Representation Learning–Based Fault Diagnosis of Angular Contact Ball Bearings for Machine Tool Spindles

Jooho Hwang[1], Dung Minh Nguyen[2], and Sungjoon Kim[3]

[1,3] *Korea Institute of Machinery and Materials, Daejeon, 34103, South Korea*
jooho@kimm.re.kr

sjkim314@kimm.re.kr

[2] *University of Science and Technology, Daejeon, 34113, South Korea*
minhdung@kimm.re.kr

ABSTRACT

As a precision mechanical component designed to reduce friction between moving parts, the rolling bearing is widely employed across various industries. Predicting the remaining useful life (RUL) of bearings is critical for preventing unexpected failures and ensuring safe and reliable equipment operation. Equally important is the role of lubrication, which not only supports proper bearing performance but also ensures the accurate operation of CNC spindles. The primary objective of predicting lubricant degradation is therefore to estimate the time at which the lubricant no longer fulfills its intended function

In this study, bearing defect frequencies and lubricant anomalies are investigated through vibration-signal analysis under different operating conditions using a representation learning approach. To be more specific, a representation framework is employed to reconstruct input signals. A target frequency band, extending from three times the rotational frequency to 10,000 Hz, is first defined. Spectral root mean square (Spectral RMS) features are then extracted from the low-to-middle frequency portion of this band, where defect-induced impulsive excitations are effectively captured. The framework is trained using labeled healthy-bearing datasets to establish an anomaly-detection threshold, which is subsequently applied during the testing phase to evaluate reconstruction loss. The threshold is defined as the 95th percentile of the reconstruction means square error obtained during training. This criterion enables the identification of anomalies in labeled fault-condition data. The effectiveness of the proposed method is demonstrated by employing two practical datasets. The results indicate that the proposed method effectively detects anomalies in unseen data and achieves robust diagnostic performance

1. INTRODUCTION

Bearings are critical machine components that are highly susceptible to failure during operation. Therefore, accurate estimation of remaining useful life (RUL) is essential for reliable and safe equipment operation.

In addition to conventional spectral analysis, data-driven methods have been widely developed for fault diagnosis and RUL prediction. These approaches learn relationships between condition-related signals, such as vibration data, and machine health states. Previous studies have applied convolutional neural networks (CNNs), bootstrap-based methods, and bidirectional gated recurrent units (Bi-GRUs) to rolling-bearing RUL prediction, showing strong feature-extraction capability without requiring detailed physical models (Huang, Huang, Li, & Peng, 2021; Shang, Tang, Zhao, Jiang, & Ran Lin, 2022). However, their performance depends heavily on the quality and representativeness of the input data. To address this limitation, physics-informed neural networks (PINNs) have been introduced to improve RUL prediction by embedding physical equations into the learning process, as shown in Gong et al., (2025); Li, Zhang, Li, & Si, (2024); Shen et al., (2021).

Lubricant degradation is another important factor in bearing failure, as lubricant performance gradually deteriorates over time until it does not fulfill its function. Recent studies have shown that data-driven methods can also be used to predict lubricant RUL in operating rolling bearings. In particular, vibration signals have been processed to highlight subtle changes in spectral characteristics caused by lubricant degradation, followed by transformer-based prediction of lubricant life, as reported by Kim, Seo, & Park, (2024).

In this study, a representation-learning framework is employed for early anomaly detection in both bearing and lubricant conditions through input reconstruction. A decision threshold is derived from the distribution of reconstruction errors during training and is then used to identify abnormal conditions during evaluation.

The remainder of this paper is organized as follows. Section 2 presents the feature-extraction method. Section 3 describes the general representation-learning framework. Section 4 applies the proposed approach to bearing-anomaly detection and lubricant-failure prediction. Finally, Section 5 concludes the paper and outlines future work.

2. PROPOSED METHODOLOGY FOR FEATURE EXTRACTION

To characterize vibration signals, representative features must first be extracted. In this study, spectral root mean square (spectral RMS) within specific frequency ranges is used as the primary features.

First, the time-domain vibration signal is transformed into the frequency domain using the Fourier transform at 1 s intervals. The resulting frequency spectrum at each interval is then divided into multi-scale frequency regions, following Mei, Jia, Zeng, Zhou, & Zhao, (2016), defined as:

$$\mathrm{Seg} = \{(\mathrm{k}-1)\mathrm{N}2^{-\mathrm{j}} \sim \mathrm{kN}2^{-\mathrm{j}}\} \tag{1}$$

where j = 0, 1, 2,…; k = 1,…, 2^j. Here N denotes the number of data points, j is the number of desired regions, and k indicates the index of regions at the jth regions. Fig 1. illustrates this multi-scale division of the frequency spectrum.

