Optimizations in Applied Machine Learning
Research Article | Volume 3 | Issue 3 | Sep 2023
Received: 2 Aug 2023 | Revised: 28 Aug 2023
Accepted: 8 Sep 2023 | Published Online: 12 Sep 2023

Analysis of Vehicle Fault Diagnosis Model Based on Causal Sequence-to-Sequence in Embedded Systems

Weidong Huang[1,2], Jiahuai Ma [3*]

[1] GAC FIAT CHRYSLER Automobiles Co., Ltd., Changsha 410100, China

[2] HYCAN Automobile Technology Co., Ltd., Guangzhou 510000, China

[2*] Corresponding, University of Florida, Herbert Wertheim College, FL, 32608, USA,

Email: maj2@ufl.edu

Abstract: The rapid development of the automotive industry has intensified the challenges faced by traditional fault diagnosis systems. This study proposes an efficient vehicle fault diagnosis model based on deep learning to improve fault identification accuracy and real-time performance, facilitating deployment in embedded systems. The model integrates a sequence-to-sequence architecture, an attention mechanism, and causal learning. The sequence-to-sequence structure captures complex time-series dependencies, while the attention mechanism enhances focus on critical features, improving fault pattern recognition. Causal learning further strengthens the model's understanding of fault relationships, enhancing diagnostic performance. Experimental evaluation on real-world vehicle datasets, including sensor data and maintenance records, demonstrates the model's superiority over state-of-the-art methods in accuracy, precision, recall, and F1 score. The results validate the model's effectiveness in complex fault scenarios and its potential for embedded system integration. This research provides a robust foundation for advancing real-time data analysis in in-vehicle diagnostic systems within the Internet of Things framework.

Keywords: *Vehicle fault diagnosis; Deep learning; Sequence-to-sequence; Attention mechanism; Causal learning; Embedded systems*

1 Introduction

With the rapid development of the automotive industry, the design and manufacturing of modern vehicles have become increasingly complex. These new vehicles are not only equipped with numerous sensors, electronic control units, and intelligent driving systems, but also integrate sophisticated network communication technologies[1]. This transformation has significantly enhanced vehicle performance and functionality, enabling advanced features such as autonomous driving, intelligent navigation, and efficient energy management. However, as these technologies continue to evolve, traditional fault diagnosis systems face growing challenges[2]. Diagnostic methods that previously relied on fixed rules and human expertise are increasingly inadequate when dealing with complex, multidimensional data. These traditional approaches lack flexibility and are unable to quickly adapt to changing driving environments, model updates, and the growing diversity of potential fault modes[3]. Recent studies have shown that improving domain adaptation techniques can significantly enhance model stability across diverse vehicle operating environments, thereby improving the practical adaptability of fault diagnosis systems[4]. Therefore, there is an urgent need to develop new diagnostic strategies to address the complexity brought by modern vehicles and ensure their reliability in terms of safety and performance.

At the same time, the rise of Connected Vehicles (CV) presents unprecedented opportunities for fault diagnosis. Through Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication, these connected vehicles can collect and transmit large amounts of operational data in real time, including vehicle speed, location, environmental conditions, and fault logs[5]. This wealth of data can not only be used to improve driving safety, optimize traffic flow, and enhance overall road efficiency, but also provides an important information foundation for improving fault diagnosis systems. However, extracting valuable fault information from this massive, complex data remains a critical problem that needs to be solved. Traditional fault diagnosis methods often struggle to operate efficiently in such dynamic data environments[6]. Therefore, leveraging advanced data processing technologies, particularly deep learning and artificial intelligence, to analyze and interpret these data will be key to enhancing fault diagnosis accuracy and real-time performance.

To address the challenges faced by current automotive fault diagnosis systems, the introduction of deep learning technology provides new possibilities for optimizing the diagnostic process[7]. Specifically, sequence-to-sequence (Seq2Seq) deep learning models have demonstrated strong pattern recognition and data processing capabilities[8]. These models can effectively handle large amounts of time-series data, capturing complex spatiotemporal features, which are crucial for understanding vehicle performance under various driving conditions[9]. Recent studies have shown that improving domain adaptation techniques can significantly enhance model stability across diverse vehicle operating environments, thereby improving the practical adaptability of fault diagnosis systems[10]. By leveraging the advantages of deep learning, we can build more intelligent and adaptive fault diagnosis systems that improve accuracy and real-time responsiveness in practical applications. Therefore, the objective of this paper is to design a deep learning model integrating attention mechanisms and causal learning to significantly improve the accuracy of vehicle fault diagnosis and enable real-time monitoring and feedback. Our research will focus on developing an encoder-decoder model based on the sequence-to-sequence architecture. In this architecture, the encoder is responsible for transforming the input sequence into a compact contextual representation, while the decoder generates the corresponding output based on this representation. During this process, we will employ self-attention mechanisms to effectively learn the relationships between features in the input sequence, ensuring that the model focuses on the most relevant information. By introducing causal learning, we will construct causal chains for fault occurrences, allowing the model not only to recognize potential faults but also to understand their root causes and impacts. This approach not only enhances fault diagnosis accuracy but also provides valuable insights, offering a more scientific basis for subsequent fault prediction and preventive measures. The goal of this research is to apply deep learning techniques to vehicle fault diagnosis, addressing the shortcomings of traditional systems. Ultimately, we hope that the model can be effectively deployed in embedded systems to provide real-time support for in-vehicle diagnostics and contribute to the development of a smarter and more efficient traffic management system.

The contributions of this paper are as follows:

1. The sequence-to-sequence (Seq2Seq) model plays a key role in vehicle fault diagnosis by efficiently handling variable-length time-series data and predicting faults in real time. Through the encoder-decoder structure, the model converts historical sensor data into context vectors, capturing the main features of the input data. The decoder then uses these features to generate corresponding fault diagnosis results. The Seq2Seq model is particularly suited for capturing temporal dependencies and trends in data, allowing it to adapt to complex fault patterns and significantly improve diagnosis accuracy and flexibility. This innovative approach provides strong support for real-time fault diagnosis, enhancing vehicle safety and reliability.

2. The attention mechanism significantly improves the performance of the vehicle fault diagnosis model, enabling it to more effectively handle long sequence data. By assigning different weights to each element in the input sequence, the attention mechanism allows the model to dynamically focus on the information most relevant to the current decoding state. This selective attention not only enhances the model's interpretability but also improves fault diagnosis accuracy. For example, when the model identifies a specific fault, it can prioritize the sensor data related to that fault, effectively capturing local features and reducing noise interference. This flexible attention strategy enables the model to provide more reliable diagnostic results under complex fault conditions.

3. Causal learning can significantly improve the accuracy of fault recognition and prediction in vehicle fault diagnosis by revealing the causal relationships between variables, helping to understand the root causes and progression of faults. This approach focuses not only on correlations but also on identifying factors that directly affect system behavior. By constructing causal chains, the model can analyze the time-series relationships between sensor data and fault occurrences, revealing how specific faults are triggered by changes in other variables. This in-depth causal understanding provides a scientific basis for preventive measures and optimization of maintenance decisions, making vehicle maintenance more efficient and intelligent.

2 Related Work

In the field of vehicle fault diagnosis, traditional fault detection methods are gradually becoming insufficient to meet the increasingly complex demands of modern automobiles as automotive technology advances[11]. The design of modern vehicles increasingly relies on electronic systems, equipped with dozens to hundreds of sensors and electronic control units, which monitor and regulate various vehicle functions such as engine performance, braking systems, and safety devices in real time[12]. Through these sensors, vehicles can generate and collect vast amounts of data, including operational status, environmental conditions, and fault logs[13]. These real-time data provide rich information for fault diagnosis but also present new challenges for data processing and analysis[14].

