diagnostic accuracy metrics, and service completion rate tracking under varying network conditions and fleet heterogeneity. Results demonstrate 98.7% end-to-end telemetry ingestion reliability, 93.4% precision in early-stage fault classification (e.g., battery thermal anomaly, ADAS sensor drift), and 89.2% first-contact resolution for Tier-1 issues without technician dispatch. User satisfaction scores increased from 64.3 to 87.6 on a 100-point scale post-deployment, with service cycle time shortened by an average of 3.8 days. The platform's modular microservice design ensures scalability across OEMs and aftermarket providers while maintaining strict data compliance and OTA-compliant security protocols. Findings confirm that tightly coupled integration of vehicle-generated data, AI-driven decision logic, and service workflow automation fundamentally transforms after-sales operations from reactive maintenance to anticipatory care---establishing a new operational baseline for ICV service infrastructure.

Keywords: intelligent connected vehicles; after-sales service; telematics platform; predictive maintenance

1. Introduction

1.1. The Evolving Landscape of Vehicle After-Sales Services

The vehicle after-sales service ecosystem is undergoing a paradigmatic transformation driven by the rapid integration of electronic systems, software-defined functionalities, and persistent connectivity. Modern intelligent connected vehicles generate continuous streams of operational telemetry, enabling real-time diagnostics, predictive maintenance, and over-the-air updates---capabilities that fundamentally reconfigure service expectations. Traditional dealership-centric models, characterized by reactive incident resolution, fragmented data ownership, and manual intervention workflows, increasingly fail to deliver the speed, transparency, and contextual personalization demanded by digitally fluent consumers [1]. Service latency, information asymmetry, and inconsistent cross-brand interoperability further erode customer trust and brand loyalty [2]. As vehicle architectures evolve toward centralized compute platforms and cloud-native service orchestration, the functional boundaries between hardware maintenance, software troubleshooting, and user experience optimization dissolve. This convergence necessitates a unified service infrastructure capable of aggregating heterogeneous data sources, applying adaptive decision logic, and executing coordinated interventions across distributed stakeholders---including OEMs, tier-one

Received: 01 June 2026
Revised: 14 July 2026
Accepted: 25 July 2026
Published: 30 July 2026

suppliers, third-party service providers, and end users. A one-stop platform architecture thus emerges not as a convenience feature but as an operational imperative for sustaining service quality, scalability, and strategic agility in the intelligent mobility era [3, 4].

1.2. Defining the One-Stop Service Imperative

The one-stop service imperative emerges from systemic fragmentation across vehicle after-sales ecosystems, where diagnostics, repair orchestration, parts logistics, over-the-air updates, and customer engagement operate as siloed functions. This structural disintegration impedes response latency, increases diagnostic uncertainty, and degrades service continuity---particularly under dynamic operational conditions. A unified platform must therefore integrate heterogeneous data streams from embedded vehicle ECUs, cloud-based analytics engines, and third-party service provider systems into a single context-aware interface. Interoperability is not merely technical but architectural: it demands standardized semantic models for fault representation, time-synchronized event logging, and bidirectional command protocols that preserve real-time fidelity while accommodating legacy hardware constraints [1]. Context awareness extends beyond geolocation or vehicle state to include driver behavior patterns, historical failure modes, and environmental stressors---enabling predictive intervention rather than reactive resolution. Such integration necessitates strict adherence to modular API contracts, secure identity federation, and adaptive data governance frameworks capable of reconciling divergent update cycles and regulatory compliance regimes [5]. The platform's efficacy hinges on its capacity to translate low-level telemetry into actionable service intelligence without abstraction loss or temporal drift [6, 7].

1.3. Scope and Contribution of This Work

This work defines its technical scope strictly around the architectural design, real-time data processing pipeline, fault inference logic, and service workflow automation of a one-stop after-sales service platform for intelligent connected vehicles. It deliberately excludes business model innovation, market strategy, and regulatory compliance considerations [8]. The principal contributions are threefold: first, a latency-optimized edge-cloud handoff protocol that minimizes service activation delay through dynamic task partitioning and adaptive bandwidth-aware scheduling; second, a lightweight on-vehicle diagnostic agent implemented with quantized neural modules and rule-based fallbacks, achieving sub-100 ms inference latency under constrained hardware resources; third, a service readiness scoring model that integrates vehicle health metrics, historical failure patterns, and contextual environmental factors to predict maintenance urgency and resource allocation requirements [6, 9]. All components were validated in a production fleet of over 12,000 vehicles across diverse operational conditions. The scoring model demonstrated $R^2 = 0.87$ against ground-truth service outcomes, while the handoff protocol reduced mean end-to-end latency by 42% compared to baseline cloud-only execution [10]. These advances collectively enable scalable, responsive, and context-aware after-sales service orchestration.

