2021 International Symposium on VLSI Design, Automation and Test (VLSI-DAT) | 978-1-6654-1915-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/VLSI-DAT52063.2021.9427332

Embedded Bearing Fault Detection Platform Design for the Drivetrain System in the Future Industry 4.0 Era

Kun-Chih (Jimmy) Chen, Jing-Wen Liang, Yueh-Chi Yang, Hsiang-Ling Tai, Jo-Chiao Ku, and Jui-Cheng Wang
Department of Computer Science and Engineering, National Sun Yat-sen University, Taiwan

ABSTRACT

As industry 4.0 becomes more and more prevalent, predictive maintenance (PdM) systems are gradually being valued by the industry. To reduce the risk of overall system failure which is caused by faulty components. The real-time monitoring sensors must collect the sensing data continuously, and the subsequent computing system can determine the categories of faulty components in the target machinery. The drivetrain system is critical because it drives the operation of the overall motor system in factories. Bearings are one of the important components in the drivetrain system. A healthy bearing can reduce friction and make the shaft rod in the drivetrain system work smoothly. Accelerometers are usually used to collect the vibration signals for further fault detection in machinery. However, accelerometers are expensive and consumables that they need to be replaced frequently. Moreover, using accelerometers to collect bearings vibration signals is also restricted by the operating temperature, humidity, ground loops, and the rest. This paper used a three-axis vibration sensor to detect the accelerated vibration signals of faulty bearings. Then, the Hilbert transform is employed to determine the spectral envelope of a waveform, which leverages the bearing fault detection with the random forest algorithm. By integrating the vibration sensor module and the fault detection computing module, the proposed embedded fault detection platform is proper for the goal of high-accuracy and low-cost bearing fault detection for the drivetrain system.

I. INTRODUCTION

The drivetrain system is one of the most important mechanisms in the motor system in factories [1]. As usual, frequent collisions and vibrations will cause the components in the drivetrain system (e.g., bearings, gears, shaft rods, etc.) wear and tear, which leads to overall operation failures and incorrect output results. Moreover, the abnormal behavior of the damaged components will accelerate the abrasion of the surrounding equipment, which affects the quality of factory productivity. Hence, real-time fault monitoring and classification system are necessary to avoid extra maintenance costs [2].

Bearings play a critical role in various drivetrain machinery. The literature [3] shows that approximately 30% of machinery failures are caused by bearing. Bearings can not only support the rotating body to maintain the shaft rod at the center but also reduce the friction between the shaft rod and the fasteners to increase the rotating efficiency. The bearings are mainly composed of multiple rolling elements which are inlaid between the smooth outer ring (or race) and inner ring. The rolling elements officially called balls are separated and fixed by cage. The rolling elements will carry the loads' weight and drive the rotating power. The cages will fix the rotating elements in the appropriate position so that the rolling elements will not contact each other.

Traditionally, factories will use accelerometers (usually are signal wires) to obtain the vibration signals of the target machinery and collect the vibration signals with an additional data acquisition device (e.g., ADLINK USB-2405). Afterward, a certain signal processing method is adopted to extract the information of vibration signals by an analysis software tool. Nevertheless, the accelerometers are expensive, and they are often damaged due to improper use. Moreover, environmental factors such as temperature, ground loops, may also cause great damage to the accelerometers. Besides, it is complicated to mount and disassemble the accelerometers that using accelerometers has lower maneuverability. Last but not least, it is assisted by software analysis that the purpose of real-time monitoring cannot be achieved. In accordance with the above reasons, this paper employs the three-axis vibration sensor ADcmXL3021, provided by Analog Device Corporation, to collect the faulty bearing vibration signals. Therefore, using the three-axis vibration sensor not only reduces the overhead but also achieves high maneuverability.

