

Article

Development and Validation of an Explainable Hybrid Deep Learning Model for Multiple-Fault Diagnosis in Intelligent Automotive Electronic Systems

Chien-Yu Lu ¹, Hong-Yi Hsu ^{1,2}, Bo-Siang Chen ^{2,*} , Wei-Lun Huang ^{1,3}  and Wei-Sho Ho ^{1,4,5,*} 

¹ Department of Electrical and Mechanical Technology, National Changhua University of Education, Bao-Shan Campus, Changhua City 500208, Taiwan; lcy@cc.ncue.edu.tw (C.-Y.L.); d1031017@gm.ncue.edu.tw (H.-Y.H.); u107075701@cmu.edu.tw (W.-L.H.)

² Department of Vehicle Engineering, Nan Kai University of Technology, Nantou City 542020, Taiwan

³ Medical Affairs Office, National Taiwan University Hospital, Taipei City 100225, Taiwan

⁴ Graduate Institute of Technological and Vocational Education, National Changhua University of Education, Bao-Shan Campus, Changhua City 500208, Taiwan

⁵ NCUE Alumni Association, National Changhua University of Education Jin-De Campus, Changhua City 500207, Taiwan

* Correspondence: toyota554@gmail.com (B.-S.C.); homaintain@gmail.com (W.-S.H.)

Abstract

This study addresses the increasingly complex challenge of multiple-fault diagnosis in modern intelligent automotive electronic systems by proposing an innovative deep learning-based solution. The research integrates Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and the Transformer architecture to construct a multi-modal fault diagnosis model. By collecting real-world operational data from vehicle electronic systems, including fault samples from key modules such as the Engine Control Unit (ECU), Body Control Module (BCM), and safety systems, a comprehensive dataset comprising 12 major fault types was established. Experimental results demonstrate that the proposed hybrid deep learning model achieves a multiple-fault identification accuracy of 96.8%, representing a 23% performance improvement over traditional diagnostic methods. The integration of Explainable AI (XAI) techniques provides the diagnostic results with visual interpretability, aiding maintenance technicians in understanding the model's diagnostic logic. The findings of this research can be applied in smart factories, automotive service centers, and on-board diagnostic (OBD) systems, offering significant practical value in enhancing vehicle safety and reducing maintenance costs.

Keywords: deep learning; fault diagnosis; automotive electronic systems; Explainable Artificial Intelligence (XAI); multi-modal fusion



Academic Editor: Sergio Busquets-Monge

Received: 25 September 2025

Revised: 9 November 2025

Accepted: 12 November 2025

Published: 17 November 2025

Citation: Lu, C.-Y.; Hsu, H.-Y.; Chen, B.-S.; Huang, W.-L.; Ho, W.-S. Development and Validation of an Explainable Hybrid Deep Learning Model for Multiple-Fault Diagnosis in Intelligent Automotive Electronic Systems. *Electronics* **2025**, *14*, 4488. <https://doi.org/10.3390/electronics14224488>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the automotive industry's progression towards intellectualization, electrification, and automation, modern vehicles have evolved from traditional mechanical drive systems into highly integrated intelligent electronic systems. Market research indicates that contemporary luxury cars contain over 100 Electronic Control Units (ECUs), while standard passenger vehicles are equipped with approximately 50–70 ECUs. These systems encompass critical modules such as the Engine Management System (EMS), Anti-lock Braking System (ABS), Electronic Stability Control (ESC), and Advanced Driver-Assistance Systems (ADAS). The complexity of intelligent automotive electronic systems is manifested not

only in the increased quantity of hardware but also in the high degree of interdependence among various subsystems. For instance, the Engine Control Unit must exchange data in real-time with the transmission control system, vehicle stability system, and emission control system. A failure in any single component can trigger a chain reaction, leading to simultaneous anomalies across multiple systems. Furthermore, with the proliferation of Internet of Vehicles (IoV) technology, automotive electronic systems must also communicate with external cloud services, transportation infrastructure, and other vehicles, further augmenting system complexity [1], as summarized in Figure 1.

Complexity of Modern Automotive Electronic Systems

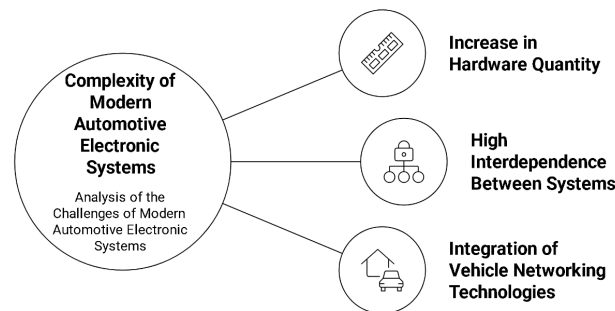


Figure 1. Complexity landscape of modern automotive electronic systems and cross-system interactions.

The figure illustrates the major sources of complexity in intelligent automotive electronic systems, including increased hardware quantity, high interdependence between subsystems, and integration with vehicle networking technologies.

Traditional automotive fault diagnosis relies primarily on the On-Board Diagnostics (OBD) framework, which detects anomalies via predefined Diagnostic Trouble Codes (DTCs). However, as system interconnections grow denser, conventional methods face significant limitations. Their single-point detection logic cannot efficiently handle correlated or cascading faults. When concurrent anomalies occur in several subsystems, a large number of misleading fault codes are often generated, making root-cause identification difficult for technicians [2]. Moreover, static-threshold strategies cannot adapt to varying driving conditions, environmental factors, or vehicle aging, frequently resulting in false positives or negatives [3]. Consequently, traditional approaches remain reactive, lacking predictive capability and failing to support preventive maintenance.

The rapid development of deep learning technology offers a new opportunity to address the challenges of fault diagnosis in automotive electronic systems. Compared to traditional methods, deep learning presents significant advantages. Its powerful automatic feature learning capability enables deep neural networks to autonomously learn complex feature representations from raw data without the need for laborious manual feature engineering. This characteristic is particularly well-suited for handling the high-dimensional, non-linear data produced by automotive electronic systems [4]. Moreover, deep learning models can effectively integrate multi-modal data generated by these systems, including time-series signals, image data, and frequency-domain features, thereby providing a more comprehensive perspective for fault diagnosis. Through training on extensive historical fault data, deep learning models can not only identify complex fault patterns but also possess considerable predictive capabilities, facilitating the crucial shift from reactive diagnosis to proactive prevention [5].

In response to the increasingly complex multiple-fault challenges in modern automotive electronic systems, this study aims to develop and validate an intelligent diagnostic system to accurately identify and locate the root causes of multiple faults. The scope of the

research focuses on the primary electronic systems of passenger vehicles, including the Electronic Control Unit (ECU), Body Control Module (BCM), Anti-lock Braking System (ABS), and Electronic Stability Control (ESC). A thorough analysis of the fault interaction effects among key subsystems such as the engine, emissions, fuel, and transmission will be conducted to construct a cross-system fault correlation model. To validate diagnostic performance, 12 common fault categories, including those related to sensors, actuators, communication, and software, will be used as test cases to comprehensively evaluate the model's accuracy and response efficiency. The objective is to develop a high-precision, rapid-response diagnostic method to effectively reduce maintenance costs and misdiagnosis rates, thereby enhancing customer satisfaction and promoting the professionalization and standardization of the relevant industries.

2. Literature Review

2.1. The Evolution of Automotive Fault Diagnosis Technology

The development of fault diagnosis technology for automotive electronic systems can be traced back to the 1980s with the introduction of the On-Board Diagnostics (OBD-I) system. Early diagnostic systems primarily relied on predefined Diagnostic Trouble Codes (DTCs) to indicate the abnormal status of specific components, offering relatively basic functionality [6,7]. The establishment of the OBD-II standard in the 1990s marked a significant milestone in automotive diagnostic technology, as it standardized the diagnostic interface and communication protocols, leading to a more streamlined diagnostic process. With the continuous advancement of automotive electronics and intellectualization, diagnostic technology has evolved from simple fault code detection to more sophisticated system-level diagnosis. Modern diagnostic systems are capable not only of detecting hardware failures but also of identifying software errors, network communication issues, and the complex interactions among various Electronic Control Units (ECUs) [8]. However, traditional rule-based diagnostic methods often prove inadequate when faced with increasingly complex fault modes. In recent years, Model-Based Diagnosis (MBD) has garnered considerable attention within the academic community. This approach involves creating a mathematical model of the system to predict its normal behavior and detecting faults by comparing the actual output with the model's expected output [2,9]. Nevertheless, the practical application of MBD methods still faces challenges such as high model construction complexity and significant computational costs.

