Multi-domain feature-based early detection of bearing faults using MLP classifier on NASA IMS dataset

Pham Van Nam*, Nguyen Vu Thang, Trinh Trong Chuong, Tran Thi Hang

Faculty of Engineering, Electrical &Automation, Hanoi University of Industry, 298 Cau Dien, Tay Tuu, Hanoi, Vietnam.
*Corresponding author: nampv@haui.edu.vn
Received 7 Apr. 2025; Revised 3 Jun. 2025; Accepted 27 Jun. 2025; Published 2 Oct. 2025.
DOI: https://doi.org/10.54939/1859-1043.j.mst.106.2025.48-54

ABSTRACT

The degradation of bearing components in industrial machinery leads to increased maintenance costs and unexpected operational downtime. This paper presents a novel methodology that integrates multi-domain statistical feature extraction spanning both time-domain and frequency-domain characteristics to enhance the precision of bearing fault detection. A Multi-Layer Perceptron (MLP) model was trained on the NASA IMS Bearing dataset, achieving a classification accuracy of 86.5% across five degradation stages. Experimental results demonstrate that the proposed method outperforms traditional classifiers such as Support Vector Machine (SVM) and Random Forest, particularly in data-scarce environments. Furthermore, the model is well-suited for deployment on resource-constrained embedded diagnostic systems. This approach offers a practical and efficient solution for predictive maintenance, contributing to the reduction of operational costs in industrial applications.

Keywords: Bearing faults prediction; NASA IMS; MLP model; Multi-domain features; Predictive maintenance.

1. INTRODUCTION

Bearings are fundamental components in modern industrial systems, ensuring the stable operation of electromechanical machinery. They serve as the primary load-bearing interfaces between rotating and stationary parts, playing a critical role in applications such as large-scale wind turbines, conveyor systems in automotive manufacturing plants, and industrial robots within automated assembly lines [1-4]. Due to their operational context, bearings are frequently subjected to harsh conditions, including high mechanical loads, elevated rotational speeds, continuous friction, and extreme ambient temperatures. Prolonged exposure to these stressors makes them vulnerable to various forms of degradation, such as inner/outer race cracks, rolling element surface spalling, and, in severe cases, complete structural failure [5-7].

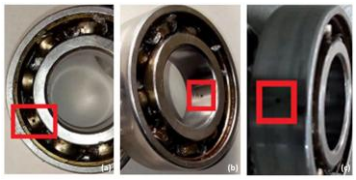

Figure 1. *Defected bearings with different faults: (a) Ball, (b) Inner race, (c) Outer race [8].*

According to a report published in *Eureka* [9], bearing failures are the leading cause of unexpected shutdowns in industrial machinery, resulting in global economic losses estimated at up to 50 billion USD annually. Notably, such failures account for nearly 40% of all unplanned downtime in medium to large-scale industrial operations. These statistics emphasize the urgent need for early detection of bearing degradation to enable predictive maintenance strategies, rather than relying on reactive maintenance following catastrophic failure. Accurate diagnosis and prognosis of bearing degradation stages not only pose a significant technical challenge but also offer substantial economic benefits, including reduced downtime, optimized maintenance schedules, and extended equipment lifespan.

Traditional approaches such as vibration spectrum analysis [10], wavelet transform [11], and thermal monitoring [12] have been widely used for bearing fault diagnosis. However, these methods are often limited in early fault detection due to noise-prone signals and a reliance on domain expertise for appropriate feature selection and threshold tuning. In the context of smart manufacturing and increasing automation, these limitations hinder their scalability and practical deployment.