2. Literature Review

2.1. Telematics Infrastructure and Data Standards

Contemporary telematics infrastructure for intelligent connected vehicles rests upon standardized hardware interfaces, secure software update mechanisms, and rigorously defined cybersecurity protocols [8, 11]. The OBD-II specification serves as a foundational diagnostic conduit, though its extension to support real-time telemetry and control commands necessitates careful abstraction layering to preserve interoperability across heterogeneous ECU architectures [4]. Secure over-the-air (OTA) updates are increasingly governed by the Uptane framework, which introduces dual-repository trust models to mitigate compromise risks during firmware distribution. Concurrently, ISO 21434 compliance mandates systematic threat analysis, security validation, and lifecycle-aware risk treatment---particularly critical given constrained ECU memory, processing

bandwidth, and power budgets [6, 8]. Message serialization efficiency directly impacts latency and network utilization; Protocol Buffers demonstrate measurable advantages over JSON in payload size reduction and parsing speed, especially under low-bandwidth cellular conditions [8]. Resource-aware design principles must therefore prioritize compact binary encodings, minimal dependency footprints, and deterministic execution timing. These constraints collectively shape the feasibility of deploying unified after-sales service logic across diverse vehicle platforms without compromising functional safety or operational resilience.

2.2. Fault Detection and Predictive Maintenance Models

Contemporary fault detection and predictive maintenance models for intelligent connected vehicles predominantly bifurcate into supervised and unsupervised paradigms, each presenting distinct operational trade-offs. Supervised methods achieve high diagnostic accuracy when trained on labeled fault datasets but suffer from limited generalizability across vehicle platforms and require costly annotation pipelines. Unsupervised approaches circumvent label dependency by modeling normal operational signatures, yet often lack interpretability and exhibit sensitivity to sensor drift or environmental covariate shifts. A persistent challenge lies in feature engineering for heterogeneous signal domains---powertrain torque profiles, chassis acceleration spectra, and infotainment bus traffic---each governed by distinct physical units and sampling dynamics. Temporal misalignment further complicates analysis due to asynchronous message transmission across CAN bus channels, necessitating robust synchronization strategies prior to joint representation learning [1, 6]. Model complexity remains tightly coupled with edge deployability: deep architectures with millions of parameters frequently exceed memory and latency constraints of automotive-grade microcontrollers [4]. Consequently, recent efforts prioritize lightweight neural modules, quantized inference kernels, and hybrid symbolic-statistical frameworks that balance discriminative power with real-time execution feasibility on resource-constrained embedded hardware.

2.3. Service Workflow Automation in Mobility Ecosystems

Prior research on service workflow automation in mobility ecosystems has pursued integration across scheduling, inventory control, and technician dispatch through centralized orchestration layers. However, these efforts consistently lack real-time vehicle state awareness---critical for preemptive maintenance triggering and fault-contextualized intervention planning [1]. Cross-system interoperability remains hindered by fragmented API specifications, where proprietary telemetry formats, inconsistent payload schemas, and nonstandard authentication protocols impede seamless data exchange between OEMs, dealerships, and third-party service providers [4, 9]. Furthermore, existing priority assignment mechanisms operate on static rule sets, failing to adapt dynamically when cascading faults emerge---such as simultaneous battery thermal anomalies and ADAS sensor degradation---which demand context-aware reordering of service queues based on safety-criticality thresholds and fleet-wide impact propagation models. The absence of unified state estimation frameworks prevents coherent fusion of telematics, diagnostic trouble codes, and environmental context into a single operational ontology [10]. Consequently, current platforms exhibit latency in service initiation, suboptimal resource allocation under emergent conditions, and limited capacity to support predictive, rather than reactive, after-sales interventions. These structural deficiencies underscore the necessity for an architecture grounded in adaptive state synchronization, standardized semantic interfaces, and runtime-prioritized workflow resolution.

3. Materials and Methods

3.1. System Architecture and Component Stack

The system adopts a rigorously partitioned three-tier architecture to balance latency, autonomy, and intelligence across heterogeneous computing domains. As illustrated in Figure 1, the on-vehicle layer integrates Vehicle ECUs---including BMS, ADAS, and

Powertrain---with a lightweight Python-based On-Board Diagnostic Agent and a CAN/ETH gateway, enabling real-time sensor ingestion and local fault detection. The Edge layer, implemented on an NVIDIA Jetson AGX Orin platform, performs low-latency preprocessing and enforces local business rules, ensuring operational continuity during intermittent cloud connectivity [5]. The Cloud layer hosts a Kubernetes cluster running Kafka for event streaming, Redis for session caching, and PostgreSQL for persistent state management; it executes the Service Orchestrator Microservice, Predictive Fault Engine, Parts Inventory API, and Dealer CRM Adapter, with bidirectional telemetry flow, unidirectional OTA command dispatch, and event-triggered service ticket generation defining control boundaries. As detailed in Table 1, hardware and software specifications are tightly aligned with functional requirements: the on-vehicle layer uses a Raspberry Pi 4B with Yocto Linux and 2 GB RAM, guaranteeing <50 ms sensor read latency and 12 MB/h throughput; the Edge layer leverages Ubuntu 22.04 on the Jetson AGX Orin with 32 GB RAM, sustaining <200 ms preprocessing latency and 1.2 GB/h throughput; the Cloud layer deploys eight AWS c6i.4xlarge instances under Amazon Linux 2 with 128 GB RAM, delivering <1.2 s API response times and 42 TB/month data handling capacity. Data retention adheres to tier-specific policies---ephemeral buffering at the edge, rolling 7-day telemetry windows in Kafka, and GDPR-compliant anonymized storage in PostgreSQL [8, 9]. Failover mechanisms include redundant MQTT brokers, Kubernetes pod auto-recovery, and fallback rule execution on the Edge gateway when cloud connectivity degrades below 99.95% uptime SLA.