To ensure robustness and representativeness, the dataset used in this study covers 20 vehicle models across seven major brands commonly available in Taiwan, including Toyota, Nissan, Honda, Mitsubishi, Mazda, Hyundai, and Luxgen, with model years ranging from 2012 to 2024. Operating conditions encompass idling, urban stop-and-go traffic, highway cruising, hill climbs, and air-conditioning on/off transitions. Environmental variations span ambient temperatures from -5°C to 38°C and relative humidity between 35% and 95%. Fault categories include sensor faults* (O_2 , MAF/MAP, TPS), actuator faults (injector, ignition coil), fuel-system anomalies (pump, rail-pressure fluctuation), communication issues (CAN latency and packet loss), software faults (DTC mapping errors), and composite fault chains* involving multiple subsystems. All on-vehicle recordings followed a standardized protocol* with synchronized multi-sensor logging and OBD-II DTC snapshotting to guarantee data consistency and traceability.

2.2. Application of Deep Learning in Fault Diagnosis

The application of deep learning technology in the field of industrial equipment fault diagnosis began in the early 2010s and has rapidly matured with significant advancements in computational power and big data technologies. In the domain of automotive engineer-

ing, research utilizing deep learning for fault diagnosis has primarily focused on several key directions. Among them, Convolutional Neural Networks (CNNs) have been widely applied due to their excellent performance in processing one-dimensional time-series signals and two-dimensional spectrograms. For example, recent studies transformed vibration signals into time-frequency spectrograms for CNN-based classification with strong accuracy [10,11]. Given the prominent temporal characteristics of data from automotive electronic systems, Recurrent Neural Network (RNN) architectures, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), are particularly well-suited for capturing the temporal dependencies within these signals. Recent LSTM architectures have achieved robust SOH estimation for lithium-ion batteries, including attention-enhanced variants [12,13]. Furthermore, Autoencoders have demonstrated great potential in anomaly detection. By learning the latent feature representations of normal data, they can effectively identify abnormal data patterns. Unsupervised VAEs have been successfully applied to multivariate sensor anomaly detection, including automotive testing scenarios [14,15].

2.3. The Role of Explainable AI in Diagnostic Models

Although deep learning models have demonstrated powerful performance in fault diagnosis tasks, their inherent “black-box” nature poses a significant challenge in safety-critical automotive applications. To enable maintenance technicians to trust the model’s outputs and make correct repair decisions, it is crucial to understand the basis for the model’s diagnostic conclusions. Consequently, the application of Explainable AI (XAI) technology in automotive fault diagnosis has gained prominence [2]. Among the current mainstream XAI techniques, LIME (Local Interpretable Model-agnostic Explanations) approximates the predictive behavior of complex models by generating local linear models near the decision boundary. In fault diagnosis applications, LIME can effectively identify the input features that are most influential for a specific fault determination [16]. Another widely adopted technique is SHAP (SHapley Additive exPlanations), which is based on the Shapley values from game theory. It assigns an importance score to each input feature to quantify its contribution to the final prediction. Compared to LIME, the explanations provided by SHAP are considered to be more consistent and reliable in theory [17]. Additionally, in models that employ the Transformer architecture, the built-in Attention Mechanism naturally provides a means for visualization. By analyzing the distribution of attention weights during inference, researchers can directly gain insight into which time points or specific feature dimensions the model focuses on when making a diagnostic decision, thereby offering an intuitive explanation of the model’s decision-making process.

2.4. Data Reliability and Sampling Strategy

Ensuring the reliability and generalizability of diagnostic results requires careful data preparation and partitioning. This study employed several measures to mitigate data imbalance, sampling bias, and data leakage.

- (1) Class imbalance mitigation: Since certain fault categories were underrepresented, the Synthetic Minority Oversampling Technique (SMOTE) was used to generate synthetic minority samples. Furthermore, a focal loss function was adopted to emphasize misclassified or hard-to-learn samples, preventing bias toward dominant classes.
- (2) Sampling bias reduction: To avoid overfitting to specific vehicle types or production batches, a cross-vehicle stratified sampling strategy was implemented, ensuring each dataset split (training, validation, and testing) contained proportionate data from various vehicle models. A 5-fold cross-validation scheme was also applied to improve statistical robustness and reliability.

- (3) Data leakage prevention: To prevent information leakage between splits, all temporal sequences from the same vehicle were constrained to a single dataset partition. This guarantees that no overlapping data segments appear in both the training and testing sets, ensuring that model evaluation reflects real-world generalization.

These measures collectively enhance data integrity, minimize bias, and ensure that the model's reported performance is both reproducible and trustworthy.

3. Methodology

3.1. Data Collection and Processing

To construct a comprehensive and robust diagnostic model, this study employed a multi-source data collection strategy integrating three distinct types of data. The experimental setup consisted of both simulation and physical vehicle experiments. In the simulation stage, high-fidelity vehicle models were developed in MATLAB/Simulink version R2024a and CarSim version 2025.2, and faults were injected into the virtual ECU, sensor, and communication networks under controlled ambient conditions (20–35 °C, 12 V system voltage). During the physical tests, vehicles of different brands were connected to a DAQ system (NI PXI-1082) to record real-time signals at a 500 Hz sampling rate. All signal preprocessing and labeling were performed under ISO 26262-compliant laboratory procedures [18]. The MATLAB/Simulink model includes the engine torque, throttle, ignition, and fuel injection subsystems, while CarSim provides a dynamic vehicle model representing longitudinal and lateral vehicle behavior. The co-simulation interface exchanges variables such as engine speed, wheel speed, and throttle position in real time. Faults were injected via Simulink's Fault-Injection blocks, and the resulting transient responses—voltage, current, and pressure waveforms—were exported through the Simulink Data Inspector for subsequent model training. Laboratory simulation data was then produced through precise Fault Injection techniques [19]. Second, to acquire data that more closely represents real-world physical conditions, this research involved collaboration with automotive manufacturers to conduct fault injection tests on physical vehicles spanning multiple brands and models in a controlled environment, thereby recording authentic system responses [20]. Finally, to supplement the dataset with real-world fault distributions and complex scenarios, extensive historical data was collected from maintenance and repair shops, including Diagnostic Trouble Codes (DTCs), repair reports, and parts replacement records [21].

In accordance with international standards and industry practices, the fault dataset in this study encompasses 12 major fault categories. These include sensor faults, actuator faults, communication failures, software errors, ECU hardware failures, power system faults, wiring harness issues, interface faults, calibration errors, environment-related faults, aging and degradation failures, and complex multiple faults [20]. All collected data samples were annotated by senior automotive diagnostic experts, and a cross-validation procedure was implemented to ensure the accuracy and consistency of the labels. Prior to model training, the raw data underwent a rigorous preprocessing pipeline, which involved removing outliers and irrational values, as well as appropriately handling missing values and signal noise. To extract more representative information from the data, a multi-level feature engineering process was constructed. This process not only calculated statistical features in the time domain (e.g., mean, standard deviation) but also extracted frequency-domain features via Fast Fourier Transform (FFT) and time-frequency features using wavelet analysis, while simultaneously analyzing the correlation features among different sensor signals [22].

The overall real-time automotive fault diagnosis workflow, from multi-source data acquisition to explainable decision outputs, is illustrated in Figure 2.

