Deep Learning–Based Predictive Maintenance of Rotating Machinery Using Vibration and Acoustic Signals

Tan Wai Ming[1], Teo Zhi Yang[2]

[1]School of Digital Technology, Wawasan Open University, George Town 10050, Penang, Malaysia
E-mail: twm1_oi@student.wou.edu.my
[2]School of Digital Technology, Wawasan Open University, George Town 10050, Penang, Malaysia
E-mail: tzy1_oi@student.wou.edu.my

Article Info

Article History:

Received Dec 18, 2025
Revised Jan 15, 2026
Accepted Feb 19, 2026

Keywords:

Predictive maintenance
Rotating machinery
Deep learning
CNN–LSTM
Vibration analysis
Acoustic signal analysis
Fault diagnosis

ABSTRACT

Predictive maintenance has emerged as a key enabler of intelligent manufacturing systems by reducing unplanned downtime and maintenance costs in rotating machinery. Traditional condition monitoring techniques rely heavily on handcrafted features and expert knowledge, which often fail to generalize under complex operating conditions. This paper presents a deep learning–based predictive maintenance framework for rotating machinery using vibration and acoustic signals. Multisensor data are collected from machinery under normal and faulty operating conditions, including bearing defects and imbalance faults. A hybrid deep learning model combining convolutional neural networks (CNN) and long short-term memory (LSTM) networks is employed to automatically learn spatial and temporal features from raw signals. Experimental results demonstrate that the proposed approach achieves superior fault detection accuracy and remaining useful life prediction performance compared to conventional machine learning methods. The results confirm the effectiveness of deep learning for real-time predictive maintenance in smart manufacturing environments.

Corresponding Author:

Tan Wai Ming,
School Of Digital Technology,
Wawasan Open University, George Town 10050, Penang, Malaysia.
E-mail: twm1_oi@student.wou.edu.my

1. INTRODUCTION

Industry 4.0 represents a major shift in modern manufacturing, driven by the integration of advanced digital technologies such as artificial intelligence, machine learning (ML) [1], and the Internet of Things (IoT) into industrial operations and maintenance systems. Maintenance practices have evolved substantially to address growing demands for machine reliability and operational efficiency. By enabling continuous data acquisition from industrial equipment, these technologies allow large volumes of machine data to be analyzed for optimizing maintenance decisions and improving equipment availability and reliability [2].

In this context, vibration signal monitoring has become a standard practice across industrial plants worldwide. Both classical machine learning techniques and deep learning models have been widely applied for fault detection using vibration data. Since vibration signals can be collected easily during machine operation and reflect real-time operating conditions, they have become a primary source of information for modern machinery fault diagnosis.

Vibration signals generated by rotating machinery are typically nonlinear in nature. This nonlinearity arises from factors such as varying load conditions [3], dynamic interactions among machine components, and changes in operating regimes. To capture the complexity of such signals, many condition monitoring studies employ entropy-based measures for feature extraction. As a result, various entropy techniques—including approximate entropy, sample entropy and multiscale entropy—have been extensively used in vibration-based fault diagnosis. Nevertheless, some recent studies have demonstrated that explicit entropy-based feature extraction is not always necessary. For example, the work reported in introduced an automatic feature-learning neural network that bypasses conventional handcrafted feature design while achieving high diagnostic accuracy.

Rotating machinery, including motors [4], pumps, compressors, and turbines, forms the backbone of mechanical and manufacturing systems. Failures in these components can cause severe production downtime, safety hazards, and increased maintenance costs. Traditional maintenance approaches, such as corrective maintenance and scheduled preventive maintenance, often prove inefficient, as they either respond only after a failure has occurred or rely on fixed intervals without considering the actual condition of the machine.

Predictive maintenance has emerged as an effective alternative to overcome these limitations by enabling continuous monitoring of equipment health and early prediction of potential faults. Vibration and acoustic signals are commonly used for this purpose due to their high sensitivity to mechanical defects such as bearing degradation, shaft misalignment, and rotor imbalance. However, conventional signal processing methods typically require manual feature extraction using techniques such as fast Fourier transform (FFT) and wavelet analysis [5]. These approaches depend heavily on expert knowledge and may struggle to maintain robustness under varying operating conditions.

