A Study on Early Fault Detection Methods for Industrial Rotating Machinery and Robots

Seung-Hwan Choi
Advanced Mobility System Group
Korea Institute of Industrial Technology
Daegu, Republic of Korea
csw1496@kitech.re.kr

Chang-Hyun Kim
Advanced Mobility System Group
Korea Institute of Industrial Technology
Daegu, Republic of Korea
limition@kitech.re.kr

Hyoeun Kwon
Advanced Mobility System Group
Korea Institute of Industrial Technology
Daegu, Republic of Korea
hekwon525@kitech.re.kr

Hiroaki Kawamoto
Institute of System and Information Engineering
University of Tsukuba
Tsukuba, Japan
kawamoto@iit.tsukuba.ac.jp

Suwoong Lee
Advanced Mobility System Group
Korea Institute of Industrial Technology
Daegu, Republic of Korea
lee@kitech.re.kr

***Abstract*—This paper presents a study on early fault detection for industrial rotating machinery and robots. Time-series vibration data collected from durability tests on three drive modules used in industrial robots were employed to compare and analyze the fault detection performance of supervised classification models and semi-supervised anomaly detection (AD) models. The classification models exhibited significantly degraded performance across all three datasets due to variations in the fault characteristics and the insufficiently distinct differences between normal and faulty conditions within each dataset. In contrast, the AD models—which incorporated weighting techniques based on time progression and the continuity of anomalies—effectively detected anomalies in the cycles immediately preceding faults across all datasets. Notably, the Transformer-based and Diffusion-based AD models demonstrated their capability for early fault detection by registering high anomaly scores in the pre-fault cycles.**

Keywords—rotating machine, vibration data, fault detection, classification, anomaly detection

I. INTRODUCTION

Recent advancements in manufacturing have rapidly accelerated the adoption of high-level automation and smart factory systems. In this process, industrial rotating machinery and robots are gradually being applied throughout production processes, enabling efficient and reliable production by handling heavy loads and performing precision assembly tasks. Moreover, these devices are capable of operating continuously 24 hours a day while maintaining consistent quality and speed, thereby establishing themselves as essential means for securing a competitive advantage [1], [2].

However, rotating machinery and robots consist of intricately interconnected mechanical components, and their operating conditions or load states can accelerate component degradation and wear. In such cases, unexpected component failures or malfunctions can lead to a complete halt of the production line or result in significant economic losses and safety hazards [3]. In particular, if driving modules—comprising elements such as gearboxes, motors, bearings, and brakes—fail, the machinery's performance can deteriorate dramatically, rendering early diagnosis and preventive maintenance indispensable [4], [5].

In the past, fault detection methods have primarily relied on extracting traditional features from sensor data. In particular, techniques such as analyzing signal characteristics in the frequency domain using the Fourier Transform (FFT) or extracting features indicative of abnormal patterns in the time-frequency domain via wavelet transform have been widely employed. Although these methods work effectively when the spectral characteristics of normal and fault conditions are clearly distinct, they encounter limitations in real industrial environments where operating conditions are complex and fault modes vary. Under such circumstances, it becomes challenging to predefine appropriate frequency bands and transformation parameters [6], [7].

To overcome these limitations, fault diagnosis techniques that utilize machine learning and artificial intelligence (AI) have recently attracted considerable attention in industrial settings. In these approaches, sensor data—such as vibration, torque, current, and temperature—collected in real time from rotating machinery or robots are analyzed to identify early signs of faults. Among these sensors, vibration sensors are particularly effective at capturing subtle changes and play a crucial role in predicting the degradation or wear of driving modules. Supervised classification models have proven effective when ample data from both normal and fault conditions are available, and they have been widely employed in previous research [8], [9].

However, in industrial settings, it is challenging to train supervised models because fault data is scarce or fault types are highly diverse. To overcome this limitation, semi-supervised approaches—which train models solely on normal data and subsequently detect anomalies as deviations from the normal range—have gained significant attention [10], [11]. In particular, anomaly detection (AD) models such as Long Short-Term Memory (LSTM), Variational Autoencoder (VAE), Diffusion, and Transformer are trained exclusively on normal data to internalize normal patterns. When new data is introduced, these models compute reconstruction errors or differences in the distribution of feature vectors, and larger

This work was supported by the Industrial Technology Innovation Program(RS-2024-00422721, Development of safe and reliable XAI-based robotic workcell safety sensors and control modules) funded by the Ministry of Trade, Industry Energy (MOTIE,Korea). And this work was supported by the Industrial Technology Innovation Program(RS-2024-00424974, Robot and robot service digital twinization framework technology) funded by the Ministry of Trade, Industry Energy (MOTIE,Korea).

errors or differences indicate that the data deviates from the normal range, thereby enabling fault detection [12], [13], [14].