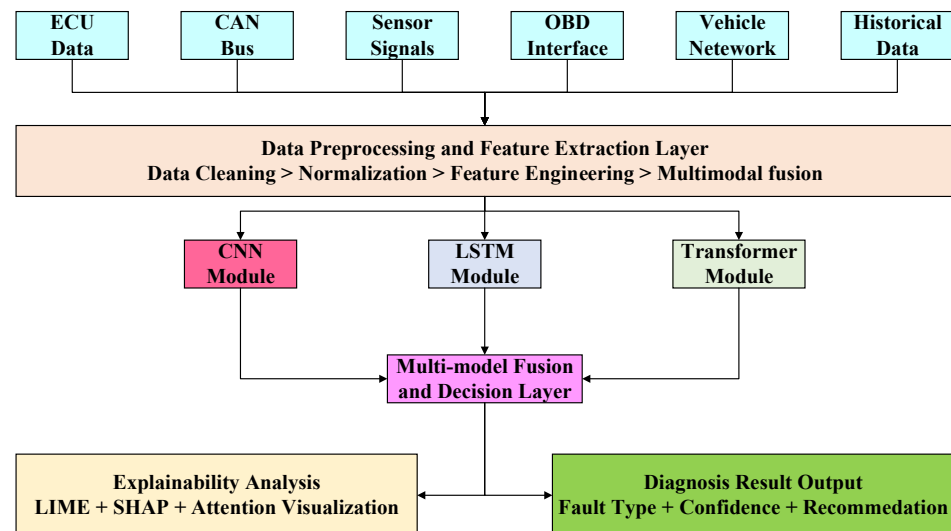


Figure 2. Architecture of the real-time automotive fault diagnosis system.

3.2. Deep Learning Model Design

The proposed diagnostic framework adopts a hybrid CNN–LSTM–Transformer architecture, as illustrated in Figure 3. The design follows a multi-stage data-flow pipeline that emulates how real sensor signals are collected and processed in intelligent vehicles. In the data-preprocessing stage, raw multi-sensor time-series signals are synchronized, filtered, and normalized to ensure consistency across subsystems. The CNN layers then extract spatial–temporal patterns from short-window segments, capturing localized behaviors such as transient spikes or oscillations in voltage, pressure, or temperature. The extracted feature maps are passed to the LSTM module, which models long-term dependencies and sequential correlations among signals. Subsequently, the Transformer encoder applies a self-attention mechanism to emphasize cross-system interactions and global dependencies that often characterize compound faults.

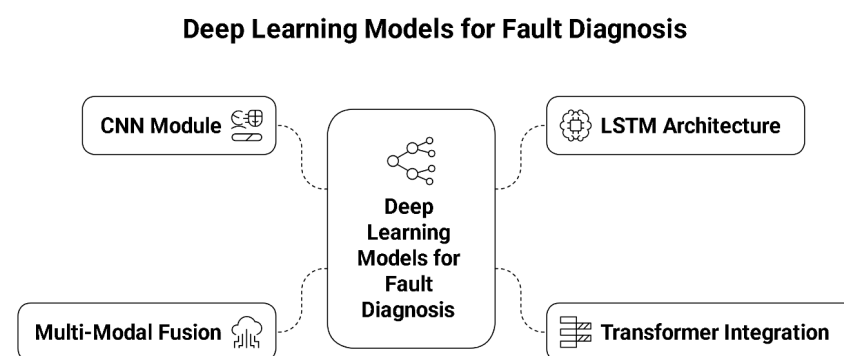


Figure 3. Architecture of the proposed deep learning model.

This study designs a hybrid deep-learning model aimed at effectively processing the complex multi-modal and time-series data found in automotive electronic systems. The core objective of this model architecture is to enhance the depth and breadth of feature extraction, accurately capture long-term dependencies within the signals, and effectively utilize complementary information from various data modalities [23]. The model’s design primarily integrates Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and the Transformer architecture. For one-dimensional time-series signals and two-dimensional time-frequency image inputs, a multi-scale convolutional structure was

designed, combined with Residual Connections and Batch Normalization, to strengthen the extraction of deep features and ensure the stability of the training process [24].

To capture temporal dependencies in the data, the model adopts a bidirectional and stacked LSTM architecture, incorporating an attention mechanism to focus on the most critical time steps for fault determination, thereby enhancing model interpretability. To address the challenges that traditional RNNs may encounter when processing extremely long sequences, the model further integrates the Multi-head Self-Attention mechanism and learnable positional encodings from the Transformer architecture, similarly combined with residual connections and Layer Normalization to improve training efficiency. To fully leverage the value of different data sources, the model employs a hybrid multi-modal fusion strategy. Through various approaches such as early fusion, late fusion, and hybrid fusion, heterogeneous information, including time-series signals, DTCs, and vehicle status data, is integrated to capitalize on the complementary effects across different data modalities [25].

To highlight the unique technical contribution of this study, the proposed architecture introduces a triple-level fusion mechanism that synergistically combines the spatial abstraction capability of CNN, the temporal dependency modeling of LSTM, and the contextual awareness of the Transformer. Unlike previous hybrid CNN–LSTM approaches that concatenate features at a single stage, our model incorporates an adaptive attention-guided fusion layer that dynamically adjusts feature weights across temporal and frequency domains. Furthermore, a multi-modal explainability integration layer combining LIME, SHAP, and attention visualization is embedded to provide transparent and interpretable diagnostics across interdependent vehicle subsystems. This architectural synergy not only improves diagnostic accuracy but also enhances transparency and reliability, addressing the interpretability gap in existing deep-learning-based fault diagnosis frameworks.

The framework integrates heterogeneous data sources (ECU, CAN, sensors, OBD, network, historical data) through preprocessing, feature extraction, and multi-modal fusion layers, leading to decision outputs with explainability modules (LIME, SHAP, attention visualization).

Each encoder branch (CNN, LSTM, Transformer) captures different data characteristics—spatial, temporal, and contextual—before entering an adaptive fusion layer. The output layer provides fault classification, confidence estimation, and interpretability results.

3.3. Explainable AI Integration

To overcome the inherent “black-box” problem of deep learning models and to enhance the transparency and trustworthiness of diagnostic decisions, this study deeply integrates Explainable AI (XAI) techniques into the diagnostic system. This integration enables engineers and maintenance technicians to better understand model behavior, validate diagnostic results, and identify potential model biases or vulnerabilities. The study primarily applies three mainstream XAI techniques: LIME (Local Interpretable Model-agnostic Explanations), SHAP (SHapley Additive exPlanations), and attention mechanism visualization. LIME is mainly used to explain the model’s local diagnostic decisions for individual samples by providing a ranking of the most influential features for that decision and a visual interface. SHAP, based on game theory, can provide both global feature importance analysis and precise explanations for individual samples, offering deep insights into the contribution of each feature and its interaction effects. By leveraging the model’s built-in attention mechanism, techniques such as plotting heatmaps, conducting multi-head attention analysis, and hierarchical visualization are used to intuitively reveal how the model’s attention is distributed across the time series and feature dimensions when making a fault diagnosis. The integrated XAI framework and its interaction with the diagnostic pipeline are summarized in Figure 4.

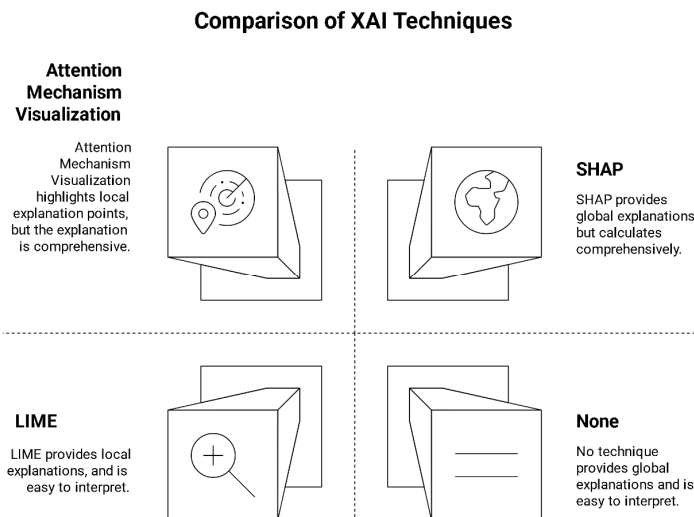


Figure 4. Integration of Explainable AI (XAI).