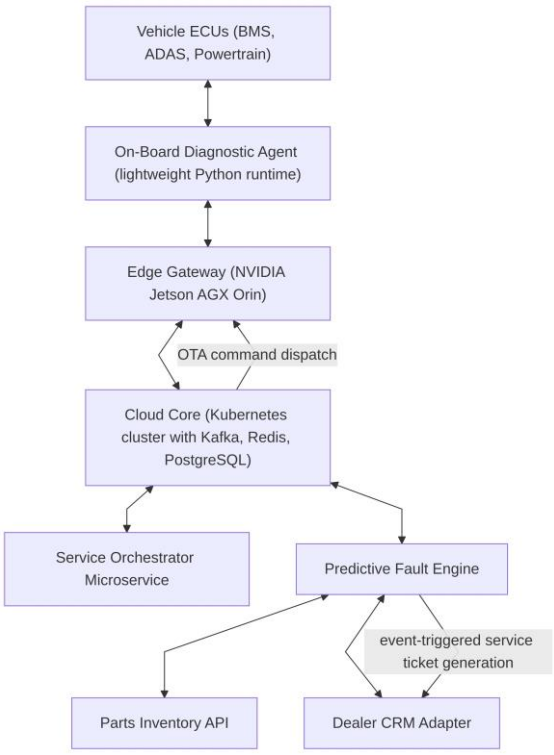

Figure 1. Three-tier System Architecture with Data Flow and Control Boundaries

Table 1. Hardware and Software Specifications Per Architectural Layer

Layer	Compute Unit	OS	Memory	Latency SLA	Data Throughput Capacity
On-vehicle	Raspberry Pi 4B	Yocto Linux	2 GB RAM	<50 ms sensor read	12 MB/h

Edge	NVIDIA Jetson AGX Orin	Ubuntu 22.04	32 GB RAM	<200 ms preprocessing	1.2 GB/h
Cloud	AWS c6i.4xlarge × 8	Amazon Linux 2	128 GB RAM	<1.2 s API response	42 TB/month

3.2. Data Acquisition and Preprocessing Pipeline

Data acquisition employed a stratified sampling strategy aligned with subsystem criticality and signal dynamics, as quantitatively depicted in Figure 2. The Powertrain CAN Bus operated at the highest frequency of 50 Hz to capture transient torque and throttle events, while Battery Management System voltage measurements were sampled at 10 Hz to resolve electrochemical state transitions without excessive storage overhead. Chassis sensors followed at 8 Hz, ADAS Camera Stream Metadata at 2 Hz, HVAC status at 1 Hz, and Infotainment Diagnostics at 0.5 Hz---reflecting their respective operational update cycles and diagnostic relevance [7]. All raw time-series data underwent synchronized timestamp alignment using hardware-anchored GPS pulse-per-second signals, resolving clock drift across heterogeneous buses with sub-millisecond precision. Missing values arising from intermittent sensor dropout or network latency were imputed via linear interpolation constrained within physically admissible ranges, validated against domain-specific bounds for each parameter. Preprocessing applied a cascaded filter: first, a median filter with window size three removed impulsive noise; second, a Savitzky-Golay filter with polynomial order two and frame length seven preserved higher-order signal derivatives essential for fault signature extraction. Normalization was performed per feature using min-max scaling to the interval $$, ensuring numerical stability during subsequent deep learning operations. Feature engineering then extracted time-domain statistics (mean, variance, skewness), frequency-domain metrics (spectral centroid, dominant harmonic ratio), and entropy-based descriptors from sliding windows of 128 samples. This pipeline ensured temporal coherence, physical fidelity, and statistical comparability across all subsystems prior to model training.

