Using Artificial Intelligence (AI) for Monitoring and Diagnosing Electric Motor Faults Based on Vibration Signals

Van – Nam Pham*
Faculty of Engineering Electrical & Automation
Hanoi University of Industry
Hanoi, Vietnam
nampv@haui.edu.vn

Quang - Huy Do Ba
Faculty of Engineering Electrical & Automation
Hanoi University of Industry
Hanoi, Vietnam
dohuy9379@gmail.com

Duc-Anh Tran Le
School of Electrical & Electronic Engineering
Hanoi University of Science and Technology
Hanoi, Vietnam
ducanhtran11082002@gmail.com

Quang - Minh Nguyen
School of Information and Communications Technology
Hanoi University of Science and Technology
Hanoi, Vietnam
qminhnguyen2211@gmail.com

Thanh - Lam Bui
Foreign Language Specialized School
Vietnam National University
Hanoi, Vietnam
thanhlam6900@gmail.com

Alberto Ernesto Coboi
Faculty of Engineering Electrical & Automation
Hanoi University of Industry
Hanoi, Vietnam
AlbertoCoboi@haui.edu.vn

***Abstract*— Detecting bearing faults in electric motors is highly crucial for improving production efficiency and reducing accidents in complex mechanical systems, which poses significant challenges for current fault diagnosis technology. This paper investigates and applies Artificial Intelligence (AI) to enhance the monitoring and diagnosis process of electric motor faults based on vibration signals. The research aims to construct a model for collecting sample data from motors with three common types of bearing faults and utilizes the Resnet-50 network to assess the accuracy of monitoring and diagnosing faults. The study conducts a vibration signal analysis to identify potential indicators of faults in electric motors. The survey results presented in the paper demonstrate the accuracy of using the Resnet-50 network in monitoring and diagnosing electric motor faults. The paper also provides essential insights into the performance of AI networks and their practical applicability in the field of industrial equipment maintenance and management.**

Keywords— Bearing fault diagnosis, industrial equipment maintenance, Resnet.

I. Introduction

In most modern factory settings, the majority of engine monitoring and alert systems are installed separately, focusing on only a limited set of electrical parameters. They often lack substantial data collection for in-depth analysis. Regular online checks and continuous monitoring of engine parameter analyses are infrequently conducted, and there is a deficiency in automated engine maintenance planning capabilities. This deficiency results in the inability to detect unforeseen breakdowns and engine malfunctions, which, when not adequately monitored, can lead to significant machinery damage, necessitating costly repairs or even engine replacements. Conversely, when conditions are continuously monitored, potential issues can be identified early, leading to simpler, quicker, and more cost-effective repairs. Consequently, it is imperative and practical to conduct research and develop systems for examining, monitoring, and predicting electric motor faults using IoT (Internet of Things technology)[18] within factory environments. Furthermore, data encompassing electrical, mechanical, and thermal parameters, among others, will be gathered and integrated with intelligent software for diagnostic purposes, leveraging artificial intelligence (AI), a hallmark of the fourth industrial revolution

The general operating principle of these systems includes the following steps:

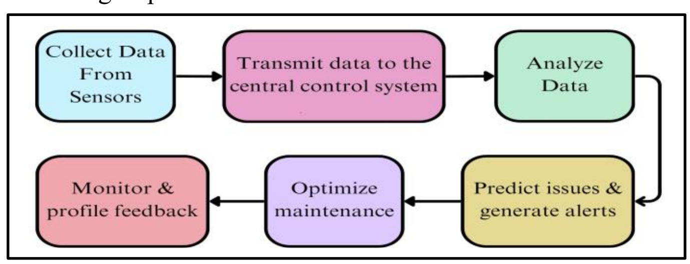

Fig. 1. General operating principle of the electric motor fault alert system.

- Collect data from sensors: Gather information from various sensors, including temperature, vibration, pressure, and current sensors affixed to electrical machinery. These sensors continuously collect data on the machines' operational conditions.
- Transmit data to the central control system: Transmit the data from these sensors to the central control system through network connections, such as the internet or local area networks. The data is directed to servers or central hubs for analysis and processing.
- Analyze data: Within the central control system, apply machine learning algorithms, artificial intelligence, and data analysis techniques to scrutinize the sensor data.

The primary objective is to identify trends, patterns, and deviations that may signify potential incidents or faults.