	$Seg = \{(k-1)N2^{-j} \sim kN2^{-j}\}$								k = 1,…,2^j
j = 0	Seg_1								k = 1
j = 1	Seg_2				Seg_3				k = 1, 2
j = 2	Seg_4		Seg_5		Seg_6		Seg_7		k = 1,2,3,4
j = 3	Seg_8	Seg_9	Seg_{10}	Seg_{11}	Seg_{12}	Seg_{13}	Seg_{14}	Seg_{15}	k = 1,2,3,4,5,6,7,8

Figure 1. The division into multi-scale frequency regions

For a selected segment of interest, the spectral RMS is calculated as follows (Kim et al., 2024):

$$\mathrm{X_{SRMS}} = \sqrt{\frac{1}{\mathrm{N_{seg}}}\sum_{\mathrm{n}\in\mathrm{Seg}} |\mathrm{X_n}|^2} \tag{2}$$

where X_n is the amplitude corresponding to frequency n within the segment in the frequency domain, and N_{seg} denotes the frequency length of this segment.

3. DATA-DRIVEN METHOD FOR ANOMALY DETECTION

3.1. Attention mechanism

The attention mechanism is a deep-learning technique that enables a model to focus on the most relevant parts of the input data. In general, it computes attention weights to different elements of an input sequence according to their importance for the target task, thereby adjusting their influence on the final representation (Vaswani et al., 2017).

The core concepts behind attention mechanism are query (Q), key (K) and value (V) vectors derived from the input data. The formal equation for Scaled Dot-product attention is defined as:

$$\mathrm{Attention(Q, K, V)} = \mathrm{softmax}\left(\frac{\mathrm{Q.K^T}}{\sqrt{\mathrm{d_K}}}\right)\mathrm{V} \tag{3}$$

Where $\mathrm{d_K}$ is the dimension of key vectors.

To enhance feature representation and model capacity, the multi-head attention (MHA) mechanism is employed. MHA allows the model to attend jointly to information from different representation subspaces at different positions, allowing it to capture relationships within the input sequence.

In MHA, the input is first linearly projected into h independent heads. Each head has its own learned weight matrices to generate distinct Q, K, V vectors. Scaled dot-product attention is then computed independently for each head. The outputs of all heads are concatenated and projected through a final learned weight matrix $\mathrm{W^0}$. The overall operation is expressed as:

$$\mathrm{MultiHead(Q, K, V)} = [\mathrm{head_1}; \ldots; \mathrm{head_h}]\mathrm{W^0} \tag{4}$$

where $\mathrm{head_i} = \mathrm{Attention}(\mathrm{QW_i^Q}, \mathrm{KW_i^K}, \mathrm{VW_i^V})$

3.2. Representation learning model

As illustrated in Figure. 2, the representation learning model consists of three stages. In the first stage, two one-dimensional convolutional layers extract local temporal features from the input sequence. The resulting feature sequences are then processed by long short-term memory (LSTM) units to capture temporal dependencies and generate hidden states over the entire sequence. Instead of using only the final hidden state of the last LSTM cell, all hidden states are passed to the MHA module in order to preserve richer temporal information. The attention-enhanced representations are then forwarded to the reconstruction stage. In the final stage, a second set of

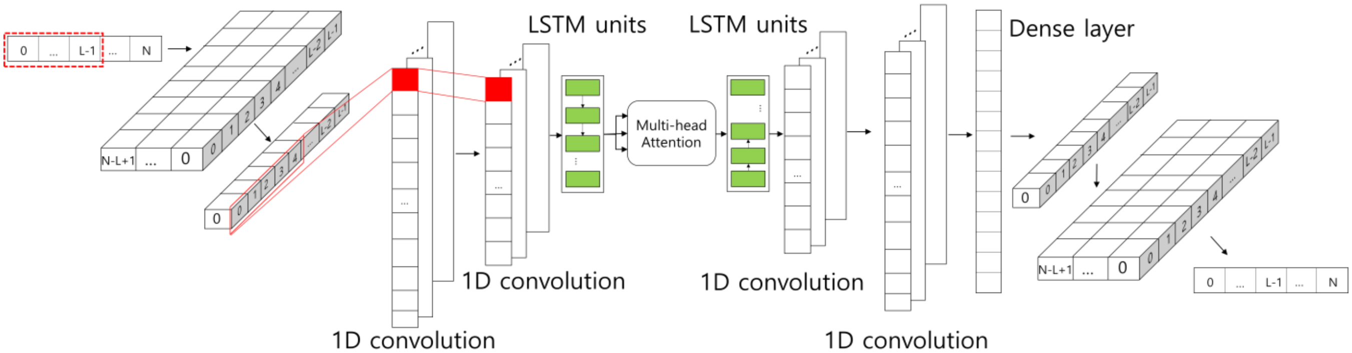
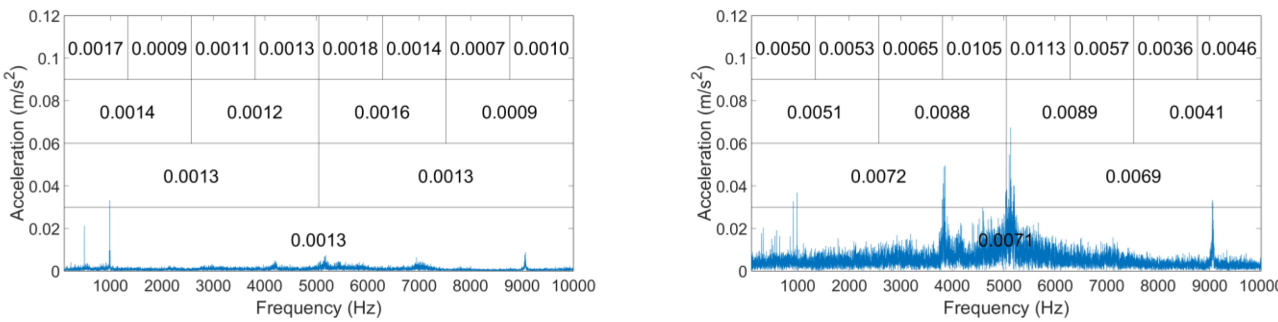


