MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 1303-1311

Deep LSTM-Based Bearing Condition Monitoring with Batch Normalization Enhancement

Sujit Kumar[a], Amit Kumar[b] , Dilkhush Rajak[c], Manoj Bhaskar[d], Pranay Kumar[e] and Brajesh Kumar[f]

[a]Department of Electrical Engineering, Government Engineering College, Banka, Bihar Engineering University Patna, Bihar 813102, India. sujitaryan0412@gmail.com

[b]Department of Electrical Engineering, Government Engineering College, Banka, Bihar Engineering University Patna, Bihar 813102, India. amitkrvineet246@gmail.com

[c]Department of Electrical Engineering, Government Engineering College, Munger, Bihar Engineering University Patna, Bihar 811202, India. dilkhush09eee@gmail.com

[d]Department of Electronics and Communication Engineering, Government Engineering College, Banka, Bihar Engineering University Patna, Bihar 813102, India. manojbhaskar89@gmail.com

[e]Department of Mechanical Engineering, Government Engineering College, Banka, Bihar Engineering University Patna, Bihar 813102, India. pranayme53@gmail.com

[f]Department of Mechanical Engineering, Government Engineering College, Banka, Bihar Engineering University Patna, Bihar 813102, India brajeshksme57@gmail.com

Received: 15.06.2025, Accepted: 03.01.2026

Abstract

Accurate diagnosis of bearing faults is critical for ensuring the reliability and safety of industrial electric motors, particularly under complex operating conditions where traditional signal processing techniques often exhibit limited robustness. To overcome these challenges, this study proposes an advanced deep learning–based diagnostic framework that integrates signal decomposition, feature optimization, and temporal modeling for intelligent bearing fault identification. Vibration signals acquired from rotating machinery are first processed using Ensemble Empirical Mode Decomposition (EEMD) to effectively suppress noise and isolate meaningful intrinsic mode functions. Relevant features are then selected based on correlation coefficient analysis, and Principal Component Analysis (PCA) is applied to reduce feature dimensionality while preserving essential fault-related information. The refined feature set is subsequently fed into a Long Short-Term Memory (LSTM) network enhanced with Batch Normalization to capture temporal dependencies and stabilize the training process. The inclusion of Batch Normalization improves convergence behavior and enhances model generalization under varying operating conditions. Experimental evaluation demonstrates that the proposed framework achieves superior diagnostic performance, reaching 100% classification accuracy and outperforming conventional fault diagnosis approaches. Owing to its robustness, low sensitivity to noise, and strong temporal learning capability, the proposed method provides an effective and reliable solution for real-time bearing fault diagnosis in industrial motor applications.

Keywords – Fault diagnosis, EEMD, PCA, Batch Normalization, LSTM

1. Introduction

Machine fault diagnosis is a critical process in industrial maintenance that identifies and classifies mechanical problems within equipment, including failures like gear damage, stator malfunctions, and bearing faults. The goal of machine fault diagnosis is to ensure that machinery remains operational and efficient, while reducing risks such as production stoppages, economic losses, and potential threats to human safety. Over the past few decades, condition monitoring systems based on expert systems and artificial intelligence (AI) have advanced significantly, providing more reliable fault detection capabilities.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 1303-1311

The introduction of machine learning (ML) techniques has revolutionized the field of machine fault diagnosis. Traditional methods often relied on manual feature extraction, but ML techniques can automatically learn and extract relevant features from raw data, making fault detection more efficient and accurate [1]. Among the various ML models used in fault diagnosis, Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) have shown remarkable success. These models typically work by transforming vibration signals from the time domain into the frequency domain using techniques such as Fast Fourier Transform (FFT). Once in the frequency domain, key features are extracted and used to classify different types of faults [2]. Research has demonstrated that incorporating multiple hidden layers in neural networks significantly improves fault detection accuracy. For example, fault classification accuracy rates in rotating machinery have been shown to increase from 88% to as high as 95% when deep learning architectures are used [3]. Principal Component Analysis (PCA) is another powerful technique used to improve the performance of machine learning models. By reducing the dimensionality of the feature space, PCA eliminates redundant data and makes the training process more efficient. This is particularly beneficial in models like multiclass SVMs, where reducing the feature space helps improve classification accuracy while reducing computational complexity. In addition to SVMs and ANNs, Random Forest (RF) algorithms have been applied for fault classification in rotating machines [4-6]. RF models use multiple decision trees to classify faults based on features extracted from methods like EEMD and wavelet decomposition. The ensemble nature of RF provides more robust and accurate fault classification compared to single decision tree models.