This paper compares the fault detection performance of a supervised classification model and a semi-supervised AD model using vibration data collected from durability tests on driving modules employed in industrial rotating machinery and robots. First, the ability of the classification model to distinguish between normal and fault data is evaluated. Then, under the assumption of limited fault data, a semi-supervised AD approach is applied to demonstrate through experiments that effective fault detection is achievable even when the fault types are ambiguous or the available data is insufficient.

The remainder of the paper is organized as follows. In Section 2, the durability test environment for driving modules and the methods for acquiring vibration data are described. Section 3 analyzes the features of the acquired data using statistical techniques. In Section 4, the vibration data is applied to both a classification model and a semi-supervised model for fault detection, and the experimental results are compared and analyzed. Finally, Section 5 presents the conclusions and future research directions.

II. Dataset of Driving Modules

In this study, a durability test environment for driving modules was constructed, as shown in Figure 1, and data were acquired from it. The driving module used in the tests was the Kaiser` KAH-25E L5BE model—a commercial product for collaborative robots—with a rated power of 500 W, a maximum average torque of 133 Nm, and a rated speed of 16.2 RPM. The vibration sensor employed was the PCB 356A17, a 3-axis accelerometer with a sensitivity of 500 mV/g, a measurement range of ±98 m/s² (peak), and a frequency range of 0.4–4000 Hz.

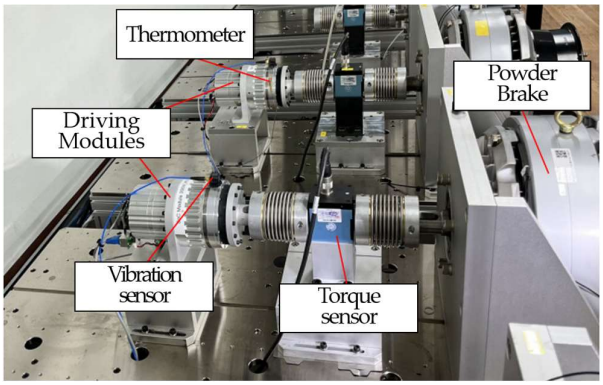

Fig. 1. Durability test environment for driving modules.

The durability test of the driving module was conducted under the conditions presented in Figure 2. A torque of 115 Nm was applied to the module, which was then rotated continuously back and forth between 0° and 360°. Each cycle lasted 20 seconds, with an acceleration/deceleration phase of 0.5 seconds (at 24 RPM), a constant-speed phase of 4.5 seconds (at 12 RPM), and a stopping phase of 1 second. The top graph in Figure 2 shows the applied torque, the middle graph displays the rotational speed, and the bottom graph illustrates the rotational position.

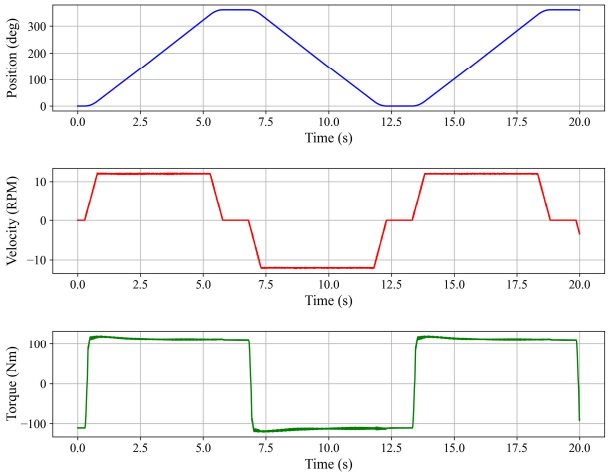

Fig. 2. Durability test conditions for driving module.

In the durability test, data were collected for 20 seconds every 10 minutes at a sampling rate of 1 kHz. Additionally, to ensure data stability, only data from the constant-speed forward rotation phase were used. The data from each 10-minute period were defined as one cycle and applied to the fault detection performance experiments.