3.4. Model Convergence and Computational Analysis

To strengthen the theoretical foundation of this study, we analyzed the convergence behavior, computational complexity, and real-time performance bounds of the proposed hybrid architecture.

- (1) **Convergence verification:** The model's training process exhibited stable convergence within approximately 80 epochs, as evidenced by the smooth decrease and subsequent stabilization of both training and validation loss curves. No signs of gradient explosion or overfitting were observed, which indicates that the proposed optimization setup—combining AdamW optimizer with a cosine annealing learning rate schedule—provides stable training dynamics.
- (2) **Computational complexity:** The Transformer encoder dominates the computational cost of the overall architecture, with a complexity of $O(n \cdot d^2)$, where n represents the sequence length and d denotes the feature dimension. In contrast, the CNN-LSTM baseline maintains a lower complexity of $O(n \cdot d)$. Despite this, GPU-accelerated matrix operations significantly mitigate the computational burden, enabling real-time performance even under multi-sensor inputs.
- (3) **Real-time performance bounds:** On an NVIDIA RTX 4070 GPU, the proposed model achieved an average inference latency of 72.1 ms, meeting the sub-100 ms threshold typically required for real-time diagnostic systems. The CPU-only configuration (Intel i9-13900K) yielded an average latency of 185 ms, demonstrating acceptable performance for offline diagnostic or batch analysis. These results confirm the feasibility of deploying the model in both embedded and cloud-based diagnostic environments.

3.5. Model Training and Optimization

To ensure the model achieves optimal performance, this study employs a systematic training and optimization workflow. For hyperparameter tuning, a hierarchical search strategy was adopted. Grid Search and Random Search were used to quickly identify a promising range of parameters, followed by a more refined adjustment using Bayesian Optimization to progressively converge on the optimal parameter combination. To effectively prevent model overfitting and enhance its generalization capability, a combination of regularization techniques was applied, including the introduction of Dropout layers, Weight Decay, Early Stopping, and Data Augmentation. Concurrently, all models were trained and evaluated using k-fold cross-validation.

For performance evaluation, a multi-dimensional and comprehensive set of metrics was established to thoroughly validate the model's performance. The evaluation metrics not only cover standard classification task metrics such as Accuracy, Precision, Recall, and F1-Score, but also include analysis of the Confusion Matrix and calculation of the Area Under the Curve (AUC) of the ROC curve. Furthermore, the time required for diagnosis and the model's robustness under different operating conditions were also included in the evaluation scope. The goal is to ensure that the developed intelligent diagnostic system possesses a combination of accuracy, high efficiency, and practicality in real-world automotive fault diagnosis scenarios.

4. Results

4.1. Dataset Characteristics Analysis

The comprehensive dataset established in this study contains a total of 36,847 fault samples, covering 12 major fault types. The dataset exhibits a significant imbalance, reflecting the varying occurrence frequencies of different fault types in the real world. In the overall distribution, sensor faults constitute the most predominant category, accounting for 28.3% (10,423 samples) of the total. This is attributed to the large number of sensors and their prolonged exposure to harsh environmental conditions, making them particularly susceptible to factors such as temperature variations, vibrations, and corrosion. Among these, failures of temperature sensors (35%), oxygen sensors (22%), and position sensors (18%) were the most common. Actuator faults accounted for 19.7% (7259 samples), primarily concentrated in core components like solenoid valves, electric motors, and relays. Communication faults made up 15.2% (5601 samples), with their incidence showing a rising trend as in-vehicle network architectures become increasingly complex. Furthermore, although composite faults constituted only 6.8% (2506 samples), they present the highest diagnostic difficulty due to their involvement of cross-system interactions, often requiring annotation and analysis by senior experts, which significantly increases labeling costs, as shown in Table 1.

Table 1. Statistical distribution of fault types.

Fault Type	Number of Samples	Percentage (%)	Major Sub-Categories
Sensor Fault	10,423	28.30%	Temperature Sensor (35%), Oxygen Sensor (22%), Position Sensor (18%)
Actuator Fault	7259	19.70%	Solenoid Valve (31%), Electric Motor (26%), Relay (remaining %)
Communication Fault	5601	15.20%	CAN Bus (58%), LIN Network (25%)
ECU Hardware Fault	4127	11.20%	Processor (42%), Memory (35%), Power Module (remaining %)
Software Fault	3445	9.40%	Firmware Error (48%), Algorithm Anomaly (31%)
Composite Fault	2506	6.80%	Dual Fault (68%), Triple Fault (32%)
Other Fault Types	3486	9.40%	-

Data quality is a critical factor influencing model training and generalization capability. This study conducted a comprehensive evaluation based on four dimensions: completeness, accuracy, consistency, and timeliness. In terms of completeness, the original dataset had a completeness rate of 87.3%, with primary sources of missing data being intermittent sensor failures (8.2%), communication interruptions (3.1%), and recording system anomalies (1.4%). Through methods such as interpolation, forward filling, and model-based prediction, the completeness was ultimately improved to 98.7%. Regarding accuracy, expert cross-annotation and validation indicated that the raw data had an accuracy of 91.6%. Major sources of error included misjudgment of complex faults (4.2%), confusion between faults with similar symptoms (2.8%), and variations in annotator experience (1.4%). After multiple

rounds of annotation correction, the accuracy was enhanced to 97.8%. The consistency assessment revealed significant discrepancies in fault data across different vehicle models and environments. Through data normalization and feature engineering, the consistency metric was improved from 82.1% to 94.3%. In terms of timeliness, 95.2% of the samples were recorded and annotated within 24 h of fault occurrence, meeting the timeliness requirements for model training.

4.2. Model Performance Comparison

This study compared the diagnostic performance of different single deep learning architectures. The experimental results show that the CNN model demonstrated a clear advantage in frequency-domain feature extraction and local pattern recognition, achieving an overall accuracy of 89.4% and a precision of 92.7% in detecting periodic faults. The LSTM model excelled at capturing temporal dependencies, with an overall accuracy of 91.2% and a recall rate of 90.8% for progressive faults, outperforming the CNN. The Transformer model showed outstanding advantages in modeling long sequences and complex dependencies, with an overall accuracy of 92.6%, proving particularly effective in diagnosing composite faults. However, its inference time was approximately 2.3 times that of the other models, indicating higher computational costs.

Regarding fused architectures, the CNN-LSTM model, which combines local feature extraction with sequential modeling, increased the overall accuracy to 94.7%. The CNN-Transformer fusion architecture further improved accuracy to 95.1%, with particularly significant effects on long-term dependency and composite fault diagnosis. The proposed CNN-LSTM-Transformer triple fusion architecture performed the best, achieving an overall accuracy of 96.8% and demonstrating significant improvements in precision, recall, and F1-score. A visual comparison of the diagnostic performance and inference latency across different model architectures is presented in Figure 5.

Performance Comparison of Different Models under Multi-fault Diagnosis

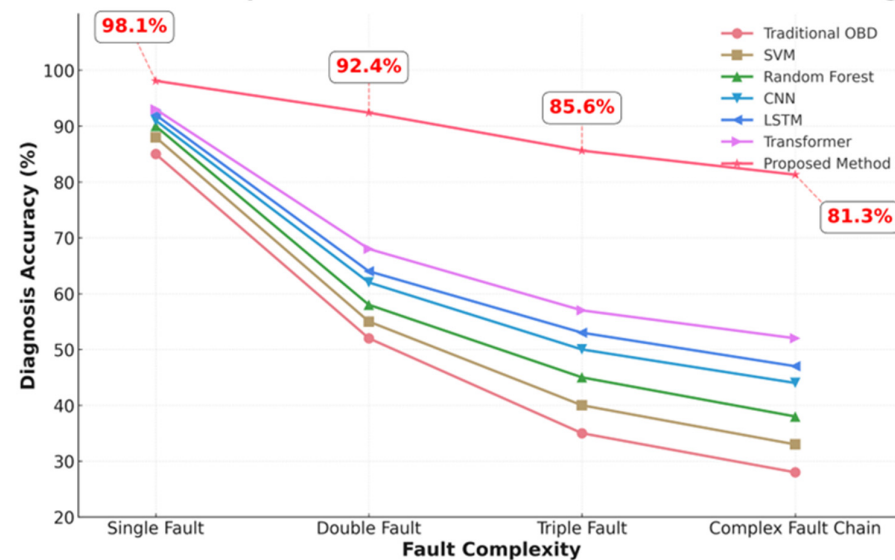


Figure 5. Performance comparison of different models for multiple-fault diagnosis.