Recent advances in artificial intelligence, particularly deep learning, have introduced new opportunities for fault diagnosis and predictive maintenance. Models such as Convolutional Neural Networks (CNN) [13], Long Short-Term Memory networks (LSTM) [14], and hybrid architectures like CNN-SVM and CNN-LSTM [15] have demonstrated high classification accuracy (typically >85%) on laboratory bearing datasets. In these approaches, vibration signals are often transformed into time–frequency representations using techniques such as Short-Time Fourier Transform (STFT), Continuous Wavelet Transform (CWT), or Empirical Mode Decomposition (EMD), and then fed into deep neural networks as either images or time series. Several studies, including those by D.H (2024) [16], Wang (2023) [17], and Han (2023) [18], have shown that combining CNN-based feature extraction with robust classifiers such as SVM significantly enhances diagnostic performance. This study proposes an efficient diagnostic framework tailored for resource-constrained industrial environments. The framework combines multi-domain statistical feature extraction with a Multi-Layer Perceptron (MLP) classifier to accurately identify different stages of bearing degradation [20].

2. METHODOLOGY

2.1. NASA IMS bearing datasets

The data acquisition process was conducted continuously over a one-month period to monitor the condition of two bearings from initial operation to failure. Each bearing was equipped with two vibration sensors oriented orthogonally (the X and Y axes), resulting in four signal channels. Data was recorded every 10 minutes, capturing one second of vibration signal at a sampling frequency of 20 kHz, corresponding to 20,000 samples per channel.

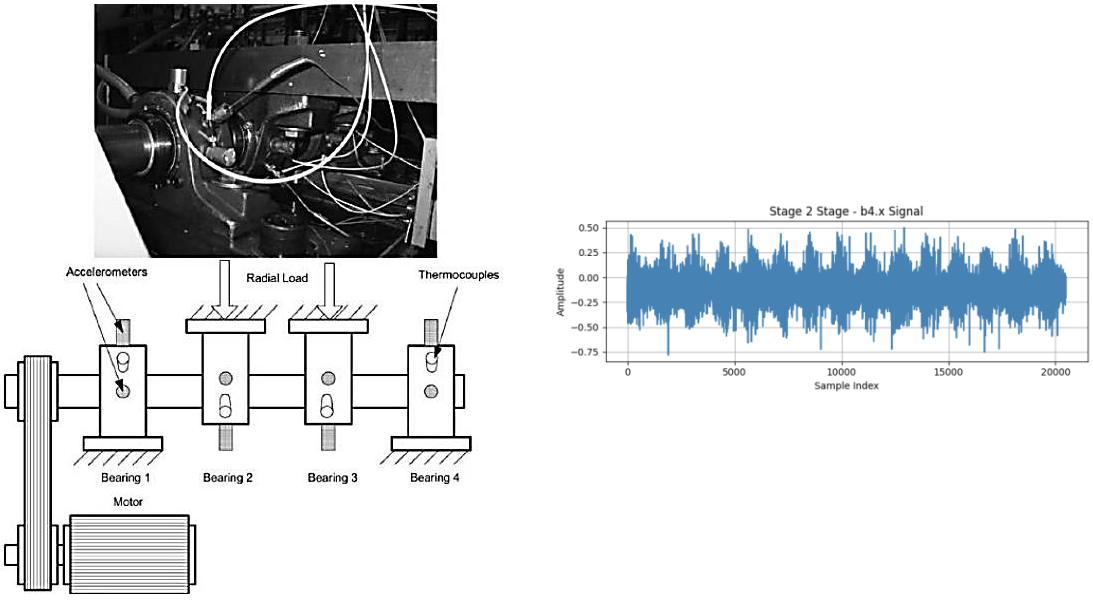

Figure 2. *Bearing test rig and sensor placement illustration [19] (left), Time-domain vibration signals of b4.x sensor (right).*

The study constructs a vibration signal dataset consisting of five distinct machine condition states: Early, representing the initial phase of operation; Normal, indicating stable performance

with no signs of failure; Suspect, reflecting slightly abnormal behavior; Failure, showing clear signs of malfunction; and Stage 2, corresponding to severe structural damage.

Illustration of vibration signals along the B4.X axis at five distinct time points during the bearing's operational lifecycle, corresponding to samples 40th (Early), 500th (Normal), 1400th (Suspect), 1700th (Failure), and 2100th (Stage 2). The variations in amplitude and waveform shape over time reflect the progressive degradation of the bearing's mechanical condition.