Recent advances in deep learning have transformed this landscape by enabling automatic feature extraction directly from raw sensor data. Deep learning architectures, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have demonstrated strong capabilities in fault diagnosis and time-series prediction tasks. Motivated by these advancements [6], this study proposes a deep learning–based predictive maintenance framework that combines vibration and acoustic signal analysis to enhance fault detection accuracy and improve health state prediction for rotating machinery.

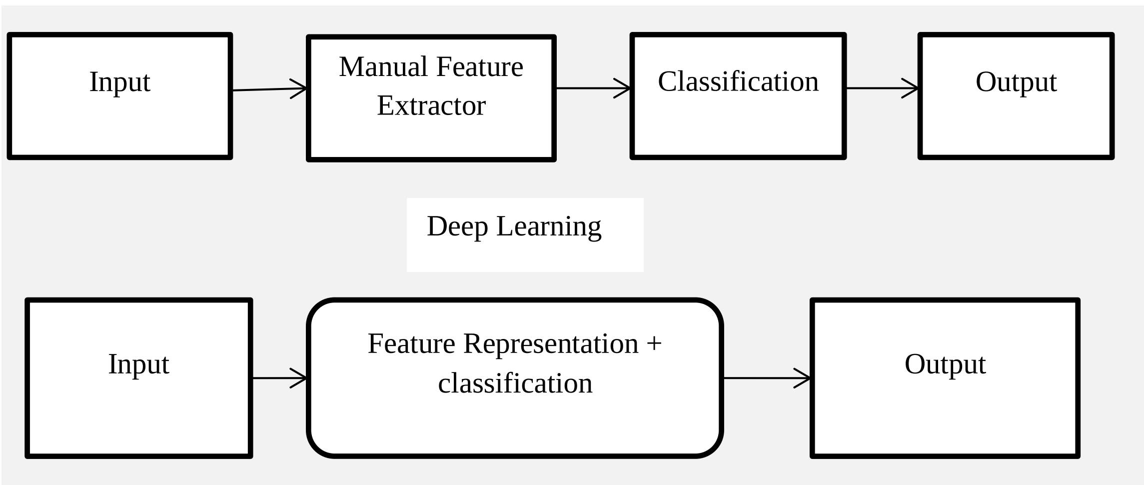

Figure 1. Difference between ML and DL

Deep learning (DL) focuses on automatic feature learning rather than relying on manually designed features to perform classification and prediction tasks. Compared to traditional shallow machine learning algorithms, DL offers distinct advantages in terms of feature representation, model architecture, and training capability (as illustrated in Figure 1 [7]). DL models typically consist of multiple layers—often dozens or even hundreds—that progressively learn hierarchical representations of the input data. These representations are generated using neural networks, which are learning frameworks composed of interconnected neurons and activation functions that enable nonlinear transformations of input samples.

1.1 Problem Statement

Rotating machinery plays a vital role in industrial operations, and unexpected failures can lead to substantial production losses, safety concerns, and increased maintenance costs. Conventional maintenance and fault diagnosis techniques usually depend on periodic inspections and handcrafted feature extraction from vibration signals. Such approaches often perform poorly in noisy environments and under varying operating conditions. In addition, fault diagnosis systems based on a single sensing modality may struggle to detect early-stage faults, thereby limiting the effectiveness of predictive maintenance strategies. Consequently, there is a growing need for an intelligent, data-driven predictive maintenance solution capable of automatically learning discriminative features, effectively integrating multisensor data, and accurately predicting machinery health conditions in real time.

1.2 Paper Organization

The rest of this paper is structured as follows. Section 2 presents a review of existing research on predictive maintenance and deep learning–based fault diagnosis methods for rotating machinery. Section 3 outlines the proposed methodology, covering data acquisition, preprocessing steps, and the CNN–LSTM–based predictive maintenance framework. Section 4 discusses the implementation details and experimental results, including performance evaluation and comparative analysis. Finally, Section 5 concludes the paper by summarizing the main findings and contributions of the proposed approach.

2. LITERATURE REVIEW

Predictive maintenance of rotating machinery has attracted significant research attention due to its critical role in improving operational reliability and minimizing unplanned failures in industrial environments. Early work in machinery condition monitoring largely depended on vibration analysis using features extracted from the time domain, frequency domain, and time–frequency domain. Commonly used statistical indicators included root mean square, kurtosis, and crest factor, while spectral features were derived using Fourier and wavelet-based techniques [8]. These methods were successfully applied to detect faults such as bearing damage and shaft misalignment; however, their effectiveness was strongly influenced by expert knowledge and manual feature design.