One of the most promising advancements in machine fault diagnosis is the development of hybrid machine learning architectures. These architectures combine multiple models to improve the accuracy and reliability of fault classification. A common approach is to use a two-stage architecture where the first stage focuses on dimensionality reduction, and the second stage performs the actual classification. For example, Recurrent Neural Network (RNN)-based variational autoencoders are often used in the first stage to reduce the dimensionality of vibration signals [7-8]. In the second stage, algorithms such as Random Forest, SVM, XGBoost, Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRUs) are applied to classify the faults based on the reduced feature set. These integrated architectures offer several advantages [9-11]. By using dimensionality reduction in the first stage, the models become more efficient, requiring less computational power while still maintaining high accuracy. The use of advanced classifiers like LSTMs and GRUs in the second stage allows the models to capture temporal dependencies in the vibration signals, further improving fault classification accuracy.

The key contributions of this paper are as follows:

1. The raw vibration signal is initially analyzed in both the time and frequency domains to capture fault-related characteristics. Ensemble Empirical Mode Decomposition (EEMD) is then applied to decompose the signal into multiple intrinsic mode functions (IMFs), enabling effective suppression of noise and non-stationary components. To retain only informative fault features, a correlation-coefficient-based selection strategy is employed, where IMFs with stronger relevance to fault conditions are preserved.
2. The selected IMFs form a high-dimensional feature set that may contain redundant information. To address this issue, Principal Component Analysis (PCA) is applied to reduce the dimensionality of the feature space while preserving the most significant fault-sensitive information. Batch Normalization is incorporated during model training to stabilize feature distributions and accelerate convergence, ensuring improved learning efficiency.
3. An LSTM-based deep learning model is subsequently developed using the optimized feature set to capture temporal dependencies inherent in vibration signals. The diagnostic performance of the proposed framework is evaluated and compared with existing machine learning approaches, demonstrating improved accuracy and computational efficiency in rolling bearing fault classification.

The paper is organized as follows: Section 1 presents an introduction to various fault diagnosis approaches. Section 2 outlines the methodology for bearing fault classification. Section 3 analyses the results and discussion. and Section 4 concludes the work.

3. Methodology

The bearing vibration signals used in this study were obtained from the data repository [15] and collected from both the fan-end and drive-end locations under healthy and various faulty operating conditions. The experiments were conducted at multiple rotational speeds, namely 1730 rpm, 1750 rpm, 1772 rpm, and 1797 rpm. These vibration signals exhibit significant noise contamination and quasi-stationary behavior, which can adversely affect fault identification accuracy. Therefore, as illustrated in Fig. 1, an Ensemble Empirical Mode Decomposition (EEMD)–based denoising strategy is employed as the initial preprocessing step. The original vibration signals are decomposed into a finite number of intrinsic mode functions (IMFs) along with a residual component. IMFs characterized by relatively low non-stationarity and strong correlation with the original signal are selected as informative features.

The resulting feature set is inherently high-dimensional, motivating the use of Principal Component Analysis (PCA) to eliminate redundant information and reduce computational complexity. Initially, eight feature components are extracted from the selected IMFs; however, to mitigate multicollinearity and improve diagnostic reliability, PCA is applied to project the feature space into a lower dimension. The eight features are effectively reduced to two principal components while preserving the majority of the signal's discriminatory information. These compact representations are subsequently provided as sequential inputs to the Batch Normalization–based LSTM model for accurate classification of bearing health conditions. The corresponding rolling bearing fault categories are summarized in Table 1.