- Predict issues and generate alerts: Based on data analysis, the system predicts potential incidents or undesirable situations in the future. When a potential incident is identified, the system generates alerts and notifies managers or technicians for timely intervention.
- Optimize maintenance: Provide in-depth insights regarding anticipated issues and the machinery's current status. This information aids in fine-tuning maintenance schedules and practices, resulting in time and resource savings.
- Monitor and provide feedback: Following the issuance of alerts, the system continues to monitor the machinery's condition and assess the outcomes of interventions. This ongoing evaluation helps gauge the solution's effectiveness and make necessary adjustments.

The remaining part of this article is organized as follows: Section 2 discusses prior related works on this issue, presenting several relevant studies to support this research. Section 3 explains experimental data collection and the development of anomaly detection algorithms. Section 4 presents the main results of algorithm development and a performance comparison study. Finally, Section 5 concludes the article and suggests feasible directions for future research.

II. Related Work

Recent literature surveys in [2-4] have highlighted the challenges and opportunities for developing robust predictive maintenance techniques based on machine learning, particularly for rotating equipment such as bearings, motors, gearboxes, and pumps. Many challenges and opportunities still await exploration in this field to enhance the accuracy of machine learning models and increase the flexibility of proposed predictive maintenance methods in the future.

Tuan A. Z. Rahman et al. [5] proposed an intelligent anomaly detection method for electric motors based on vibration signals combined with AI algorithms. They developed an unsupervised learning model for two different types of motors within the same category: a new experimental motor and an old industrial motor. The model's performance in anomaly detection for both types of motors was extensively studied, and the results showed that it had the highest anomaly detection capability for standardized motor conditions using mapped features. However, they currently utilized only data from normal motor conditions due to a lack of information about fault conditions.

M. Masood Tahir et al. [6] presented a solution using vibration signal features like RMS, Mean, Variance, skewness, kurtosis, median, range, etc., for model training, similar to many other studies. During the data preprocessing phase, this paper introduced an approach called Median-based Outlier Detection (MOD) to detect outliers (data samples in which features are affected by external factors, not due to faults, and exclude them from the training process to improve model performance). However, this paper did not address the classification of similar fault types with different fault sizes.

In recent years, there has been a growing focus on utilizing deep learning (DL) methods [7-11] to address the mentioned issues, with many DL approaches applied to bearing fault diagnosis. While Convolutional Neural Networks (CNN) are classic DL structures for image classification [7,8,19], various DL models [9,10,11,20] have found wide applications in fault detection tasks. However, deeper DL models often face the problem of gradient vanishing, requiring performance sacrifices during training. To overcome these challenges, Residual Network (ResNet) was introduced [12]. ResNet employs residual connections, enabling the learning of residual functions from input rather than complex mappings from input to output. This innovation has significantly improved the performance of deep neural networks, leading to the introduction of various ResNet-based models [13,14]. In this research paper, we propose the use of ResNet-50 to train with a dataset containing three common bearing fault types constructed from our experimental model.

III. Methodology

A. Data Collection

Conducting real-time, automated diagnostics on engines operating in practical conditions presents significant limitations due to the complexity of the problem. Vibration signals and machinery noise during production often exhibit high levels of noise and variability influenced by environmental conditions. Careful selection of appropriate feature values for signal recognition can enhance the efficiency of the recognition model.

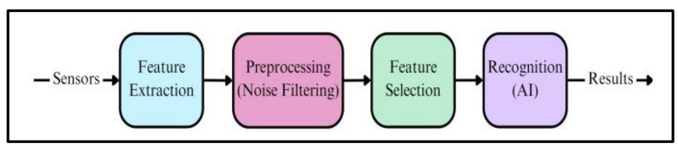

Fig. 2. The general model of the engine fault detection system

In this paper, we have developed a motor model to address the problem of identification and training a deep learning network. Here is a basic description of this model:

- Mechanical Component: A 3-phase, 2HP motor with 2 bearings and an electromagnetic brake assembly to simulate load and torque effects.
- Control Component: A 2.5KW inverter controls the motor's rotational speed, ranging from 1500 to 1750 RPM.
- Vibration Measurement Equipment: Vibration data is collected, stored on an SD card, and transmitted using WIFI communication standards for AI model training, with sampling frequencies of 6 KHz and 12 KHz.