Figure 2. Representation learning framework



a) b)

Figure 3. Spectral RMS of bearing 3 (dataset 1): a) at the beginning and b) at the ending period

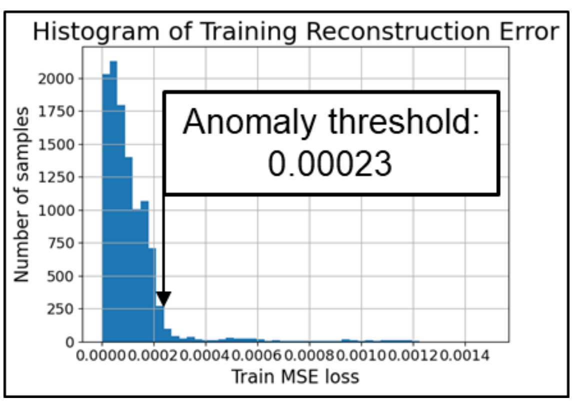


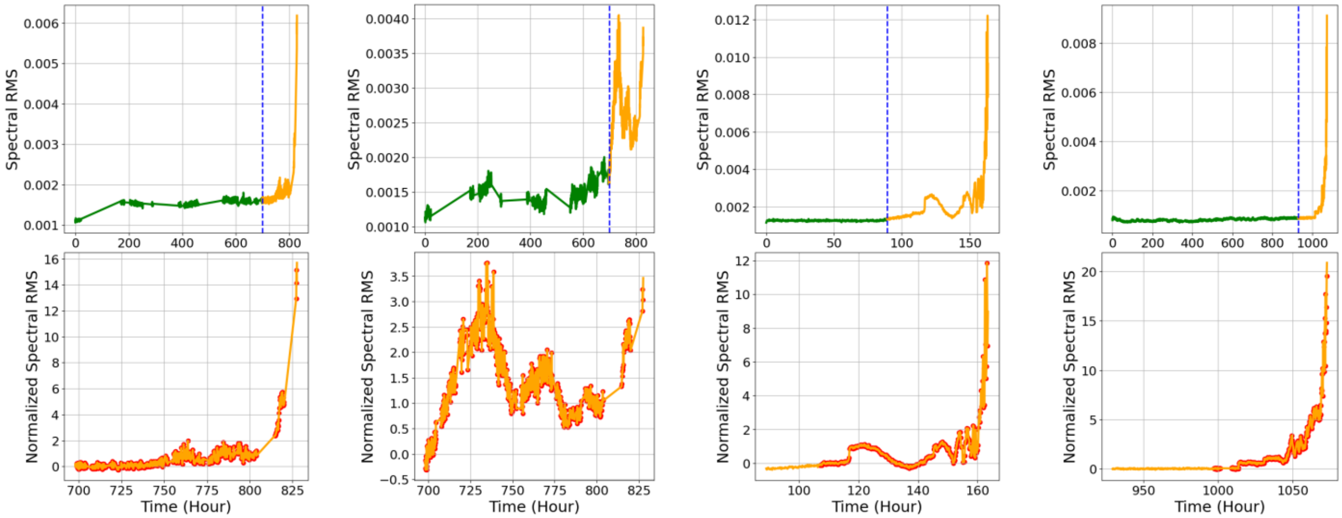
Figure 4. Training reconstruction error and statistical threshold selection

LSTM units reconstruct the latent representation back into a temporal sequence. Two convolutional layers further refine the reconstruction by recovering local temporal patterns. Finally, a fully connected layer maps the reconstructed sequence back to a single feature dimension, producing the output signal.

