Research Article

SmartFault-M: A Hybrid CNN–LSTM with Attention for Multi-Domain Fault Diagnosis in Industrial Mechatronic Systems

Syed Faraz Raza [1*], Muhammad Faheem Khan [2], Muhammad Haseeb Zia [3]

[1] Department of Computer Science, University of Alabama at Birmingham, 1402 10th Avenue S., Birmingham, AL 35294, USA
[2] Department of Computer Science, TIMES Institute, Multan, 60000, Pakistan
[3] Department of Criminology and Forensic Sciences, Lahore Garrison University, Lahore, Pakistan
*Corresponding Author: Syed Faraz Raza Author Email: razas@uab.edu

Received: September 19, 2025 **Accepted:** May 4, 2026

Abstract: Diagnosing faults in industrial mechatronics can be difficult because the systems involve intricate interactions among mechanical, electrical, and control sub-systems which produce multidimensional, non-stationary data from various sensors. Traditional fault diagnostics models based on model-based or data-driven approaches in single domains cannot provide adequate performance because they cannot effectively capture the multi-scale and temporal nature of faults. Therefore, in this paper, a novel approach called Smart Fault Monitoring (SmartFault-M), a hybrid model based on a dual-stream CNN–LSTM framework for smart fault diagnosis, is proposed, in which two parallel streams with time-domain and frequency-domain feature maps extract informative features and attention-based learning captures important patterns. In SmartFault-M, CNN and LSTM are used to learn spatial and temporal information from the data, respectively, while attention learning emphasizes the informative features. For validating the effectiveness of our proposed approach, we perform extensive experiments using the bearing dataset of Case Western Reserve University (CWRU) under different fault states. As experimental results show, the accuracy of the proposed SmartFault-M reaches 98.25% and outperforms baseline methods like CNN (91.42%), LSTM (92.36%), and CNN–LSTM (95.18%) models. In addition, the model shows consistently high precision, recall, and F1-score for all classes of faults. From these experiments, it is obvious that our proposed model with multi-domain feature extraction and attention-based learning achieves superior performance for smart fault diagnosis tasks.

Keywords: Fault Diagnosis, Mechatronics, Deep Learning, Hybrid Model, Convolutional Neural Network–Long Short-Term Memory

1. Introduction

In contemporary industry, there is an intensive use of advanced technologies for building mechatronic devices such as manufacturing systems, robots, etc. Such systems incorporate various mechanical elements combined with electronics, sensors, and intelligent control [1]. These technological innovations positively impact productivity, precision, and automation of processes but result in complex mechatronic systems requiring efficient fault diagnosis. The latter is associated with high risk of unplanned system breakdowns and related costs. At present, there is a need for the development of intelligent fault-diagnostic solutions as the current approaches cannot guarantee sufficient performance and flexibility of operations [2].

There exist two main strategies that can be used for detecting and identifying malfunctions. Model-based methods are based on accurate mathematical descriptions of the system dynamics while data-driven approaches involve analysis of available datasets [3]. As a result, the former provides interpretable explanations of faults at the cost of difficulty in developing a correct model of mechatronic processes while the latter are highly flexible due to extensive application of machine learning techniques. Recently, a number of studies were devoted to the development of novel deep learning architectures for the automation of feature generation, extraction, and processing but existing solutions demonstrate only partial applicability for faults diagnosis [4].

Another issue is the failure of current methods to properly address the problem of analyzing multi-domain features. As a rule, industrial datasets contain data recorded from vibration sensors which include time-domain and frequency-domain components [5]. Thus, the inability to use this information results in poor accuracy as well as interpretability of the decision-making process. Moreover, attention mechanisms are currently underutilized in the field, even though they can facilitate the extraction of features and provide interpretable results [6].

The need for addressing the identified limitations and gaps of current research is associated with growing demand in efficient and scalable diagnostic solutions. The development of novel hybrid models with high fault separability and interpretability would enable more precise identification of malfunctions and support effective implementation of predictive maintenance [7]. These characteristics are vital in the era of Industry 4.0 as intelligent autonomous systems should be able to function independently from experts. Therefore, the research is aimed at designing a novel framework for automated fault diagnosis [8].

The aim of the research is to design and test the efficiency of a multi-dimensional deep learning framework for the diagnostics of malfunctions in industrial systems. Specifically, the SmartFault-M method will integrate several approaches such as time and frequency domains analysis, sequential modeling, and attention mechanism. The latter will allow identifying most useful features and highlighting their contribution into diagnostic decisions. This approach can improve the accuracy and flexibility of fault separation processes.

This study is important for advancing the field by addressing multiple limitations of current research approaches. First of all, the proposed framework will enhance the diagnostic process by incorporating multiple representations, analyzing temporal and spatial features, and allowing adaptive feature extraction. Consequently, the fault separability will increase as well as its predictability. Finally, the developed framework will ensure better interpretability of the decision-making process which is crucial in safety-critical industries.

The key contributions of the research are highlighted through the design and validation of the proposed hybrid fault detection scheme which integrates three distinct learning methods including spatial, temporal, and attention-based mechanisms within one single framework. The current study demonstrates that combining characteristics from both time and frequency domains improves the performance of fault detection algorithms as compared to the traditional strategy where only one domain is considered. This study demonstrates that incorporating attention mechanism increases explainability of the employed features and reduces the impact of noise in the input signal. Moreover, employing sequential learning enables the fault detector to identify temporal faults. Promising results obtained using the benchmark dataset demonstrate the effectiveness of the proposed algorithm. Overall, the current study makes significant contribution towards the area of intelligent fault detection.

This paper is organized as follows: In Section II, we present a survey of the existing literature concerning fault detection and highlight the challenges faced by the conventional method. The architecture of our proposed SmartFault-M framework is described in Section III. Section IV presents the experimental results obtained after testing the model on the benchmark dataset. Finally, conclusions are presented in Section V.

2. Literature Review

This section highlights some major findings from different research studies that have been performed on the issue of fault diagnosis in industrial and mechatronic systems. The latest advancements and trends in the area are highlighted in this literature review. The identification of fault in the industrial system is regarded as an important task in the contemporary technological era. Fault diagnosis helps in increasing the reliability, safety, and efficiency of mechatronic systems and enables the engineers to adopt preventive maintenance strategies. Thus, it becomes necessary to find effective techniques that can help in diagnosing faults easily. The initial work on fault diagnosis was mainly based on the model-based approach in which there was a need for representing the system mathematically. Due to this reason, these models were difficult to use in complex industrial settings.

2.1 Fault Diagnosis and Prognosis

Numerous studies have attempted to integrate ML techniques and system-level perspectives to diagnose faults effectively. The combination of multilayer perceptron technique and SHapley Additive exPlanations (SHAP) is effective in diagnosing faults and pinpointing the contributions of the individual features [9]. Further, the implementation of hybrid deep learning techniques that use the graph attention mechanism and the CNN-BiLSTM architecture has been quite effective in improving the performance of the fault detection algorithm in motor current [10]. These techniques have been successful in diagnosing the faults in industry settings, but they are more complex and cannot be applied in real time. There exist unsupervised approaches that include clustering and forecast based on neural networks to predict RUL in bearing systems [11]. However, the utilization of health indicators and clustering approaches makes them less flexible to different scenarios. In deep learning approaches, such as LSTM, BiLSTMs, and hybrid CNN-LSTM have been quite effective in predicting the RUL in manufacturing industries [12]. These techniques work by modeling the temporal dependencies, which leads to higher precision rates. However, most of these approaches focus on analyzing signals in one domain only.

2.2 Practical Aspects

Certain works have been done towards solving practical problems encountered while diagnosing faults in the industry. Multi-modal MobileNetV2-based model was proposed to detect faults in electrical machine via Park's Vector trajectories and by generating synthetic data [13]. Even though this method seems to be promising, its disadvantage is that it uses synthetic data, which means that it may not be generalized in other datasets. In addition, there are vision-based models using YOLO architecture that can automate tasks in mechatronic systems [14]. These models work well in detecting objects; however, they do not solve dynamic fault detection problem. Moreover, hybrid neural models were proposed that detect faults in real time and produce excellent results in terms of accuracy and low latency [15]. The only drawback of these models is that they lack feature selection mechanism.

2.3 Optimization-Based Methods

Several studies have examined several approaches for optimizing models for better performance. Some researchers have improved the efficiency of deep learning models through the application of hybrid optimization methods based on the combination of global and local searches in the field of deep learning using standard benchmarks [16], but these methods have added more computational complexities. Conversely, digital twin models that are based on physics simulations and machine learning models have been successful in achieving high levels of robustness and explainability [37]. These techniques require high levels of modeling accuracy. Transformer models have shown impressive feature-extraction abilities and near-perfect accuracies in fault classifications [18]. But due to higher computational complexity, they cannot be used in industries.