As machine learning techniques evolved, data-driven approaches began to replace traditional rule-based methods for machinery fault diagnosis and health assessment [9]. Conventional machine learning algorithms, such as support vector machines, artificial neural networks, and k-nearest neighbor classifiers, were used to identify machine health conditions based on manually extracted features. Although these approaches achieved improved diagnostic accuracy, their robustness was often limited when operating conditions changed or when signals were contaminated by noise. Moreover, the dependence on handcrafted features remained a major obstacle for reliable real-world implementation.

The introduction of deep learning has fundamentally changed predictive maintenance research by enabling automatic feature learning directly from raw sensor measurements [10]. Convolutional neural networks (CNNs) have been widely adopted for vibration-based fault diagnosis because of their ability to capture localized patterns and fault-related signatures in time-series data and time–frequency representations. Numerous studies have reported that CNN-based models outperform traditional machine learning techniques, particularly in complex and variable operating environments. To further enhance feature learning, researchers have transformed vibration signals into representations such as spectrograms and scalograms before feeding them into deep networks.

Despite their effectiveness in spatial feature extraction [11], CNNs have limited capability in modeling long-term temporal dependencies associated with progressive machinery degradation. To address this limitation, recurrent neural networks—especially long short-term memory (LSTM) networks—have been employed for sequential data modeling and remaining useful life prediction. LSTM-based approaches have demonstrated strong performance in capturing degradation patterns and forecasting failure evolution. More recently, hybrid CNN–LSTM architectures have gained attention, as they combine spatial feature extraction with temporal dependency learning, leading to improved predictive maintenance outcomes.

Beyond vibration-based monitoring, acoustic emission and sound signals have emerged as valuable non-contact sources of information for machinery condition assessment. Acoustic signals can reveal early-stage fault characteristics that may not be [12] immediately observable in vibration data. Recent deep learning–based studies have shown promising results in diagnosing faults using acoustic measurements, particularly for bearings and electric motors. However, acoustic signals are highly sensitive to ambient noise, which limits their reliability when used independently in industrial settings.

To address the limitations of single-sensor systems, multimodal sensor fusion strategies have been proposed. By jointly analyzing vibration and acoustic data, researchers have reported improved fault detection accuracy and enhanced robustness. Deep learning models are especially well suited for multimodal fusion, as they can learn complementary and discriminative representations from heterogeneous sensor inputs. Nevertheless, much of the existing literature primarily focuses on fault classification, with less emphasis on broader predictive maintenance objectives such as early fault detection and remaining useful life estimation.

In addition, most reported studies are conducted under controlled laboratory conditions, with limited attention given to real-time deployment and industrial scalability. Challenges related to model interpretability, computational efficiency, and adaptability to varying operating conditions remains largely unresolved. As a result, there is a clear research gap in developing an integrated deep learning–based predictive maintenance framework that effectively fuses vibration and acoustic signals while addressing practical considerations for reliable and scalable industrial deployment.

3. METHODS AND MATERIALS

This study introduces a deep learning–driven predictive maintenance framework for rotating machinery that utilizes both vibration and acoustic signal analysis. The proposed methodology includes data acquisition from a rotating machinery test platform, signal preprocessing, automatic feature learning using deep neural networks, and predictive maintenance analysis. A unified CNN–LSTM architecture is employed as the core model, enabling the extraction of spatial fault-related features and the learning of temporal degradation patterns from multi-sensor data simultaneously.

3.1 Experimental Setup and Data Collection

Experimental data were obtained from a laboratory-scale rotating machinery test rig designed to replicate typical industrial operating conditions. The system comprises an electric motor connected to a rotating shaft supported by rolling element bearings. Multiple machine health conditions were introduced, including normal operation and fault scenarios such as bearing defects and shaft imbalance. Vibration measurements were collected using piezoelectric accelerometers mounted near the bearing housing, while acoustic data were recorded using a high-sensitivity microphone positioned at a fixed distance from the rotating assembly. Both sensors were synchronized to ensure accurate temporal correspondence between vibration and acoustic signals. Data were sampled at a high frequency to capture fault-related information and stored for subsequent analysis. All experiments were conducted under constant speed and load conditions to ensure consistency and to clearly isolate fault-induced signal variations.