The identification states are divided into four categories:

- N (normal state): The motor operates normally without any damage.

- MI (Misalignment damaged): Misalignment occurs when the shaft is not aligned properly or not in the correct plane. Misalignment can generate uneven torque and cause vibration.
- IR (inner raceway damage): The bearing's inner raceway is damaged.
- OR (outer raceway damage): The bearing's outer raceway is damaged.

The shaft bearings had inner and outer raceway damage at a 2 mm diameter point. Vibration signals were recorded using an accelerometer at the specified location, as shown in Figure 3. The signals were recorded for approximately 2 minutes at a 12 kHz sampling frequency.

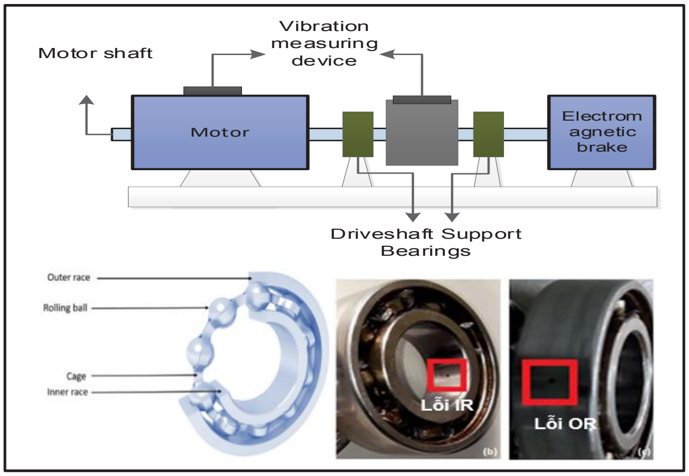

Fig. 3. Test model, device placement, and bearing faults location.

The measuring device is attached to the motor casing. All data files are in MATLAB Data (*.mat) format, collected at rates of 6000 and 12000 samples per second. Speed and torque data are obtained using a torque encoder and recorded manually by a motor controller.

The collected sample dataset, detailed in the table below:

TABLE I. DATA COLLECTED FROM THE EXPERIMENTAL MODEL

Type of error	*Sampling frequency (Hz)*	*Number of samples*	*Time per sample (minutes)*
N	12000	4	2
MI	12000	4	2
IR	12000	4	2
OR	12000	4	2

The following is an image of the collected data set after being converted to time domain(s).

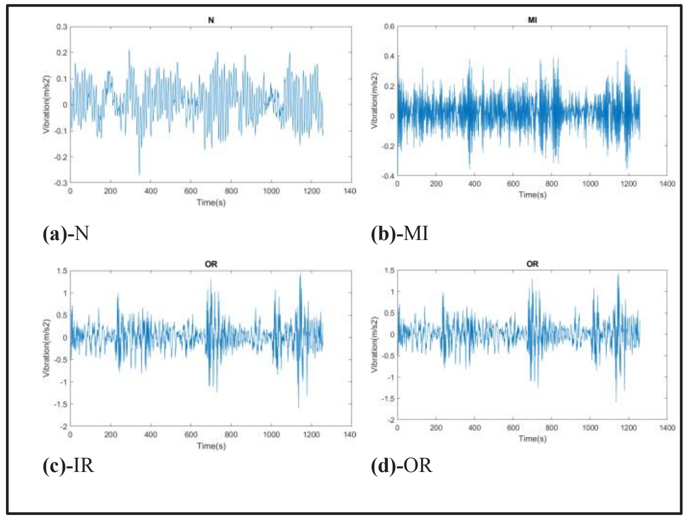

Fig. 4. Representation of the collected data sample set in the time domain (s).

B. ResNet-50

ResNet-50 is a widely recognized and influential deep learning architecture, particularly in the field of computer vision. It belongs to the family of ResNet (Residual Neural Network) models, which are designed to address the challenge of training very deep neural networks by introducing residual learning. Conventional neural networks encounter challenges when it comes to training extremely deep structures because of the vanishing gradient issue and performance degradation with increasing network depth. The vanishing gradient problem hinders effective learning, particularly in the initial layers of the network [15].