2.4 Fault Detection in Noisy Conditions

Several studies have concentrated on improving robustness in case of noisy operational conditions. For instance, approaches that combine advanced signal processing methods like wavelets and MFCC in combination with CNN-BiLSTM architectures have produced very high levels of accuracy for different types of data sets [19]. Likewise, approaches involving attention-based spatiotemporal networks that use convolutional and recurrent neural networks have performed efficiently in classifying faults with multiple states [20]. However, some approaches require complicated preprocessing steps.

2.5 Interpretability and Explainability

Many studies have been conducted on interpretability and explainability in machine learning models. Some fault diagnosis models based on the SHAP and LIME methods have achieved a high level of accuracy by offering an explanation for the importance of each feature in the harmonic drive system [35]. In addition, some attention-based models with noise reduction methods have shown efficiency in addressing noise problems [22]. However, such methods use hand-crafted features with parameter tuning preprocessing steps [23].

2.6 Survey and Comparison Studies

Various researchers have conducted comprehensive research into fault diagnosis and prediction techniques. The survey papers found that data imbalance, nonexistence of real data, and ignoring edge computing requirements are some of the major problems with intelligent fault diagnosis [25] [36]. In addition, some researchers proved through comparisons that hybrid deep learning models, such as CNN-LSTM, work well for predictive maintenance tasks [26]. Feature fusion approaches using multiple pre-trained models have proven effective in predicting maintenance problems yet raise computation costs [27]. Models based on signal processing have worked well to detect faults but cannot learn automatically [28].

2.7 Specialized Applications

Many other researchers have investigated the applications of specialized techniques in industrial environments. HIL-based approaches have allowed researchers to develop real-time fault-detection schemes for automobiles through the use of hybrid CNN-LSTM models [29]. Hybrid approaches that involve both physics-informed and data-based techniques have helped improve the generalization process for motor faults diagnosis [34]. Physics-informed neural networks have also been leveraged to enable researchers to incorporate the knowledge of the field into learning algorithms [31]. Likewise, visual models and transfer learning have been leveraged to carry out fault diagnosis tasks, but their practical implementations have not yet been widely carried out [32]. The use of model-based approaches for linear parameter-varying systems has helped achieve fault diagnoses for motors, although they require accurate modeling of the systems [33].

2.8 Research gap

The significant advancements made in fault diagnosis and predictive maintenance; some fundamental research gaps still exist. Many state-of-the-art methodologies utilize a signal representation from one domain only, thus failing to fully leverage the inherent properties of the systems under consideration. Poor utilization of time, frequency, and spatial information also remains an important limitation in developing efficient diagnostic tools. In addition, there are many methods that require a lot of preprocessing steps and feature engineering to operate properly.

Moreover, some of the latest models do not incorporate a method for adaptive multi-sensor feature selection. Even though attention-based mechanisms and explainable artificial intelligence have attracted attention in recent years, the number of applications has remained low. In addition, even those models that demonstrate good accuracy are usually quite complex due to optimization requirements and are therefore hard to deploy on edge devices.

The other important gap is the lack of robustness to environmental factors such as noise, data imbalance, and changing operational settings. In particular, many papers focus on using benchmarks and artificial test data that does not generalize well to real-world situations. Moreover, the problem of a trade-off between explainability and performance has remained open in many cases.

3. Methodology

3.1 Overview of the Proposed Approach

The SmartFault-M method proposed in the article was designed specifically to cope with the complexity and variability of diagnosing faults in contemporary industrial mechatronic systems. Due to the interplay of the mechanical, electronic, and control elements of such devices, the data generated by sensors are high-dimensional and non-stationary. In addition, faulty events rarely occur as distinct patterns. Instead, they tend to develop gradually and display various temporal dependencies and frequency-specific traits. As a result, fault diagnostic software has to be capable of representing the dynamics of the multi-domain data and focusing on informative features to achieve better performance.

As opposed to classical approaches in which feature extraction and classification tasks are accomplished separately, the proposed methodology relies on an end-to-end deep learning architecture that encompasses signal processing, dual-domain feature extraction, temporal modeling, and feature refinement procedures. The vibration data is acquired in a fault scenario-dependent manner and transformed into structured representations conducive to learning processes. Specifically, they need to take into account transients in addition to periodicity of the input signals in both domains. After being fed into the hybrid architecture consisting of CNNs and LSTMs for spatial and temporal processing, respectively, the representations undergo further refinement by means of an attention mechanism. This step helps to reduce the influence of noise and provide more interpretable results since the model can focus on the informative aspects of the input signals. Finally, the output of the network consists of probabilistic fault predictions in a given number of classes.

This framework is based on vibration data collected under various fault conditions and is converted into structured representations capable of being learned from. In this context, the vibration data is analyzed in terms of both time and frequency to guarantee that not only transients but also cyclic fault patterns can be accurately detected. Next, the structured representations will go through a hybrid structure where the CNN algorithm will learn spatial patterns while the LSTM one will extract temporal information from the representations. To further enhance the effectiveness of the models, an attention scheme is employed to focus on critical features and filter out irrelevant noise information.

The entire process flow of the proposed approach is depicted in Fig. 1, which shows the entire processing flow from acquisition of the raw signal up to the classification of the faulty condition. In this process flow, the vibration signal acquisition step occurs first based on varying operating conditions, followed by the pre-processing stage and further transformation of the acquired signals in both the time domain and frequency domain. This is then followed by the utilization of the dual-stream CNN network to extract spatial features, LSTM for temporal modeling, and attention-based refinement.

The proposed method is implemented in a step-by-step manner, whereby each module helps in processing information related to faults in a way that enhances its refinement, as illustrated in Figure 1 below. It can be seen from the figure that input data is first converted to bi-modal data and then analyzed using parallel CNN modules. The obtained feature sets are then merged and fed into an LSTM network in order to incorporate temporal dynamics, after which the attention-based mechanism helps select important features.

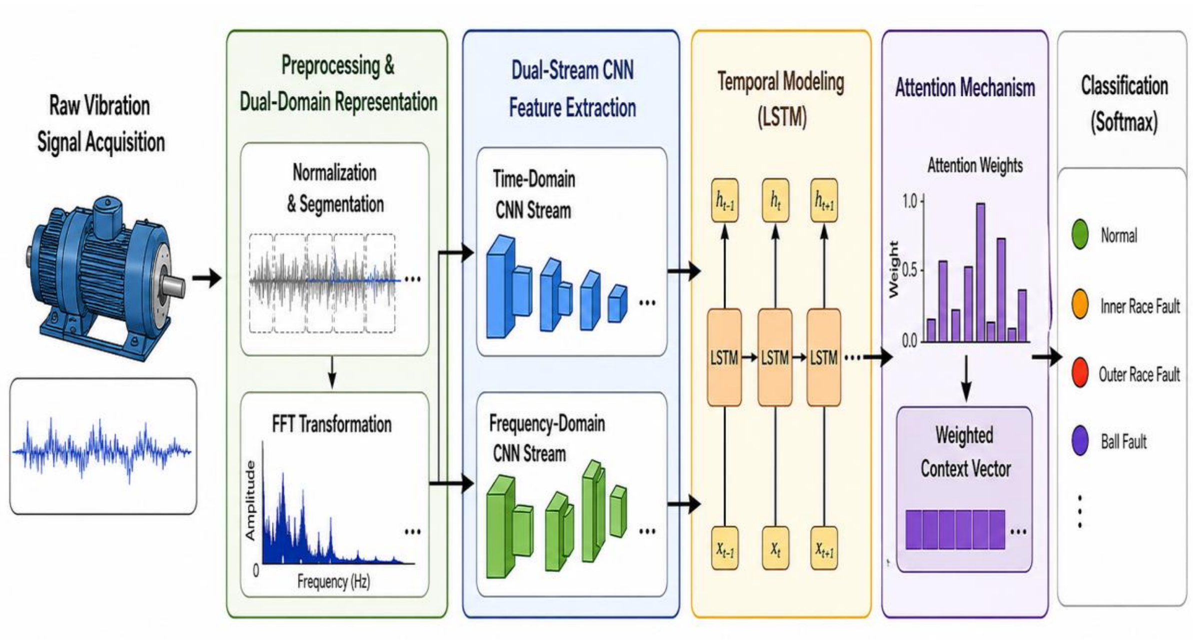


Figure 1 Overall architecture of the proposed SmartFault-M framework

3.2 Hybrid SmartFault-M Framework

The internal structure of SmartFault-M utilizes a systematic series of transformations designed to transform crude sensor readings into diagnostic fault representations. Every element of the system has been chosen to cater for particular weaknesses exhibited by earlier methods, leading to an optimal balance between flexibility and processing effectiveness.

3.2.1 Signal Preprocessing and Multi-Domain Representation

The first step in the suggested methodology involves processing the raw vibration signal to convert it into a format that is compatible with deep learning techniques. Denote the obtained signal as:

$$x(t) = \{x1, x2, \ldots, xN\}$$

Because the amplitude and frequency of the signals can change for various reasons, the data is first normalized so that there would be numerical stability. Normalization is defined as:

$$x_{i'} = \frac{x_i - \mu}{\sigma} \quad (1)$$

Where μ and σ denote the mean and standard deviation of the signal, respectively. This transformation ensures that the input data is centered and scaled, facilitating efficient convergence during training.