3.2 Signal Preprocessing

The raw vibration and acoustic signals include noise and non-informative frequency components that can degrade model performance. To address this, several preprocessing steps were applied prior to deep learning analysis. First, the continuous signals were divided into fixed-length segments to generate multiple samples suitable for neural network training. Band-pass filtering was applied to the vibration signals to retain frequency ranges associated with rotating machinery faults, while acoustic signals were filtered to suppress background and environmental noise. Signal normalization was then performed to scale the data within a consistent range, promoting numerical stability and faster convergence during model training. Finally, the processed vibration and acoustic segments were temporally aligned and combined to form synchronized multisensor input samples.

3.3 Deep Learning–Based Feature Extraction and Modeling

A hybrid convolutional neural network (CNN) and long short-term memory (LSTM) architecture is adopted as the primary modeling approach in this work. The CNN component performs automatic feature extraction directly from the preprocessed vibration and acoustic signals. Convolutional layers employ multiple kernels to identify localized fault signatures, such as impulsive vibration patterns and abnormal acoustic responses, while pooling layers reduce dimensionality and retain the most relevant features. Feature representations derived from the vibration and acoustic channels are then merged to facilitate effective multisensor fusion.

The fused feature set is subsequently passed to an LSTM network, which models temporal dependencies across sequential time windows. This structure enables the capture of long-term degradation behavior and evolving fault dynamics, which are essential for predictive maintenance applications. By integrating CNN-based spatial feature learning with LSTM-based temporal modeling, the proposed framework removes the need for manual feature engineering and offers a comprehensive representation of machinery health conditions.

3.4 Model Training and Validation

The CNN–LSTM model is trained using labeled datasets representing different health states of the rotating machinery. A supervised learning strategy is employed, with fault classification as the primary task. The dataset is partitioned into training, validation, and test sets to ensure objective and reliable performance evaluation. Categorical cross-entropy is used as the loss function, and the Adam optimizer is applied to update network weights during training. To mitigate overfitting, regularization techniques such as dropout are incorporated into the network architecture. Model performance is assessed using metrics including accuracy, precision, recall, and F1-score, providing a comprehensive evaluation of fault detection effectiveness.

3.5 Predictive Maintenance Strategy

Once trained, the CNN–LSTM model is deployed to continuously analyze incoming vibration and acoustic signals and predict the operational health of the rotating machinery. Early identification of abnormal signal patterns enables timely maintenance actions before severe failures occur. The temporal modeling capability of the LSTM component supports tracking of gradual degradation trends, facilitating condition-based maintenance planning. Overall, the proposed deep learning–based predictive maintenance strategy improves system reliability, reduces maintenance costs, and supports intelligent decision-making in smart manufacturing environments.