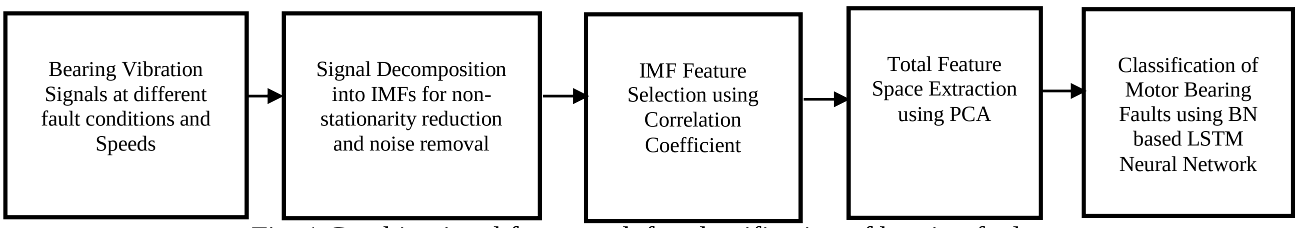

Fig. 1 Combinational framework for classification of bearing faults

Table 1 Rolling Bearing State

Bearing State (Approx Motor speed (rpm) = 1730,1750,1772,1797		
No.	Fault Diameter (inches)	Fault Location
1	-	Normal Condition (NC) (Class 0)
2	0.007	Inner Race Fault (IRF007) (Class 1)
3	0.021	Inner Race Fault (IRF021) (Class 2)
4	0.007	Outer race Fault (ORF007) (Class 3)
5	0.007	Outer Race Fault @ (6:00)^{am} (ORF007@6) (Class 4)
6	0.014	Outer Race Fault @ (12:00)^{am} (ORF014@12) (Class 5)

3.1 Data Pre-processing and Feature Preparation using EEMD

The raw vibration signals acquired from the fan-end (FE) and drive-end (DE) bearings under different operating conditions and rotational speeds are first examined in both the time and frequency domains. Frequency-domain analysis of vibration signals is widely adopted for bearing fault diagnosis, as it enables the identification of defect-related frequency components. This is achieved by applying the

Fast Fourier Transform (FFT) to time-domain signals, thereby revealing fault signatures in the frequency–amplitude spectrum.

In this work, the original vibration signals are transformed into the frequency domain prior to further processing. Ensemble Empirical Mode Decomposition (EEMD) is then applied to the FE and DE bearing signals collected at four distinct rotational speeds under healthy and faulty conditions. Through EEMD, each vibration signal is adaptively decomposed into a set of intrinsic mode functions (IMFs) along with a residual component, effectively mitigating mode mixing and reducing noise interference. High-frequency components are progressively separated across lower-order IMFs, with early IMFs capturing dominant fault-related frequencies, albeit with some noise contamination, while higher-order IMFs provide improved isolation of characteristic frequency components.

For subsequent analysis, eight vibration signals of length 15,000 samples are considered for each bearing condition. Each signal is decomposed into 14 IMF components, resulting in a high-dimensional feature representation suitable for noise suppression and non-stationary signal handling. Consequently, the constructed IMF-based feature dataset has a dimensional structure of [6×14×8×15000], where 6 denotes the bearing health conditions, 14 represents the number of IMFs, 8 corresponds to the vibration signals, and 15,000 indicates the sample length of each signal.

3.2 Feature Selection and Extraction

Any well-developed classification model relies heavily on being trained with relevant and significant features. It is crucial to avoid the capture of insignificant patterns by the model due to noise, which underscores the importance of appropriate feature selection. While EEMD offers benefits, it also increases the number of input signals, posing a challenge. To address this, the correlation relationship between all decomposed IMF signals and the actual raw signals is computed to select the best de-noised and highly correlated IMFs with the original signal. The highest IMF coefficient values are selected as the best features for fault classification. This process is repeated for all types of fault conditions at different speeds. Ultimately, the application of EEMD and feature selection using correlation coefficients yields a set of 8 IMF features, each with a sample length of 15,000 for six bearing conditions [6*8*15000].