Traditional fault detection methods are often rule-based or rely on expert knowledge, depending on manually set thresholds and rules. These methods may be effective for simple faults, but as the complexity of vehicle systems increases, the flexibility and adaptability of traditional methods become insufficient[15]. Many fault patterns are multidimensional, involving the interaction of multiple sensors and systems, so relying solely on experience or fixed rules is inadequate to adapt to the changing driving environment and potential fault modes in real-time[16]. In the field of vehicle fault diagnosis, traditional methods such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Autoencoders (AE), Variational Autoencoders (VAE), clustering, and density-based methods have been widely applied in recent years due to their effectiveness in handling time-series data[17]. Convolutional Neural Networks (CNN)[18] are particularly suitable for handling structured data, as they can effectively capture local features by extracting spatial features from sensor data through convolutional layers. When dealing with time-series data with translational invariance, CNNs can automatically learn complex patterns in the data, making them especially effective for processing combinations of image and time-series data (such as video streams). However, CNNs perform poorly when handling long-term dependencies in sequential data, as they primarily focus on local features, potentially ignoring global information in long sequences. Additionally, in vehicle fault diagnosis, faults often involve complex relationships across multiple sensors and time dimensions, limiting the applicability of CNNs. Recurrent Neural Networks (RNN)[19] are designed for sequential data and can maintain temporal dependencies, making them suitable for analyzing temporal features in time-series data. In fault diagnosis, RNNs can predict based on past sensor data, identifying potential fault patterns. However, RNNs suffer from issues like vanishing or exploding gradients when dealing with long time-series data, making training more difficult. Moreover, RNNs may lack the learning capacity to handle complex fault patterns, especially when multiple sensor data need to be analyzed together. Autoencoders (AE)[20] and Variational Autoencoders (VAE)[21], which learn low-dimensional representations of data, can effectively perform dimensionality reduction and feature extraction, making them suitable for anomaly detection. Autoencoders typically rely on unsupervised learning but may not capture all the complex relationships in the data, affecting the accuracy of anomaly detection. VAEs excel in generative models and can generate new samples similar to the input data, improving the recognition of anomalous conditions. However, these models still require labeled data to better distinguish between normal and anomalous states in vehicle fault diagnosis. Clustering and density-based methods are representative of unsupervised learning, capable of identifying outliers in data without labels, making them suitable for handling large-scale sensor data. Clustering methods group similar data together, helping to identify anomalous behavior that differs from the majority of data[22]. However, clustering methods are sensitive to parameter settings, and results may vary with different parameters, leading to instability. Density-based methods have higher computational complexity when dealing with high-dimensional data, which may impact real-time fault diagnosis efficiency[23]. In summary, traditional deep learning methods have shown good potential for application in vehicle fault diagnosis, effectively processing and analyzing large amounts of time-series data. However, these methods have certain limitations, especially as vehicle systems become more complex. Future research may need to combine these traditional methods with emerging technologies, such as sequence-to-sequence models and attention mechanisms, to improve the accuracy and real-time performance of fault diagnosis.

To address these challenges, researchers have begun to explore fault diagnosis methods based on deep learning. In recent years, several advanced methods have been proposed for vehicle fault diagnosis. Sequence-to-sequence (Seq2Seq)-based methods are widely applied to fault diagnosis, effectively handling time-series data through the encoder-decoder architecture[24]. Studies have shown that this method can capture temporal dependencies in the data, achieving accurate predictions of future faults. However, traditional Seq2Seq models may face context information loss when handling long sequences, affecting prediction performance. Many researchers have introduced attention mechanisms into Seq2Seq models to enhance the model’s ability to focus on key information in the input sequence[25]. This method has shown excellent performance in fault diagnosis, as it can dynamically adjust the model’s focus on different sensor data, improving diagnostic accuracy and interpretability. However, excessive reliance on attention mechanisms may lead to bias in feature selection, affecting overall performance. Causal learning aims to model the causal relationships between variables, providing a deeper understanding of fault diagnosis[26]. By identifying and constructing causal chains, researchers can reveal the direct influence relationships between faults, offering a scientific basis for fault prediction. The advantage of this method lies in its causal reasoning ability, but constructing accurate causal models remains a challenge in practical applications, especially in cases with sparse or noisy data.

Despite the progress made with these methods in fault diagnosis, several shortcomings remain. First, the issue of context loss in traditional Seq2Seq models when handling long time-series data has not been effectively resolved. Second, existing attention mechanisms may lead to bias in feature selection, which affects the model's generalization ability. Additionally, causal learning faces challenges in constructing causal relationship models, especially when data is sparse or noisy, limiting its wide application. In conclusion, combining the strengths of these methods while addressing their limitations and proposing a new fault diagnosis model will be an important direction for future research.

3 Method

The overall algorithm diagram of the vehicle fault diagnosis model is shown in Figure 1.

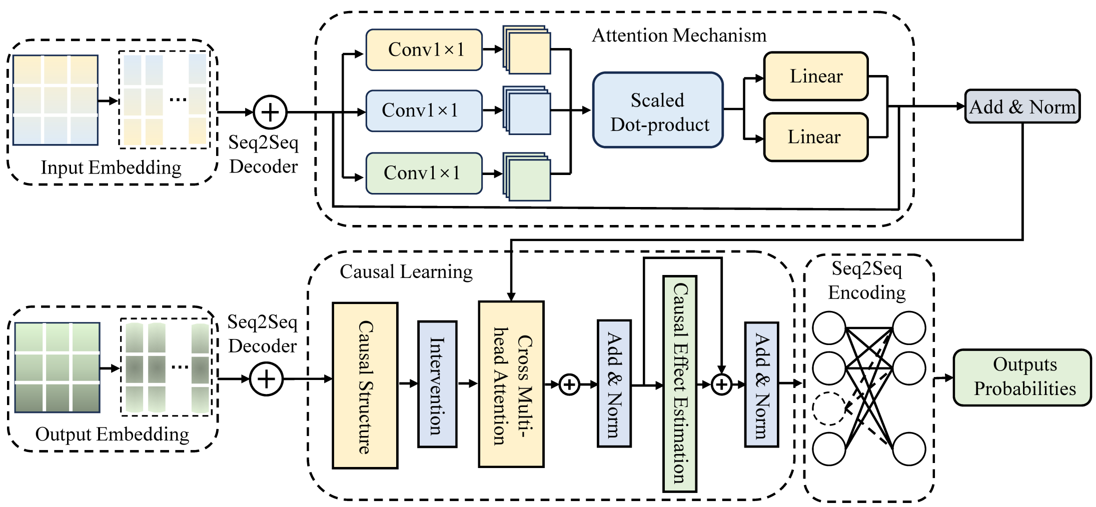

Figure 1. Overall algorithm architecture.

3.1 Seq2Seq

Seq2Seq (Sequence-to-Sequence) model is a deep learning framework designed to transform an input sequence into an output sequence. It has been widely applied in tasks such as machine translation, speech recognition, and text summarization. For fault diagnosis, the Seq2Seq model can process sequence data to predict fault development trends and complete fault diagnosis reports[27]. The architecture diagram is shown in Figure 2.