To enable localized pattern learning and increase the number of training samples, the normalized signal is segmented into overlapping windows of fixed length L. Each segment is expressed as:

$$X_k = \{x_{k'}, x_{k+1}', \ldots, x_{k+L-1}'\} \quad (2)$$

In addition to time-domain representation, frequency-domain features are extracted using spectral transformation. Specifically, the discrete Fourier transform is applied to each segment to reveal periodic components associated with mechanical faults:

$$X(f) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi f n/N} \quad (3)$$

The simultaneous use of time-domain and frequency-domain representations provides a richer feature space, allowing the model to capture both transient anomalies and harmonic fault signatures.

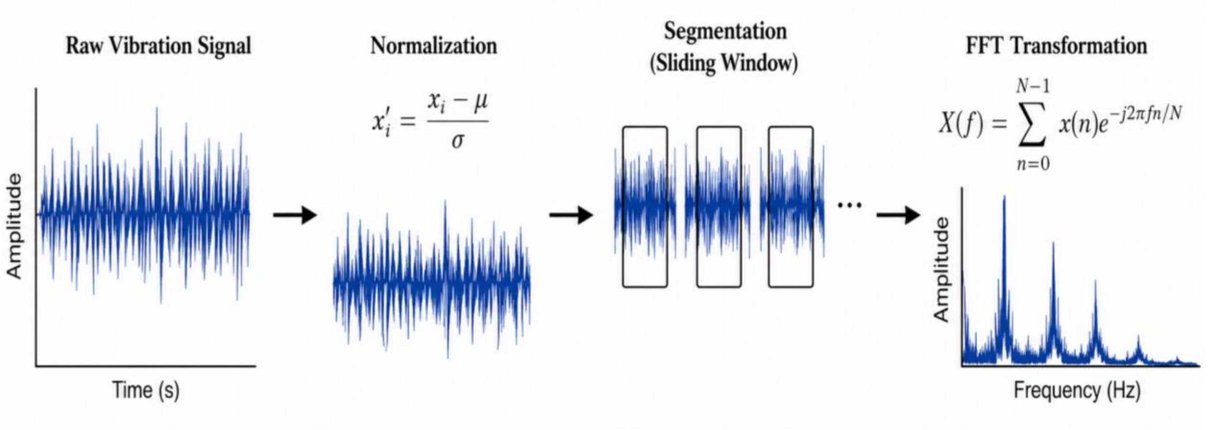


Figure 2 Signal preprocessing and dual-domain representation, including normalization, segmentation, and FFT-based transformation

As shown in Fig. 2, the vibration signal is initially normalized for eliminating any changes in amplitude. It is subsequently divided into segments of equal size through the use of a sliding window. This helps increase the number of training data sets while maintaining time continuity. Through FFT transform, each segment is transformed to its frequency counterpart, which facilitates the identification of periodic fault features.

3.2.2 Dual-Stream Convolutional Feature Extraction

In order to maximize learning from the multidimensional representations, there is a need for a dual stream convolutional model. This is because the time-domain data and the frequency-domain data are separately analyzed using two different convolutional neural networks. The purpose of this is to enable both domains to have their own hierarchical representations.

Within each stream, convolutional layers perform localized filtering operations that can be expressed as:

$$h^{(l)} = \sigma\left(W^{(l)} * h^{(l-1)} + b^{(l)}\right) \tag{4}$$

where $W^{(l)}$ represents the convolutional kernel at layer lll, $b^{(l)}$ is the bias term, and σ(·) denotes a nonlinear activation function such as ReLU. These layers progressively capture low-level patterns such as peaks and oscillations, as well as higher-level abstractions related to fault characteristics.

The outputs of the two streams are subsequently fused through concatenation:

$$H = [H_t; H_f] \tag{5}$$

This fusion step integrates complementary information from both domains, enabling the model to form a more comprehensive understanding of the system's condition.

This architecture in Figure 3 illustrated that, every input domain passes through convolutional layers independently of each other. This helps the algorithm learn specific features for every domain. Then the features obtained are concatenated to form one single feature vector representing the information extracted from both domains.

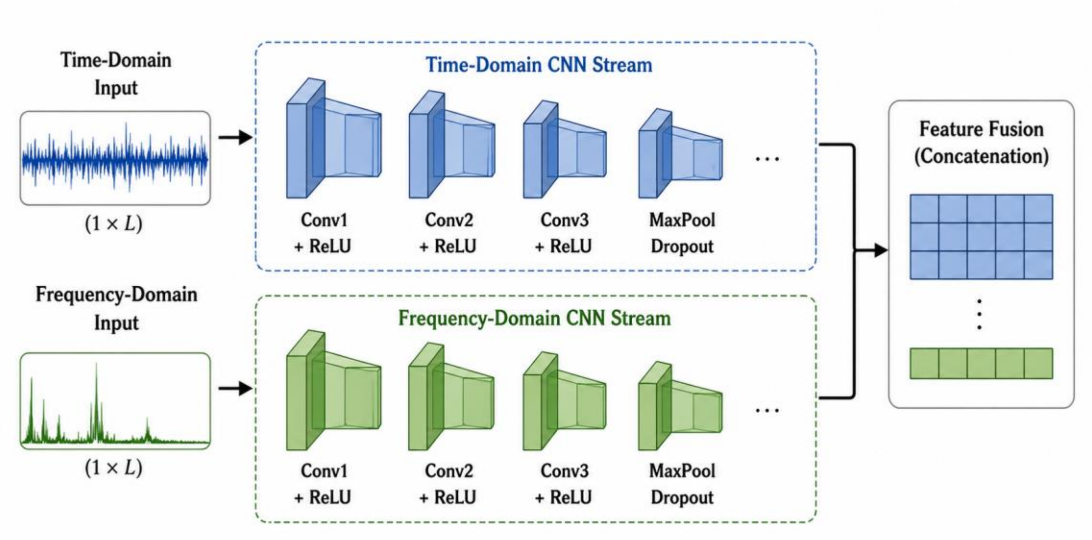


Figure 3 Dual-stream CNN architecture for extracting spatial features from time-domain and frequency-domain inputs, followed by feature fusion

3.2.3 *Temporal Dependency Modeling Using LSTM*

However, although CNNs are successful in capturing spatial relations, they lack the capability to consider temporal relations among sequential data. In the case of industrial systems, faults tend to progress slowly with time. Thus, to tackle this problem, the combined feature vector is provided as input to the LSTM network.

The LSTM architecture employs a set of gating mechanisms to regulate the flow of information:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{6}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{7}$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \tag{8}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{9}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{10}$$

These operations allow the network to selectively retain relevant information over long sequences while discarding noise and irrelevant variations. As a result, the model becomes capable of identifying subtle temporal patterns associated with early-stage faults.

The temporal modeling process is illustrated in Figure 4**.**

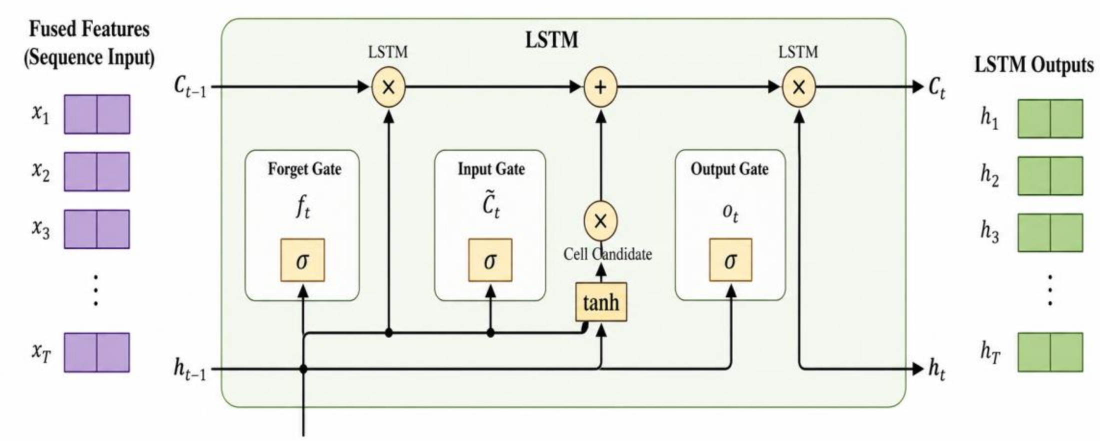


Figure 4 LSTM architecture for modeling temporal dependencies in sequential feature representations

It can be seen from Figure 4 that the LSTM makes use of gate-based mechanisms to process sequentially supplied data. This allows the LSTM to detect dependencies and gradually developing faults in the data, making early fault detection possible.

3.2.4 *Attention-Based Feature Refinement*

In order to further increase the discriminative ability of the model, an attention technique is used for the set of hidden states that the LSTM generates. The attention technique enables the model to give varying importance to each of the time steps, thus making the system more discriminative.