PCA was applied to the initial feature space of [6* 8*15000] to reduce dimensionality and eliminate data redundancy. All selected IMFs were reduced to two principal components, as they captured most of the variance in the data, resulting in a reduced data size of [6* 2*15000] for each bearing condition. The reduced feature vectors for all six conditions were then used as input to train the BN Based LSTM classifier mode

3.3 LSTM model

Relevant features are extracted from preprocessed data using the EEMD technique and the Correlation Coefficient. These features are then fed into a neural network to identify the health state of roller bearings. Figure 2 shows the architecture of the combined LSTM model, featuring both a feature extractor and a classifier module. The reduced features fed into the LSTM layer, which captures temporal dependencies in the data, thereby improving fault detection accuracy.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 1303-1311

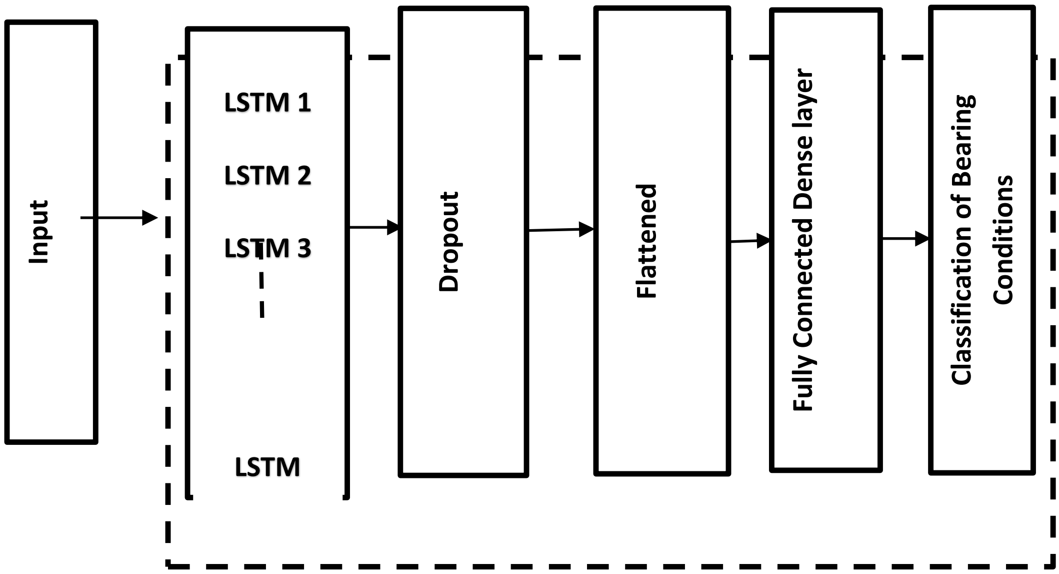

Fig. 2 Combinational framework of proposed model

3.4 Combinational Framework of Bearing Condition

The fault diagnosis algorithm is divided into two sections. The first is to capture the dynamic information from vibrational data and the second is to develop a deep learning classifier model for classifying the various types of bearing faults under different conditions. The following steps has been taken for rolling bearing fault diagnosis using the BN-based LSTM model. The details are as follows

1. Sensor data are first collected from the bearing system under different operating and fault conditions. The acquired signals are subsequently preprocessed and normalized to a fixed range between 0 and 1 to ensure uniform scaling and numerical stability.
2. Ensemble Empirical Mode Decomposition (EEMD) is applied to the vibration signals to perform adaptive time–frequency analysis, enabling effective handling of non-stationary characteristics and noise components.
3. Relevant fault-sensitive features are selected using a correlation coefficient–based method, retaining only those intrinsic mode functions (IMFs) that exhibit strong correlation with the original vibration signals.
4. Principal Component Analysis (PCA) is employed to reduce the dimensionality of the selected IMF feature set, preserving the most informative components associated with six different bearing fault categories while eliminating redundant features.
5. The prepared dataset is partitioned into training, validation, and testing subsets to ensure unbiased model development and performance assessment.
6. Batch Normalization is incorporated into the learning framework to stabilize feature distributions, accelerate convergence, and enhance the generalization capability of the neural network.
7. A Batch Normalization–enhanced Long Short-Term Memory (LSTM) neural network is trained using the optimized feature set to learn temporal dependencies and classify bearing fault conditions.
8. Model performance is evaluated using multiple metrics, including classification accuracy, training and validation loss, confusion matrix analysis, receiver operating characteristic (ROC) curves with area-under-the-curve (AUC) values, and additional performance indices.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 1303-1311