ResNet was introduced by Kaiming He et al. in their paper "Deep Residual Learning for Image Recognition" in 2015 [15]. The key innovation was the introduction of residual blocks, which enable the training of extremely deep networks by incorporating shortcut connections or skip connections. They have introduced the concept of residual learning, which is applied to multiple layers within the ResNet framework. Residual blocks in ResNet are effective when the input and output data have the same dimensions. Moreover, ResNet blocks come in two varieties, with two layers for ResNet-18 and ResNet-34 networks and three layers for ResNet-50 and ResNet-101 networks. The initial two layers of the ResNet architecture are reminiscent of GoogleNet, involving a 7x7 convolution operation and 3x3 max-pooling with a stride of 227 [16].

ResNet-50 utilizes a bottleneck design in each of its residual blocks. This design comprises three convolutional layers with kernel sizes of 1x1, 3x3, and 1x1, which serves to reduce computational complexity while maintaining representational capacity. The name 'ResNet-50' stems from the fact that this architecture consists of 50 layers. It's a deep structure that incorporates multiple residual blocks, each with varying filter counts. A residual block is composed of two 3x3 convolutional layers and a shortcut connection. This shortcut connection bypasses one or more layers and directly combines the input with the output, forming what is referred to as the residual. A visual representation of the ResNet-50 architecture is provided in Figure 5.

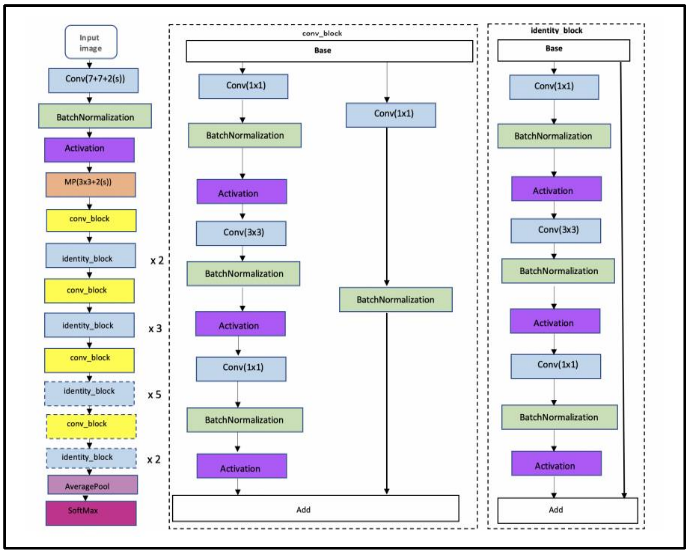

Fig. 5. Resnet-50 architecture.

A neural network gains knowledge by utilizing backpropagation. In the ResNet50 model, the upper layers remain adaptable, allowing them to learn through backpropagation, while the lower layers are kept fixed. The process of modifying the weights during backpropagation in these upper layers is known as fine-tuning. Fine-tuning the upper layers of ResNet50 is necessary because there's no assurance that their statistical properties, like mean and variance, will align with those of our specific dataset [17].

IV. Experiment and Results

A. Experimental Dataset

This research validates the proposed approach using self-generated experimental data. The data acquisition system includes an accelerometer, a measurement module, a chassis, and LabVIEW software, creating a fully automated data collection system. The software handles vibration signal presentation, analysis, and gathering. The dataset includes both typical and defective bearings with internal ring defects, external ring defects, and shaft misalignment in three failure scenarios. An accelerometer positioned above the bearing records vibration signals at a 12 kHz sampling frequency. Detecting machine vibration serves as an early warning system for unfavorable bearing conditions, particularly critical for high-power electric motors that require a warm-up period.

B. Identify the Headings

Using the self-created dataset, we developed a MATLAB program to convert raw data into frequency spectra. Initially, we constructed a time vector 't' ranging from 0 to 1 second, with intervals of 1/Fs, aligned with the sampling frequency (Fs) of 12000 Hz. Subsequently, we utilized the Fast Fourier Transform (FFT), an algorithm derived from the combination of discrete and continuous Fourier transforms, to analyze the signals and obtain their corresponding frequencies. The discrete Fourier transform (DFT) can be described as follows:

$$\Upsilon(k) = \sum_{j=1}^{n} X(j) W_n^{(j-1)(k-1)} \tag{1}$$

Where $W_n = e^{(-2\pi i)/n}$ Fourier transform is represented as:

$$\Upsilon(\omega) = \int_{-\infty}^{+\infty} x(t) e^{-i\omega t}\, dt \tag{2}$$

With $t/T = (j-1)/n$ and $\omega t = 2\pi(j-1)(k-1)/n$, we can represent the sampled signal data using the sinc function:

$$x(t) comb(t) = \sum_{j=1}^{n} X(j)\delta(t-(j-1)\Delta t) \tag{3}$$

Accordingly, four frequency spectra are plotted, each corresponding to one of the three fault charts and one for the normal chart as shown in Figure 6. The MATLAB script reads the raw data files, which are then divided into arrays of 256 samples for the signal image transformation. The final step involves converting this signal data into spectral images, highlighting key features from the original data.