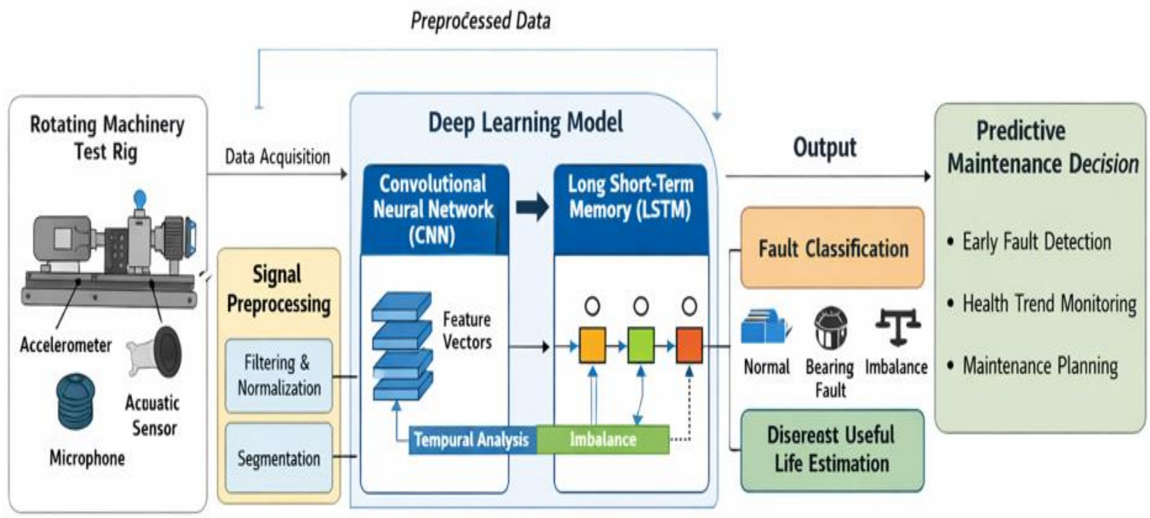

Figure 2. CNN–LSTM–Based Predictive Maintenance System Architecture

Figure 2 presents the overall architecture of the proposed deep learning–based predictive maintenance system for rotating machinery that utilizes both vibration and acoustic signals. The framework begins with data acquisition from a rotating machinery test setup equipped with accelerometers and acoustic sensors to record vibration and sound data under various operating conditions. The collected raw signals are then processed through a preprocessing stage that includes noise reduction, signal segmentation, and normalization to improve data quality and prepare the inputs for deep learning analysis.

The processed multisensor data are subsequently supplied to a convolutional neural network (CNN), which performs automatic extraction of fault-relevant spatial features from the vibration and acoustic inputs. The extracted feature representations are then passed to a long short-term memory (LSTM) network, which captures temporal dependencies and tracks machinery degradation behavior over time. This combined CNN–LSTM structure enables simultaneous learning of localized fault characteristics and long-term health evolution.

At the output stage, the model generates fault classification results, distinguishing between operating states such as healthy condition, bearing defects, and imbalance faults. Furthermore, the temporal information learned by the network supports degradation assessment and remaining useful life estimation. These outputs are used to support predictive maintenance actions, including early fault identification, continuous condition monitoring, and optimized maintenance scheduling, ultimately enhancing system reliability and minimizing unplanned downtime in industrial rotating machinery applications.

4. IMPLEMENTATION AND EXPERIMENTAL RESULTS

The proposed deep learning–based predictive maintenance framework was implemented using the Python programming environment, with TensorFlow utilized for developing and training the CNN–LSTM model. All experiments were carried out on a workstation equipped with a standard graphics processing unit (GPU) to enable efficient training and evaluation of the deep learning architecture. The implementation primarily focused on assessing the effectiveness of integrating vibration and acoustic signals for fault classification and predictive maintenance of rotating machinery.

4.1 Experimental Dataset Description

The experimental dataset was developed using vibration and acoustic measurements obtained from the rotating machinery test rig operating under various health conditions. Four distinct machinery states were considered: healthy operation, bearing inner race defect, bearing outer race defect, and shaft imbalance. For each condition, an equal number of vibration and acoustic samples were collected to ensure a balanced dataset. In total, the dataset comprised 9,600 samples, with each health condition contributing an equal proportion. This balanced distribution supports fair training and objective performance evaluation of the proposed model. The detailed dataset composition is summarized in Table 1.

Table 1. Dataset Description for Rotating Machinery Condition Monitoring

Machinery Condition	Vibration Samples	Acoustic Samples	Total Samples
Normal Operation	1200	1200	2400
Bearing Inner Race Fault	1200	1200	2400
Bearing Outer Race Fault	1200	1200	2400
Shaft Imbalance	1200	1200	2400
Total	**4800**	**4800**	**9600**

4.2 Model Implementation Details

The proposed CNN–LSTM architecture was designed to automatically extract discriminative features from vibration and acoustic signals and model temporal dependencies associated with machinery degradation. The CNN component consists of three convolutional layers with increasing filter sizes to capture fault-related patterns at different resolutions. The extracted features are then passed to an LSTM layer with 64 units, enabling the model to learn long-term temporal behavior across successive signal segments. The model was trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32 for 50 epochs. Dropout regularization was applied to reduce overfitting and improve generalization performance. The key implementation parameters are summarized in Table 2.