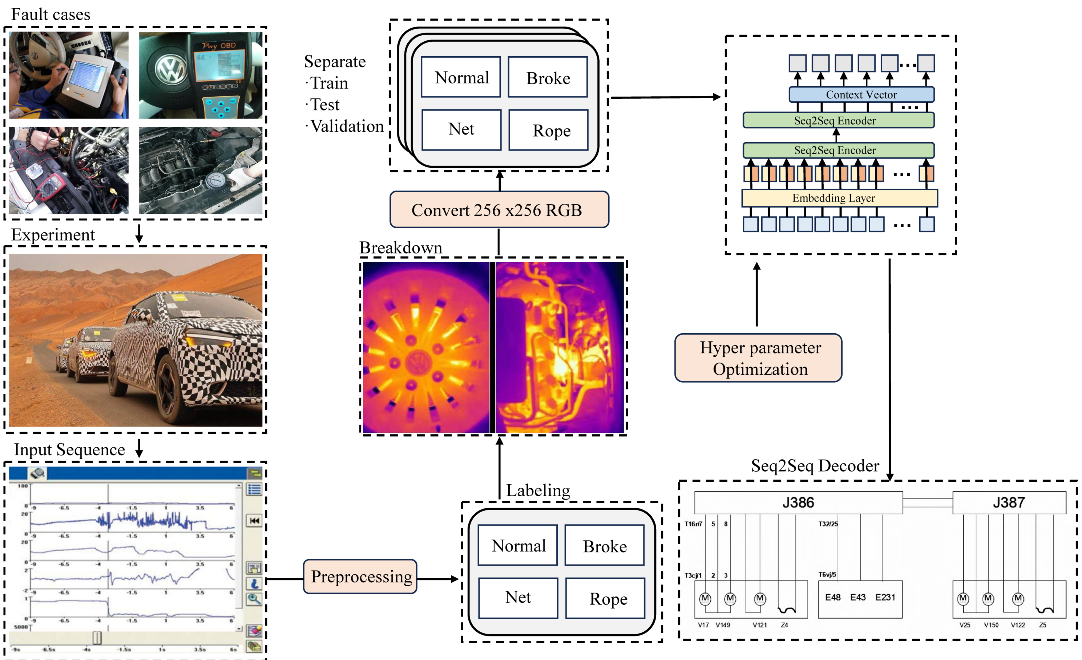

Figure 2. Structure diagram of Seq2Seq.

The Seq2Seq model typically consists of two main components: an encoder and a decoder. For the encoder part, the input sequence: $X = (x_1, x_2, \ldots, x_T)$ is fed into the encoder. The encoder is an RNN network that generates a hidden state sequence: $h = (h_1, h_2, \ldots, h_T)$. The final hidden state h_T is used to create the context vector c, summarizing the entire input sequence. For the decoder part, the context vector c is used as the initial input to generate the output sequence: $Y = (y_1, y_2, \ldots, y_{T'})$. The decoder is also an RNN network that produces the output step by step until the final stop token is generated.

For each time step t, the encoder's hidden state h_t is updated via the recurrent neural network as follows:

$$h_t = RNN(x_t, h_{t-1}) \tag{1}$$

The RNN's update steps are as follows:

$$\begin{gathered} i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \\ f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \\ o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \\ \tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \\ c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t h_t = o_t \odot \tanh(c_t) \end{gathered} \tag{2}$$

Here, σ is the sigmoid activation function. The variables i_t, f_t, and o_t represent the input gate, forget gate, and output gate, respectively.

The decoder initializes with the final hidden state h_T of the encoder as the initial state. The output sequence is $Y = (y_1, y_2, \ldots, y_{T'})$. The hidden state is updated as:

$$s_t = RNN(y_{t-1}, s_{t-1}, c_t) \tag{3}$$

The output prediction is:

$$y_t = \mathrm{softmax}(W_s s_t + b_s) \tag{4}$$

In vehicle fault diagnosis, the input sequence can be sensor data or event logs, while the output sequence can be fault types or diagnostic reports. The Seq2Seq model learns the mapping relationship between the input and output, enabling accurate prediction of complex faults.

3.2 Attention Mechanism

The attention mechanism is a technique that improves the performance of Seq2Seq models, especially when processing long sequences. It allows the model to focus on different parts of the input sequence when generating each output step[28]. The architecture diagram is shown in Figure 3.

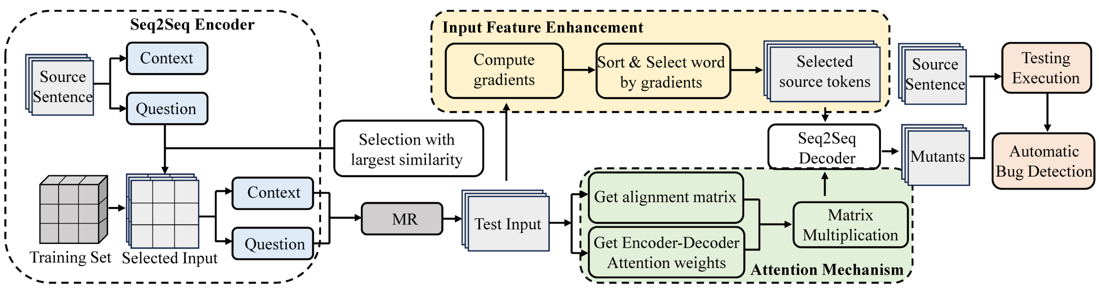

Figure 3. Structure diagram of attention mechanism.

In the Seq2Seq model, the encoder maps the input sequence into a series of hidden states: $h = (h_1, h_2, \ldots, h_T)$. The decoder generates the output step-by-step while attending to relevant parts of the input sequence through the attention mechanism. For each decoding step t, the alignment score between the decoder's current hidden state s_{t-1} and each encoder hidden state h_i is calculated as:

$$e_{t,i} = \mathrm{score}(s_{t-1}, h_i) \tag{5}$$

Common scoring functions include:
Dot Product:

$$e_{t,i} = s_{t-1}^T h_i \tag{6}$$

Bilinear:

$$e_{t,i} = s_{t-1}^T W_a h_i \tag{7}$$

MLP (Multi-Layer Perceptron):

$$e_{t,i} = v_a^T \tanh(W_a[s_{t-1}; h_i]) \tag{8}$$

The attention weights are computed using softmax:

$$\alpha_{t,i} = \frac{\exp(e_{t,i})}{\sum_{k=1}^{T} \exp(e_{t,k})} \tag{9}$$

The context vector c_t is computed as a weighted sum of encoder hidden states:

$$c_t = \sum_{i=1}^{T} \alpha_{t,i} h_i \tag{10}$$

The context vector c_t provides relevant input information for generating the next output. The decoder uses c_t, along with the previous hidden state, to generate the next output:

$$s_t = RNN(y_{t-1}, s_{t-1}, c_t) \tag{11}$$

$$y_t = \text{softmax}(W_s s_t + b_s) \tag{12}$$

The attention mechanism allows the model to dynamically focus on different parts of the input sequence, which greatly enhances performance, especially for long sequences. By using attention weights, it is possible to interpret which parts of the input the model focused on when generating a particular output, which is very useful for identifying key features in fault diagnosis. In vehicle fault diagnosis, the attention mechanism helps the model more accurately identify fault characteristics, improving diagnostic accuracy and efficiency.

3.3 Causal Learning

Causal learning is a method for identifying and understanding causal relationships in data. Unlike correlation analysis, causal learning focuses on understanding the cause-and-effect relationships between variables. This is particularly important in fault diagnosis and root-cause analysis[29]. The architecture diagram is shown in Figure 4. Causal learning is deeply integrated with the Seq2Seq decoder and attention mechanism to construct an efficient and interpretable fault diagnosis framework. After the input data is processed by the Seq2Seq model to capture temporal features, the causal learning module builds a causal structure graph to model the relationships between variables and performs intervention analysis and causal effect estimation to identify key fault features. Additionally, by combining with the multi-head cross-attention mechanism, causal learning enhances the focus on critical features, seamlessly integrating causal inference results with the decoding process, ultimately producing accurate and interpretable diagnostic outcomes.