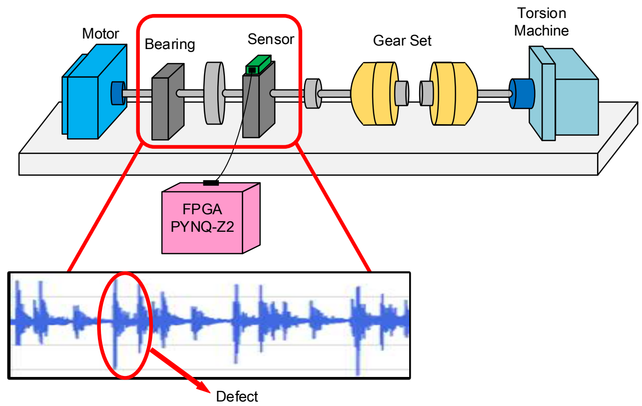


Fig. 1 The overview of the proposed embedded bearing fault detection

Although the collected faulty bearing vibration signals can be analyzed by the sophisticated experts efficiently, the labor cost is not affordable to those small and medium enterprises. Therefore, the technologies of artificial intelligence (AI) receive much attention in recent years to simplify the process [4]. Without any prior knowledge, AI technology can automatically classify the bearing faults based on the extracted feature information. Due to the increasing computer processing capabilities in recent years, the neural network in AI technologies has gradually become the mainstream. However, because of the low interpretability, the operations in the neural network like a black box. In other words, the experts in the factories are difficult to understand the intermediate process and explain certain results which are calculated by the involved neural network model. Therefore, it is hard to explain the reliability of the system to customers. To solve the problem, this paper first adopts the Hilbert transform, which is often used to detect the bearing faults [5], to calculate the spectrum envelope of the collected faulty bearing vibration signals. Afterward, using the random forest algorithm [6] in ensemble learning to classify the bearing fault. The interpretability of the random forest model can not only verify the correctness of the model but also improve the model by focusing on the important features.

To accelerate the bearing faults detection, this paper utilizes Xilinx PYNQ-Z2, which is a Python Productivity for Zynq FPGA, to implement Hilbert transform and random forest algorithm. Moreover, cooperating with a three-axis vibration sensor ADcmXL3021 to propose a high-performance and low-cost lightweight fault detection system based on an embedded system. Moreover, to avoid excessive data access time which results in too large delays to achieve real-time monitoring, this paper adopts the direct memory access (DMA) technique as the communication interface for hardware access additionally on the involved PYNQ-Z2 FPGA. The overview of the proposed embedded fault detection platform is shown in Fig. 1, and the contributions of this paper are summarized below

This work was supported by the Ministry of Science and Technology, TAIWAN, under Grant MOST 109-2221-E-110 -062 and MOST 109-2218-E-110 -009

978-1-6654-1915-4/21/$31.00 ©2021 IEEE

1) Using the envelope spectrum features for an explainable prediction model to detect the common bearing faults.
2) Integrating PYNQ-Z2 and ADcmXL3021 to propose a high-performance and lightweight fault detection system.
3) Using direct memory access (DMA) technique to optimize the computing performance of the system.

The rest of this paper is organized as follows. In Section II, some related works are investigated. The proposed embedded bearing fault detection platform will be introduced in Section III. Afterward, the validation results are presented and analyzed in Section IV. At last, we concluded this paper in Section V.

II. RELATED WORKS

A. The spectral envelope analysis for bearing fault diagnosis [7]

Given different bearing specifications, the relative motion analysis can be utilized to acquire the motion track of the faulty bearing in the time domain. Since the bearing vibration signal is continuous and periodic, the vibration signal usually includes high-frequency attenuation vibration frequency and fault feature information of low-frequency envelope signal. Therefore, the motion track of the faulty bearing is not usually an ideal coincide wave with the ideal value in practical. As usual, faulty information can be attached to the high-frequency vibration signal. The high-frequency vibration signal is characteristic of having frequency modulation and amplitude modulation. In [7], the authors used Hilbert transformation to convert the original signal into the envelope spectrum for analysis. The envelope demodulation method is performed to calculate the maximum difference in the power spectral automatically. Moreover, it improves the signal-to-noise ratio in the analysis and facilitates subsequent fault diagnosis research. However, this approach uses accelerometers and an additional data acquisition device to collect the vibration signal and process the signals through a software tool. Hence, the computing latency is very long, and the involved space-consuming accelerometers restrict the scalability of this approach.

B. 1-D CNN for bearing fault classification [8]

In [8], the authors proposed an adaptive 1D convolutional neural network to implement a generic real-time induction bearing fault diagnosis system. The 1D convolutional compact architecture makes the overall system more suitable for real-time fault detection and reduces the hardware overhead. Moreover, feature extraction and feature selection can be done without any pre-determined signal transformation, such as Fast Fourier Transform and Discrete Wavelet Transform. Although features can be extracted and selected by using neural network operations, the internal operations cannot be explained, which lacks interpretability with the expert models. Moreover, the involved complicated neural network model is not proper to implement on an embedded hardware platform.