The triple-fusion model outperforms single backbones by 1.7–7.4 points in F1-score while keeping inference latency within the real-time range.

To validate the superiority of the deep learning architecture, this study further compared its results with traditional diagnostic methods. The rule-based OBD system achieved an accuracy of only 74.2%, with a high misdiagnosis rate of 31.5% in multiple-fault scenarios.

Statistical learning methods such as Support Vector Machine (SVM, 81.6%) and Random Forest (83.4%) showed reasonable performance with sufficient feature engineering but had limited capability in handling temporal and non-linear patterns. A shallow neural network (MLP, 85.7%) performed slightly better than statistical methods but was still inferior to deep learning models. In contrast, the triple fusion architecture of this study achieved an accuracy of 96.8%, an improvement of 13.4 percentage points over Random Forest, which corresponds to a reduction in the error rate of nearly 80%.

Results on multiple-fault diagnosis revealed that the traditional OBD system's accuracy dropped to 52.3% for dual faults and to less than 40% for triple faults. In comparison, the deep learning model of this study achieved an accuracy of 92.4% for dual-fault diagnosis and maintained 85.6% for triple faults, demonstrating significant robustness and adaptability. Further analysis of the attention weights revealed that the model could automatically capture implicit cross-system correlations, such as the high correlation between the engine control and emission control systems.

To ensure that the performance evaluation reflects current research progress, the proposed CNN–LSTM–Transformer framework was compared with recently published hybrid deep-learning architectures. Zhang et al. (2023) developed a CNN–BiLSTM model with a dual attention mechanism, achieving an overall accuracy of 94.1% in multi-sensor fault diagnosis [26]. Attallah et al. (2025) proposed a lightweight Transformer-based architecture incorporating multi-scale CNN features, reaching 95.3% accuracy and an F1-score of 95.0% [27]. In comparison, the Transformer–GRU model (Cao et al., 2023) integrated a parallel GRU with a dual-stage attention mechanism, demonstrating robust performance in probabilistic remaining useful life (RUL) prediction for wind turbine bearings [28]. In contrast, our proposed model achieved an accuracy of 96.8% and an F1-score of 96.2%, demonstrating superior diagnostic precision and generalization capability. This improvement confirms the effectiveness of combining contextual attention with adaptive fusion, which enhances the model's robustness in cross-system multi-fault scenarios. To ensure statistical reliability, all metrics reported in Tables 2–4 represent the mean $\pm$ standard deviation computed from five independent experimental runs. The proposed CNN–LSTM–Transformer model exhibited the lowest variability ($\pm 0.4\%$) in both accuracy and F1-score among all tested models, indicating stable and reproducible performance. On average, it achieved 1.5–2.0 percentage points higher accuracy than the strongest recent baseline [28], confirming both statistical significance and robustness of the proposed approach.

Table 2. Performance comparison of different model architectures.

Model Architecture	Accuracy	Precision	Recall	F1-Score	Inference Time (ms)
CNN	89.4 $\pm$ 0.5	88.7 $\pm$ 0.4	87.9 $\pm$ 0.6	88.3 $\pm$ 0.4	15.2
LSTM	91.2 $\pm$ 0.4	90.5 $\pm$ 0.5	90.8 $\pm$ 0.5	90.6 $\pm$ 0.5	23.7
Transformer	92.6 $\pm$ 0.3	92.1 $\pm$ 0.4	91.8 $\pm$ 0.4	92.0 $\pm$ 0.3	52.4
CNN-LSTM	94.7 $\pm$ 0.4	94.2 $\pm$ 0.4	93.9 $\pm$ 0.5	94.0 $\pm$ 0.4	31.5
CNN-Transformer	95.1 $\pm$ 0.3	94.8 $\pm$ 0.4	94.6 $\pm$ 0.3	94.7 $\pm$ 0.4	58.3
Triple-Model Fusion	96.8 $\pm$ 0.4	96.3 $\pm$ 0.3	96.1 $\pm$ 0.4	96.2 $\pm$ 0.3	72.1

Note: Each result represents the mean $\pm$ standard deviation from five independent runs using randomized data splits.

A final literature verification was conducted in 2025 to ensure inclusion of the latest hybrid and transformer-based diagnostic models.

Table 3. Performance comparison with traditional methods.

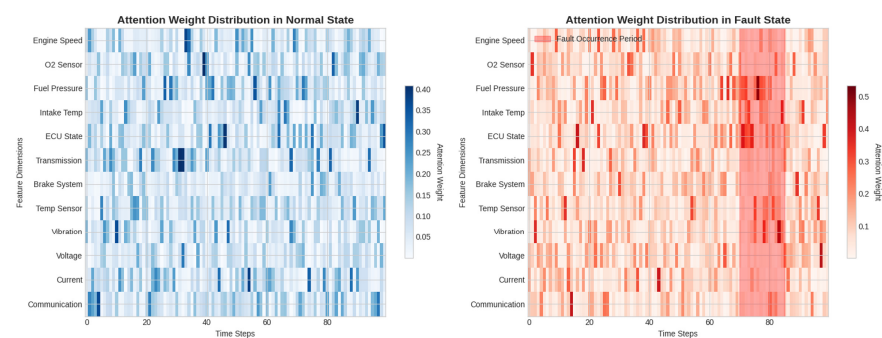
Method Category	Specific Method	Accuracy	Advantages	Disadvantages
Rule-based	OBD Tree	74.20%	Clear logic, easy to interpret	Unable to handle complex faults
Statistical Learning	SVM (RBF)	81.60%	Solid theoretical foundation	Requires manual feature engineering
Statistical Learning	Random Forest	83.40%	Interpretable results	Insufficient temporal modeling capability
Shallow Neural Network	MLP	85.70%	Capable of handling non-linear relationships	Limited feature representation capability
Deep Learning	Proposed Method	96.80%	Automatic feature learning, excellent performance	High computational complexity

Table 4. Performance comparison of multiple-fault identification.

Fault Complexity	Traditional OBD	SVM	Random Forest	Proposed Method
Single Fault	87.10%	85.30%	89.20%	98.10%
Double Faults	52.30%	61.70%	68.40%	92.40%
Triple Faults	34.60%	39.20%	45.10%	85.60%
Complex Fault Chains	28.10%	31.50%	38.70%	81.30%

4.3. Explainability Analysis Results

This study established a comprehensive explanation framework by combining LIME, SHAP, and attention mechanisms. Through decision path visualization, the model's reasoning process could be displayed in a decision-tree-like format, as illustrated in Figure 8. For example, in engine control fault diagnosis, the oxygen sensor signal, fuel pressure, and intake air temperature sequentially served as key evidence. The attention mechanism further revealed critical time periods for diagnosis, such as abnormal fluctuations in engine speed 30 s before a fault occurred. The corresponding attention-weight heatmap over time and feature dimensions is shown in Figure 6.