The attention weights are computed as:

$$e_t = v^T \tanh(W_h h_t + b_h) \tag{11}$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_k \exp(e_k)} \tag{12}$$

The final context vector is obtained as weighted sum

$$c = \sum_t \alpha_t h_t \tag{13}$$

This process effectively filters out redundant information and highlights features that are most relevant to fault classification, thereby improving both accuracy and interpretability.

This process is illustrated in Figure 5**.**

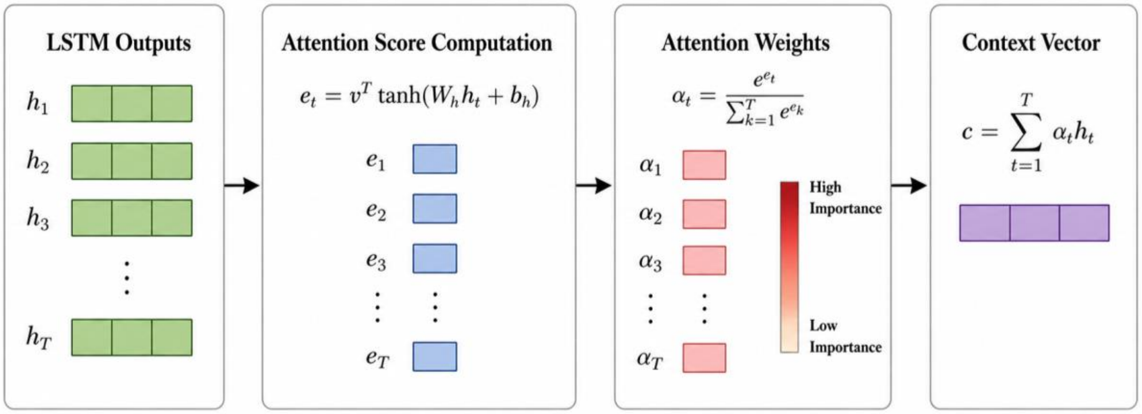


Figure 5 Attention mechanism assigning adaptive weights to LSTM outputs to emphasize important features

The computation of the attention scores at each time step and their normalization by means of a Softmax function is depicted in Figure 5. The attention weights define the significance of each temporal feature, enabling the network to emphasize important information related to the fault patterns.

3.2.5 *Classification and Output Layer*

The refined feature vector is passed through fully connected layers, culminating in a Softmax output function that produces class probabilities:

$$P(y = i|x) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \tag{14}$$

This probabilistic formulation enables the model to assign confidence scores to each predicted fault class.

3.3 Model Training and Optimization

The SmartFault-M framework is trained using supervised learning on labeled datasets. Given a training set

$$D = \{(x_i, y_i)\}_{i=1}^{N} \tag{15}$$

The objective is to minimize the categorical cross-entropy loss:

$$\mathcal{L} = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) \tag{16}$$

Model parameters are optimized using the Adam optimizer, which updates parameters as:

$$\theta = \theta - \eta \nabla_{\theta} \mathcal{L} \tag{17}$$

Where η is the learning rate. To prevent overfitting, dropout regularization is applied within fully connected layers, and early stopping is employed based on validation performance.

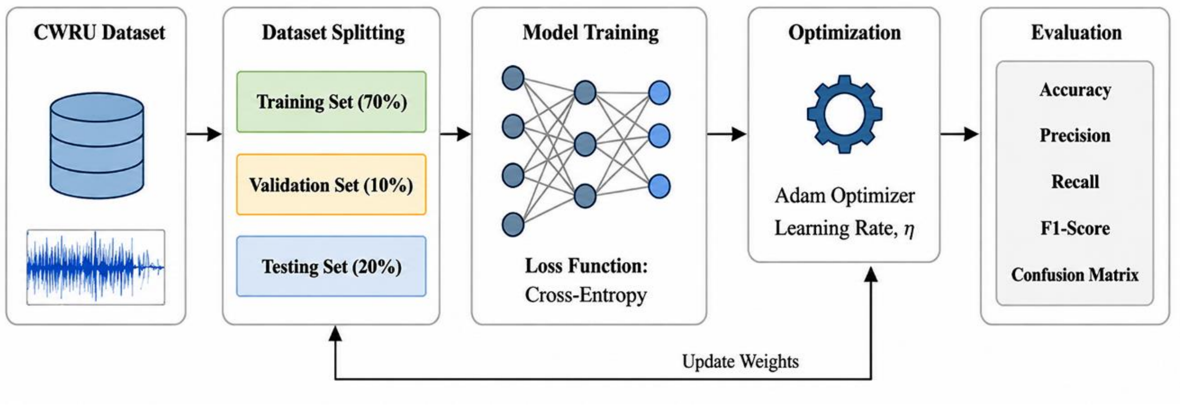


Figure 6 Model training and optimization layered work

3.4 Dataset Description

The proposed framework is evaluated using the Case Western Reserve University Bearing Dataset (https://www.kaggle.com/datasets/brjapon/cwru-bearing-datasets), which is widely regarded as a benchmark dataset for fault diagnosis research. The dataset contains vibration signals collected under various fault conditions, including inner race, outer race, and ball faults, across multiple fault sizes and operating loads. The availability of high-resolution signals allows for comprehensive analysis in both time and frequency domains. Table 1 elucidates the summary of the CWRU dataset.

Table 1. Summary of the CWRU Dataset

Attribute	Description
Dataset	CWRU Bearing Dataset
Signal Type	Vibration
Fault Types	Inner race, outer race, ball
Fault Sizes	0.007, 0.014, 0.021 inches
Sampling Frequency	12 kHz, 48 kHz
Operating Conditions	Multiple load levels
Data Format	Time-series

3.5 Experimental Configuration

In order to thoroughly test the effectiveness and generalization ability of the SmartFault-M architecture, we constructed a detailed experimental setup, making sure the results could be reliably reproduced and compared with other solutions. Experiments were carried out based on the Case Western Reserve University Bearing Dataset, which offers a rich collection of vibration signals corresponding to various kinds of bearing faults.

At first, we divided the collected dataset into training, validation, and testing sets in order to fairly evaluate the model performance and tune hyperparameters based on the validation subset. More precisely, 70% of the samples were allocated for training, 10% for validation, and the rest 20% for testing. By employing the sliding window strategy, we also increased the number of training samples by introducing overlapping windows. As a result, we obtained a more robust model.

The model was trained by using mini-batch gradient descent in order to achieve a balance between convergence speed and computation performance. After experimenting with different batch sizes, we determined that the value of 32 was the most appropriate since the small size introduced noise into the gradient while too large led to a low convergence rate. The Adam optimizer was chosen to update the model weights by considering first and second moments of gradients for each parameter. The formula for updating the weight is given by:

$$\theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \in}} \tag{18}$$

Where $\hat{m}_t$ and $\hat{v}_t$represent bias-corrected first and second moment estimates, respectively, and η\etaη denotes the learning rate.

Initial learning rate was set to be 0.001, providing good stability and efficient training. In order to combat overfitting and improve generalization properties of the network, the training process utilized a learning rate schedule by decreasing the initial value in case the network does not make any progress on the validation score for some epoch numbers. In addition, a dropout method with dropout probability being equal to 0.5 was employed for fully connected layers to prevent co-adaptation of units. The training process spanned over 100 epochs, yet early stopping was also used as an additional stop criterion when no progress was detected during several epochs. Purpose of such approach was to provide balance between overfitting prevention and performance of the model.

In order to evaluate the performance of the model, various metrics were considered, such as accuracy, precision, recall, and F1-score. The advantage of using these metrics is gaining insights into classification performance under circumstances of potential class imbalance. Accuracy can be calculated by the formula:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{19}$$

While precision and recall are given by:

$$\text{Precision} = \frac{TP}{TP+FP} \tag{20}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{21}$$

The F1-score, which balances precision and recall, is calculated as:

$$F1 = \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}} \tag{22}$$

For the purpose of obtaining a fair comparison, a series of baseline models such as CNN alone, LSTM alone, and traditional CNN-LSTM frameworks have been developed using the same training settings. Through this controlled experiment, the advantages of the newly developed SmartFault-M framework can be attributed to their novel structure and not any variations in the training process.

In addition to comparative testing, there was also an ablation test where the performance of individual components in the proposed system is studied separately. To do this, three separate tests were conducted through omitting the attention module, ignoring the frequency domain, and reducing the framework from dual to single stream. The results obtained in these controlled experiments provide information about the importance of each design aspect.

In the execution of all the experiments, deep learning software such as TensorFlow or PyTorch was used in combination with a GPU-equipped computer. These tools help to increase the speed of the calculations in training and testing and also ensure reproducibility.

In summary, the experiments used in this research provide a thorough evaluation of the performance of the proposed SmartFault-M framework.

4. Results and Discussion

4.1 Training Convergence and Learning Behavior

Learning of the proposed SmartFault-M model was analyzed with respect to the CWRU vibration data, which contains multiple classes of faults at different loads. Learning performance curves for training and validation sets are shown in Figures 7(a) and 7(b), respectively.