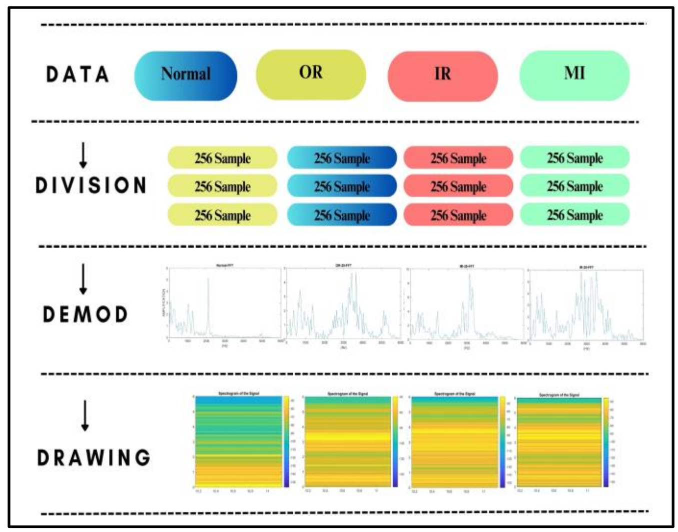

Fig. 6. Analyze the raw data using Short-time Fourier Transform (STFT).

After the conversion of raw data into spectral domain images, we proceeded with the individual classification and labeling of error image files. This process resulted in an image dataset comprising a total of 11.9k images per faulty class. These images were then distributed in a ratio of 70% for training, 20% for validation, and 10% for testing, for each class.

C. Training datasets to each Model

The division of the training dataset into four distinct classes is of paramount importance. Three of these classes are dedicated to faulty bearings (IR, OR, MI), while one class represents normal bearings. It is imperative that each class possesses a diverse dataset, and the inclusion of 11.9k images per faulty class is essential. This ensures that the model comprehensively learns variations associated with different faults, necessitating meticulous data collection.

Similarly, the 'Normal' class requires a sufficient number of images for effective discrimination between normal and faulty bearings.

The inclusion of 11.9k images for this class ensures that the model develops a robust understanding of the typical characteristics exhibited by normal bearings. The creation of these four classes, each with a substantial and diverse dataset, serves as a foundational element for the model's future performance in bearing classification and prediction tasks.

D. *Results of fault diagnosis and comparative analysis post-training for each model*

Table II shows the exact percentage loss of the ResNet-50 model after training.

TABLE II. SUMMARY OF RESNET-50

Model	The Accuracy of Training Model			
	Test Accuracy	***Val_accuracy***	***Val_loss***	***Test loss***
Resnet-50	96.527%	93.381%	0.1953	0.3530

The provided table offers a comprehensive evaluation of test and validation performance metrics [21] for various computer vision. The training and evaluation outcomes of the ResNet50 model are reported in the table. The model exhibits strong performance in various aspects, indicating its optimal capabilities.

Test Accuracy: The model achieves an impressive test accuracy of 96.527%. This metric reflects the model's ability to correctly classify images in an unseen dataset, which is a crucial measure of its generalization capacity.

Validation Accuracy: The validation accuracy, at 93.381%, is another positive indicator of the model's performance. This metric is crucial during training as it helps monitor the model's performance on a separate dataset, making it an essential tool for preventing overfitting.

Validation Loss: The low validation loss value of 0.1953 suggests that the model generalizes well during training. Lower validation loss values indicate that the model is not overfitting and is effectively learning from the training data.

Test Loss: The test loss is 0.3530, which is slightly higher than the validation loss. This difference can be expected, as the test loss measures the model's performance on a completely unseen dataset. The proximity of the test loss to the validation loss suggests that the model maintains its performance on new, unseen data.