III. Feature Analysis of Data

Driving modules used in industrial environments can experience wear, degradation, and various anomalies due to prolonged operation or exposure to harsh conditions. In particular, the vibration signals from these driving modules provide a crucial indicator for the early detection of such abnormalities.

Figure 3 shows the vibration data obtained from the durability tests. Under identical test conditions, three driving modules were tested: Module 1 failed after 882 cycles (147 hours), Module 2 after 3,606 cycles (601 hours), and Module 3 after 1,190 cycles (198.3 hours).

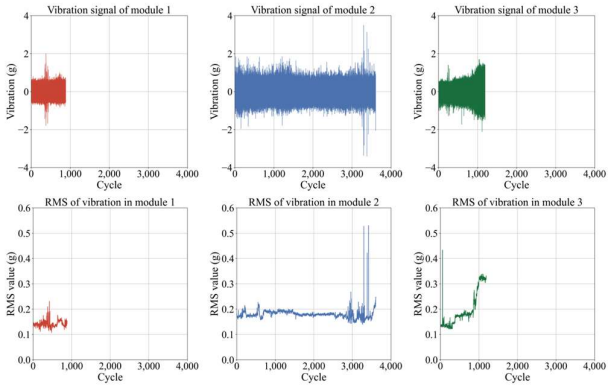

Fig. 3. Vibration data collected from durability test.

Module 1's vibration signal undergoes a brief period of abrupt fluctuations in the mid-phase, but in the latter phase, it maintains a level similar to the initial phase without any further sharp increases or spikes. In Module 2, the vibration signal rapidly intensifies toward the end, with large spikes observed, and the RMS value also rises markedly just before the fault occurs. Meanwhile, Module 3 exhibits a short spike in the early phase, remains relatively stable until the mid-

phase, and then experiences a rapid increase in vibration in the latter phase, leading to a substantial rise in the RMS value.

In this paper, various features of the vibration signals generated by each module are analyzed to confirm the changes between normal and pre-fault conditions. To achieve this, both traditional frequency analysis methods—namely, FFT (or PSD) analysis—and the Discrete Wavelet Transform (DWT), which allows for simultaneous time-frequency analysis, were employed to analyze how the frequency bands of each module vary.

Figure 4 presents a comparison of the power spectral density (PSD) obtained via frequency transformation for the average signals in both the normal and pre-fault intervals (cycles 1–10, 101–110, 201–210, last 20–11, and last 10–1) for Modules 1, 2, and 3. The horizontal axis represents frequency (Hz) and the vertical axis shows PSD (dB/Hz), with the PSD curves for each interval overlaid.

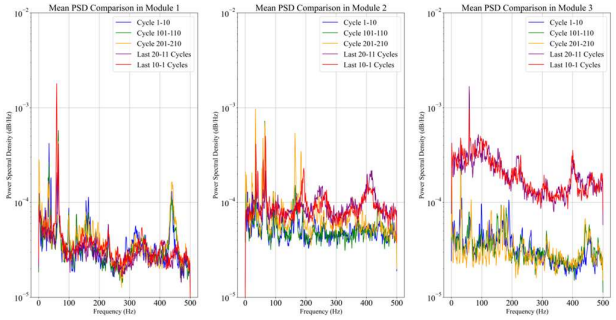

Fig. 4. Analysis of vibration data features using PSD.

For module 1, the PSD is notably high in the low-frequency region (around 60 Hz). However, when comparing normal and pre-fault intervals, there is no substantial overall difference. In the high-frequency range (around 440 Hz), normal cycles show slightly higher PSD values, but other frequency bands exhibit relatively stable distributions with no significant changes, indicating that there is no marked distinction between normal and pre-fault states. For module 2, both normal and pre-fault cycles present similar curves in the low-frequency range (0–100 Hz), maintaining a stable pattern without distinct spikes. Although cycles 201–210 display comparatively higher PSD values in the mid-frequency range (around 200 Hz), the overall fluctuation from low to high frequencies remains minimal. However, in the high-frequency range (around 400 Hz), a relatively large difference appears between normal and pre-fault cycles, suggesting that vibration energy increases in certain sections. For module 3, the PSD in the pre-fault cycles is significantly higher than that in the normal cycles across the entire frequency range, and some segments also exhibit substantially larger fluctuations. This implies that, in the pre-fault interval, vibration energy broadly rises from low to high frequencies. Compared to the other modules, module 3 shows a more pronounced change during the pre-fault stage.