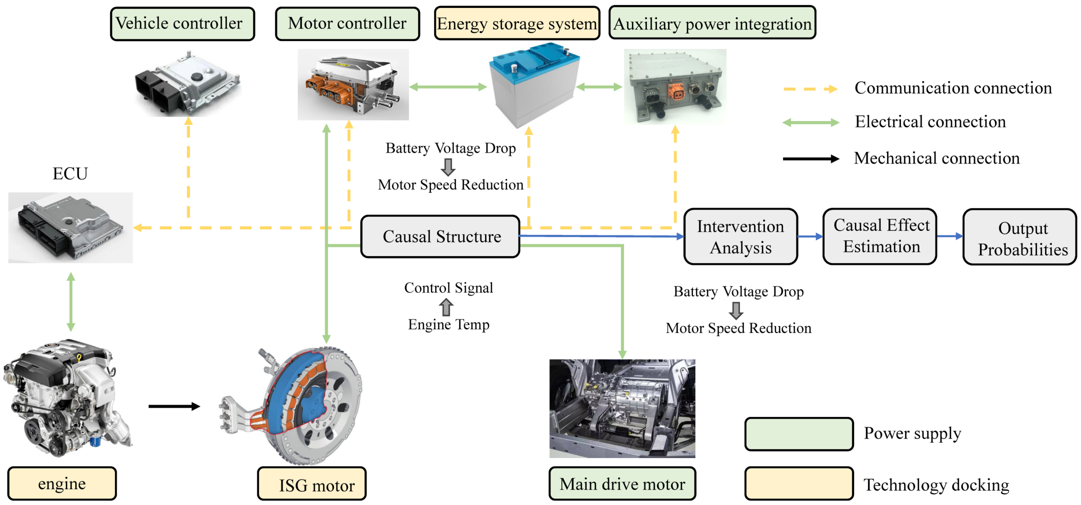

Figure 4. Structure diagram of causal learning.

Causal learning often uses structural equations to express the causal relationships between variables. The causal relationship between variables is expressed as:

$$X_i = f_i(\text{Parents}(X_i), \epsilon_i) \tag{13}$$

Here, $\text{Parents}(X_i)$ represents the direct causal variables of X_i, and ϵ_i represents random noise. Through interventions, causal effects can be quantified. The key idea is to intervene on a variable and predict the resulting changes in other variables:

$$P(Y \mid \text{do}(X = x)) \tag{14}$$

This represents the probability distribution of Y when X is set to x through intervention.

Using domain knowledge and data-driven methods (such as randomized experiments), causal graphs can be constructed. This method is used to estimate the causal effect of interventions. For example, to answer the question "What will happen if we do...?", causal learning helps to identify the root cause of faults.

The Average Causal Effect (ACE) measures the average causal effect of variable X on Y:

$$ACE = \text{E}[Y \mid \text{do}(X = x_1)] - \text{E}[Y \mid \text{do}(X = x_0)] \tag{15}$$

Here, E represents the expected value. Through causal learning, it is possible to identify and isolate key causal relationships, reduce the complexity of dimensionality, and improve the effectiveness and efficiency of diagnosis. Causal learning enables a deeper understanding of the causal relationships between variables and provides a practical tool for explaining faults compared to traditional correlation analysis.

4 Experiment

4.1 Experimental Environment

This study's experimental environment is based on modern deep learning frameworks and aims to evaluate the performance of a proposed vehicle fault diagnosis model based on a sequence-to-sequence (Seq2Seq) model, attention mechanisms, and causal learning. The proposed vehicle fault diagnosis model focuses on identifying four typical fault types: Normal (no issues), Broke (mechanical damage), Net (electrical or communication faults), and Rope (overheating or overload). Input features include sensor data such as engine temperature, RPM, oil pressure, current, and voltage, as well as time-series features like state change trends and anomalies. The experiment is conducted on a server equipped with high-performance computing resources, including an NVIDIA Tesla V100 GPU (16GB VRAM) and an Intel Xeon processor (16 cores). This configuration supports large-scale parallel processing of data and efficient training of deep neural networks, making it particularly suitable for handling complex time-series data and the training requirements of deep learning models. The operating system used is Ubuntu 20.04 LTS, ensuring system stability and compatibility. To build the deep learning experimental environment, the TensorFlow 2.x framework is used, along with Keras for model construction and training. Keras provides a simple API interface, facilitating rapid prototyping of deep learning models. Throughout the experiment, all data processing and model training are performed on the GPU to accelerate the training process. Additionally, Python 3.8 and related deep learning libraries are employed for data processing, model evaluation, and result visualization. To ensure the model's robustness and the stability of the training process, multiple rounds of training and cross-validation methods are used, along with hyperparameter tuning. Specific parameters are shown in Tables 1 and 2. Table 1 presents the basic information of the dataset, including the number of samples, input features, and sequence length, while Table 2 lists the relevant configurations for model training, including the number of layers in the encoder and decoder, hidden layer dimensions, learning rates, and other key parameters.

Table 1. Dataset Parameters

Parameter Name	Description	Value or Range
Total Samples	Number of samples in each dataset	10,000 records
Input Features	Includes sensor data, usage data, etc.	100 features
Sequence Length	Number of time steps in the input sequence	30-60

Table 2. Model Training Parameters

Parameter Name	Description	Value or Range
Encoder Layers	Number of LSTM or GRU layers in the encoder	2-3 layers
Decoder Layers	Number of LSTM or GRU layers in the decoder	2-3 layers

Parameter Name	Description	Value or Range
Hidden Layer Dimensions	Dimensions of hidden layers in the encoder and decoder	128-256
Learning Rate	Learning rate for the optimizer	0.001
Batch Size	Number of samples used in each iteration	32-64
Training Epochs	Total number of training epochs	50
Optimizer	Optimization algorithm used for training	Adam
Dropout Rate	Dropout probability to prevent overfitting	0.3-0.5
Activation Function	Non-linear activation function for hidden layers	ReLU
Regularization Method	Regularization method to prevent overfitting	L2 regularization
Evaluation Metrics	Metrics used to assess model performance	Accuracy, Precision, Recall, F1 score

4.2 Experimental Data

- **UCI Vehicle Data**

The UCI Vehicle Data[30] is a public dataset provided by the UCI Machine Learning Repository, primarily used for vehicle fault diagnosis and prediction. This dataset contains sensor data from different vehicles, including engine temperature, speed, fuel consumption, battery voltage, and other information. By using these multi-dimensional sensor data, researchers can train models to identify and predict potential faults in vehicle systems. The dataset is time-series data, making it suitable for testing fault diagnosis models based on time-series analysis.

- **Ford GoBike**

The Ford GoBike dataset[31] comes from the Ford GoBike bike-sharing system and includes user riding data such as bike rental duration, location, and rental/return station information. While it is not a traditional vehicle fault dataset, the dataset contains detailed records on the usage of shared bikes. Researchers can analyze this data to predict fault patterns, usage frequency, and failure trends. The dataset is suitable for exploring prediction models related to transportation vehicles, particularly in the analysis of the operation and maintenance of shared transportation systems.

- **CACHET**

The CACHET (Comprehensive Assessment of the Health Effects of Transport) dataset[32] focuses on the health effects of vehicle use, particularly health data related to vehicle operation. This dataset includes a large amount of sensor data related to vehicle performance, driving behavior, and environmental conditions. The unique feature of the CACHET dataset is that it includes health factors, such as the physiological data of drivers, which allows researchers to explore health risks associated with vehicle use and potential fault warning mechanisms.