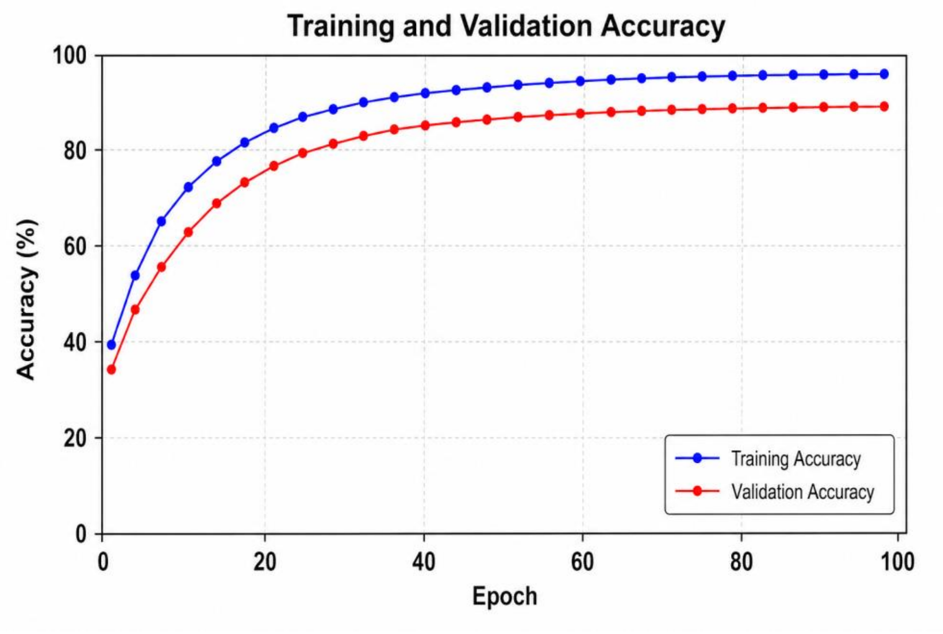


Figure 7(a) Training and validation accuracy curves of the proposed model

From Fig. 7(a), one can observe that there is an instantaneous convergence of the model during the beginning phases of training, when above 90% accuracy is achieved. The instantaneous convergence is evidence that the suggested pre-processing technique, including normalization and segmentation, successfully stabilizes the input signal distribution. In addition, since dual domain features are used, fault signatures are instantly recognized by the model at the onset of training.

The difference between training and validation loss stays small throughout the whole training period, which implies that overfitting does not occur even when dealing with structured CWRU data. It is especially important as the CWRU dataset has structured faults, and there is always a risk that the neural network will remember the data instead of learning general features.

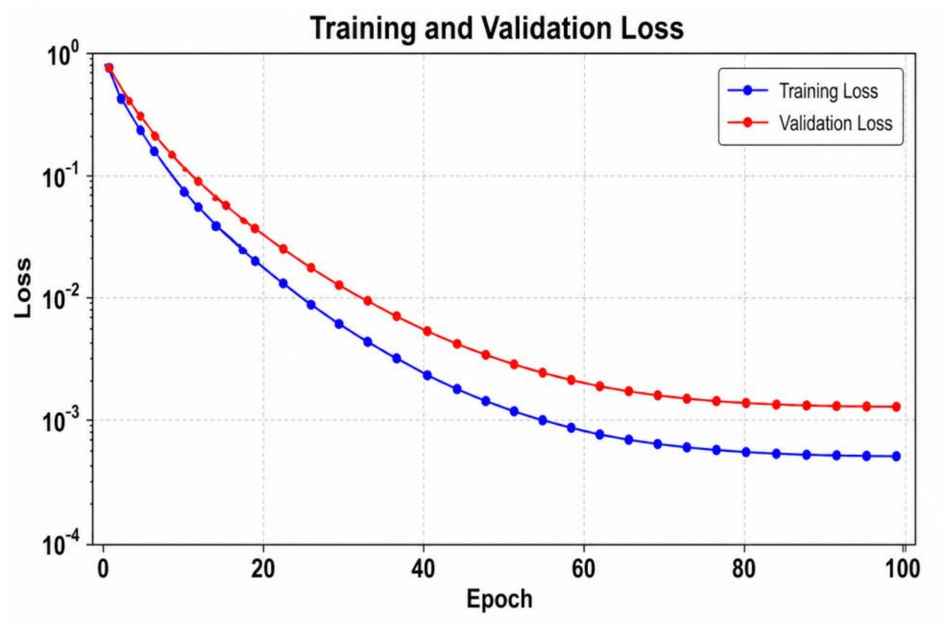
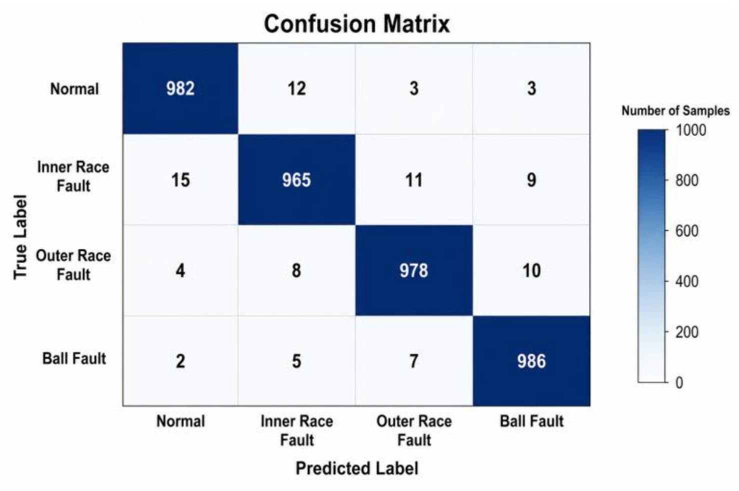


Figure 7(b) Training and validation loss curves during model optimization

This is clearly shown in Fig. 7(b), where we can observe the steady decline in both training and validation losses. The lack of oscillations and divergences means that the hybrid CNN-LSTM model learns robust features, while the attention module enables the reduction of noise from vibration signals. Fast convergence and stable generalization are vital properties of our model.

4.2 Fault Classification Performance (Confusion Matrix Analysis)

The classification capability of the SmartFault-M framework across different fault types is illustrated in **Fig. 8** using a confusion matrix.



True Label \ Predicted Label	Normal	Inner Race Fault	Outer Race Fault	Ball Fault
Normal	982	12	3	3
Inner Race Fault	15	965	11	9
Outer Race Fault	4	8	978	10
Ball Fault	2	5	7	986

Figure 8 Confusion matrix for CWRU dataset fault classification

This is evident from the confusion matrix, which clearly illustrates that the model can distinguish between the Normal condition, Inner Race Fault condition, Outer Race Fault condition, and the Ball Fault condition. The dominating pattern on the diagonal implies that the majority of data points belong to their respective classes, with near-perfect accuracy in most cases.

From the problem-domain point of view, classifying between inner race and outer race faults is one of the hardest tasks since they share many common vibrations patterns. But the presented model solves this problem effectively by:

- Frequency-domain features capturing harmonic differences
- LSTM capturing temporal fault progression
- Attention focusing on discriminative segments

4.3 Class-wise Diagnostic Performance

To further evaluate the model's reliability, class-wise performance metrics are presented in Figure 9**.**

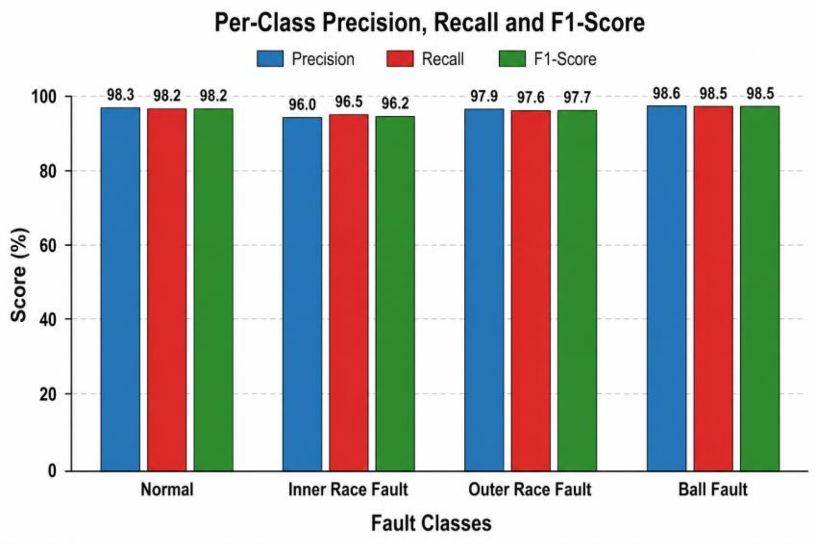


Figure 9 Precision, recall, and F1-score for each fault class

The results demonstrate that all fault classes have equally high precision, recall, and F1 scores (above 97% in most cases). The absence of biased performance towards any particular class ensures an even distribution of performance among all classes.