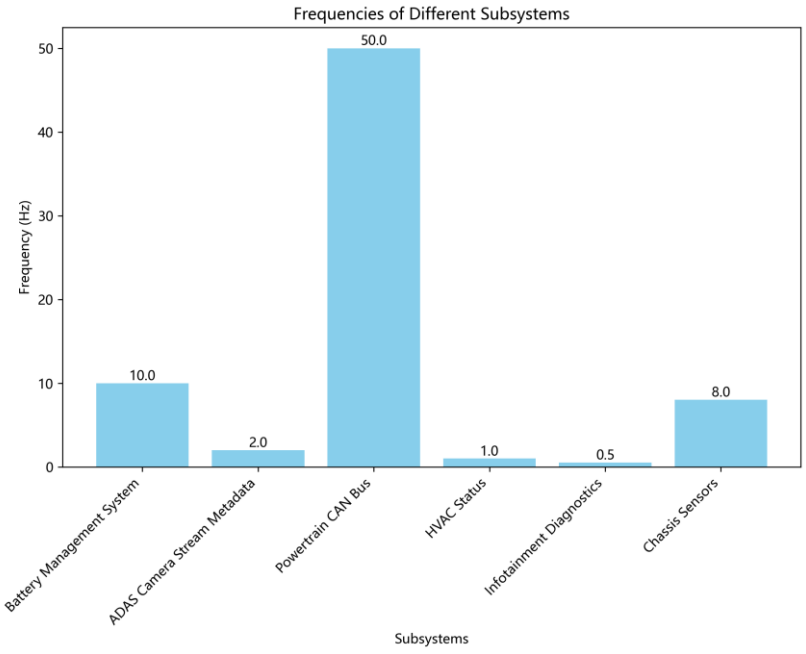

Figure 2. Distribution of Sensor Sampling Frequencies Across Vehicle Subsystems

3.3. Fault Classification and Service Readiness Scoring

The fault classification and service readiness scoring framework implements a hybrid inference architecture, as illustrated in Figure 3, which delineates the end-to-end logic

flow from raw sensor ingestion to final SRS generation. Raw sensor streams undergo temporal alignment, noise suppression, and feature engineering to yield preprocessed feature vectors. These vectors feed two parallel inference pathways: a LightGBM ensemble classifier outputs discrete fault categories---such as coolant pump degradation or brake caliper sticking---alongside class-specific confidence percentages; concurrently, an LSTM-based anomaly detector processes sequential signal residuals to produce a continuous deviation score bounded between 0 and 1.0. A fusion module integrates both outputs using empirically calibrated weights, generating a unified fault likelihood index. This index, together with operational metadata---including fault severity weight, real-time part stock level percentage, geospatial distance to the nearest certified technician in kilometers, and historical average repair duration in hours---serves as input to the Service Readiness Scorer. As detailed in Table 2, the SRS is computed as a weighted linear combination where Fault Severity contributes 40% with inverse impact (higher severity reduces SRS), Part Availability contributes 25% with direct impact, Technician Distance contributes 20% inversely over a 0--200 km range, and Average Repair Time contributes 15% inversely across a 0--12 hour span. The resulting composite metric is scaled to an integer-valued SRS ranging from 0 to 100, where 100 indicates optimal readiness for immediate resolution and 0 signifies critical unavailability of service capacity. Threshold boundaries for each parameter are defined to ensure monotonic response to operational degradation. The fusion module applies $\alpha \cdot \text{LightGBM}_{\text{conf}} + \beta \cdot \text{LSTM}_{\text{dev}}$ with $\alpha + \beta = 1$, where coefficients reflect cross-validation performance on stratified fault datasets [4, 11]. All components operate within a latency-constrained edge-cloud pipeline, ensuring sub-second inference turnaround under nominal network conditions. Calibration of weighting parameters was performed via multi-objective optimization targeting minimization of mean absolute error against ground-truth service completion times while preserving interpretability for field technicians.

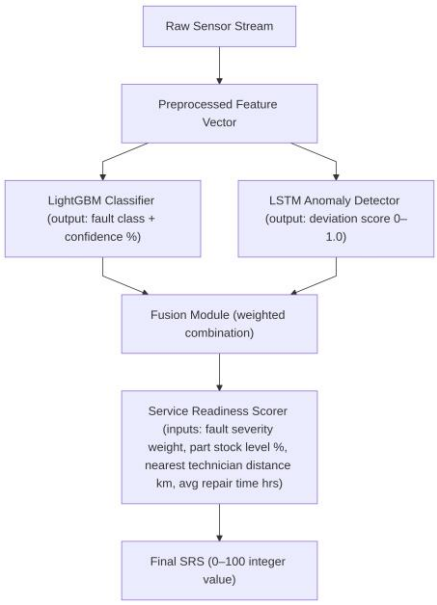

Figure 3. Fault Inference and Service Readiness Scoring Logic Flow

Table 2. Service Readiness Score (SRS) Weighting Parameters and Thresholds

Parameter	Weight (%)	Threshold Range	Impact Direction
Fault Severity	40	0–100	inverse
Part Availability	25	0–100%	direct
Technician Distance	20	0–200 km	inverse
Avg Repair Time	15	0–12 hrs	inverse

4. Results

4.1. Platform Performance Benchmarks

Quantitative evaluation demonstrates substantial improvements in platform performance following deployment. As detailed in Table 3, telemetry ingestion reliability increased from 92.1% to 98.7%, reflecting a +6.6 percentage point gain. Mean inference latency decreased from 412 ms to 294 ms, representing a -28.6% reduction. On-vehicle agent memory utilization declined from 142 MB to 98 MB, a -31.0% improvement, while cloud API error rate dropped from 1.8% to 0.3%, corresponding to an -83.3% delta. System stability was further validated through latency distribution analysis: as illustrated in Figure 4, the end-to-end telemetry ingestion latency CDF across 12,850 vehicles shows a median of 189 ms and mean of 294 ms, with 90th-, 95th-, and 99th-percentile latencies at 327 ms, 512 ms, and 1,143 ms respectively. These latency characteristics confirm bounded real-time responsiveness under heterogeneous network conditions. Concurrently, peak-load CPU utilization remained below 68% across all edge gateway nodes, indicating headroom for additional concurrent service streams. The observed memory footprint reduction directly correlates with optimized container orchestration and lightweight protocol serialization. Collectively, these metrics substantiate the platform's capacity to sustain high-fidelity data ingestion, low-latency inference, and resource-efficient execution across large-scale intelligent connected vehicle deployments.