- **Nissan Vehicle Data**

The Nissan Vehicle Data[33] is provided by Nissan Motor Corporation and includes real-time sensor data and fault records from Nissan vehicles. This dataset contains information on vehicle acceleration, temperature, voltage, RPM, fault codes, and more. It is suitable for developing efficient vehicle fault diagnosis systems. The Nissan Vehicle Data set is notable for its large volume and variety, covering vehicle condition monitoring data from various driving scenarios. By analyzing this data, researchers can identify different types of vehicle faults and predict maintenance needs.

4.3 Evaluation Metrics

- **Accuracy**

Accuracy is a fundamental metric for evaluating the overall classification performance of a model. It represents the proportion of correctly classified samples out of the total samples. In vehicle fault diagnosis, accuracy reflects the proportion of correct predictions in all diagnostic results. The formula is:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

Where:
TP (True Positive): The number of samples that are truly faulty and correctly diagnosed as faulty.
TN (True Negative): The number of samples that are truly non-faulty and correctly diagnosed as non-faulty.
FP (False Positive): The number of samples that are truly non-faulty but incorrectly diagnosed as faulty.
FN (False Negative): The number of samples that are truly faulty but incorrectly diagnosed as non-faulty.

- **Precision**

Precision focuses on the proportion of samples predicted as faulty by the model that are actually faulty. In vehicle fault diagnosis, precision is particularly important because false alarms may lead to unnecessary repair costs. High precision means the model can accurately predict faults, reducing the likelihood of false positives and improving the reliability of diagnostics. The formula is:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{17}$$

Improved precision ensures better accuracy in fault prediction.

- **Recall**

Recall measures the proportion of actual faults that the model correctly identifies as faulty. In vehicle fault diagnosis, recall reflects the model's ability to detect faults, especially in early fault prediction. A higher recall means the model can identify more potential faults, reducing the risk of missed diagnoses, which is critical for enhancing vehicle safety and reliability. The formula is:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{18}$$

In real-time fault monitoring and diagnosis, recall directly affects the timeliness and comprehensiveness of fault detection.

- **F1 Score**

The F1 score is the harmonic mean of precision and recall, taking both into account. In fault diagnosis tasks, the F1 score is particularly important because it finds a balance between tolerance for false positives and false negatives. The F1 score is especially useful for imbalanced datasets, as it helps prevent the model from being biased toward one class, providing a more balanced evaluation. The formula is:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{19}$$

A high F1 score indicates that the model excels in both accuracy and completeness, optimizing both fault identification precision and recall.

4.4 Experimental Comparison and Analysis

In this section, we compare the performance of the proposed method with other state-of-the-art models across four datasets. The comparison includes several key performance indicators, such as accuracy, precision, recall, and F1-score. These metrics are essential for evaluating the effectiveness of fault diagnosis systems in vehicles.

Table3. Comparison of relevant indicators of the proposed method with other methods on four datasets.

Model	UCI Vehicle Data				Ford GoBike			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
Min et al. [34]	89.35	89.64	89.84	89.74	89.62	86.69	86.45	86.57
Zhang et al. [35]	90.22	90.89	87.45	89.14	89.67	87.18	90.14	88.64
Guo et al. [36]	87.43	87.38	88.78	88.07	91.14	87.58	87.62	87.60
Surendran et al. [37]	89.83	91.84	90.31	91.07	91.49	88.56	88.70	88.63
Shi et al. [38]	87.15	91.06	88.13	89.57	91.68	87.45	86.28	86.86
Mao et al. [39]	88.65	88.97	89.02	88.99	88.05	90.58	90.89	90.73
Ours	93.54	92.46	93.67	93.06	92.49	93.26	94.24	93.75
Model	CACHET				Nissan Vehicle Data			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
Min et al. [34]	87.43	88.97	91.63	90.28	88.36	89.89	88.38	89.13
Zhang et al. [35]	87.32	89.70	88.04	88.86	89.36	91.57	89.74	90.65
Guo et al. [36]	87.55	90.49	90.78	90.63	89.13	86.98	89.96	88.44
Surendran et al. [37]	89.72	88.95	89.10	89.02	87.43	90.42	88.27	89.33
Shi et al. [38]	90.32	89.00	91.68	90.32	88.27	90.92	88.33	89.61
Mao et al. [39]	88.61	90.21	88.26	89.22	89.22	88.19	86.23	87.20
Ours	93.79	94.27	93.26	93.76	92.54	92.48	93.71	93.09

From Table 3, we observe that the proposed method consistently outperforms all other models in every metric across all four datasets. On the UCI Vehicle Data, the accuracy of the proposed model (93.54%) exceeds that of the second-best method (Zhang et al., 90.22%) by a significant margin. The same trend is seen on the Ford GoBike dataset, where the proposed method achieves an accuracy of 92.49%, surpassing other models by 1-4 percentage points. This consistent superiority across both datasets highlights the robustness of the proposed method in handling vehicle fault diagnosis tasks. Notably, the proposed model also leads in precision, recall, and F1-score, which indicates its balanced performance in both minimizing false positives and detecting true faults effectively. Figure 4 provides a visual comparison of these results.

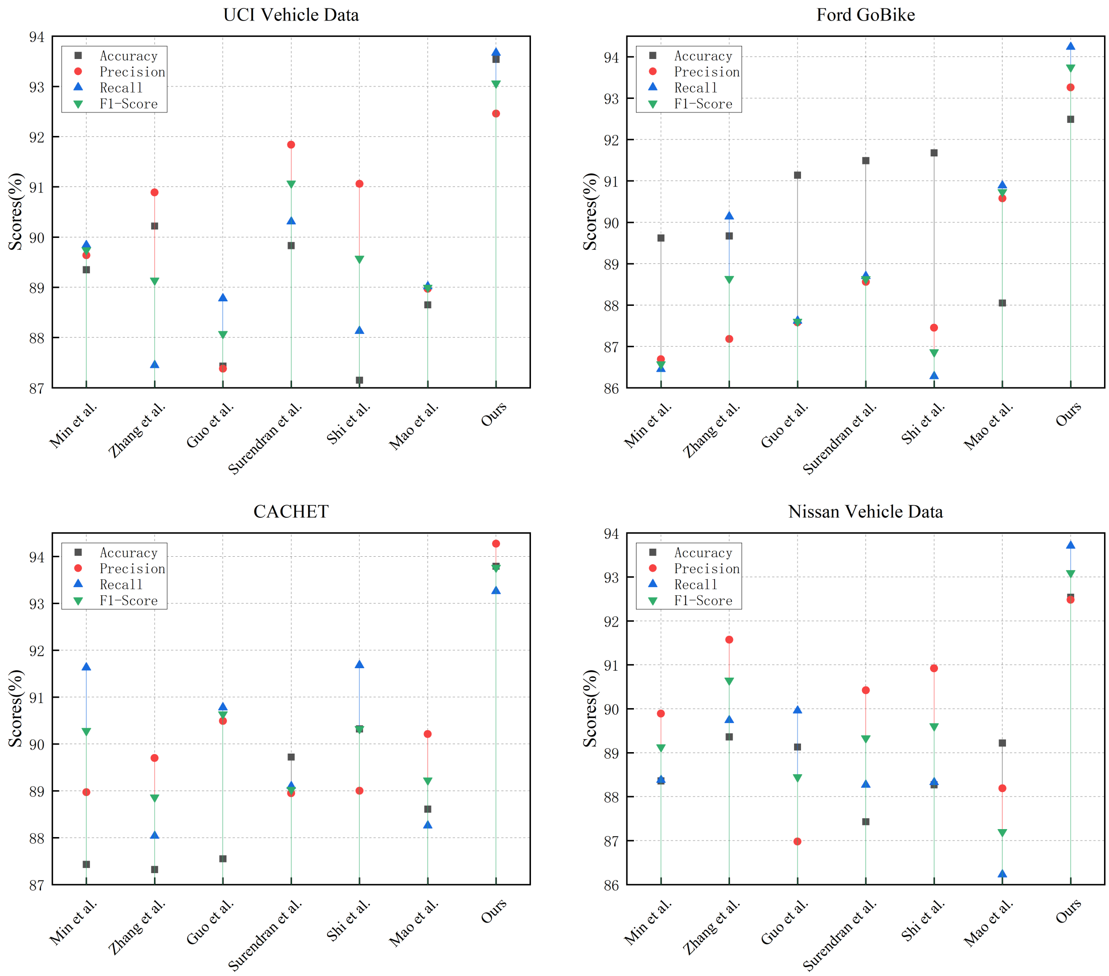

Figure 4. Visual comparison of relevant indicators on four datasets.