Following the training of the dataset using Resnet50, the subsequent step involved the utilization of the test dataset comprising four distinct classes, each consisting of up to 2000 images per class. The primary objective was to evaluate the accuracy of the AI models post-training. The evaluation results are conveyed through the presentation of the confusion matrix below:

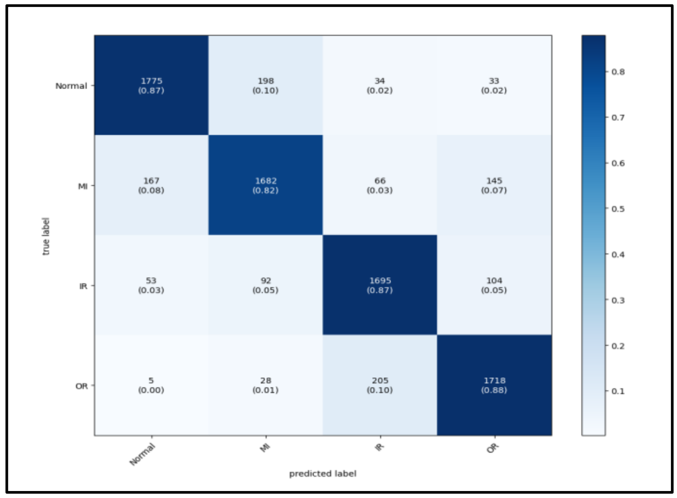

Fig. 7. Resnet50 Confusion matrix.

In conclusion, the ResNet50 model demonstrates optimal capabilities in image classification tasks. It exhibits high test accuracy and low validation and test loss values, indicating its ability to generalize well and make accurate predictions on unseen data. These results suggest that the model is well-trained and can be considered reliable for various image recognition applications.

E. *Evaluate the model's accuracy once more using a different dataset:*

The chart below describes the number of misclassified images in a test set of 2000 images using the ResNet50 model, showcasing its performance in different categories:

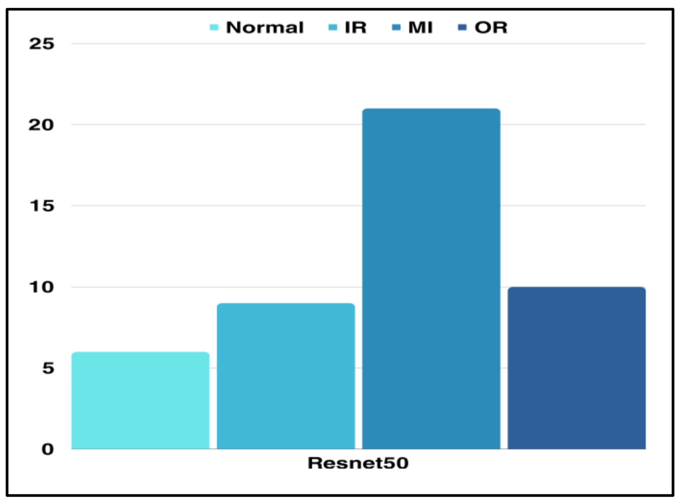

Fig. 8. The chart displays the count of misclassified images.

The ResNet50 model was evaluated on a diverse test set of 2000 images, and the results are presented in the chart. The model's performance varies across different domains, as indicated by the misclassification counts in the respective categories. Notably, the ResNet50 model demonstrates a relatively strong performance in the 'Normal' category, with only 6 misclassified images out of 2000. This underscores its robustness in standard image recognition tasks.

In the 'IR' category, the model performs moderately well with 9 misclassified images, indicating its ability to adapt to different image modalities. However, in more specialized domains such as 'MI' , the model shows room for improvement, with 21 misclassified images. Furthermore, in the 'OR' category, the model exhibits a reasonable performance with 10 misclassified images, suggesting its suitability for object recognition tasks.

In conclusion, the ResNet50 model excels in standard image recognition tasks but may require further fine-tuning or domain-specific adaptations for optimal performance in specialized domains. Understanding the model's strengths and weaknesses in different domains is essential for selecting the appropriate tool for specific image analysis applications.