2.2. Feature extraction

2.2.1. Time-domain feature extraction

To enhance the learning efficiency and generalization capability of the AI model, the raw vibration signals are transformed into representative statistical features. This transformation helps eliminate noise and highlight underlying trends or anomalies in the signal. The extracted features include: minimum value, maximum value, mean, median, skewness, kurtosis, and root mean square (RMS).

Table 1. *Time-domain feature extraction formulas for vibration signal analysis.*

No.	Feature Name	Formula
1	Mean	$\text{Mean} = \frac{1}{N}\sum_{i=1}^{N} x_i$
2	Skewness	$\text{Skew} = \frac{N}{(N-1)(N-2)}\sum_{i=1}^{N}\left(\frac{x_i - \text{Mean}}{\text{SD}}\right)^3$
3	Kurtosis	$\text{Kurtosis} = \frac{N(N+1)}{(N-1)(N-2)(N-3)}\sum_{i=1}^{N}\left(\frac{x_i - \text{Mean}}{\text{SD}}\right)^4 - \frac{3(N-1)^2}{(N-2)(N-3)}$
4	Root mean square	$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$

Where: N is the number of data points in the signal sample; x_i is the value of the i-th data point in the sample.

2.2.2. Frequency-domain feature extraction

To detect frequency components characteristic of bearing faults, the signal is transformed into the frequency domain using Fast Fourier Transform (FFT). The features include:

Table 2. *Frequency-domain characteristic formulas for bearing fault diagnosis.*

No.	Feature name	Fomula
1	Fundamental Train Frequency	$FTF = \frac{1}{2} f_r \left(1 - \frac{B_d}{P_d}\cos\ \alpha\right)$
2	Ball Pass Frequency Outer	$BPFO = \frac{N_b}{2} f_r \left(1 - \frac{B_d}{P_d}\cos\ \alpha\right)$
3	Ball Pass Frequency Inner	$BPFI = \frac{N_b}{2} f_r \left(1 + \frac{B_d}{P_d}\cos\ \alpha\right)$
4	Ball Spin Frequency	$BSF = \frac{P_d}{2B_d} f_r \left(1 - \left(\frac{B_d}{P_d}\cos\ \alpha\right)^2\right)$

Where: f_r – Rotational frequency; P_d – Pitch diameter; B_d– Ball diameter; $\cos\ \alpha$ – Cosine of the contact angle (α).

To enhance feature resolution in the frequency domain, the FFT spectrum of the signal is divided into four distinct bands, and the sum of amplitudes within each band is calculated to represent the energy distribution: Low Frequency Power (below 1,250 Hz), Medium Frequency Power (1,250–2,600 Hz), High Frequency Power (2,600–6,000 Hz), and Very High Frequency Power (above 6,000 Hz). During the feature extraction process, each sensor generates 15 features, including 07 time-domain features and 08 frequency-domain features. Since each bearing is monitored by two sensors (X-axis and Y-axis), the total number of features obtained for each bearing is 30.

2.3. Machine learning for motor fault diagnosis

This study employs three machine learning algorithms for classification performance evaluation: Support Vector Machine (SVM), applied with a One-vs-One strategy and class weighting to enhance generalization and handle class imbalance; Random Forest, an ensemble of decision trees trained on random data subsets to reduce overfitting and highlight important features; and Multi-Layer Perceptron (MLP), a neural network with fully connected layers capable of learning nonlinear relationships from extracted feature vectors.

3. RESULTS AND DISCUSSION

3.1. Implementation and environment setup

The study extracts information from raw NASA IMS vibration signal data into time-domain and frequency-domain features. The dataset consists of 2,156 vibration signal samples with a sampling frequency of 20 kHz, representing 5 operating states of the engine: early, normal, suspect, failure, and stage 2, for each sensor. There are 4 sensors on two bearings, B3 and B4, resulting in a total of 8,624 samples. The data is split into 70% for training, 15% for validation, and 15% for testing.