Next, we examine the comparison of training indicators, as shown in Table 4. The table presents key training metrics such as the number of parameters, inference time, and training time for the various models on each dataset. These indicators provide insights into the computational efficiency and scalability of the methods.

Table4. Comparison of training indicators on four datasets.

Model	UCI Vehicle Data Parameters (M)	UCI Vehicle Data Inference Time(ms)	UCI Vehicle Data Trainning Time(s)	Ford GoBike Parameters(M)	Ford GoBike Inference Time(ms)	Ford GoBike Trainning Time(s)
Min et al. [34]	363.35	349.77	275.70	365.73	378.20	236.13
Zhang et al. [35]	382.90	376.91	207.81	352.67	360.52	300.09
Guo et al. [36]	353.83	366.64	213.75	383.75	353.37	223.53
Surendran et al. [37]	400.11	352.18	205.15	350.51	373.50	225.80
Shi et al. [38]	376.79	343.17	208.60	371.96	375.60	277.39
Mao et al. [39]	360.41	333.23	272.33	398.31	382.76	285.55
Ours	347.62	312.84	168.74	343.47	322.64	187.52
Model	CACHET Parameters (M)	CACHET Inference Time(ms)	CACHET Trainning Time(s)	Nissan Vehicle Data Parameters(M)	Nissan Vehicle Data Inference Time(ms)	Nissan Vehicle Data Trainning Time(s)
Min et al. [34]	403.24	354.13	285.53	360.72	371.56	257.41
Zhang et al. [35]	374.91	343.92	204.23	391.26	313.66	255.18
Guo et al. [36]	387.27	361.94	200.73	375.87	302.38	239.21
Surendran et al. [37]	393.40	388.08	295.90	371.69	349.91	225.54
Shi et al. [38]	366.50	380.12	196.18	390.11	342.29	301.71
Mao et al. [39]	408.51	366.72	201.00	371.32	362.79	244.44
Ours	341.54	336.74	182.91	353.09	282.76	215.43

Table 4 presents a comparison of training indicators, such as the number of parameters, inference time, and training time, which provide insights into the efficiency of the models in addition to their performance. While the proposed method delivers the best performance, it also demonstrates computational efficiency. For instance, it has fewer parameters compared to other models, such as those by Min et al. and Zhang et al., while maintaining high performance. This suggests that the proposed method is more parameter-efficient, making it potentially easier to deploy in resource-constrained environments. In terms of training time, the proposed method shows significant improvement, with training times of 168.74 seconds for UCI Vehicle Data and 187.52 seconds for Ford GoBike. These values are lower than those of many other models, such as Surendran et al. and Mao et al., indicating that the proposed model not only achieves superior performance but also requires less training time. This efficiency is particularly valuable when working with large-scale datasets and in scenarios requiring real-time fault detection. Figure 5 provides a visual comparison of these results.

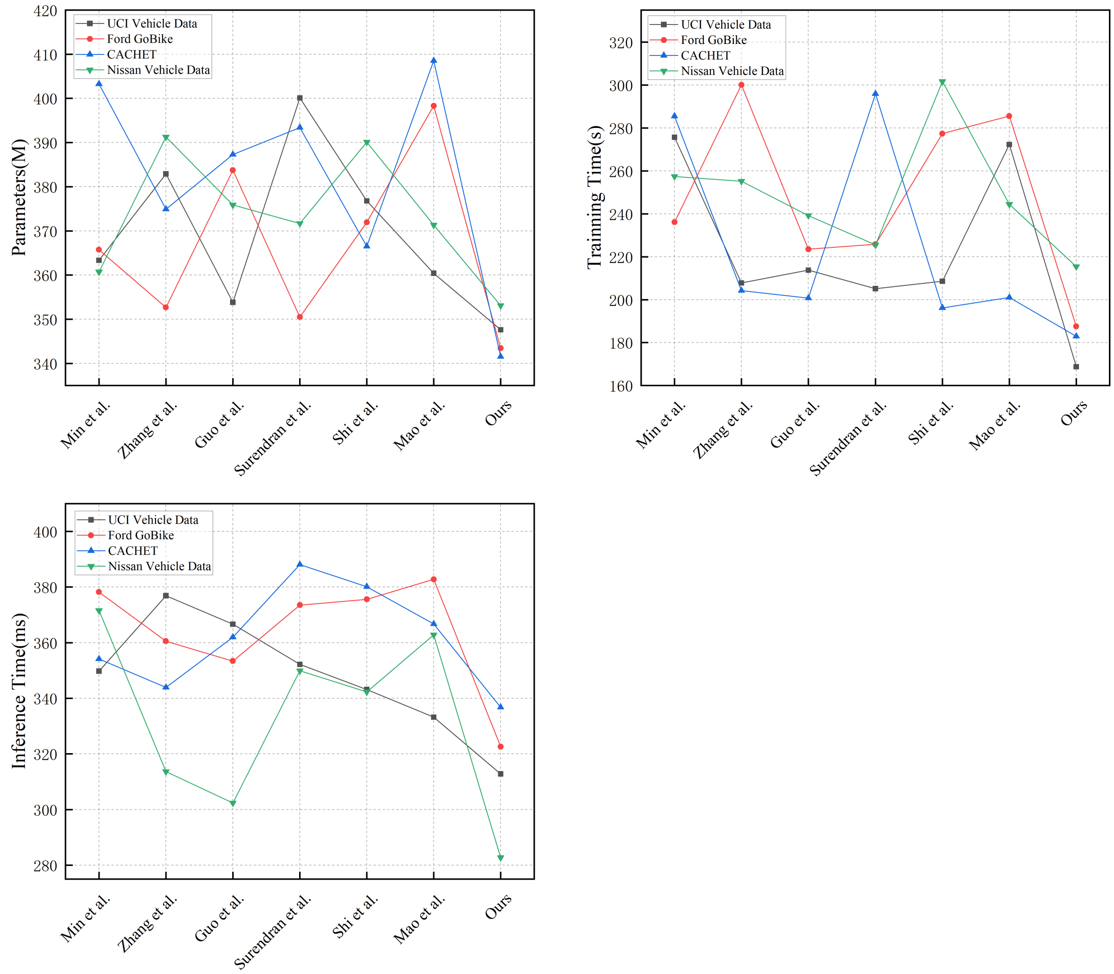

Figure 5. Visual comparison of training indicators.

Furthermore, the results from the ablation experiments, summarized in Table 5, offer additional insights into the contributions of different components of the proposed model. The ablation study includes the baseline model, which is progressively enhanced by adding the Seq2Seq model, the attention mechanism (Att), and the combination of both (Seq2Seq-Att).

Table5. Ablation experiments on four datasets.