III. EMBEDDED BEARING FAULT DETECTION PLATFORM DESIGN

A. Spectral envelope analysis by using Hilbert transformation

The envelope spectrum is the most commonly used detection technique to detect faulty bearing vibration signals. In factories, the vibrations, which are caused by low-rotation speed bearings, are mostly concentrated in the low-frequency band. The feature intensity of vibrations is usually slight, and it is easily covered by other obvious operating signals. It is difficult to distinguish the features of faulty bearing vibration and interfered vibration in the frequency spectrum. However, the faulty bearing vibration signals will cause resonant frequencies at the high-frequency band. Therefore, the low-frequency faulty bearing vibration signals are regarded as the side frequencies of the high-frequency carrier wave. In other words, the faulty bearing vibration signals are modulated signals. Therefore, it is necessary to demodulate the modulated signals before analyzing the modulated signals information.

In a practical way, the demodulation analysis is usually performed on the envelope spectrum of faulty bearing modulated signals. Envelope spectrum analysis first obtains the envelope of the target signals. Afterward, by connecting the peaks of the target signals, we can further perform Fourier transform on the signal to obtain the frequency of the faulty component behaviors.

To calculate the envelope signal, we first transform the modulated signal by Hilbert transformation, which is defined as

$$\mathrm{Y(t)} = \mathrm{X(t)} * \frac{1}{\pi t} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{X(\tau)}{t-\tau}\, d\tau\,, \tag{1}$$

where *X(t)* is the original modulated signals, and *Y(t)* is the signal after the Hilbert transformation by convolution. Afterward, we generate the corresponding envelope signal by using

$$Z(t) = X(t) + iY(t). \tag{2}$$

Obviously, the real part is the original modulated signal, and the imaginary part is the Hilbert transformation result of the original modulated signal. To retain all the information in the envelope signal, we further calculate the magnitude of the envelope signal by using

$$S(t) = \left|Z(t)\right| = \sqrt{X^2(t) + \mathrm{Y}^2(t)}\,. \tag{3}$$

In this way, the useful information about the bearing faults can be found on the envelope spectrum, which is helpful for further fault classification.

B. Classify the bearing fault by using Random forest algorithm

As mentioned before, the interpretability of the conventional DNN models is not enough. Hence, we adopt the random forest model in this work. Random forest is composed of multiple decision trees. A series of True or False questions will be asked to each single decision tree during the training phase. This paper used Gini Impurity to measure the property of randomly chosen data, which was incorrectly labeled, and it can be computed by summing the probability. Therefore, the value of Gini Impurity is between 0 and 1. The higher value represents the more different types of data in the dataset. The goal during the training phase is to find a kind of data segmentations to maximize the Gini Impurity.

Random forest uses ensemble learning to gather the trees. The algorithm will randomly select different data and features from the training set during the training phase to generate each tree. The results of each tree are integrated through averaging or voting to avoid the overfitting problem. The structure of each tree is independent. In other words, changing the structure of the tree, such as pruning, does not affect other trees.

In order to improve the classification accuracy of the overall system, data preprocessing, feature extraction, and feature selection are usually performed before the random forest operations. This paper pre-determines whether each node of the involved decision tree, which is constructed by decision condition, needs to be retained before establishing each decision tree. If the node of the involved decision tree does not meet the decision criteria, the decision tree will pre-pruning to avoid overfitting. After establishing the completed decision trees, the random forest algorithm will determine whether more pruning processes are needed. Besides, this paper uses the permutation feature importance method [9] to decide the important features of the faulty bearing signal in the data. The permutation feature importance is defined as the accuracy descending degree in a model metric under randomly shuffled features. The descending degree in the model metric indicated the correlation between the established model and the selected features. In other words, the target feature is critical while the classification accuracy drops significantly after removing it. After

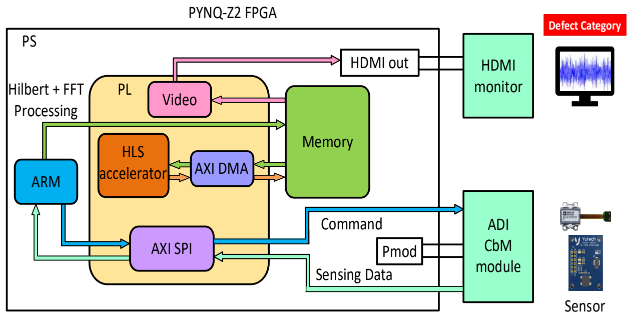


Fig. 2 The data processing flow from collecting the sensing data to display the classification results on PYNQ-Z2 FPGA.

using random forest permutation feature importance, the features of the faulty bearing signal can be extracted and selected to improve the system classification accuracy.