Table 3. *Dataset split for training the model.*

Class	Total	Train (70%)	Validation (15%)	Test (15%)
Early	1,396	977	209	210
Normal	4,180	2,926	627	627
Suspect	1,530	1,071	230	229
Failure	884	619	133	132
Stage 2	634	443	95	96
Total	8,624	6,036	1,294	1,294

3.2. Impact of hidden layer size on MLP accuracy

To determine the optimal architecture of the MLP model, the study evaluates its performance under varying hidden layer sizes, incrementally testing the number of neurons from 2 to 100.

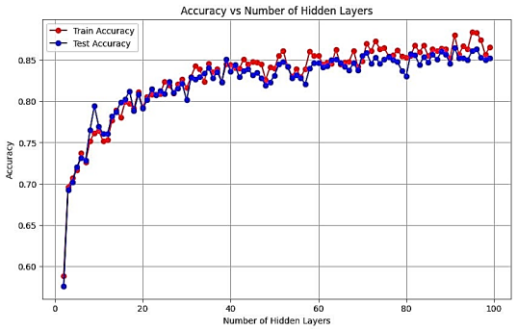

Figure 3. *Classification accuracy of the MLP model with varying numbers of hidden neurons.*

As illustrated in the chart, the classification accuracy increases with the number of neurons, reaching its peak at 91 neurons. Although performance shows a slight upward trend when increasing the number of neurons to 100, the improvement is marginal. To optimize computational efficiency and meet the prediction requirements on the device, the study selected 91 neurons as the optimal number, ensuring a balance between accuracy and computational performance.

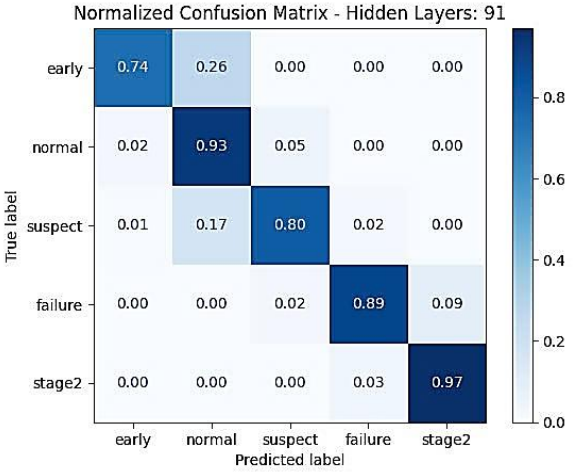

Figure 4. *Confusion matrix of the proposed MLP.*

Figure 4 illustrates the confusion matrix of the proposed model (MLP). Most of the misclassifications occur between the 'suspect' and 'normal' classes due to overlapping signal features (similarities between the two). On the other hand, the model shows no confusion between the 'normal' and 'stage 2' states, demonstrating its ability to distinctly differentiate between the degradation stages.

3.3. Classification model performance

Table 4. *Performance comparison between the proposed method and baseline models.*

Model	Accuracy on testing data	Training time (s/epoch)
SVM (OvO)	84.2%	12.5
MLP	86.5%	8.0
Random Forest	88.7%	9.2

The result provides an overview of the test performance and training time for popular classification models such as SVM, Random Forest, and MLP. Among these, MLP stands out due to its good integration with embedded devices, with a training time of only about 8 seconds per epoch. SVM achieves a test accuracy of 84.2%, while Random Forest performs better with 88.7%.

The results of the three methods SVM (OvO), MLP (1 hidden layer), and Random Forest show similar accuracy. However, MLP (1 hidden layer) is the optimal choice for mobile devices with limited resources due to its low computational complexity, minimal memory requirements, and fast deployment capability. SVM (OvO) and Random Forest may become cumbersome as the data size and number of classes increase.