Model	UCI Vehicle Data				Ford GoBike			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
baseline	79.24	78.37	78.26	78.31	79.52	81.16	80.29	80.72
+Seq2Seq	85.24	85.42	86.47	85.94	85.53	88.61	87.18	87.89
+Att	89.17	88.06	90.54	89.28	88.24	90.32	90.92	90.62
+Seq2Seq-Att	93.54	92.46	93.67	93.06	92.49	93.26	94.24	93.75
Model	CACHET				Nissan Vehicle Data			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
baseline	79.41	80.64	79.32	79.97	79.64	78.23	78.37	78.30
+Seq2Seq	83.48	84.67	85.27	84.97	84.54	86.37	85.28	85.82
+Att	89.72	88.32	89.95	89.13	89.84	89.47	90.42	89.94
+Seq2Seq-Att	93.79	94.27	93.26	93.76	92.54	92.48	93.71	93.09

Table 5 further elaborates on the contribution of different model components through an ablation study. The baseline model shows a lower performance in terms of accuracy, precision, recall, and F1-score. The addition of the Seq2Seq model provides a significant boost in all metrics, reflecting the importance of sequence modeling in fault diagnosis. When the attention mechanism (Att) is incorporated, the model's ability to focus on critical features is further enhanced, leading to improved performance, especially in recall and F1-score. The combination of Seq2Seq and attention mechanisms (Seq2Seq-Att) results in the highest performance across all datasets, confirming that both components contribute substantially to the model's overall effectiveness. On the UCI Vehicle Data, for instance, the Seq2Seq-Att model achieves an accuracy of 93.54%, a marked improvement over the baseline’s 79.24%. The increase in recall also indicates that the proposed model is better at identifying true faults, which is crucial for safety-critical applications. Figure 6 visually depict these trends.

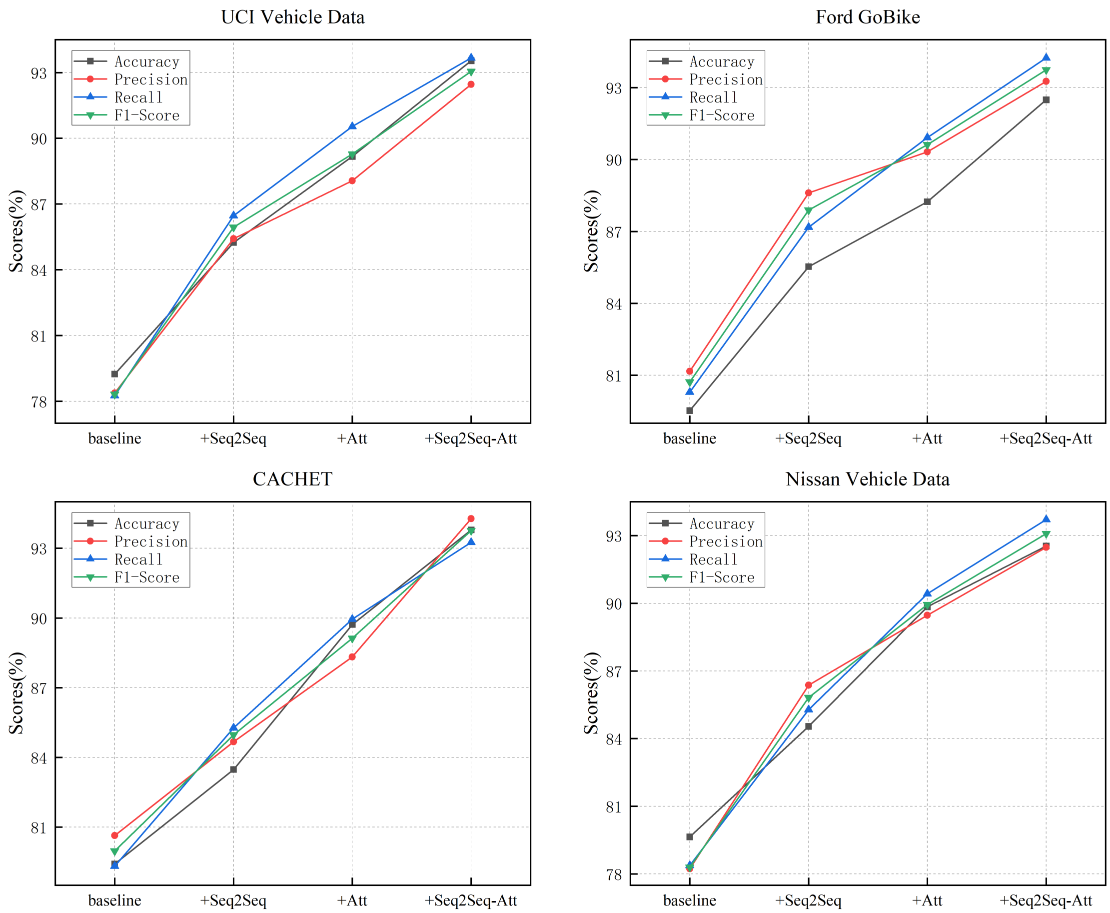

Figure 6. Visual comparison of ablation experiments on four datasets.

5 Conclusion

In this paper, we proposed a novel vehicle fault diagnosis model based on a sequence-to-sequence (Seq2Seq) architecture with attention mechanisms. The model aims to improve the accuracy and efficiency of diagnosing potential faults in vehicles by leveraging time-series sensor data. Our approach was tested on four different datasets: UCI Vehicle Data, Ford GoBike, CACHET, and Nissan Vehicle Data, and the results demonstrated its superiority over existing state-of-the-art methods. The experimental results show that our proposed method significantly outperforms other models across all datasets in key performance metrics, including accuracy, precision, recall, and F1-score. The model's ability to accurately diagnose faults and predict maintenance needs was consistently better, making it a highly effective tool for real-world vehicle monitoring. In addition to its high performance, our approach is computationally efficient, with lower training times and fewer parameters compared to other methods, which makes it suitable for real-time applications in resource-constrained environments. Through the ablation study, we also highlighted the crucial roles of the Seq2Seq model and attention mechanism in improving fault detection. The combination of these components allowed the model to achieve superior performance, especially in detecting rare faults and minimizing false negatives, which is critical for ensuring vehicle safety and reliability. In conclusion, the proposed vehicle fault diagnosis model not only demonstrates state-of-the-art performance but also offers practical advantages in terms of efficiency and scalability.

Future work will explore further enhancements, such as incorporating additional sensor modalities or improving the interpretability of the model, to further enhance its applicability in diverse real-world scenarios.

Funding
Not applicable

Author Contributions
Weidong Huang and Jiahuai Ma contributed to conceptualization, methodology, and investigation. Jiahuai Ma supervised the project, conducted formal analysis, and reviewed the manuscript. Both authors participated in writing and approved the final manuscript.

Institutional Reviewer Board Statement
Not applicable

Informed Consent Statement
Not applicable

Data Availability Statement
Not applicable

Conflict of Interest
The authors declare no conflict of interest.

Reference

[1] M. A. Rahim, M. A. Rahman, M. M. Rahman, A. T. Asyhari, M. Z. A. Bhuiyan, and D. Ramasamy, "Evolution of IoT-enabled connectivity and applications in automotive industry: A review," Vehicular Communications, vol. 27, p. 100285, 2021.

[2] J. S. Liang, "A process-based automotive troubleshooting service and knowledge management system in collaborative environment," Robotics Computer-Integrated Manufacturing, vol. 61, p. 101836, 2020.

[3] A. Theissler, J. Pérez-Velázquez, M. Kettelgerdes, and G. Elger, "Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry," Reliability engineering system safety, vol. 215, p. 107864, 2021.

[4] G. Zhang, T. Zhou, and Y. Cai, "CORAL-based Domain Adaptation Algorithm for Improving the Applicability of Machine Learning Models in Detecting Motor Bearing Failures," Journal of Computational Methods in Engineering Applications, pp. 1-17, 2023.