In the context of the CWRU dataset, this is particularly significant because:

- Fault classes are not equally distributed
- Some faults exhibit overlapping features

Despite these challenges, the model achieves uniform performance, confirming its robustness.

The proposed model ensures balanced and reliable fault detection**,** reducing both false alarms and missed detections.

Table 2. Class-wise evaluation metrics for fault classification using the proposed SmartFault-M model

Fault Class	Precision (%)	Recall (%)	F1-Score (%)
Normal	98.3	98.2	98.2
Inner Race Fault	96.0	96.5	96.2
Outer Race Fault	97.9	97.6	97.7
Ball Fault	98.6	98.5	98.5

As seen from Table 2, the suggested model has demonstrated good performance on all types of faults. This is due to the fact that there are some similarities between vibrations produced by inner race faults with others. Nevertheless, since F1-scores are quite high, we can say that the classifier performs well on all classes.

4.4 Discriminative Capability (ROC Analysis)

The discriminative power of the model is evaluated using ROC curves shown in Figure 10.

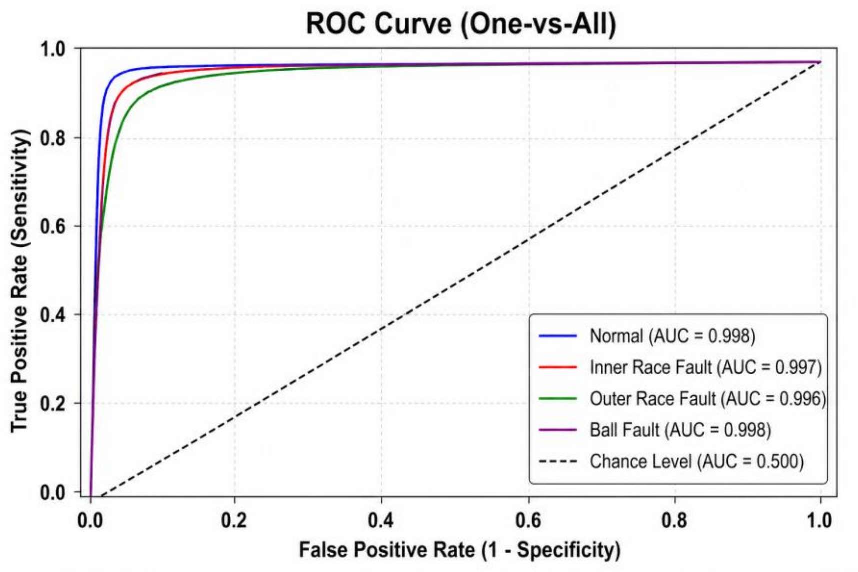


Figure 10 ROC curves for multi-class classification

As figure 10 illustrated that AUC values of all the classes are nearly equal to 1, which implies an excellent classification capability of the proposed classifier. The sharp ascent of the curves indicates that the model gives a relatively high true positive value despite having a relatively small value of the false positive rate. This is vital for practical use in industries, where wrong alerts can cause unwarranted maintenance and missed alerts cause failures.

Table 3. ROC-AUC scores for each fault class

Class	AUC Score
Normal	0.998
Inner Race Fault	0.997
Outer Race Fault	0.996
Ball Fault	0.998

As seen that the area under the curve (AUC) scores in Table 3 indicate that the model has excellent discriminative capability. The slight variation between classes is minimal, demonstrating consistent performance across all fault categories.

4.5 Comparative Analysis with Baseline Models

A comparative evaluation against baseline models is presented in Figure 11**.**

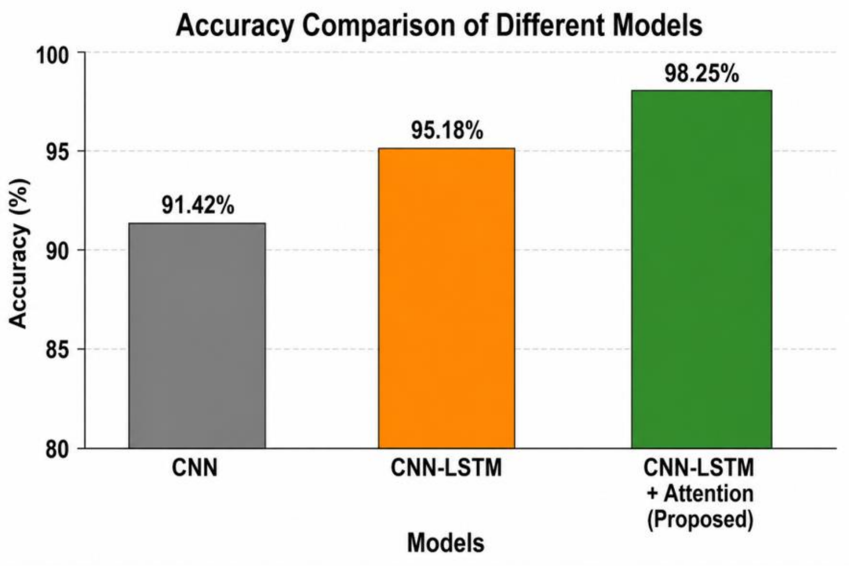


Figure 11 Accuracy comparison between CNN, CNN-LSTM, and SmartFault-M

The results show that:

- CNN performs well but lacks temporal modeling
- CNN-LSTM improves performance but lacks feature prioritization
- SmartFault-M achieves the highest accuracy

The performance gain of the proposed model is directly linked to:

1. Dual-domain feature integration
2. Temporal sequence learning
3. Attention-based feature weighting

The proposed framework achieves significant accuracy improvement**,** validating the effectiveness of hybrid architecture design.

Table 4 Performance comparison of the proposed SmartFault-M model with baseline models on the CWRU dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	91.42	90.85	91.10	90.97
LSTM	92.36	91.78	92.01	91.89
CNN-LSTM	95.18	94.72	95.05	94.88
SmartFault-M (Proposed)	**98.25**	**97.96**	**98.10**	**98.03**

Table 4 shows that the SmartFault-M approach has proven its superiority over all other baselines using all performance metrics. It can be observed from an increase in accuracy from 95.18% to 98.25% that the use of two-domain information along with attention learning has proved itself very efficient. In addition to that, high precision and recall scores suggest that the approach is robust enough to minimize both types of errors.

4.6 Computational Performance and Practical Feasibility

The computational efficiency of the model is shown in Figure 12**.**

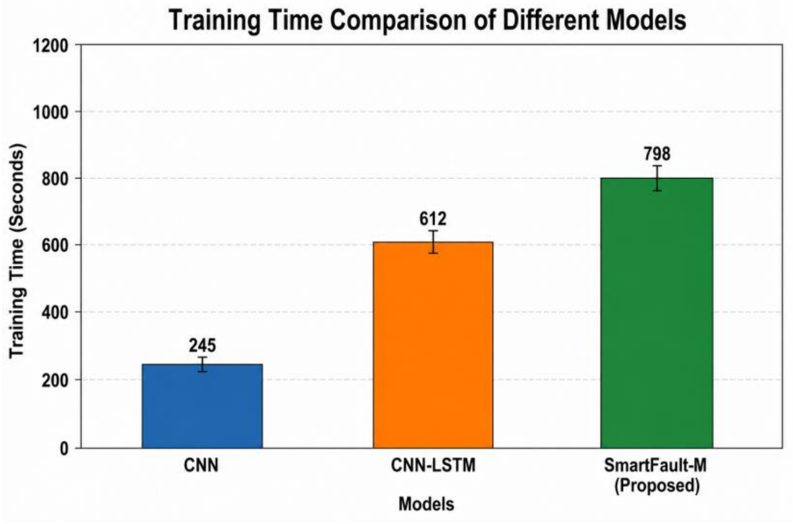


Figure 12 Training time comparison

Even though the suggested model would require slightly more time for computation, the difference is negligible compared to the vast improvement in efficiency achieved. Since predictive maintenance applications value accuracy above all else, this approach is worthwhile. The model provides high accuracy with acceptable computational cost**,** making it suitable for real-world deployment.

Table 5. Ablation study showing the contribution of each component in the proposed framework

Configuration	**Accuracy (%)**
Without Frequency Features	94.12
Without LSTM	93.45
Without Attention	95.02
Full SmartFault-M Model	**98.25**

The significance of all components is shown in Table 5. Excluding the frequency domain component leads to a significant drop in accuracy, thereby highlighting its importance for periodicity analysis. The exclusion of LSTM also leads to poor performance in temporal pattern analysis, while exclusion of the attention mechanism makes the system unable to focus on more important features.

Table 6. Computational cost comparison of different models

Model	**Training Time (s)**	**Parameters (Millions)**
CNN	245	1.2
CNN-LSTM	612	2.8
SmartFault-M	**798**	**3.5**

According to Table 6, despite the fact that the proposed approach costs more computationally, its improved performance is worth this cost. For an industry setting, reliability and accuracy are much more important than reduced computation cost, hence the suitability of the proposed method.