Figure 5 presents the results of applying the Discrete Wavelet Transform (DWT) to the average signals of both normal and pre-fault cycles for each module. The horizontal axis indicates the sample index, and the vertical axis represents the wavelet coefficient value. Levels 1 and 2 primarily capture lower-frequency or larger-scale information, while Levels 3 and 4 mainly reflect mid- to high-frequency information.

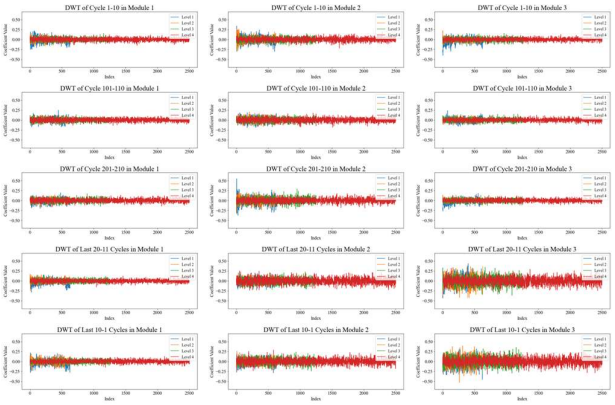

Fig. 5. Analysis of vibration data features using DWT.

For Module 1, the Level 1 and 2 coefficients (blue and orange) remain relatively stable within a narrow range across the entire interval, and the Level 3 and 4 coefficients (green and red) also exhibit relatively consistent behavior without large amplitude fluctuations. Similarly, for Module 2, neither the normal nor the fault cycles show any noticeable amplitude spikes or fluctuations in the Level 1–4 coefficients, maintaining a generally stable pattern in both the low-frequency and mid-/high-frequency ranges. In contrast, for Module 3, the amplitude range of the Level 2 and 3 coefficients becomes somewhat wider in the pre-fault cycles compared to the other two modules, and greater variation is observed in the mid-/high-frequency bands.

IV. Experiments and Results for Fault Detection

A. Fault Detection using Classification Model

Supervised classification models have been widely applied not only to image-based tasks but also to fault detection in industrial settings where ample normal and fault data are available. A fault detection experiment was conducted using classification models: for each module, the first 100 cycles of normal operation and the 100 cycles immediately preceding failure were used for training. Training inputs comprised raw time-series signals, FFT-based power spectral density data, and DWT-transformed features. Fault detection was performed with an Artificial Neural Network (ANN) and a one-dimensional Convolutional Neural Network (1D-CNN), and their performances were compared and analyzed.

The ANN model consumes one-dimensional feature vectors and passes them through four fully-connected layers of 512 → 256 → 128 → 64 units (each with ReLU and 50% dropout), ending in a single-unit output layer that yields the raw logit. Training uses the AdamW optimizer (lr = 0.001, weight decay = 0.0001). The 1D-CNN model takes raw 1,000-sample segments as (batch, 1, 1000) inputs, applies two Conv1d + MaxPool1d blocks (1→64, then 64→128 channels; kernel sizes 5→2), flattens the result, and feeds it through a 256-unit dense layer (ReLU + 50% dropout). A final sigmoid-activated neuron produces the fault probability. Training uses BCELoss with AdamW (lr = 0.001, weight decay = 0.0001).

Table 1 presents the fault detection results on vibration data using the classification model. Out of the three modules, two modules were used as training data while the remaining one was used as test data.

First, when modules 2 and 3 were used for training and module 1 was held out for testing, overall accuracy remained

around 50%. This low performance likely stems from an unclear separation between normal and fault patterns in module 1, which prevented the model—trained only on modules 2 and 3—from reliably detecting faults. Next, with modules 1 and 3 for training and module 2 for testing, accuracy improved slightly but remained unsatisfactory. Although module 2's data show some distinguishing features, their distribution does not fully match that of modules 1 and 3, leading to degradation in certain models.

Finally, training on modules 1 and 2 and testing on module 3 yielded much better results: FFT-ANN and 1D-CNN achieved high accuracy and F1 scores, while Raw-ANN and DWT-ANN lagged behind. This indicates that module 3's data offer a clearer normal-vs-fault separation, which the former models captured more effectively. In summary, 1D-CNN delivered the best overall performance, followed by the FFT-based ANN, with Raw-ANN and DWT-ANN showing relatively lower accuracy.