This framework includes an input layer, hidden layers, dense layers, a softmax layer, and an output layer. The extracted data is fed into the neural network, with each LSTM layer comprising 50 neurons. The LSTM model is trained with the following hyperparameters: Adam optimizer for its efficiency and adaptive learning rate, mean squared error loss function, and a batch size of 50 for balanced memory usage and convergence. A dropout rate of 0.2 is used to mitigate overfitting, while the model undergoes 50 epochs of training with a learning rate of 0.01. Softmax activation is employed for classification. The LSTM network is developed and trained using Python with the Keras package and TensorFlow 1.0, leveraging their robust features for effective model development and training.

4. Results and Discussion

This study develops a Batch Normalization–based stacked LSTM (BN-LSTM) network for multi-class rolling bearing fault classification and systematically evaluates its performance. A confusion matrix is used to compare predicted and true bearing conditions, where correct classifications lie along the diagonal and misclassifications appear off-diagonal. As shown in Fig. 3, the BN-LSTM accurately classifies bearing classes 0–5, enabling the computation of key performance metrics. Owing to its ability to automatically learn temporal fault features, the model eliminates the need for manual feature extraction, with quantitative results summarized in Table 2. Further validation using ROC analysis (Fig. 4) shows curves strongly biased toward the true positive rate, confirming the high sensitivity, robustness, and reliability of the proposed BN-LSTM framework for multi-class bearing fault diagnosis.

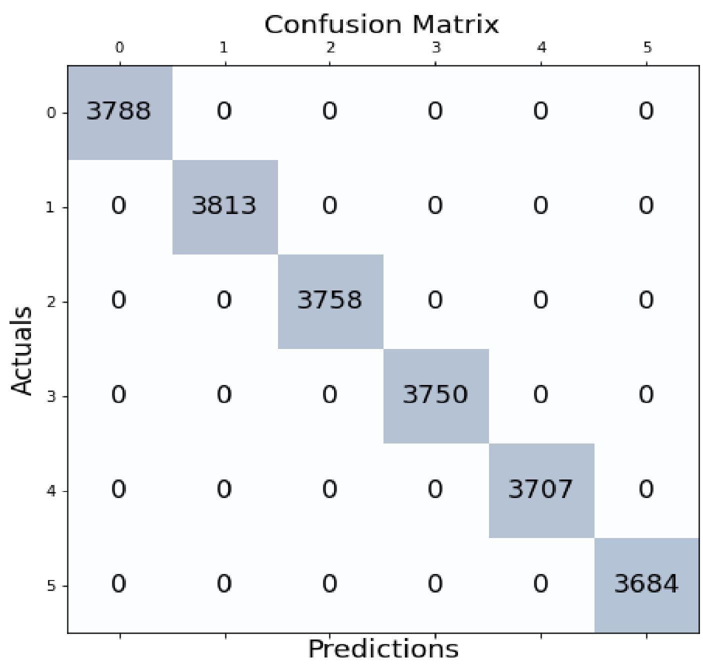

Fig. 3 Confusion Matrix of Proposed BN based LSTM

The classification accuracy thus obtained using BN based LSTM model has been compared with other machine learning and deep learning models used in the literature. Table 2 shows the classification accuracy of the proposed model compared with other existing work cited in the literature.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 35 Issue 2, 2025, Pages: 1303-1311

Table 2 Comparison of Classification Accuracy

Methods	Testing Accuracy (%)
SVM [16]	87.45
CNN + RF [16]	99.73
CNN [16]	99.66
VAE + RF [3]	98.19
VAE + Neural Network [3]	96.97
XGBoost [3]	94
Random Forest [3]	55.5
Neural Network [3]	85
1D-CNN-LSTM [17]	97.69
Attention LSTM [17]	84.73
Stacked LSTM deep network (Reduced features) [18]	98.88
CNN-LSTM Neural Network Without BN	89.55
BN based LSTM Model (Proposed)	**100**