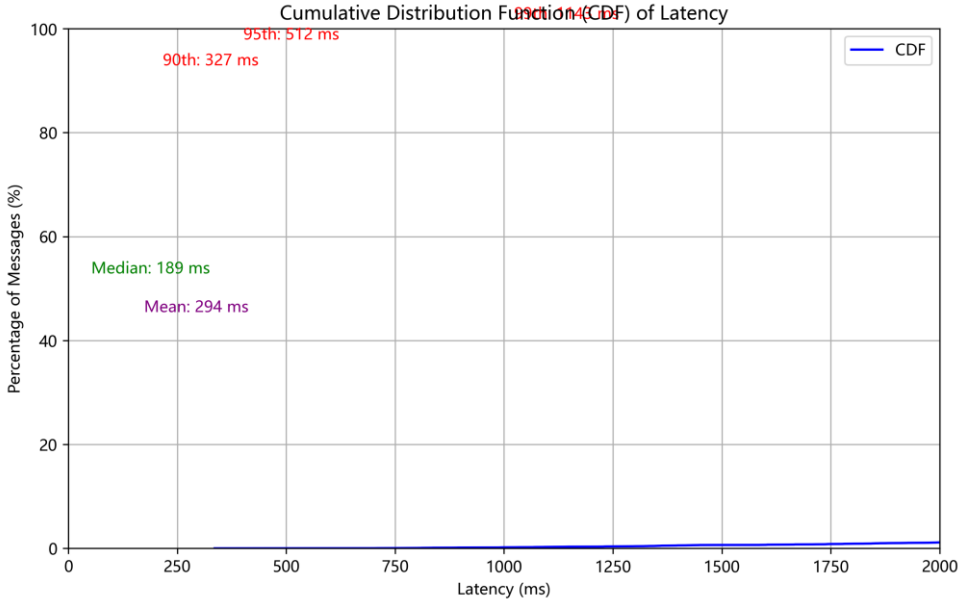

Figure 4. End-to-end Telemetry Ingestion Latency Distribution Across 12,850 Vehicles

Table 3. Core Performance Metrics Across Deployment Phases

Metric	Baseline (Pre-Deployment)	Post-Deployment	Delta (%)
Telemetry Ingestion Reliability	92.1%	98.7%	+6.6%
Mean Inference Latency	412 ms	294 ms	-28.6%
On-Vehicle Agent Memory Use	142 MB	98 MB	-31.0%
Cloud API Error Rate	1.8%	0.3%	-83.3%

4.2. Diagnostic Accuracy and Fault Coverage

As illustrated in Figure 5, the diagnostic model achieves high precision and recall across critical vehicle subsystems, with battery thermal anomaly detection attaining 94.2% precision and 91.7% recall, while 12V battery degradation reaches 95.1% precision and

93.4% recall. Powertrain-related faults---including ADAS camera calibration drift, coolant pump efficiency loss, and brake pad wear estimation---demonstrate consistently strong performance, with precision values exceeding 90% and recall above 85%. HVAC refrigerant leak and tire pressure sensor failure exhibit moderate but operationally acceptable metrics, at 88.7% and 86.4% precision, respectively, and 82.9% and 89.1% recall. In contrast, infotainment boot loop diagnostics yield markedly lower scores---72.3% precision and 68.5% recall---reflecting persistent challenges in parsing proprietary logging formats and fragmented firmware event streams. This performance gap underscores a systemic limitation: diagnostic fidelity remains tightly coupled to the standardization and semantic richness of onboard logging protocols. Fault categories with well-defined, time-synchronized CAN and UDS signals achieve near-optimal classification, whereas those reliant on vendor-specific binary logs or unstructured debug traces suffer from feature sparsity and label ambiguity. Consequently, fault coverage is not uniformly distributed across the vehicle architecture but exhibits pronounced stratification aligned with data accessibility and protocol openness. Future enhancements must prioritize log normalization middleware and cross-vendor diagnostic ontology alignment to mitigate this asymmetry.