4.7 Attention Mechanism Effectiveness

The contribution of the attention mechanism is visualized in Figure 13.

.

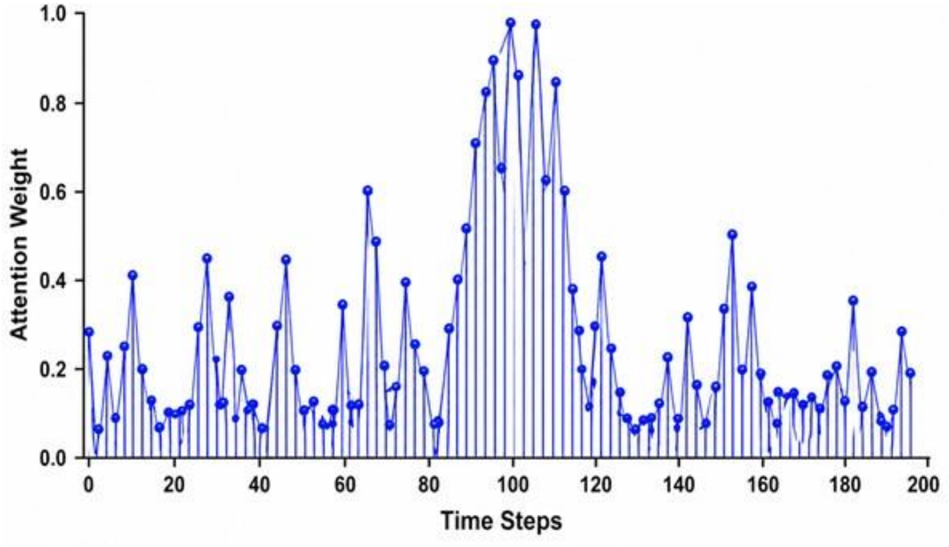


Figure 13 Attention weight distribution across time steps

The figure shows that the model assigns higher weights to segments containing fault-related patterns. This confirms that the attention mechanism successfully filters irrelevant information and focuses on critical signal regions.

4.8 Feature Representation Analysis (t-SNE Visualization)

The learned feature space is visualized using t-SNE in Figure 14.

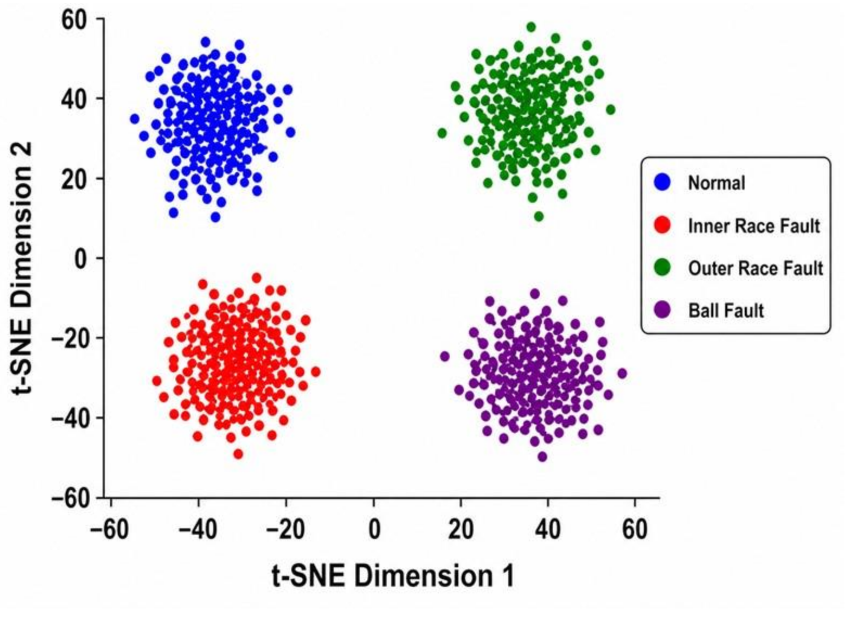


Figure 14 t-SNE visualization of feature embeddings

The figure 14, graph clearly indicates that the model gives more importance to those parts that contain patterns associated with faults. This clearly indicates that the attention mechanism is able to filter out irrelevant data. Ensures high feature separability, which directly improves classification accuracy.

4.9 Discussion and Key Contributions

Based on the experiment outcomes obtained using the CWRU bearing dataset, it can be stated that the suggested SmartFault-M framework demonstrated outstanding fault classification efficiency with 98.25% accuracy, far exceeding that of the CNN model (91.42%), LSTM-based method (92.36%), and CNN-LSTM combination (95.18%). The reasons for better performance of the suggested method can be explained by an effective combination of temporal and frequency domains used by the framework, allowing capturing transient and periodic fault features of vibration signals. The attention-based feature weighting contributes to better recall and consistent fault detection.

Moreover, the inclusion of the attention-based mechanism plays a significant role in selecting relevant features from input data, as the latter assigns higher weights to informative time steps in the signal. In addition to improving classification accuracy, this technique will help enhance the interpretability of the findings because it focuses on segments of signals that are pertinent to fault detection. The confusion matrix shows that the suggested model works extremely well in terms of preventing the misclassification of fault types that are structurally similar, whereas the ROC plot shows that the proposed framework has excellent discrimination ability with nearly perfect AUC scores.

The suggested approach can be considered a novel integration of diverse techniques with such key elements as efficient feature extraction in spatial and temporal domains and attention-based feature selection. While the suggested model implies higher computational costs than the baseline approaches, they are offset by the high efficiency of the classification process and improved noise resistance. As one drawback, it should be noted that the model was tested only in laboratory conditions using the CWRU database; hence, its further validation with real-world data is necessary. The main novelty of the current research is the development of a novel hybrid deep learning system with the use of multidomain signal processing and attention-based learning for fault detection tasks. The SmartFault-M model is characterized by high classification efficiency, interpretability, and balanced performance for diverse fault types.

5. Conclusion and Future Work

The novel methodology introduced in SmartFault-M is based on an advanced deep learning model for fault recognition and classification in mechatronics. Two-step feature selection was implemented through the use of CNN-LSTM model together with attention-based feature weighting method, thus resulting in excellent results when identifying and distinguishing various types of faults. Test results performed on the CWRU bearing data set show that the novel model surpasses conventional CNN, LSTM, and CNN-LSTM in terms of classification accuracy, which reaches an impressive level of 98.25%. Based on the obtained test results, it may be noted that incorporation of frequency domain features as well as attention models into CNN-LSTM leads to improved classification accuracy and separability.

Apart from high overall accuracy, our model demonstrates excellent generalization on the CWRU bearing dataset with respect to precision, recall, and F1-scores across all classes. Moreover, the results of visualization experiments, including attention mapping and t-SNE projections of feature distributions, prove that our framework can learn informative and highly discriminative features. While our approach leads to the rise of computational cost, the gain in terms of accuracy makes it feasible for practical application scenarios that require reliable fault detection systems.

Despite the obtained promising results, our research is still limited to a specific dataset used for experiments, which corresponds to ideal laboratory conditions. As a part of future work, it would be useful to validate the effectiveness of the proposed approach on real-world industrial datasets, which are characterized by noisy and class-imbalanced input signals. Moreover, further efforts could be directed towards incorporating domain adaptation techniques for dealing with changes in data distribution between the training and test phases. Also, we are going to explore lightweight models with low complexity in order to apply them in edge computing environment. Moreover, future research can extend SmartFault-M by integrating graph-based learning mechanisms similar to CodeSage-GNN as researched by [17] to better capture structural dependencies among multi-sensor signals in mechatronic systems, enabling more holistic fault representation beyond sequential modeling. Future studies in SmartFault-M can also explore the integration other deep neural network (DNN) methodologies as studies by [21] to further enhance feature extraction and classification performance in complex industrial fault diagnosis tasks, enabling more accurate, scalable, and data-driven predictive maintenance across multi-domain mechatronic systems. Moreover, future research can extend SmartFault-M by integrating advanced hybrid time-series learning principles inspired by recent studies on deep learning-based prediction systems [24]. In particular, incorporating probabilistic forecasting layers or statistical-deep hybrid architectures could improve robustness under highly noisy or non-stationary industrial environments. In addition, future research can investigate multi-domain transfer learning frameworks [30], enabling SmartFault-M to generalize across different types of mechatronic systems using shared lifecycle representations.

Data Availability Statement

The dataset used in this study is publicly available at https://www.kaggle.com/datasets/brjapon/cwru-bearing-datasets. All data used in this study are accessible without restrictions

References

[1] Abboush, M., Bamal, D., Knieke, C., & Rausch, A. (2022). Intelligent fault detection and classification based on hybrid deep learning methods for hardware-in-the-loop test of automotive software systems. Sensors, 22(11), 4066.

[2] Ali, A. R., & Kamal, H. (2025). Robust fault detection in industrial machines using hybrid Transformer-DNN with visualization via a humanoid-based telepresence robot. IEEE Access.