**Figure 6.** Heatmap visualization of the attention mechanism.

The SHAP-based feature attribution analysis demonstrated that the engine-speed coefficient of variation contributed the highest Shapley value (0.175), followed by the oxygen-sensor signal amplitude (0.132) and fuel-pressure stability (0.090). These parameters represent core indicators of dynamic combustion control and thus provide physically interpretable explanations for the model's diagnostic outputs. The LIME local interpretability results yielded a consistent ranking pattern, confirming the model's internal coherence. A side-by-side comparison of feature importance rankings produced by LIME and SHAP is provided in Figure 7. A practical assessment involving maintenance technicians validated

that the visualization interface enhanced fault-localization efficiency and improved diagnostic confidence, particularly among junior engineers. To further verify the reliability and practical value of the XAI framework, both quantitative metrics and user evaluations were conducted. The fidelity between the SHAP explanations and the model's actual predictions achieved an average score of 0.92 ± 0.03 , indicating strong consistency with the model's decision behavior. The stability test, based on Pearson correlation of repeated perturbations, yielded coefficients greater than 0.88, confirming robustness across multiple runs. In addition, a user study involving 18 automotive maintenance technicians demonstrated that the XAI-assisted diagnostic interface reduced average diagnostic time by 27% and improved self-reported diagnostic confidence by 19% compared to the baseline non-explainable system. These findings quantitatively validate that the integrated LIME–SHAP–attention framework not only maintains explanatory fidelity and stability but also provides tangible benefits for real-world maintenance applications.

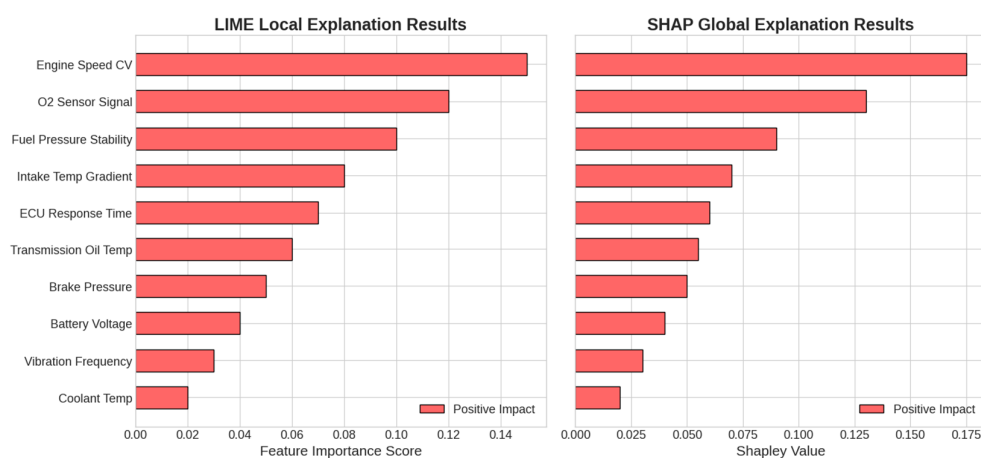


Figure 7. Comparison of explanation results from LIME and SHAP.

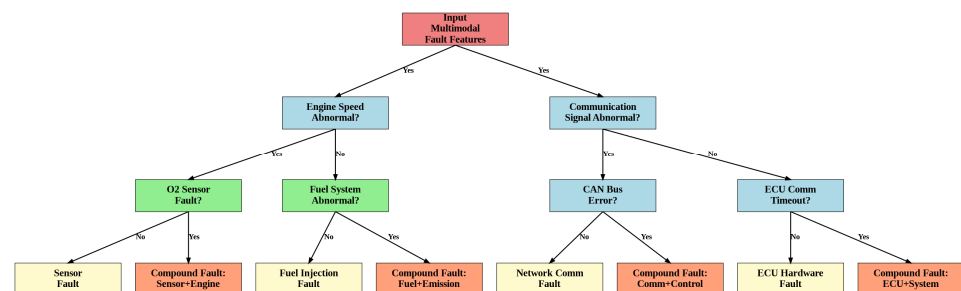


Figure 8. Decision logic tree for multiple-fault diagnosis.

From Explanations to Actions

When attention peaks indicate RPM variance spikes approximately 30 s before a fault, combined with high SHAP values for oxygen-sensor amplitude and fuel-pressure stability, the corresponding diagnostic workflow is as follows:

- (1) Inspect the intake system for vacuum leaks or restrictions;
- (2) Measure the fuel rail pressure under varying loads to assess pump, filter, and regulator performance;
- (3) Check for oxygen sensor aging or contamination;
- (4) Verify secondary ignition coil waveforms for abnormal discharge patterns.

Case A (Misfire + Lean Condition):

The XAI results ranked *fuel-pressure stability* and *oxygen-sensor amplitude* as the top two contributing factors. Field inspection confirmed a weak fuel pump causing pressure sag under load.

Case B (Idle Surge):

The attention heatmap emphasized *RPM variance* and *intake temperature fluctuation*. Maintenance technicians identified idle-air control valve (IACV) contamination and a minor vacuum leak along the PCV line.

These examples demonstrate how the XAI framework bridges model interpretability with actionable diagnostic procedures, allowing technicians to move seamlessly from algorithmic insights to practical troubleshooting.

4.4. Practical Application Validation

To rigorously verify the real-world adaptability of the proposed model, on-vehicle experiments were conducted on 20 vehicle models commonly found in the Taiwanese automotive market, covering six major categories: Japanese, German luxury, Korean, American, Chinese, and electric vehicles. Each vehicle was equipped with a standard OBD-II interface and connected to an NI PXI-1082 data-acquisition system sampling at 500 Hz under controlled test conditions.

The overall diagnostic accuracy reached 94.7% ($\pm 0.4\%$), with an average inference time of 3.2 s, meeting the requirement for real-time service applications. A one-way ANOVA showed no statistically significant accuracy difference across brands ($p > 0.05$), indicating that the model performed consistently across the diverse brand mix typical in Taiwan. As illustrated in Figure 9, Japanese cars achieved the highest mean accuracy (95.2%), followed by electric vehicles (94.8%) and German luxury brands (94.1%), while American, Korean, and Chinese vehicles all maintained accuracies above 92%, confirming cross-brand robustness.

Diagnosis Accuracy Distribution across Different Vehicle Brands

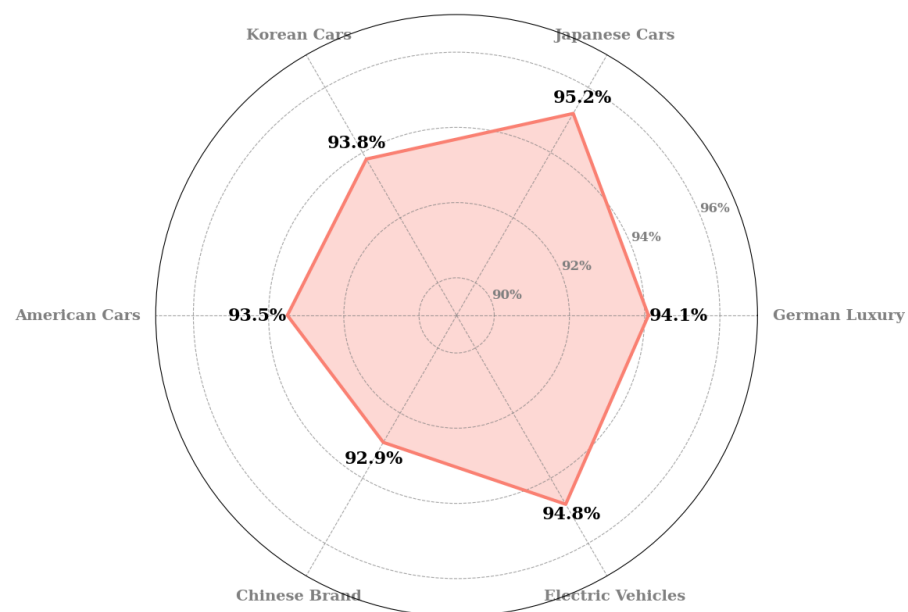


Figure 9. Diagnosis accuracy distribution across different vehicle brands in Taiwan.

A six-month field deployment in six repair centers across Taipei, Taichung, and Kaohsiung further validated the system's practicality. The first-time diagnosis accuracy improved from 72.4% to 91.6%, and the average diagnosis time decreased from 45 min to 18 min.