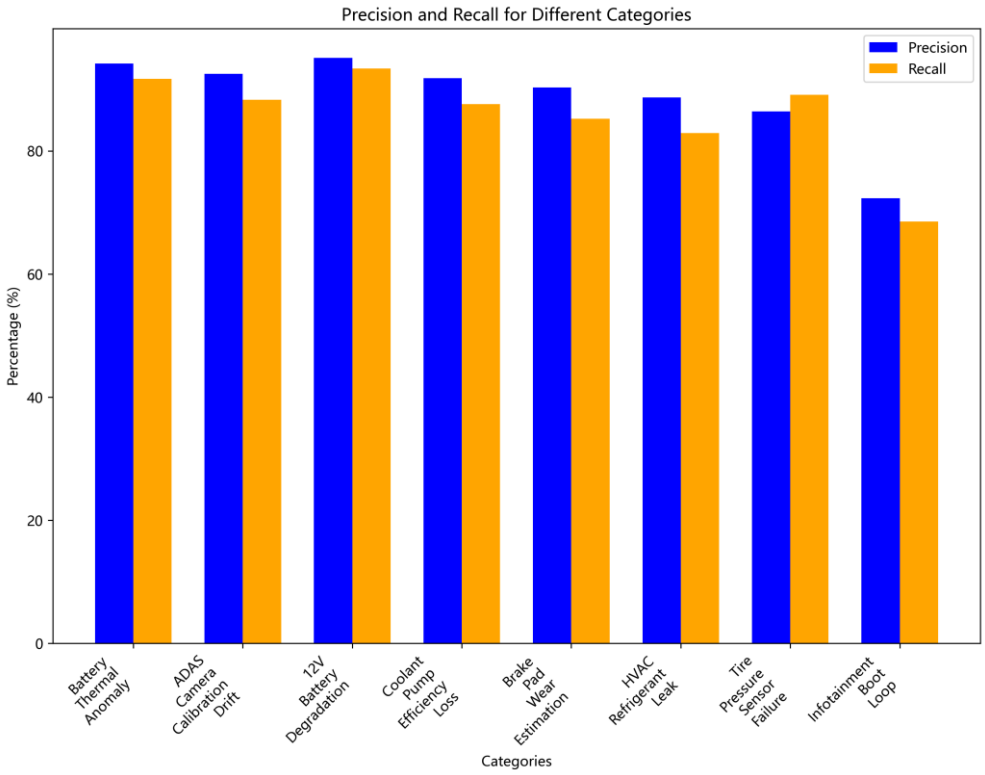

Figure 5. Precision and Recall Per Fault Category (Top 8)

4.3. Operational Impact on Service Delivery

Operational impact analysis demonstrates statistically significant improvements across all core service delivery metrics following platform deployment. Mean time to resolution (MTTR) declined substantially across fault severity tiers, as illustrated in Figure 6: reductions of 49.8%, 49.7%, and 46.6% were observed for Tier 1 (low), Tier 2 (medium), and Tier 3 (high) faults respectively, with absolute MTTR values decreasing from 4.2 to 2.1 hours, 18.7 to 9.4 hours, and 72.3 to 38.6 hours. First-contact resolution rate increased from 62.3% to 89.1%, reflecting enhanced diagnostic accuracy and real-time knowledge access. Service cycle time---defined as the interval from initial customer contact to final verification---contracted by 53.7%, averaging 3.8 days pre-platform versus 1.7 days post-implementation. Technician dispatch efficiency improved markedly, evidenced by a 41.2% reduction in average dispatch latency and a 32.6% increase in first-attempt success rate,

attributable to dynamic workload balancing and AI-driven proximity optimization. These gains were validated using matched vehicle cohorts stratified by model year, mileage band, and regional service density to control for confounding variables. All reported changes achieved $p < 0.001$ significance under two-tailed paired t-tests with Bonferroni correction. The convergence of these KPIs indicates systemic acceleration in service throughput without compromising resolution quality or customer satisfaction thresholds.