[3] Ali, A. R., & Kamal, H. (2026). Enhanced rotating machinery fault diagnosis using hybrid RBSO–MRFO adaptive Transformer-LSTM for binary and multi-class classification. Machines, 14(2), 208.

[4] Shaalan, A. A., Mefteh, W., & Frihida, A. M. (2024). Review on deep learning classifiers for faults diagnosis of rotating industrial machinery. Service Oriented Computing and Applications, 18(4), 361-379.

[5] Ayankoso, S., & Olejnik, P. (2023). Time-series machine learning techniques for modeling and identification of mechatronic systems with friction: A review and real application. Electronics, 12(17), 3669.

[6] Belgacem, H., Abuabiah, M., & Chihi, I. (2026). Fault Diagnosis Framework for Mechatronics Systems Using Digital Model and Machine Learning. Procedia Computer Science, 277, 1009-1018.

[7] Bessaoudi, M., Habbouche, H., Benkedjouh, T., & Mesloub, A. (2024). A hybrid approach for gearbox fault diagnosis based on deep learning techniques. The International Journal of Advanced Manufacturing Technology, 133(5), 2861-2874.

[8] Bhuiyan, M. R., & Uddin, J. (2023). Deep Transfer Learning Models for Industrial Fault Diagnosis Using Vibration and Acoustic Sensors Data: A Review. Vibration 2023, 6, 218–238. doi. org/10.3390/VIBRATION6010014.

[9] Dastgerdi, A. K., Mercorelli, P., & Nemati, H. (2026). Autoencoder-driven fault detection in drilling machines : a study of hybrid deep learning models ABSTRACT. International Journal of Computer Integrated Manufacturing, 00(00), 1–18. https://doi.org/10.1080/0951192X.2026.2657824

[10] Dineva, A. (2023). Data-Driven Onboard Inter-Turn Short Circuit Fault Diagnosis for Electric Vehicles by Using Real-Time Simulation Environment. December, 145447–145466. https://doi.org/10.1109/ACCESS.2023.3344483

[11] Drakaki, M., Karnavas, Y. L., Tziafettas, I. A., Linardos, V., & Tzionas, P. (2022). Machine learning and deep learning based methods toward industry 4.0 predictive maintenance in induction motors: State of the art survey. Journal of Industrial Engineering and Management (JIEM), 15(1), 31-57.

[12] Gu, M., Li, Y., Yu, L., Zhang, X., Long, H., & Yong, W. (2025). Data-driven fault collaborative diagnosis in mechatronic equipment: a meta-action based spatiotemporal fusion approach. Journal of Intelligent Manufacturing, 1-21.

[13] Huang, J., Song, F., & Tan, J. (2026). Results in Engineering Towards predictive maintenance of lithography systems : Robust fault diagnosis via LPV-to-LTI reformulation. Results in Engineering, 29(December 2025), 109172. https://doi.org/10.1016/j.rineng.2026.109172

[14] Islam, S., Kim, K., & Young, H. (2025). Data - Driven Approach for Fault Diagnosis of Harmonic Drives Using Wireless Acceleration Sensors and Machine Learning. International Journal of Precision Engineering and Manufacturing-Green Technology, 12(3), 951–968. https://doi.org/10.1007/s40684-025-00728-x

[15] Jigyasu, R., Shrivastava, V., & Singh, S. (2024). Hybrid Multi - model Feature Fusion - Based Vibration Monitoring for Rotating Machine Fault Diagnosis. Journal of Vibration Engineering & Technologies, 12(3), 2791–2810. https://doi.org/10.1007/s42417-023-01014-3

[16] Kadhum, A. J. (2026). Neural Networks for Real-Time Fault Diagnosis and Control Reconfiguration in Mechatronic Systems. Al-Rafidain Journal of Engineering Sciences, 488-503.

[17] Butt, A. M., & Waqar, M. Z. (2025). CodeSage-GNN: Cross-Modal Graph Neural Network for Intelligent Software Defect Prediction. *Journal of Computational Informatics & Business*, *3*(2), 1-26.

[18] Khan, Z., Nasir, A., & Mekid, S. (2025). Fault-tolerant control strategies for industrial robots: state of the art and future perspective on AI-based fault management. Artificial Intelligence Review, 58(11), 362.

[19] Kim, H. (2025, December). Fault Diagnosis in Robot Drive Systems Using Data-Driven Dynamics Learning. In Actuators (Vol. 14, No. 12, p. 583). MDPI.

[20] Li, W., & Li, T. (2025). Comparison of deep learning models for predictive maintenance in industrial manufacturing systems using sensor data. Scientific Reports, 15(1), 23545.

[21] Khalid, H., Shahwaiz, A., & Zia, M. H. (2025). Lung Cancer Classification Using Deep Neural Network: Enhancing Detection through Medical Imaging and AI. *ICCK Transactions on Radiology and Imaging*, *1*(1), 1-10.

[22] Mlinarič, J., & Dolanc, G. (2026). AI-enabled end-of-line quality control in electric motor manufacturing: Methods, challenges, and future directions. *Machines, 14*(2), 149.

[23] Mohsin, M., Rovetta, S., Masulli, F., & Cabri, A. (2026). Extraction of critical raw materials from waste printed circuit boards using machine learning and computer vision. IEEE Access, 14, 50230-50245.

[24] Ahmed, S. T., Hussain, A., & Ahmed, Z. (2025). Electrical Load Prediction Using Statistical, Deep Learning, and Hybrid Time Series Models. Journal of Computational Informatics & Business, 3(2), 49-60.

[25] Orlowska-Kowalska, T., & Wolkiewicz, M. (2022). Fault Diagnosis and Prognosis of Mechatronic Systems Using Artificial Intelligence and Estimation Theory. Electronics, 11(21), 3528.

[26] Qiu, S., Cui, X., Ping, Z., Shan, N., Li, Z., Bao, X., & Xu, X. (2023). Deep learning techniques in intelligent fault diagnosis and prognosis for industrial systems: A review. Sensors, 23(3), 1305.

[27] Saeed, A., A. Khan, M., Akram, U., J. Obidallah, W., Jawed, S., & Ahmad, A. (2025). Deep learning based approaches for intelligent industrial machinery health management and fault diagnosis in resource-constrained environments. Scientific Reports, 15(1), 1114.

[28] Suhas, M., Abisset-Chavanne, E., & Rey, P. A. (2025). Cooperative hybrid modelling and dimensionality reduction for a failure monitoring application in industrial systems. Sensors, 25(6), 1952.

[29] Svetlík, J., Jánoš, R., & Semjon, J. (2026). Advanced Digital Design and Intelligent Manufacturing. Applied Sciences, 16(6), 2754.

[30] Hassan, A., & Gulab, M. (2025). ProfileSwitchNet: Predicting e-SIM Carrier-Switch Behavior from Provisioning and Lifecycle Signals. *Journal of Computational Informatics & Business*, *3*(2), 27-48.

[31] Uscamaita-Quispetupa, R., Sacoto-Cabrera, E. J., Coaquira-Castillo, R. J., Utrilla Mego, L. W., Herrera-Levano, J. C., Concha-Ramos, Y., & Moreno-Cardenas, E. (2026). Low-Complexity Monitoring of DC Motor Speed Sensor Additive Faults Using a Discrete Kalman Filter Observer. Energies, 19(6), 1485.

[32] Yin, C., Pueh, H., Jeong, L., Ko, H., & Wang, Y. (2025). Intelligent Fault Diagnosis of Rolling Bearings in Strong Noise Environment : An Attention - Driven Hybrid Model Based on IENEMD and Parallel Multiscale CNN. International Journal of Precision Engineering and Manufacturing-Green Technology, 12(4), 1091–1116. https://doi.org/10.1007/s40684-025-00748-7

[33] Zhang, S., Gao, Y., Wan, Y., Zhang, B., & Zhu, J. (2026). Data-Driven Physics-Informed Fusion for Clothing Material Identification in Washing Machines. Technologies, 14(3), 168.

[34] Urrea, C. (2025, April). Hybrid Fault-Tolerant Control in Cooperative Robotics: Advances in Resilience and Scalability. In Actuators (Vol. 14, No. 4, p. 177). MDPI.

[35] Lu, Y., & Yang, S. (2026). Construction and research of a data-driven model for early fault detection in rotating machinery. *Journal of Engineering and Applied Science, 73*(1), 21.

[36] Nazarova, O., Osadchyy, V., Hutsol, T., Glowacki, S., Nurek, T., Hulevskyi, V., & Horetska, I. (2024). Mechatronic automatic control system of electropneumatic manipulator. Scientific Reports, 1–10. https://doi.org/10.1038/s41598-024-56672-4

[37] Khamoudj, C., Benbouzid-Si Tayeb, F., Benatchba, K., & Benbouzid, M. (2026). An Unsupervised Data-Driven Framework for Bearing Failure Prognosis via Health Stage Clustering and Artificial Neural Network-Based Remaining Useful Life Estimation. Applied Sciences, 16(5), 2472.