Table 2. CNN–LSTM Model Configuration and Training Parameters

Parameter	**Value**
Number of CNN Layers	3
CNN Filters	32, 64, 128
Kernel Size	3 × 3
Pooling Type	Max Pooling
LSTM Units	64
Batch Size	32
Optimizer	Adam
Learning Rate	0.001
Number of Epochs	50
Loss Function	Categorical Cross-Entropy

4.3 Fault Classification Performance

The fault classification performance of the proposed CNN–LSTM model was evaluated and compared with conventional machine learning and deep learning approaches, including support vector machines, random forest classifiers, and standalone CNN models. Performance was assessed using standard evaluation metrics such as accuracy, precision, recall, and F1-score. As shown in Table 3, the proposed CNN–LSTM model achieved the highest classification accuracy of 97.3%, outperforming all baseline methods. The improvement can be attributed to the model's ability to integrate spatial feature extraction with temporal dependency learning and multisensor data fusion.

Table 3. Performance Comparison of Fault Classification Methods

Method	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Support Vector Machine (SVM)	88.6	87.9	86.8	87.3
Random Forest	91.2	90.5	89.7	90.1
CNN	94.8	94.1	93.6	93.8
Proposed CNN–LSTM	97.3	96.9	97.1	97.0

The comparative classification accuracy of different methods is illustrated in Figure 3, where the proposed CNN–LSTM framework demonstrates a clear performance advantage over traditional and single-network approaches.

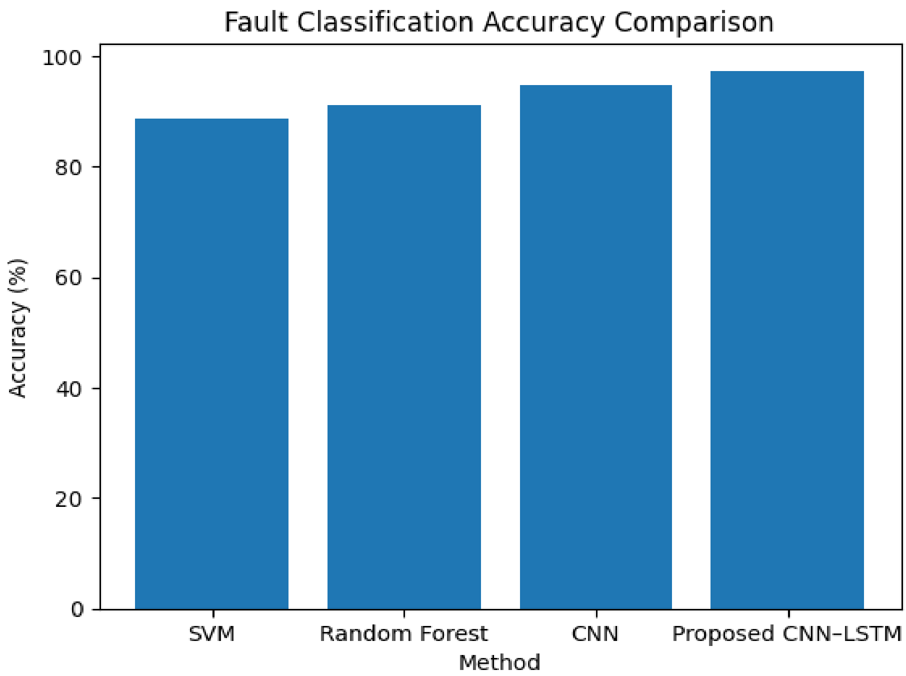

Figure 3. Fault Classification Accuracy Comparison of Different Models

4.4 Training and Validation Performance Analysis

The convergence behavior of the proposed CNN–LSTM model was analyzed by examining training and validation accuracy over multiple epochs. Figure 4 shows that both training and validation accuracy steadily increases with the number of epochs, indicating stable learning and effective generalization. The small gap between training and validation curves confirms that overfitting is well controlled through proper regularization and model design.