V. CONCLUSION

Deep learning algorithms offer numerous advantages for machinery fault diagnosis, including engines, transformers, and cutting machines. We anticipate a rapid increase in studies diagnosing roller bearing faults using deep learning. This paper's primary contribution lies in creating a dataset of three types of bearing faults and testing it with various state-of-the-art deep learning networks. The findings highlight the advantages of the ResNet-50 network in this context, opening new possibilities for classification tasks. In the future, we plan to expand and refine our dataset to enhance its comprehensiveness. Specifically, we will build and collect data with many different bearing sizes. Furthermore, we will find solutions to solve problems related to our big data for real-time transmission to IOT systems. This helps engineers remotely monitor and diagnose machine errors

REFERENCES

[1] Çınar, Z. M., Abdussalam Nuhu, A., Zeeshan, Q., Korhan, O., Asmael, M., & Safaei, B. (2020). Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. *Sustainability*, *12*(19), 8211.

[2] Rai, A., & Upadhyay, S. H. (2016). A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribology International*, *96*, 289-306.

[3] Saufi, S. R., Ahmad, Z. A. B., Leong, M. S., & Lim, M. H. (2019). Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review. *Ieee Access*, *7*, 122644-122662.

[4] Zhao, Z., Wu, J., Li, T., Sun, C., Yan, R., & Chen, X. (2021). Challenges and opportunities of AI-enabled monitoring, diagnosis & prognosis: A review. *Chinese Journal of Mechanical Engineering*, *34*(1), 1-29.

[5] Rahman, T. A., Chek, L. W., & Ramli, N. Intelligent Vibration-based Anomaly Detection for Electric Motor Condition Monitoring.

[6] Tahir, M. M., Khan, A. Q., Iqbal, N., Hussain, A., & Badshah, S. (2016). Enhancing fault classification accuracy of ball bearing using central tendency based time domain features. IEEE Access, 5, 72-83.

[7] Chen, Z.; Li, C.; Sanchez, R.V. Gearbox fault identification and classification with convolutional neural networks. Shock. Vib. 2015, 2015, 390134.

[8] Zhao, J.; Yang, S.; Li, Q.; Liu, Y.; Gu, X.; Liu, W. A new bearing fault diagnosis method based on signal-to-image mapping and convolutional neural network. Measurement 2021, 176, 109088.

[9] Gao, Y.; Liu, X.; Huang, H.; Xiang, J. A hybrid of FEM simulations and generative adversarial networks to classify faults in rotor-bearing systems. ISA Trans. 2021, 108, 356–366.

[10] Liu, H.; Zhou, J.; Xu, Y.; Zheng, Y.; Peng, X.; Jiang, W. Unsupervised fault diagnosis of rolling bearings using a deep neural network based on generative adversarial networks. Neurocomputing 2018, 315, 412–424.

[11] Wang, R.; Jiang, H.; Li, X.; Liu, S. A reinforcement neural architecture search method for rolling bearing fault diagnosis. Measurement 2020, 154, 107417.

[12] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

[13] Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.

[14] Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

[15] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (CVPR) (pp. 770-778).

[16] Sarwinda, D., Paradisa, R. H., Bustamam, A., & Anggia, P. (2021). Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. Procedia Computer Science, 179, 423-431.

[17] Reddy, A. S. B., & Juliet, D. S. (2019, April). Transfer learning with ResNet-50 for malaria cell-image classification. In 2019 International Conference on Communication and Signal Processing (ICCSP) (pp. 0945-0949). IEEE.

[18] A.Ghasempour, "Internet of Things in Smart Grid: Architecture, Application, Services, Key Technologies, and Challenges," Inventions Journal,vol.4,no.1,pp.1-12,2019, https://doi.org/10.3390/inventions4010022.

[19] Pham Van Nam and Tran Hoai Linh. "Electrocardiogram (ECG) Circuit Design and Using the Random Forest for ECG Arrhythmia Classification." International Conference on Engineering Research and Applications, 1 Dec. 2022, page range 477-494.

[20] Vuong Hoang Nam, Hoai Linh Tran, and Van Nam Pham. "Multiple Neural Network Integration Using a Binary Decision Tree to Improve the ECG Signal Recognition Accuracy." International Journal of Applied Mathematics and Computer Science, 2014, Vol. 24, No. 3, 647–655.

[21] A.Ghasempour and M.Martinez-Ramon, " Short-Term Electric load Prediction in Smart Grid using Multi-Output Gaussian Processes Regression" IEEE Kansas Power and Energy Conference (IEEE KPEC),2023