4. CONCLUSIONS

This study presents a bearing fault detection method based on multi-domain feature extraction, achieving an accuracy of 86.5% on the NASA IMS dataset. The proposed approach offers low latency and minimal resource requirements, making it suitable for deployment in embedded and resource-constrained environments. Among the evaluated classifiers, including SVM (OvO), MLP with a single hidden layer, and Random Forest, the MLP is identified as the most practical choice

for mobile applications due to its simple architecture, low computational complexity, and reduced memory consumption. While the method yields promising results under controlled experimental conditions, it is acknowledged that real-world scenarios may involve high noise levels and dynamic system behaviors that could affect model performance. To address these challenges, future work will focus on integrating Kalman filtering to enhance noise robustness, and incorporating advanced signal processing techniques such as wavelet-based denoising or empirical mode decomposition to improve feature quality. In addition, adaptive learning strategies and domain adaptation methods will be considered to maintain model performance in varying operating environments. The system will also be extended to support multi-sensor fusion, and uncertainty-aware modeling approaches may be explored to further improve diagnostic reliability. These enhancements aim to increase the robustness and generalizability of the proposed method, thereby facilitating its broader application in practical industrial settings.

REFERENCES

[1]. Hart, E., Clarke, B., Nicholas, G., Kazemi Amiri, A., Stirling, J., Carroll, J., Dwyer-Joyce, R., McDonald, A., and Long, H. *"A review of wind turbine main bearings: design, operation, modelling, damage mechanisms and fault detection"*, Wind Energ. Sci., 5, 105–124, (2020), https://doi.org/10.5194/wes-5-105-2020.

[2]. Deepika, C., Taj, K., and Bedar, P. *"Automation in production systems: enhancing efficiency and reducing costs in mechanical engineering"*, Nanotechnology Perceptions, (2024), https://doi.org/10.62441/nano-ntp.vi.3895.

[3]. Zimroz, R., Bartelmus, W., Barszcz, T., and Urbanek, J. *"Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings"*, Mechanical Systems and Signal Processing, 46, 16–27, (2014), https://doi.org/10.1016/j.ymssp.2013.12.020.

[4]. Purarjomandlangrudi, A., Nourbakhsh, G., Ghaemmaghami, H., and Tan, A. *"Application of anomaly technique in wind turbine bearing fault detection"*, 1984–1988, (2014).

[5]. Gu, J., and Huang, M. *"Fault diagnosis method for bearing of high-speed train based on multitask deep learning"*, Shock and Vibration, 1–8, (2020), https://doi.org/10.1155/2020/8840040.

[6]. Buchaiah, S., and Shakya, P. *"Bearing fault diagnosis and prognosis using data fusion based feature extraction and feature selection"*, Measurement, 188, 110506, (2021), https://doi.org/10.1016/j.measurement.2021.110506.

[7]. El Laithy, M., Wang, L., Harvey, T. J., Vierneusel, B., Correns, M., and Blass, T. *"Further understanding of rolling contact fatigue in rolling element bearings – A review"*, Tribology International, 140, 105849, (2019), https://doi.org/10.1016/j.triboint.2019.105849.

[8]. Saini, M. K., and Aggarwal, A. *"Detection and diagnosis of induction motor bearing faults using multi wavelet transform and naive Bayes classifier"*, Int. Trans. Electr. Energy Syst., 28, (2018), https://doi.org/10.1002/etep.2526.

[9]. Fanning, P. *"High-quality bearings reduce downtime and save cost"*, Eureka, (2025), https://www.eurekamagazine.co.uk/content/technology/high-quality-bearings-reduce-downtime-and-save-cost/.

[10]. Saruhan, H., Sarıdemir, S., Çiçek, A., and Uygur, I. *"Vibration analysis of rolling element bearings defects"*, Journal of Applied Research and Technology, 12, 384–395, (2014), https://doi.org/10.1016/S1665-6423(14)71620-7.

[11]. Nizwan, C. K. E., Ong, S. A., Yusof, M. F. M., and Baharom, M. Z. *"A wavelet decomposition analysis of vibration signal for bearing fault detection"*, IOP Conference Series: Materials Science and Engineering, 50, 012026, (2013), https://doi.org/10.1088/1757-899X/50/1/012026.