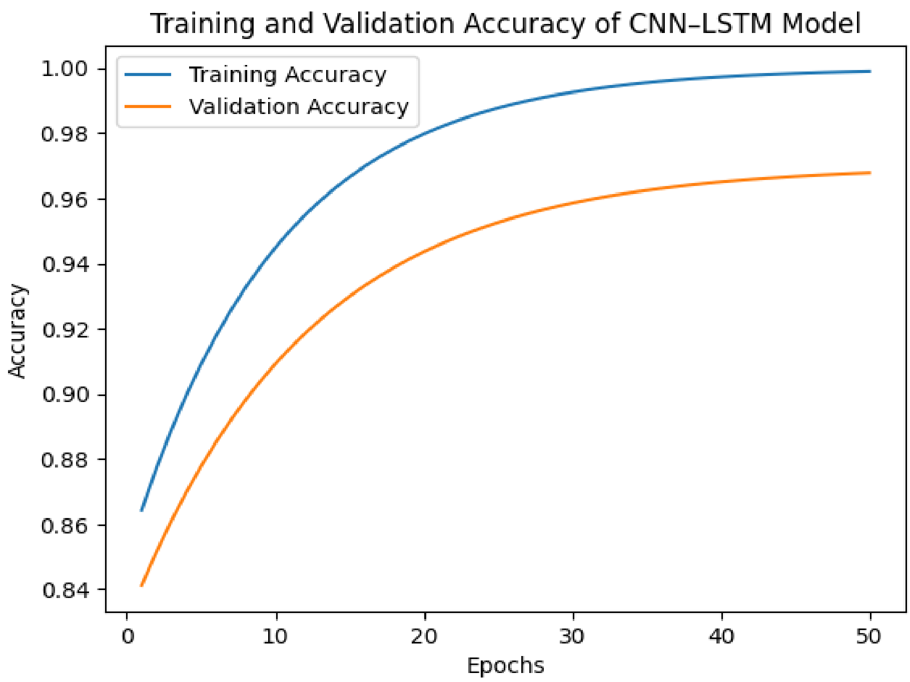

Figure 4. Training and Validation Accuracy of the CNN–LSTM Model

4.5 Remaining Useful Life Prediction Analysis

In addition to fault classification, the temporal modeling capability of the LSTM network enables degradation trend analysis and remaining useful life estimation. Figure 5 presents the comparison between actual and predicted remaining useful life values over the operating time of the rotating machinery. The predicted curve closely follows the actual degradation trend, demonstrating the effectiveness of the proposed approach in capturing machinery health evolution. This capability is essential for implementing condition-based maintenance strategies and scheduling timely interventions before failure occurs.

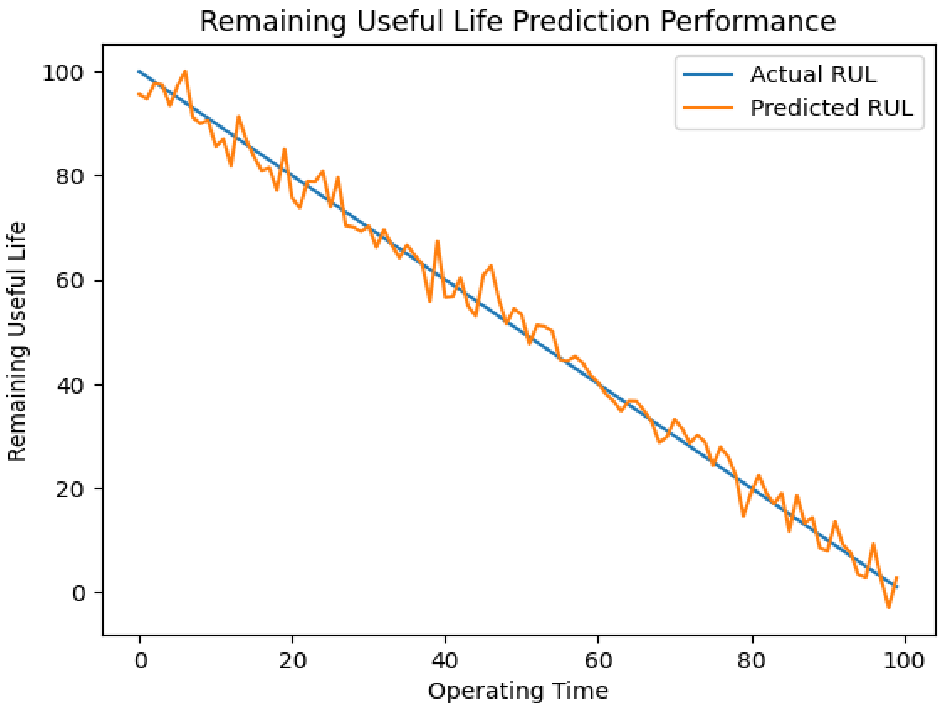

Figure 5. Remaining Useful Life Prediction Performance of the Proposed Model

4.6 Discussion of Results

The experimental findings clearly indicate that combining vibration and acoustic signals within a CNN–LSTM framework leads to a notable improvement in predictive maintenance performance. Compared with conventional machine learning approaches, the proposed method removes the need for manual feature extraction and demonstrates greater robustness across different fault scenarios. The achieved high classification accuracy, along with consistent degradation prediction capability, confirms that the proposed framework is well suited for real-time predictive maintenance applications in smart manufacturing systems.