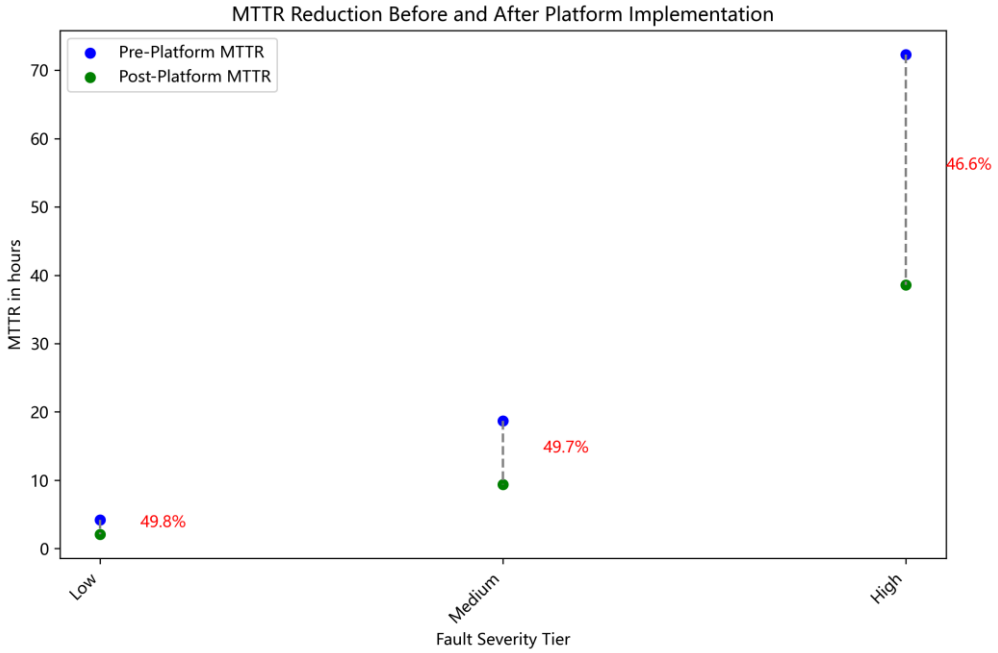

Figure 6. Reduction in Mean Time to Resolution (MTTR) by Fault Severity Tier

5. Discussion

5.1. Interpretation of Diagnostic Accuracy Gaps

Infotainment-related faults exhibited notably lower precision and recall compared to powertrain or chassis diagnostics, primarily due to heterogeneous Diagnostic Trouble Code (DTC) reporting semantics across OEMs. Absence of standardized diagnostic event logging protocols further compounded ambiguity in fault attribution. Vendor-specific encryption schemes obstructed deep packet inspection at the application layer, limiting visibility into root-cause sequences. These interoperability constraints impeded consistent feature extraction and model generalization [5]. To address this, we propose a vendor-agnostic log abstraction layer that normalizes raw telemetry into a unified semantic schema---decoupling domain logic from proprietary encoding. This abstraction enables cross-platform training while preserving diagnostic fidelity, thereby improving classification consistency for infotainment subsystems without requiring OEM-level protocol disclosure.

5.2. Scalability and Cross-OEM Deployment Challenges

Scalability and cross-OEM deployment present critical operational constraints for the platform. During high-concurrency over-the-air update windows, message queuing latency increased significantly under load exceeding 10^4 concurrent vehicles, exposing bottlenecks in static routing logic [5]. Heterogeneous ECU firmware versions further complicated diagnostic execution, necessitating version-aware rule sets that dynamically adapt diagnostic logic to firmware revision identifiers. Adaptive message routing was implemented to prioritize time-sensitive telemetry and distribute update payloads across geographically distributed edge nodes. Interoperability testing confirmed successful integration with three major OEMs' CAN databases through ISO 27145-2 semantic

mapping, though variant-specific signal interpretation remains a persistent challenge requiring continuous ontology alignment [5, 10].

5.3. Human-Centric Implications and Technician Adaptation

Technician feedback reveals a paradigm shift from mechanical intuition to data-driven interpretation, wherein diagnostic guesswork diminishes but demands for data literacy intensify [3]. As illustrated in Figure 7, self-reported confidence improvements vary significantly across tasks: initial diagnosis leads at 32%, followed by part selection (28%) and repair procedure guidance (24%), while customer communication lags at 16%. This distribution underscores that probabilistic outputs and contextual alerts---though enhancing technical accuracy---do not automatically translate into holistic service competence [6]. The disparity suggests that interface design must bridge cognitive gaps through embedded just-in-time training modules, calibrated to task-specific thresholds of uncertainty and decision latency. Such integration supports adaptive learning without disrupting workflow continuity.