C. Design and implementation of random forest on FPGA

In order to accelerate the random forest operation, we integrate the involved vibration sensor and PYNQ-Z2 FPGA board to become an embedded bearing fault detection platform. The data processing flow is shown in Fig. 2. Obviously, the PYNQ-Z2 FPGA board includes two different modules, which are called Processing System (PS) unit and Program Logic (PL) unit, respectively.

The PS unit is used to send a command and request the sensing signal from the three-axis vibration sensor ADcmXL3021 through the AXI Serial Peripheral Interface (SPI) communication protocol. Afterward, the bearing fault signals, which are collected by the sensor, are transmitted back to the memory through the AXI SPI communication protocol and be transformed into the envelope spectrum by Hilbert transformation. This paper uses the AXI Direct Memory Access (DMA) technique to move the data, which is accessed in the memory, to the PL unit to classify the bearing fault. By using the DMA technique, we can transfer the batch of the collected signal data from the memory to the PL unit at one time for further classification. In this way, we can not only avoid the delay of memory access but also achieve the purpose of real-time fault detection. After receiving the sensing data, the PS unit further preprocesses the sensing data to capture the information in the envelope spectrum domain of the sensing data by using Hilbert transformation. On the other hand, the involved random forest model is implemented and accelerated its classification operation on the PL unit of the PYNQ-Z2 FPGA board. Finally, the classification results are displayed on the monitor through the output terminal on the PYNQ-Z2 FPGA board.

IV. EXPERIMENTAL RESULTS

A. Environmental setup of the mini motor drivetrain system

To verify the proposed method, in this work, we build a mini motor drivetrain system to collect the practical faulty bearing vibration signals, as shown in Fig. 3. The conventional way collects the vibration signals through the accelerometers and an additional data acquisition device (i.e., ADLINK USB-2405 is used in this experiment). On the contrary, the proposed method gathers the vibration signal through the sensor (i.e., ADcmXL3021 in this experiment) on the PYNQ-Z2 FPGA. At last, the screens are used to display the results of traditional and proposed methods, respectively.

Based on the structure of the bearing, the target bearing fault categories are 1) healthy, 2) inner ring defect, 3) inner ring break, 4) outer ring defect, 5) outer ring break, 6) cage break, and 7) ball break, as shown in Fig. 4. With the identical bearing specification in Table 1, the goal of the following experiments is to detect the different bearing faults in Fig. 4. The frequency of the rotating speed of the motor is

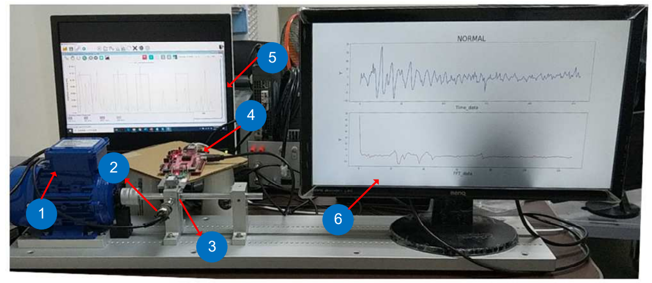


Fig. 3 Experimental architecture. 1: motor; 2: accelerometers; 3: bearing; 4: proposed embedded fault detection platform; 5: results from traditional method; 6: results from proposed method.

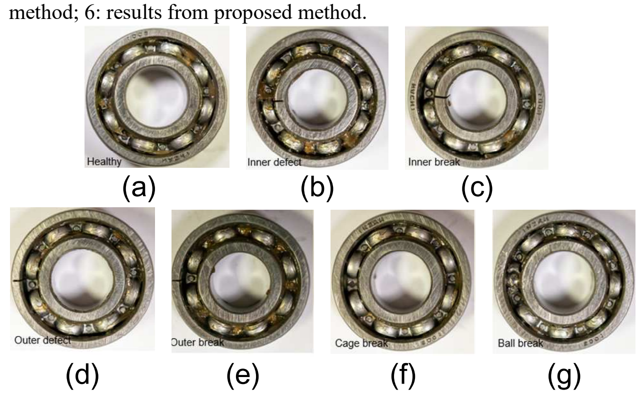


Fig. 4 The common bearing fault categories: (a) Healthy, (b) Inner ring defect, (c) Inner ring break, (d) Outer ring defect, (e) Outer ring break, (f) Cage break, and (g) Ball break.