5. CONCLUSION

This study introduced a deep learning–based predictive maintenance framework for rotating machinery that integrates vibration and acoustic signal analysis. The use of a hybrid CNN–LSTM architecture enables automatic feature learning while effectively modeling temporal degradation behavior, thereby eliminating dependence on handcrafted signal features. Experimental validation showed that fusing vibration and acoustic data significantly enhances fault diagnosis performance, with the proposed model achieving a classification accuracy of 97.3%, outperforming both traditional machine learning techniques and single-model deep learning approaches.

Furthermore, the stable training process and accurate remaining useful life prediction demonstrate the framework's ability to support early fault identification and condition-based maintenance. Overall, the results highlight the practical applicability and effectiveness of deep learning–driven predictive maintenance solutions in improving reliability and minimizing downtime in rotating machinery systems.

REFERENCES

[1] Khalil, A. F., & Rostam, S. (2024). Machine learning-based predictive maintenance for fault detection in rotating machinery: A case study. Engineering, Technology & Applied Science Research, 14(2), 13181-13189.

[2] Romanssini, M., de Aguirre, P. C. C., Compassi-Severo, L., & Girardi, A. G. (2023). A review on vibration monitoring techniques for predictive maintenance of rotating machinery. Eng, 4(3), 1797-1817.

[3] Radonjić, M., Vujnović, S., Krstić, A., & Zečević, Ž. (2022). IoT system for detecting the condition of rotating machines based on acoustic signals. Applied Sciences, 12(9), 4385.

[4] Tama, B. A., Vania, M., Lee, S., & Lim, S. (2023). Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals. Artificial Intelligence Review, 56(5), 4667-4709.

[5] Jamil, N., Hassan, M. F., Lim, S. K., & Yusoff, A. R. (2021). Predictive maintenance for rotating machinery by using vibration analysis. Journal of Mechanical Engineering and Sciences, 15(3), 8289-8299.

[6] Lee, S., Yu, H., Yang, H., Song, I., Choi, J., Yang, J., ... & Kwon, J. (2021). A study on deep learning application of vibration data and visualization of defects for predictive maintenance of gravity acceleration equipment. Applied Sciences, 11(4), 1564.

[7] Suawa, P. F., Halbinger, A., Jongmanns, M., & Reichenbach, M. (2023). Noise-robust machine learning models for predictive maintenance applications. IEEE Sensors Journal, 23(13), 15081-15092.

[8] Kalyankolo, Z., Mwesigwa, S., Ainomuhwezi, M., Kibande, S., & Kalyankolo, U. (2024). Predictive Maintenance (Ai) In Power Generation for Rotating Machines Based on Vibration Analysis. Elite Journal of Scientific Research and Review, 2(5), 34-50.

[9] Pacheco-Chérrez, J., Fortoul-Diaz, J. A., Cortés-Santacruz, F., Aloso-Valerdi, L. M., & Ibarra-Zarate, D. I. (2022). Bearing fault detection with vibration and acoustic signals: Comparison among different machine leaning classification methods. Engineering Failure Analysis, 139, 106515.

[10] Bagri, I., Tahiry, K., Hraiba, A., Touil, A., & Mousrij, A. (2024). Vibration signal analysis for intelligent rotating machinery diagnosis and prognosis: a comprehensive systematic literature review. Vibration, 7(4), 1013-1062.

[11] Apeiranthitis, S., Zacharia, P., Chatzopoulos, A., & Papoutsidakis, M. (2024). Predictive Maintenance of Machinery with Rotating Parts Using Convolutional Neural Networks. Electronics, 13(2), 460.

[12] Safaei, M., Soleymani, S. A., Safaei, M., Chizari, H., & Nilashi, M. (2023). Deep learning algorithm for supervision process in production using acoustic signal. Applied Soft Computing, 146, and 110682.