TABLE I. FAULT DETECTION RESULTS USING CLASSIFICATION

Dataset		Metric	Fault Detection Performance using Classification Model			
Train	*Test*		*Raw-ANN*	*FFT-ANN*	*DWT-ANN*	*1D-CNN*
2, 3	1	Accuracy	0.5546	0.5042	0.5156	0.5128
		F1-Score	0.2957	0.0327	0.1047	0.0580
1, 3	2	Accuracy	0.5370	0.5839	0.5750	0.6578
		F1-Score	0.5690	0.6658	0.4669	0.7213
1, 2	3	Accuracy	0.5333	0.8544	0.5972	0.8344
		F1-Score	0.5492	0.8729	0.6253	0.8551

The experimental results demonstrated that classification models are not effective for fault detection in driving modules with diverse characteristics. It was observed that for a classification model to achieve high performance, not only is sufficient training data required, but the features distinguishing normal and fault conditions must also be clearly defined.

B. Fault Detection using Anomaly Detection Model

In real industrial environments, the operating conditions and work processes are highly variable, which leads to complex fault types and data characteristics. This complexity imposes limitations on fault detection methods based on classification models, as they require clearly defined labels and sufficient data. To address these challenges, this paper proposes a fault detection approach based on semi-supervised AD models.

This paper employs representative AD models—LSTM, VAE, Transformer, and Diffusion—for fault detection. For performance comparison, data from two out of three modules were used for training, while the remaining module's data was used for testing. In addition, only the first 100 cycles of data from each module were used for training the AD models, and fault detection was performed on all cycles of the test data.

In this study, to effectively detect pre-fault anomalies in time-series vibration data, weights based on time progression and anomaly continuity were applied. The time progression weight $\omega_t(i)$ was computed using Equation (1) to assign increasing weight as the probability of failure rises over time.

$$\omega_t(i) = exp\left(\frac{i}{N-1} ln(\alpha)\right), i = 0,1,\dots,N-1 \quad (1)$$

Here, i represents the current cycle, N denotes the total number of cycles, and α is the rate at which the weight increases over time.

Additionally, since consecutive anomaly detections are considered to be more indicative of impending faults than isolated anomalies, a continuity weight $\omega_c(i)$ is applied as shown in Equation (2).

$$\omega_c(i) = \begin{cases} 1 & if\ i = 1\ or\ a_i - a_{i-1} > \delta \\ \beta \cdot \omega_c(a_{i-1}) & if\ a_i - a_{i-1} \le \delta \end{cases} \quad (2)$$

Here, a_i denotes the index of the cycle where an anomaly occurs, δ represents the number of cycles used to assess continuity, and β is the rate of weight increase for consecutive anomalies.

For each cycle i, the reconstruction error $e(i)$ is computed, and by calculating and multiplying the time progression weight and the continuous anomaly weight, the final weighted error $E_\omega(i)$ is derived as shown in Equation (3).

$$E_\omega(i) = e(i) \times \omega_t(i) \times \omega_c(i) \quad (3)$$

The final $Anomaly\ Scorer(i)$ was obtained as shown in Equation (4) by normalizing the final weighted error $E_\omega(i)$ by dividing it by the maximum value of $E_\omega(j)$, yielding a score in the range [0,1].

$$Anomaly\ Score(i) = \frac{E_\omega(i)}{max\{E_\omega(j)\}}, j = 0,1,\dots,N-1 \quad (4)$$

Figure 6 shows the results of visualizing the anomalies detected and their corresponding anomaly scores by applying an AD-based fault detection model to the vibration data of each module.