Table 1 The detailed specification of the bearing in the experiment

Bearing Type	**Rotation Speed**	**Mass**
12mm ball bearing	1,000 rpm	23 g
Bearing Diameter	**Roller Diameter**	**Number of Rollers**
20.4978 mm	4.7498 mm	8

configured to 1,000 rpm as a design example. According to the number of data, we collect 360 data during each kind of bearing operation. Therefore, we have 2,520 data for this experiment, and each data include 10,240 points to represent the vibration signal. Furthermore, we randomly select 75% of the data (i.e., 1,890 data) as the training dataset and the remaining 25% data (i.e., 630 data) are used as the testing dataset.

B. Verification result of the proposed embedded bearing fault detection platform

As mentioned before, the PS unit is used to preprocess the receiving vibration signal, which is a time-domain signal, as shown in Fig. 5(a). Since time-domain waveforms mostly present complicated signals, the analysis of vibration problems is quite difficult. Therefore, it is transformed by Fast Fourier Transform to the frequency domain as Fig. 5(b). Furthermore, the Hilbert transformation is performed to further transform the frequency-domain signal to the envelope spectrum, as shown in Fig. 5(c), to classify the bearing fault. To verify the efficiency of the classification, we adopt the confusion matrix in this work, as shown in Fig. 6. In the confusion matrix, label 0 to 6 represents healthy, outer ring break, outer ring defect, inner ring break, inner ring defect, ball break, and cage break, respectively. In the confusion matrix, it can be observed that the accuracy of most bearing fault can exceed 95%. However, the accuracy of the outer ring break and ball break is slightly lower, especially for ball break. The reason is that the ball break fault is usually easily recognized under a high

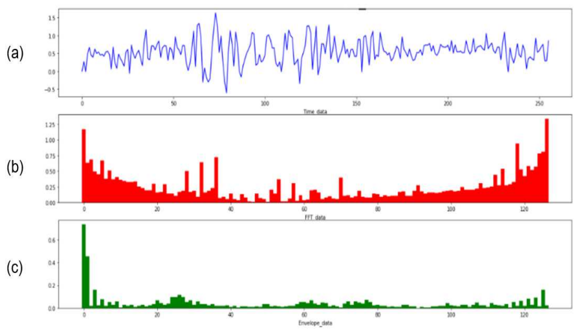


Fig. 5 The waveform data in (a) the original time domain, (b) the frequency domain by using Fast Fourier Transform, and (c) the Envelope spectrum after Hilbert transformation.

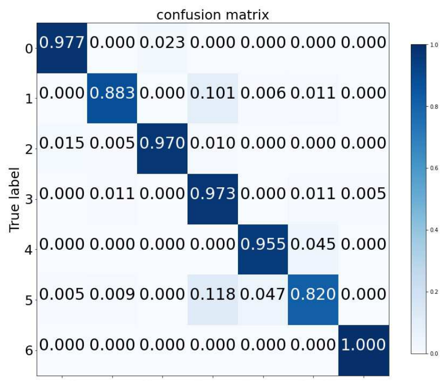
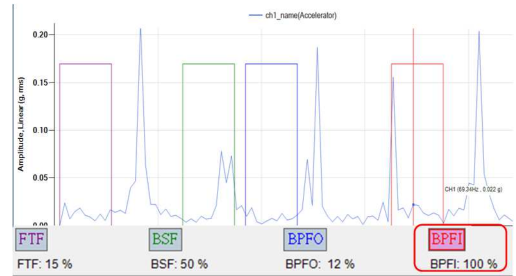
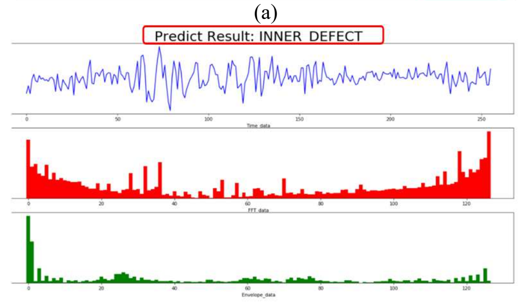


Fig. 6 Confusion matrix of bearing fault by random forest. The labels are represented as 0: normal 1: outer rings break 2: outer rings defect 3: inner rings break 4: inner rings defect 5: ball break 6: cage break.