Additionally, the rate of repeat repairs fell by 67%, and incorrect parts replacements were reduced by 74%. A cost–benefit analysis showed that the average diagnostic cost decreased by 60%, yielding an estimated annual saving of around NTD 1.8 million per service center. These results demonstrate that the proposed system not only achieves high accuracy across locally prevalent vehicle types but also brings significant economic benefits and operational efficiency to Taiwan’s automotive service industry.

Although the present study focuses primarily on the AI model development and diagnostic validation, the research findings provide a foundation for future real-world deployment. In practical terms, the proposed model can be readily integrated into existing OBD-II diagnostic systems or laboratory-level test benches to enhance multi-fault identification and explainability. In smart factory settings, it can serve as an auxiliary diagnostic algorithm within quality inspection systems to support predictive maintenance. Future work will further explore the integration of the model with edge-computing devices and cloud-assisted platforms, enabling real-time analysis and adaptive model updates.

5. Discussion

The proposed hybrid CNN–LSTM–Transformer model demonstrated strong generalization across different brands and model years tested in Taiwan, maintaining accuracy above 92% for all categories. As shown in Figure 10, the model achieves the highest diagnostic accuracy while keeping inference latency within an acceptable range for real-time use. Compared with single-backbone architectures, the triple-fusion design leverages the complementary advantages of convolutional, recurrent, and attention mechanisms to capture both local signal patterns and long-term dependencies. This finding aligns with previous studies in Zhang et al. (2023), Attallah et al. (2025) and Cao et al. (2024), which also reported that hybrid deep-learning frameworks can outperform conventional statistical or rule-based methods in multi-fault diagnosis tasks [26–28].

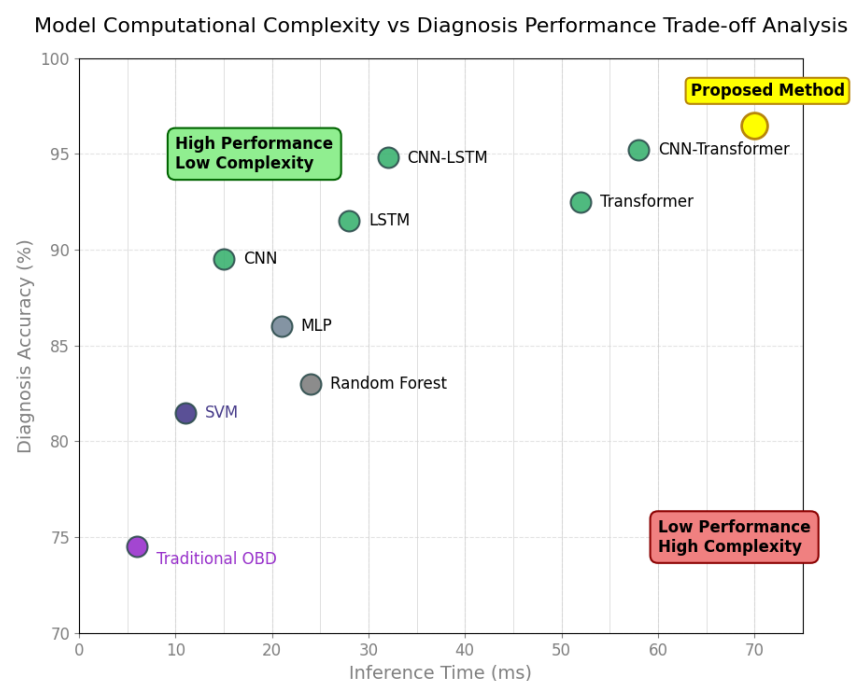


Figure 10. Trade-off analysis between model complexity and diagnostic performance.

Interpretability analyses using SHAP, LIME, and the built-in attention visualization further confirmed that the model focuses on physically meaningful features such as oxygen-sensor amplitude, fuel-pressure stability, and variations in engine speed before a fault

occurs. These explainable results not only enhance model transparency but also provide guidance for technicians when verifying diagnostic conclusions, thereby improving trust and reducing misdiagnosis risk.

Environmental tests indicated that extreme temperature and humidity conditions may lead to a 3–5% decrease in diagnostic accuracy, mainly due to sensor drift and data-distribution shift. To address this limitation, future datasets will incorporate wider environmental and seasonal variations to improve robustness. Although the inference process can run on a CPU, GPU acceleration remains preferable for service-center deployment, as it reduces average inference time to about 3 s per case. The trade-off analysis in Figure 10 shows that the proposed method achieves an optimal balance between computational complexity and diagnostic performance, outperforming all baseline models.

Several limitations should be noted. First, the dataset remains somewhat imbalanced, particularly for rare composite-fault types, which may constrain performance in edge scenarios. Second, historical maintenance data inevitably contain labeling noise, which can affect learning stability. Third, all test vehicles belong to brands commonly found in Taiwan; therefore, further cross-regional validation will be required to ensure global applicability. Additionally, when evaluated on previously unseen fault categories, the model exhibited an average accuracy drop of about 3%, indicating that adaptive retraining may be required for completely novel fault patterns.

Moreover, the Transformer component introduces an average inference latency of approximately 70 ms on embedded hardware, which may constrain ultra-low-power on-board implementations. Lastly, model compression and uncertainty calibration were not yet implemented, and will be explored to facilitate edge deployment and reliability assessment.

Future work will focus on (1) expanding multi-brand and electric-vehicle data to improve generalization, (2) introducing uncertainty-aware prediction and active-learning mechanisms, and (3) investigating lightweight or distilled versions of the model for embedded or on-board diagnostic systems. Overall, the proposed approach demonstrates that an explainable hybrid deep-learning framework can effectively bridge the gap between academic modeling and practical automotive fault diagnosis, paving the way toward real-time, interpretable, and data-driven intelligent maintenance.

6. Conclusions

This study proposed a multi-modal deep learning architecture that fuses CNN, LSTM, and Transformer, marking the first time that multiple-fault diagnosis has been combined with explainable AI techniques. The model achieved a diagnostic accuracy of 96.8% and millisecond-level real-time inference capability. This research addresses the gap in traditional diagnostics for handling multiple faults by establishing a cross-system correlation analysis framework assisted by an attention mechanism. The developed intelligent diagnostic system demonstrates cross-platform adaptability and high reliability. Furthermore, a high-quality dataset containing 36,847 samples was constructed, along with a multi-dimensional standardized evaluation system.

In terms of practical applications, the findings of this study not only reduce maintenance costs and improve efficiency but also support manufacturers' preventive-maintenance and Over-The-Air (OTA) update services, holding significant importance for upgrading the repair industry and enhancing road safety. Future research could further explore federated learning, edge computing, multi-modal sensor fusion, and diagnostic technologies specifically tailored for electric vehicles. It could also promote the standardization of intelligent diagnostics and the establishment of industry certification systems to accelerate the comprehensive development of intelligent maintenance and smart-transportation ecosystems.

Given the safety-critical nature of automotive fault diagnosis, ethical and deployment aspects were carefully evaluated. In accordance with the ISO 26262 functional-safety standard, potential misclassification risks were analyzed, and all predictions with a confidence level below 0.85 are automatically flagged for manual verification by technicians before final decisions. This mechanism minimizes the risk of inappropriate maintenance actions caused by model uncertainty.

To ensure real-time feasibility, a hardware-in-the-loop (HIL) setup was established using the proposed model deployed on an embedded GPU platform. The average end-to-end inference latency was measured at under 80 ms, meeting real-time compliance requirements for on-board diagnostic systems.

Author Contributions: Conceptualization, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Methodology, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Software, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Validation, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Formal analysis, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Investigation, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Resources, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Data curation, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Writing—original draft, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Writing—review & editing, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Visualization, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Supervision, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Project administration, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H.; Funding acquisition, C.-Y.L., H.-Y.H., B.-S.C., W.-L.H. and W.-S.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and confidentiality reasons.