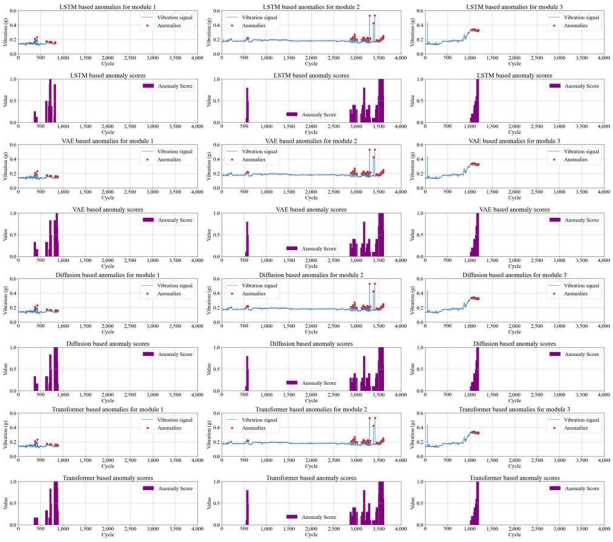

Fig. 6. Fault Detection Results Using an AD Model.

For Module 1, the features of the vibration signals in the normal and fault states were not clearly distinguishable, and significant vibration fluctuations occurred in the mid-cycle. As a result, although all AD models detected relatively high anomalies during the mid and later cycles, the pre-fault cycles exhibited the highest anomaly scores, thereby enabling effective fault detection. For module 2, although significant fluctuations in the vibration data were observed in the early

cycles, leading to somewhat elevated anomaly detections in the AD models, high anomaly scores were recorded in the later and pre-fault cycles, effectively enabling fault detection. Lastly, Module 3 exhibited a clear difference in vibration amplitude between the normal and fault states, with all AD models assigning high anomaly scores in the pre-fault segment, thereby achieving effective fault detection. Overall, the performance comparison among the models confirmed that the Diffusion and Transformer models were the most effective at detecting anomalies immediately preceding faults.

Table 2 shows the fault detection performance of the AD models for each dataset. For each dataset, the data were segmented into the last 200, 100, 50, and 10 cycles prior to fault, and an anomaly score was calculated for each cycle to compare how effectively pre-fault anomalies are detected.

TABLE II. ANOMALY SCORES OF AD MODELS BY DATASET

Dataset		Model	Fault Detection Performance (Anomaly Score)			
Train	*Test*		*200 cycle*	*100 cycle*	*50 cycle*	*10 cycle*
2, 3	1	LSTM	0.2262	0.2273	0.1667	0.5000
		VAE	0.2778	0.3636	0.5278	0.5000
		Diff.	0.2778	0.3939	0.5556	0.5833
		Transf.	0.3000	0.4333	0.5333	0.3333
1, 3	2	LSTM	0.4952	0.8455	0.9333	0.8000
		VAE	0.4857	0.8455	0.9333	0.8000
		Diff.	0.4905	0.8545	0.9333	0.8000
		Transf.	0.5250	0.9200	0.9200	0.6000
1, 2	3	LSTM	0.3000	0.5300	0.7000	1.0000
		VAE	0.3000	0.5300	0.7200	1.0000
		Diff.	0.3000	0.5300	0.7200	1.0000
		Transf.	0.3000	0.5100	0.6000	1.0000

When module 1 was used as the test dataset, the overall anomaly scores were somewhat lower than for the other modules. However, in the final 100 cycles before the fault, the highest scores appeared in the order of Transformer, Diffusion, VAE, LSTM; and in the final 50-cycle segment, the order was Diffusion, Transformer, VAE, LSTM—enabling effective fault detection. The slightly lower scores in the last 10 cycles are presumed to result from the module continuing to operate for a period even after the fault occurred, so those particular scores may be disregarded. In module 2, from the final 100 cycles onward, all models produced high anomaly scores above 0.84, with the Transformer model exhibiting especially prominent performance. At the final 50-cycle mark, every model again recorded anomaly scores exceeding 0.92, thereby demonstrating stable fault detection. For module 3, the Transformer model generally exhibited slightly lower anomaly scores compared to the other models; however, starting from the final 50 cycles, its scores exceeded 0.6, and during the last 10 cycles, the score reached a maximum of 1, thereby demonstrating highly effective fault detection. Overall, the Transformer model detected faults earlier than the others by registering high anomaly scores at earlier cycles, while the Diffusion model also reliably identified faults across all datasets, supported by its consistently high anomaly scores.