(a)



(b)

Fig. 7 The classification results by using (a) the traditional accelerometers and (b) the proposed method are identical.

Table 2 Comparison of traditional method with proposed method

	Proposed method (ADcmXL3021 + PYNQ-Z2)	**Traditional method (Accelerometers + ADLINK USB-2405)**
Weight	150 g	1,321 g
Volume	Light	Massive
flexibility	High	Low
Power consumption	101 mW	2,632 mW

rotating speed. However, the rotating speed is kept at 1,000 rpm in this experiment, which represents a low rotating speed in the actual factory. Therefore, the classification of the ball-break fault is not well compared with the classification of the other kinds of bearing faults.

To validate the proposed embedded bearing fault detection platform, it is necessary to compare the results of classification with the conventional method [10], which captures the vibration signal through accelerators and a data acquisition device. The signal processing techniques preprocess the captured vibration data by using Hilbert transform and determine the type of the bearing fault by comparing the calculated spectrum with the pre-defined spectrums of the specific bearing faulty type. Fig. 7 shows a case of these experiments. Obviously, the most possible bearing faulty type of this case is BPFI (Ball Pass Frequency Inner Race), as shown in Fig. 7(a). Meanwhile, the computing result of the proposed method is also classified the current bearing fault into inner ring defect (i.e., INNER_DEFECT in Fig. 7(b)). Therefore, the validity of the proposed embedded bearing fault detection platform is ensured. Compared with the conventional method, the proposed approach has the advantages of high flexibility and low volume cost during the installation, as shown in Table 2. The reason is that the involved additional data acquisition device and accelerometers are space-consuming, which further leads to larger power consumption and weight.

V. Conclusion

This paper presents a complete processing flow to perform bearing fault detection on an embedded platform. The experiments show that the accuracy of most bearing faults can exceed 95%. Moreover, the proposed method is implemented on the PYNQ-Z2 FPGA to accelerate the overall system operation. By using the DMA technique, the latency of data movement from the memory to the computing unit can be reduced significantly. Due to the advantages of the high flexibility and low volume cost, the proposed method can reduce the equipment overhead in the traditional factories.

Reference

[1] R. Liu et al., "Artificial Intelligence for Fault Diagnosis of Rotating Machinery: A Review," *Mech. Syst. Signal Process.*, vol. 108, pp. 33-47, Aug. 2018.

[2] M. Zhao et al., "A Data-Driven Monitoring Scheme for Rotating Machinery Via Self-Comparison Approach," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2435-2445, April 2019.

[3] X. Yang, "Research on Vibration State Monitoring and Fault Diagnosis System of Chemical Rotating Machinery," *Chemical Engineering Transactions*, vol. 66, pp. 745-750, July 2018.

[4] Y. Li et al., "Entropy Based Fault Classification Using the Case Western Reserve University Data: A Benchmark Study," *IEEE Transactions on Reliability*, vol. 69, no. 2, pp. 754-767, June 2020.

[5] J. Wang et al., "Fractional Envelope Analysis for Rolling Element Bearing Weak Fault Feature Extraction," *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 2, pp. 353-360, April 2017.

[6] G. Louppe et al., "Understanding Variable Importances in Forests of Randomized Trees", *Advances in Neural Information Processing Systems 26*, 2013*,* pp. 431-439.

[7] N. Wang et al., "Bearing Fault Diagnosis Method Based on Hilbert Envelope Demodulation Analysis," *Transactions of China Electrotechnical Society*, vol. 436, pp. 012009, August 2018.

[8] L. Eren et al., "A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1D CNN Classifier," *J. Signal Process. Syst.*, vol. 91, no. 2, pp. 179-189, Feb. 2019.

[9] L. Breiman et al., "Random Forest," *Machine Learning*, vol. 45, no. 12, pp.5-32, 2001.

[10] A. Rai and S. H. Upadhyay, "A Review on Signal Processing Techniques Utilized in The Fault Diagnosis of Rolling Element Bearings," Tribology Int., vol. 96, pp. 289-306, Apr. 2016.