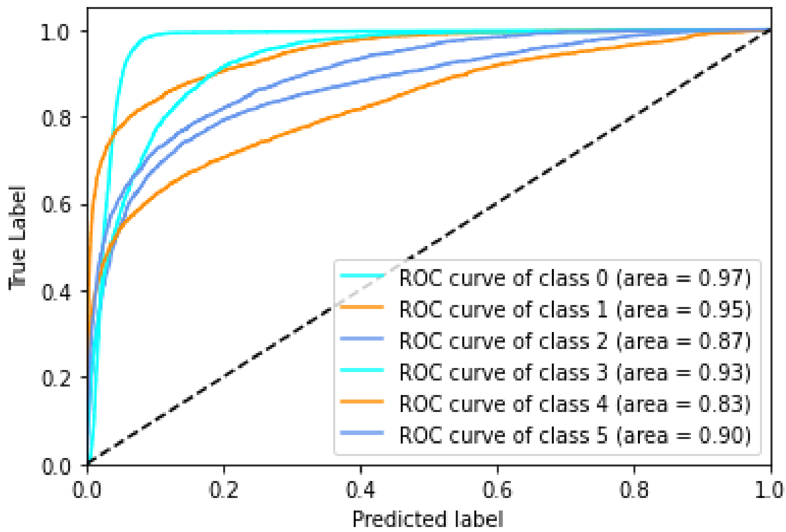

Fig. 4 ROC Curves of Classes Predicted by BN Based LSTM Model

The Area Under Curve of ROC for each Class 0, Class 1, Class 2, class 3, class 4 and class 5 deduced as 0.97, 0.95 0.87, 0.93, 0.83 and 0.90 respectively (Fig. 4) also validates the model's ability to correctly classify the condition that bearings belongs to. Above all, confirms the suitability of the proposed BN based LSTM model for classification of bearing condition from the raw vibration data.

6. Conclusion

Rolling element bearings are critical components in rotating machinery and are prone to degradation under harsh operating conditions. This work proposes an intelligent bearing fault classification framework using a Batch Normalization–based Long Short-Term Memory (BN-LSTM) network. Ensemble Empirical Mode Decomposition (EEMD) is employed to handle noise and non-stationary vibration signals by extracting fault-relevant intrinsic mode functions, while Principal Component Analysis (PCA) reduces feature dimensionality and improves computational efficiency. The proposed method achieves 100% classification accuracy, demonstrating superior performance over conventional approaches. The high accuracy is attributed to LSTM's ability to capture long-term temporal dependencies and the stabilizing effect of Batch Normalization. The combined EEMD–PCA preprocessing further enhances robustness, making the framework suitable for practical industrial condition monitoring..

ACKNOWLEDGMENT

The authors would like to thank Case Western Reserve University for the provided motor bearing dataset on its website.

Declarations

Funding There is no funding provided for this research.

Conflicts of interest / Competing interests We authors declare that there are no conflicts of interests relevant to this manuscript.

References

[1] Kankar, Pavan Kumar, Satish C. Sharma, and Suraj Prakash Harsha. Fault diagnosis of ball bearings using machine learning methods. Expert Systems with applications. 2011;38(3): 1876-1886.Available from https://doi.org/10.1016/j.eswa.2010.07.119

[2] Zhang, Yahui, Taotao Zhou, Xufeng Huang, Longchao Cao, and Qi Zhou. Fault diagnosis of rotating machinery based on recurrent neural networks. Measurement. 2021;171:108774. Available from https://doi.org/10.1016/j.measurement.2020.108774

[3] Yang Huang, Chiun-Hsun Chen, and Chi-Jui Huang. Motor Fault Detection and Feature Extraction using RNN-based Variational Autoencoder. IEEE Access. 2019; 7: pp. 139086-139096. Available from https://doi.org/10.1109/ACCESS.2019.2940769

[4] Xu, Gaowei, Min Liu, Zhuofu Jiang, Dirk Söffker, and Weiming Shen. Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning. Sensors. 2019; 19(5): 1088. Available from https://doi.org/10.3390/s19051088

[5] Jiang, Fan, Zhencai Zhu, Wei Li, Yong Ren, Gongbo Zhou, and Yonggen Chang. A fusion feature extraction method using EEMD and correlation coefficient analysis for bearing fault diagnosis. Applied Sciences. 2018;8:1621. Available from https://doi.org/10.3390/app8091621.