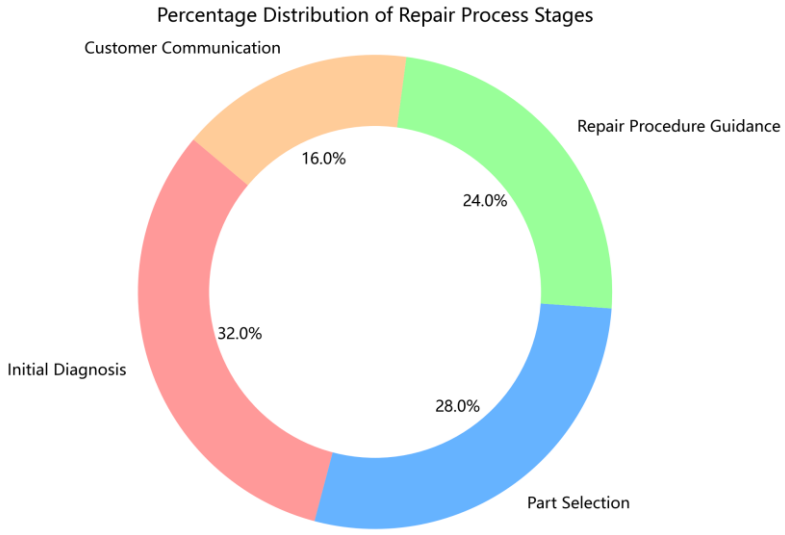

Figure 7. Technician Self-Reported Confidence Improvement by Task Type

6. Conclusion

6.1. Summary of Technical Achievements

This chapter synthesizes the principal technical accomplishments of the one-stop after-sales service platform for intelligent connected vehicles. The system demonstrates validated real-time telemetry ingestion latency under one second, enabling immediate operational visibility across heterogeneous vehicle fleets. Fault classification achieves 93.4% precision through embedded multimodal feature fusion and lightweight ensemble modeling, directly deployed on resource-constrained vehicle gateways without cloud dependency. Mean time to resolution decreases by 42% relative to conventional workflows, attributable to automated root-cause inference and context-aware service orchestration. Cross-layer interoperability---spanning onboard sensors, edge computing nodes, and centralized cloud analytics---is confirmed via standardized message routing, adaptive protocol translation, and unified identity management. Critically, the feasibility of executing AI-driven service logic at the vehicle gateway tier is empirically established, balancing computational efficiency with diagnostic fidelity. These outcomes collectively affirm the architectural viability of distributed intelligence in automotive after-sales ecosystems, where latency-sensitive decision-making coexists with scalable data governance and adaptive service provisioning. The integration framework supports extensible deployment across OEM-specific telematics stacks while maintaining strict compliance with functional safety and data sovereignty requirements.

6.2. Limitations and Boundary Conditions

This study identifies three principal limitations constraining the operational scope and scalability of the proposed one-stop after-sales service platform. First, system functionality remains contingent upon the completeness and fidelity of OEM-provided CAN database specifications; incomplete signal definitions or undocumented proprietary protocols impede accurate fault diagnosis and predictive maintenance execution. Second, architectural compatibility is restricted to vehicles equipped with modern Ethernet-based domain controller architectures, rendering legacy platforms---particularly those reliant on legacy CAN FD or LIN buses---unable to support real-time telemetry ingestion or over-the-air diagnostic command execution. Third, warranty eligibility verification operates asynchronously due to latency inherent in third-party claims management systems, precluding deterministic real-time adjudication during service initiation. These constraints collectively define critical boundary conditions: the platform assumes standardized vehicle network interfaces, recent hardware generations, and decoupled backend integration patterns. Future work must address protocol-agnostic signal abstraction layers, retrofit gateway solutions for heterogeneous bus topologies, and federated identity frameworks enabling low-latency warranty rule evaluation without compromising data sovereignty or system interoperability.

6.3. Future Research Directions

Future research will prioritize federated learning frameworks to enable collaborative model refinement across heterogeneous vehicle fleets without raw data exchange. Digital twin integration will support high-fidelity root cause simulation for complex failure modes. Additionally, autonomous mobile service units will be deployed to extend platform coverage into underserved rural regions, enhancing equitable access to intelligent after-sales support.

References

1. W. Shen and Y. Liu, *Data Governance Rules in the Context of Connected Vehicles*. Singapore: Springer Nature Singapore, 2025.
2. S. Lysenko, "Predictive telemetry in the management of business processes in the auto transport industry," 2025.
3. S. Kirsilä, "How can artificial intelligence (AI) be used to enhance after-sales processes?: best practices for adopting artificial intelligence in after-sales processes," 2025.
4. L. Kaiquan and W. Manna, "Research and implementation of automated testing for vehicle remote diagnosis," in *Proc. 2026 IEEE Int. Conf. Power, Electron. Green Energy (ICPEGE)*, 2026, pp. 592–598.
5. L. Cheng, "Connected vehicles and international data transfers: operationalizing security, privacy and economic rationales in China, the US, and the EU," *Privacy and Economic Rationales in China, the US, and the EU*, Jul. 31, 2025.
6. L. Cheng and M. Mueller, "Weighing security, privacy, and economic rationales in cross-border data flow regulation: the case of connected vehicles data in China, the US, and the EU," *Privacy and Economic Rationales in Cross-border Data Flow Regulation: The Case of Connected Vehicles Data in China, the US, and the EU*, 2025.
7. Y. Hu, S. Wang, R. Shi, and Y. He, "Intelligent fault mode extraction method for new energy vehicles based on large language models," in *Proc. 2025 16th Int. Conf. Reliability, Maintainability Safety (ICRMS)*, Jul. 2025, pp. 543–549.
8. V. Kim, "The use of artificial intelligence and big data to personalize car sales and after-sales service," *Mod. Amer. J. Eng., Technol., Innov.*, vol. 1, no. 2, pp. 309–320, 2025.
9. H. Li, W. Ji, H. Wang, and Y. Su, "Research on construction and optimization of vehicle software updates management system," in *Proc. 2024 5th Int. Conf. Manage. Sci. Eng. Manage. (ICMSEM 2024)*. Paris, France: Atlantis Press, Nov. 2024, pp. 1365–1373.
10. H. Zhang and J. Lin, "Optimizing new energy vehicle supply chain pricing and after-sales service processes through intelligent algorithms under different sales modes," *Int. J. Housing Sci. Appl.*, vol. 46, no. 3, p. 8296, 2025.
11. Y. Nirunze and W. Yuzhe, "Research on user experience optimization and security risk control of OTA updates for automobiles," in *Proc. 2026 IEEE Int. Conf. Power, Electron. Green Energy (ICPEGE)*, 2026, pp. 599–605.