[12]. Jakubek, B., Grochalski, K., Rukat, W., and Sokol, H. *"Thermovision measurements of rolling bearings"*, Measurement, 189, 110512, (2021), https://doi.org/10.1016/j.measurement.2021.110512.

[13]. Zhou, S., Lin, L., Chen, C., Pan, W., and Lou, X. *"Application of convolutional neural network in motor bearing fault diagnosis"*, Computational Intelligence and Neuroscience, 2022, 1–11, (2022), https://doi.org/10.1155/2022/9231305.

[14]. Wang, Y., and Cheng, L. *"A combination of residual and long-short-term memory network for bearing*

fault diagnosis based on time-series model analysis", Measurement Science and Technology, 32, (2020), https://doi.org/10.1088/1361-6501/abaa1e.

[15]. Agarap, A. F. *"An architecture combining convolutional neural network (CNN) and support vector machine (SVM) for image classification"*, arXiv, (2017), https://arxiv.org/abs/1712.03541.

[16]. Pham, V.-N., Do, Q.-H., and Tran, L. D.-A. *"Using artificial intelligence (AI) for monitoring and diagnosing electric motor faults based on vibration signals"*, 2024 International Conference on Information Networking (ICOIN), (2024), https://doi.org/10.1109/ICOIN.2024.104567.

[17]. Wang, X., Meng, R., Wang, G., Liu, X., Liu, X., and Lu, D. *"The research on fault diagnosis of rolling bearing based on current signal CNN-SVM"*, Measurement Science and Technology, 34, (2023), https://doi.org/10.1088/1361-6501/acefed.

[18]. Han, T., Zhang, L., Yin, Z., and Tan, A. *"Rolling bearing fault diagnosis with combined convolutional neural networks and support vector machine"*, Measurement, 177, 109022, (2021), https://doi.org/10.1016/j.measurement.2021.109022.

[19]. Qiu, H., Lee, J., Lin, J., and Yu, G. *"Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics"*, Journal of Sound and Vibration, 289, 1066–1090, (2006), https://doi.org/10.1016/j.jsv.2005.03.007.

[20]. Teler, K., Skowron, M., and Orłowska-Kowalska, T. *"Implementation of MLP-based classifier of current sensor faults in vector-controlled induction motor drive"*, IEEE Transactions on Industrial Informatics, 20, 4, 5702–5713, (2024), https://doi.org/10.1109/TII.2023.3336348.

TÓM TẮT

Phát hiện sớm lỗi vòng bi dựa trên việc trích xuất đặc trưng đa miền sử dụng bộ phân loại MLP trên bộ dữ liệu NASA IMS

Sự hao mòn và suy giảm hiệu suất của vòng bi trong máy móc công nghiệp dẫn đến chi phí bảo trì tăng cao và thời gian ngừng hoạt động không mong muốn. Nghiên cứu này đề xuất một phương pháp mới kết hợp trích xuất đặc trưng thống kê đa miền (thời gian và tần số) với thuật toán phân cụm K-means thích ứng để nâng cao độ chính xác phát hiện lỗi. Mô hình Multi-Layer Perceptron (MLP) được huấn luyện trên tập dữ liệu NASA IMS Bearing, đạt độ chính xác 86.5% trong việc phân loại 5 giai đoạn suy thoái của vòng bi. Kết quả thực nghiệm cho thấy phương pháp đề xuất vượt trội so với SVM và Random Forest trong điều kiện dữ liệu hạn chế, đồng thời có thể triển khai trên thiết bị kiểm tra di động, nhỏ gọn với tài nguyên hạn chế. Nghiên cứu này cung cấp giải pháp hiệu quả cho bảo trì dự đoán, giúp giảm thiểu chi phí vận hành trong công nghiệp.

Từ khoá: Dự đoán lỗi vòng bi; NASA IMS; Đặc trưng đa miền; Mô hình MLP; Bảo trì dự đoán.