V. CONCLUSION AND FUTURE WORK

This paper presents a study on early fault detection in industrial rotating machinery and robots. Based on time-series vibration data collected from durability tests on drive modules, the fault detection performance of a supervised classification model and a semi-supervised anomaly detection (AD) model was compared and analyzed. Experimental results revealed that when the fault data features are not distinctly differentiated or pronounced across all three datasets, the supervised classification model fails to achieve satisfactory performance, thereby limiting its applicability in real industrial environments. In contrast, the AD model, which is trained solely on normal data and incorporates weighting techniques based on time progression and the continuity of anomalies, effectively detects pre-fault anomalies. Notably, among the AD models, the Transformer and Diffusion models exhibited outstanding performance in fault detection.

In future research, additional datasets from various operating environments and drive conditions will be collected to further validate the performance of the AD models in depth. In addition, the study aims to enhance the models' performance and reliability to contribute to reducing maintenance costs and ensuring safety in industrial settings.

REFERENCES

[1] P. Kumar,S. Khalid, and H. S. Kim, "Prognostics and health management of rotating machinery of industrial robot with deep learning applications—A review", *Mathematics*, vol. 11, no. 13, pp. 3008, July 2023.

[2] M. J. Kim, K. M. Ku, MD. S. Islam, M. J. Chung, H. Y. Kim and K. Kim, "Fault diagnosis of industrial robots using CNN and vibration data, *Journal of the Semiconductor & Display Technology*, vol. 23, no. 3, pp.127-134, Sep 2024.

[3] X. Chen, R. Yang, Y. Xue, C. Yang, B. Song, and M. Zhong, "A novel momentum prototypical neural network to cross-domain fault diagnosis for rotating machinery subject to cold-start" *Neurocomputing*, vol.555, pp.126656, Oct 2023.

[4] A. H. Sabry, U. A. B. U. Amirulddin, "A review on fault detection and diagnosis of industrial robots and multi-axis machines", *Results in Engineering*, vol. 23, pp, 102397, Sep 2024.

[5] X. Zhang, K. P. Rane, I. Kakaravada and M. Shabaz, "Research on vibration monitoring and fault diagnosis of rotating machinery based on internet of things technology" *Nonlinear Engineering*, vol.10, pp. 245-254, Oct 2021, DIO:10.1515/nleng-2021-0019.

[6] Z. Liu, J. Zhang, X. He, Q. Zhang; G. Sun, and D. Zhou, "Fault diagnosis of rotating machinery with limited expert interaction: A multicriteria active learning approach based on broad learning system", *IEEE Transactions on Control Systems Technology*, vol. 31, Mar 2023.

[7] H. O. A. Ahmed and A. K. Nandi, "Vibration image representations for fault diagnosis of rotating machines: A review", *Machines*, vol. 10, pp.1113 2022.

[8] N. Sikder,A. S. M. Arif, M. M. M. Islam, A. Nahid, "Induction motor bearing fault classification using extreme learning machine based on power features", *Arabian Journal for Science and Engineering*, vol. 46, pp. 8475-8491, Mar 2021.

[9] H. Pu, S. Teng, D. Xiao, L. Xu, Y. Qin, and J. Luo, "Compound fault diagnosis of rotating machine through label correlation modeling via graph convolutional neural network", *IEEE Transactions on Instrumentation and Measurement*, vol. 73, Dec 2023.

[10] Y. Xu, X. Lu, T. Gao, and R. Meng, "A self-supervised multiview contrastive learning network for the fault diagnosis of rotating machinery under limited annotation information", *IEEE Transactions on Instrumentation and Measurement*, vol. 74, Jan 2025.

[11] J. Cohen, and J. Ni, "Semi-supervised learning for anomaly classification using partially labeled subsets", *J. Manuf. Sci. Eng*, vol. 144, pp. 061008, 2022.

[12] N. Chen, H. Tu, X. Duan, L. Hu, and C. Guo, "Semisupervised anomaly detection of multivariate time series based on a variational autoencoder", *Applied Intelligence*, vol. 53, pp. 6074–6098, 2023.

[13] P. Yu, M. Ping, J. Ma, and J. Cao, "Unsupervised signal anomaly transformer method: Achieving bearing life anomaly detection without the need for failure samples", *Engineering Applications of Artificial Intelligence*, vol. 136, pp. 108940, Oct 2024.

[14] W. Hu, G. Frusque, T. Wang, F. Chu, and O. Fink, "Classifier-Free Diffusion-Based Weakly-Supervised Approach for Health Indicator Derivation in Rotating Machines: Advancing Early Fault Detection and Condition Monitoring", Sep 2024, arXiv:2409.01676v1.