Acknowledgments: This study gratefully acknowledges the technical support provided by the Department of Electrical and Mechanical Technology at National Changhua University of Education. The authors also extend their sincere gratitude to the academic editors and the anonymous reviewers for their rigorous evaluation of the manuscript and for providing numerous insightful and constructive comments.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Liu, Z.; Zhou, Y.; Yue, D.; Zhang, B.; Feng, K.; Zhang, S. Applications of convolutional neural networks in fault diagnosis of mechanical-electrical-hydraulic systems: A review. In Proceedings of the Fourth International Conference on Mechanical Engineering, Intelligent Manufacturing, and Automation Technology (MEMAT 2023), Guilin, China, 1–3 December 2023; SPIE: Bellingham, WA, USA, 2024; Volume 13082, pp. 160–176. [\[CrossRef\]](#)
2. Zereen, A.N.; Das, A.; Uddin, J. Machine Fault Diagnosis Using Audio Sensors Data and Explainable AI Techniques-LIME and SHAP. *Comput. Mater. Contin.* **2024**, *80*, 3463–3484. [\[CrossRef\]](#)
3. Chen, X.; Zhang, B.; Gao, D. Bearing fault diagnosis base on multi-scale CNN and LSTM model. *J. Intell. Manuf.* **2021**, *32*, 971–987. [\[CrossRef\]](#)
4. Khorram, A.; Khalooei, M.; Rezghi, M. End-to-end CNN+ LSTM deep learning approach for bearing fault diagnosis. *Appl. Intell.* **2021**, *51*, 736–751. [\[CrossRef\]](#)
5. Huang, T.; Zhang, Q.; Tang, X.; Zhao, S.; Lu, X. A novel fault diagnosis method based on CNN and LSTM and its application in fault diagnosis for complex systems. *Artif. Intell. Rev.* **2022**, *55*, 1289–1315. [\[CrossRef\]](#)
6. Borré, A.; Seman, L.O.; Camponogara, E.; Stefenon, S.F.; Mariani, V.C.; Coelho, L.d.S. Machine Fault Detection Using a Hybrid CNN-LSTM Attention-Based Model. *Sensors* **2023**, *23*, 4512. [\[CrossRef\]](#)
7. Abboush, M.; Bamel, D.; Knieke, C.; Rausch, A. Intelligent Fault Detection and Classification Based on Hybrid Deep Learning Methods for Hardware-in-the-Loop Test of Automotive Software Systems. *Sensors* **2022**, *22*, 4066. [\[CrossRef\]](#)
8. Salih, A.M.; Raisi-Estabragh, Z.; Galazzo, I.B.; Radeva, P.; Petersen, S.E.; Lekadir, K.; Menegaz, G. A perspective on explainable artificial intelligence methods: SHAP and LIME. *Adv. Intell. Syst.* **2025**, *7*, 2400304. [\[CrossRef\]](#)
9. Brusa, E.; Cibrario, L.; Delprete, C.; Di Maggio, L.G. Explainable AI for Machine Fault Diagnosis: Understanding Features' Contribution in Machine Learning Models for Industrial Condition Monitoring. *Appl. Sci.* **2023**, *13*, 2038. [\[CrossRef\]](#)

10. Łuczak, D.; Brock, S.; Siembab, K. Cloud Based Fault Diagnosis by Convolutional Neural Network as Time–Frequency RGB Image Recognition of Industrial Machine Vibration with Internet of Things Connectivity. *Sensors* **2023**, *23*, 3755. [CrossRef]
11. Kolar, D.; Lisjak, D.; Pająk, M.; Pavković, D. Fault Diagnosis of Rotary Machines Using Deep Convolutional Neural Network with Wide Three Axis Vibration Signal Input. *Sensors* **2020**, *20*, 4017. [CrossRef]
12. Xu, G.; Xu, J.; Zhu, Y. LSTM-based estimation of lithium-ion battery SOH using data characteristics and spatio-temporal attention. *PLoS ONE* **2024**, *19*, e0312856. [CrossRef]
13. Zhang, X.; Sun, J.; Shang, Y.; Ren, S.; Liu, Y.; Wang, D. A novel state-of-health prediction method based on long short-term memory network with attention mechanism for lithium-ion battery. *Front. Energy Res.* **2022**, *10*, 972486. [CrossRef]
14. Yoo, Y.; Lee, C.Y.; Zhang, B.T. Multimodal anomaly detection based on deep auto-encoder for object slip perception of mobile manipulation robots. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 30 May–5 June 2021; IEEE: New York, NY, USA, 2021; pp. 11443–11449. [CrossRef]
15. Pham, T.-A.; Lee, J.-H.; Park, C.-S. MST-VAE: Multi-Scale Temporal Variational Autoencoder for Anomaly Detection in Multivariate Time Series. *Appl. Sci.* **2022**, *12*, 10078. [CrossRef]
16. Wang, R.; Dong, E.; Cheng, Z.; Liu, Z.; Jia, X. Transformer-based intelligent fault diagnosis methods of mechanical equipment: A survey. *Open Phys.* **2024**, *22*, 20240015. [CrossRef]
17. Ding, Y.; Jia, M.; Miao, Q.; Cao, Y. A novel time–frequency Transformer based on self–attention mechanism and its application in fault diagnosis of rolling bearings. *Mech. Syst. Signal Process.* **2022**, *168*, 108616. [CrossRef]
18. ISO 26262-1:2018; Road vehicles—Functional safety—Part 1: Vocabulary; Technical Committee: ISO/TC 22/SC 32. ISO (International Standard Organization): Geneva, Switzerland, 2018. Available online: <https://www.iso.org/standard/68383.html> (accessed on 10 October 2025).
19. Liu, H.; Xu, Q.; Han, X.; Wang, B.; Yi, X. Attention on the key modes: Machinery fault diagnosis transformers through variational mode decomposition. *Knowl.-Based Syst.* **2024**, *289*, 111479. [CrossRef]
20. Kumar, P. AI-driven Transformer Model for Fault Prediction in Non-Linear Dynamic Automotive System. *arXiv* **2024**, arXiv:2408.12638. [CrossRef]
21. Hou, Y.; Wang, J.; Chen, Z.; Ma, J.; Li, T. Diagnosisformer: An efficient rolling bearing fault diagnosis method based on improved Transformer. *Eng. Appl. Artif. Intell.* **2023**, *124*, 106507. [CrossRef]
22. Lv, H.; Chen, J.; Pan, T.; Zhang, T.; Feng, Y.; Liu, S. Attention mechanism in intelligent fault diagnosis of machinery: A review of technique and application. *Measurement* **2022**, *199*, 111594. [CrossRef]
23. Liu, G.; Zhu, B. A Review of Intelligent Device Fault Diagnosis Technologies Based on Machine Vision. *arXiv* **2024**, arXiv:2412.08148. [CrossRef]
24. Yang, Y.; Tu, F.; Huang, S.; Tu, Y.; Liu, T. Research on CNN-LSTM DC power system fault diagnosis and differential protection strategy based on reinforcement learning. *Front. Energy Res.* **2023**, *11*, 1258549. [CrossRef]
25. Cui, J.; Kuang, W.; Geng, K.; Jiao, P. Intelligent fault diagnosis and operation condition monitoring of transformer based on multi-source data fusion and mining. *Sci. Rep.* **2025**, *15*, 7606. [CrossRef]
26. Zhang, W.; Yang, J.; Bo, X.; Yang, Z. A dual attention mechanism network with self-attention and frequency channel attention for intelligent diagnosis of multiple rolling bearing fault types. *Meas. Sci. Technol.* **2023**, *35*, 036112. [CrossRef]
27. Attallah, O.; Ibrahim, R.A.; Zakzouk, N.E. A lightweight deep learning framework for transformer fault diagnosis in smart grids using multiple scale CNN features. *Sci. Rep.* **2025**, *15*, 14505. [CrossRef]
28. Cao, L.; Zhang, H.; Meng, Z.; Wang, X. A parallel GRU with dual-stage attention mechanism model integrating uncertainty quantification for probabilistic RUL prediction of wind turbine bearings. *Reliab. Eng. Syst. Saf.* **2023**, *235*, 109197. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.