[5] P. Arthurs, L. Gillam, P. Krause, N. Wang, K. Halder, and A. Mouzakitis, "A taxonomy and survey of edge cloud computing for intelligent transportation systems and connected vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 7, pp. 6206-6221, 2021.

[6] T. Zhang et al., "Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions," ISA transactions, vol. 119, pp. 152-171, 2022.

[7] T. Z. G. Zhang, and W. Huang, "Research on Fault Diagnosis of Motor Rolling Bearing Based on Improved Multi-Kernel Extreme Learning Machine Model," Artificial Intelligence Advances, vol. 5(1), 41–48, 2023, doi: https://doi.org/10.30564/aia.v5i1.7489.

[8] N. Marzo i Grimalt, "Natural Language Processing Model for Log Analysis to Retrieve Solutions For Troubleshooting Processes," ed, 2021.

[9] A. Trilla, N. Mijatovic, and X. Vilasis-Cardona, "Towards learning causal representations of technical word embeddings for smart troubleshooting," International Journal of Prognostics Health Management, vol. 13, no. 2, 2022.

[10] Z. Zhou, J. Wu, Z. Cao, Z. She, J. Ma, and X. Zu, "On-Demand Trajectory Prediction Based on Adaptive Interaction Car Following Model with Decreasing Tolerance," in 2021 International Conference on Computers and Automation (CompAuto), 2021: IEEE, pp. 67-72.

[11] A. Abid, M. T. Khan, and J. Iqbal, "A review on fault detection and diagnosis techniques: basics and beyond," Artificial Intelligence Review, vol. 54, no. 5, pp. 3639-3664, 2021.

[12] J. Jhung, H. Suk, H. Park, and S. Kim, "Hardware accelerators for autonomous vehicles," in Artificial Intelligence and Hardware Accelerators: Springer, 2023, pp. 269-317.

[13] Y. Hao, Z. Chen, J. Jin, and X. Sun, "Joint operation planning of drivers and trucks for semi-autonomous truck platooning," Transportmetrica A: Transport Science, pp. 1-37, 2023.

[14] Y. Chi, Y. Dong, Z. J. Wang, F. R. Yu, and V. C. Leung, "Knowledge-based fault diagnosis in industrial internet of things: a survey," IEEE Internet of Things Journal, vol. 9, no. 15, pp. 12886-12900, 2022.

[15] M. H. Lipu et al., "Intelligent algorithms and control strategies for battery management system in electric vehicles: Progress, challenges and future outlook," Journal of Cleaner Production, vol. 292, p. 126044, 2021.

[16] Z. Zhu and H. Zhao, "A survey of deep RL and IL for autonomous driving policy learning," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 9, pp. 14043-14065, 2021.

[17] M. R. Bachute and J. M. Subhedar, "Autonomous driving architectures: insights of machine learning and deep learning algorithms," Machine Learning with Applications, vol. 6, p. 100164, 2021.

[18] C. Du et al., "Research on fault diagnosis of automobile engines based on the deep learning 1D-CNN method," Engineering Research Express, vol. 4, no. 1, p. 015003, 2022.

[19] R. Khaliqi and C. Iulian, "A Deep Learning Approach To Vehicle Fault Detection Based On Vehicle Behavior," ed, 2023.

[20] S. Zhang, F. Ye, B. Wang, and T. G. Habetler, "Semi-supervised bearing fault diagnosis and classification using variational autoencoder-based deep generative models," IEEE Sensors Journal, vol. 21, no. 5, pp. 6476-6486, 2020.

[21] A. He and X. Jin, "Deep variational autoencoder classifier for intelligent fault diagnosis adaptive to unseen fault categories," IEEE Transactions on Reliability, vol. 70, no. 4, pp. 1581-1595, 2021.

[22] X. Li, X. Li, and H. Ma, "Deep representation clustering-based fault diagnosis method with unsupervised data applied to rotating machinery," Mechanical Systems Signal Processing, vol. 143, p. 106825, 2020.

[23] K. K. Sharma, A. Seal, A. Yazidi, and O. Krejcar, "A new adaptive mixture distance-based improved density peaks clustering for gearbox fault diagnosis," IEEE Transactions on Instrumentation

Measurement, vol. 71, pp. 1-16, 2022.

[24] S. Du, T. Li, Y. Yang, and S.-J. Horng, "Multivariate time series forecasting via attention-based encoder–decoder framework," Neurocomputing, vol. 388, pp. 269-279, 2020.

[25] X. Wang, Z. Cai, Y. Luo, Z. Wen, and S. Ying, "Long time series deep forecasting with multiscale feature extraction and Seq2seq attention mechanism," Neural Processing Letters, vol. 54, no. 4, pp. 3443-3466, 2022.

[26] S. Jia, Y. Li, X. Wang, D. Sun, and Z. Deng, "Deep causal factorization network: A novel domain generalization method for cross-machine bearing fault diagnosis," Mechanical Systems Signal Processing, vol. 192, p. 110228, 2023.

[27] W. Yu, I. Y. Kim, and C. Mechefske, "Analysis of different RNN autoencoder variants for time series classification and machine prognostics," Mechanical Systems Signal Processing, vol. 149, p. 107322, 2021.

[28] Y. Li, M. Du, and S. He, "Attention-based sequence-to-sequence model for time series imputation," Entropy, vol. 24, no. 12, p. 1798, 2022.

[29] P. Wu and J. Liu, "Learning causal temporal relation and feature discrimination for anomaly detection," IEEE Transactions on Image Processing, vol. 30, pp. 3513-3527, 2021.

[30] S. Samuelsen, B. Shaffer, J. Grigg, B. Lane, and J. Reed, "Performance of a hydrogen refueling station in the early years of commercial fuel cell vehicle deployment," International Journal of Hydrogen Energy, vol. 45, no. 56, pp. 31341-31352, 2020.

[31] D. A. Pandya, "Station-graph: bike flow prediction using spatio-temporal graph convolutional network," 2020.

[32] F. Colarusso, P. C. Cheng, G. Garcia Garcia, A. Stockdill, D. Raggi, and M. Jamnik, "A novel interaction for competence assessment using micro-behaviors: Extending CACHET to graphs and charts," in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, pp. 1-14.

[33] M. Ragone, V. Yurkiv, A. Ramasubramanian, B. Kashir, and F. Mashayek, "Data driven estimation of electric vehicle battery state-of-charge informed by automotive simulations and multi-physics modeling," Journal of Power Sources, vol. 483, p. 229108, 2021.

[34] H. Min et al., "A fault diagnosis framework for autonomous vehicles with sensor self-diagnosis," Expert Systems with Applications, vol. 224, p. 120002, 2023.

[35] X. Zhang, J. Hong, and X. Xu, "Fault diagnosis of real-scenario battery systems based on modified entropy algorithms in electric vehicles," Journal of Energy Storage, vol. 63, p. 107079, 2023.

[36] Z. Guo, Y. Hao, H. Shi, Z. Wu, Y. Wu, and X. Sun, "A fault diagnosis algorithm for the dedicated equipment based on the CNN-LSTM mechanism," Energies, vol. 16, no. 13, p. 5230, 2023.

[37] R. Surendran, O. I. Khalaf, C. A. Tavera Romero, and Continua, "Deep learning based intelligent industrial fault diagnosis model," Computers, Materials, vol. 70, no. 3, 2022.

[38] J. Shi, D. Peng, Z. Peng, Z. Zhang, K. Goebel, and D. Wu, "Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks," Mechanical Systems Signal Processing, vol. 162, p. 107996, 2022.

[39] W. Mao, W. Feng, Y. Liu, D. Zhang, and X. Liang, "A new deep auto-encoder method with fusing discriminant information for bearing fault diagnosis," Mechanical Systems Signal Processing, vol. 150, p. 107233, 2021.