[6] Xiwen Qin, Qiaoling Li, Xiaogang Dong, and Siqi Lv. The Fault Diagnosis of Rolling Bearing Based on Ensemble Empirical Mode Decomposition and Random Forest. Shock and Vibration. 2017;2017. Available from https://doi.org/10.1155/2017/2623081.

[7] Ma, Jianpeng, Chengwei Li, and Guangzhu Zhang. Rolling bearing fault diagnosis based on deep learning and autoencoder information fusion. Symmetry. 2022: 14(1):13. Available from https://doi.org/10.3390/sym14010013.

[8] Rui Zhao, Jinjiang Wang, Ruqiang Yan and Kezhi Mao. Machine health monitoring with LSTM networks. Proceedings by 10th international conference on sensing technology (ICST). IEEE. 2016;1-6. Available from https://doi.org/10.1109/ICSensT.2016.7796266.

[9] Wu, QingE, Tao Zong, Wenfang Cheng, Yong Li, and Penglei Li. Network Construction for Bearing Fault Diagnosis Based on Double Attention Mechanism. Computational Intelligence and Neuroscience. 2022; 2022. Available from https://doi.org/10.1155/2022/3987480.

[10] Zhu, Yingmou, Hongming chen, Wei Meng, Qing Xiong, and Yongjian Li. A wide kernel CNN-LSTM-based transfer learning method with domain adaptability for rolling bearing fault diagnosis with a small dataset. Advances in Mechanical Engineering. 2022; 14(11):16878132221135745. Available from https://doi.org/10.1177/1687813222113.

[11] He, Feng, and Qing Ye. A bearing fault diagnosis method based on wavelet packet transform and convolutional neural network optimized by simulated annealing algorithm. Sensors. 2022; 22(4):1410. Available from https://doi.org/10.3390/s22041410.

[12] Sun, Haibin and Zhao, Shichao. Fault diagnosis for bearing based on 1DCNN and LSTM. Shock and Vibration. 2021; 2021:1-17. Available from https://doi.org/10.1155/2021/1221462.

[13] Sun, HaiBin and Fan, YueGuang. Fault diagnosis of rolling bearings based on CNN and LSTM networks under mixed load and noise. Multimedia Tools and Applications. 2023; 82(28): 42543-43567. Available from https://doi.org/10.1007/s11042-023-15325-w.

[14] Chen, Hongming and Meng, Wei and Li, Yongjian and Xiong, Qing. An anti-noise fault diagnosis approach for rolling bearings based on multiscale CNN-LSTM and a deep residual learning model. Measurement Science and Technology. 2023;34(4):045013. Available from https://doi.org/10.1088/1361-6501/acb074.

[15] Case Western Reserve University bearing data center. http://engineering.case.edu/bearingdatacenter.

[16] Xu, Gaowei, Min Liu, Zhuofu Jiang, Dirk Söffker, and Weiming Shen. Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning. Sensors. 2019; 19(5):1088. Available from https://doi.org/10.3390/s19051088.

[17] Xie, Yuan and Zhang, Tao and others. Fault diagnosis for rotating machinery based on convolutional neural network and empirical mode decomposition. Shock and Vibration. 2017; 2017. Available from https://doi.org/10.1155/2017/3084197.

[18] Kumar, Sujit and Ganga, D. Combinational Framework for Classification of Bearing Faults in Rotating Machines. Journal of Computing and Information Science in Engineering. 2023; 24(2): 021012. Available from https://doi.org/10.1115/1.4062453.