Instead of using the one-dimensional data directly, the spectral RMS sequences were segmented into overlapping subsequences using a sliding window approach with a window length L. These subsequences were then reshaped

Table 1. Structure and Parameters of Representation learning model (Total number of Parameters: 668,409)

Layers	Parameters	Output shape
Conv1d	kernel = 5 dilation = 1 padding = 'same' stride = 1	(1, 256, L)
Conv1d	kernel = 5 dilation = 1 padding = 'same' stride = 1	(1, 128, L)
LSTM	batch_first = 'true' dropout = 0.1	(1, L, 64)
Multi-head-Attention		(1, L, 64)
LSTM	batch_first = 'true' dropout = 0.2	(1, L, 128)
Conv1d	kernel = 7 dilation = 1 padding = 'same' stride = 1	(1, 256, L)
Conv1d	kernel = 7 dilation = 1 padding = 'same' stride = 1	(1, 60, L)
Linear		(1, L, 1)



a) b) c) d)

Figure 5. Spectral RMS and anomaly detection results of failing bearings at different datasets: a). Bearing 3 (dataset 1); b). Bearing 4 (dataset 1); c). Bearing 1 (dataset 2); d). Bearing 3 (dataset 3)

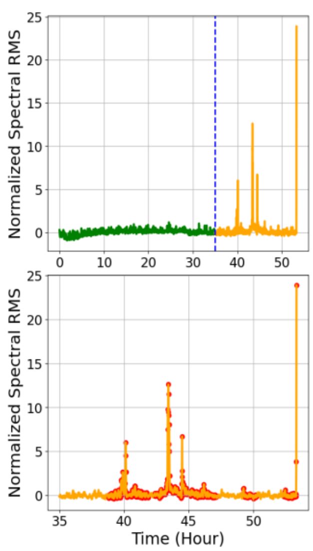


Figure 6. Spectral RMS of two lack of lubrication datasets and anomaly detection results of dataset 1

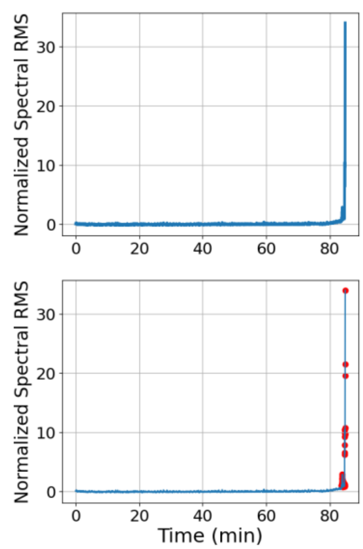


Figure 7. Spectral RMS of two lack of lubrication datasets and anomaly detection results of dataset 2

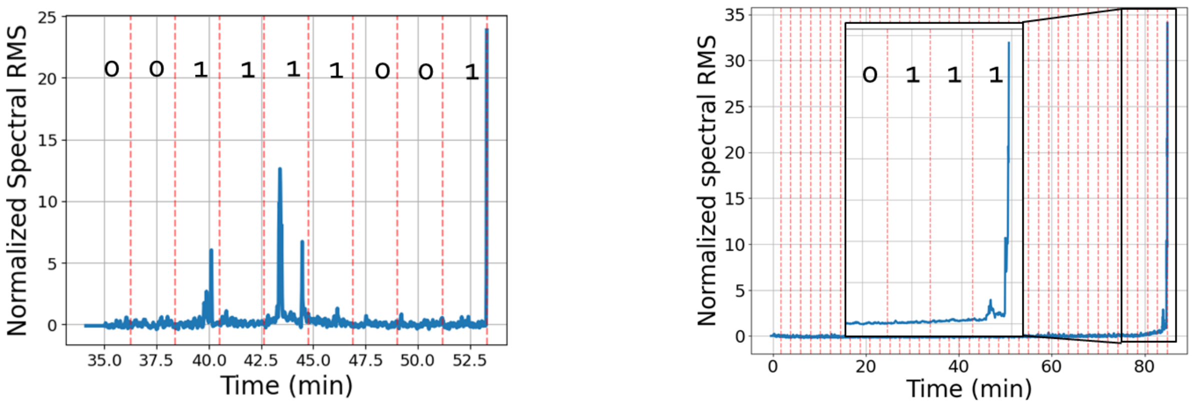
into a three-dimensional form to introduce a dimension compatible with CNN architecture.

In this study, mean square error (MSE) is employed to establish a statistical threshold, which is set as the 95th percentile of the training reconstruction error distribution. The MSE is defined as:

$$\mathrm{MSE}(\hat{y}) = \frac{1}{L}\sum_{i}^{L}(y_i - \hat{y}_i)^2 \tag{5}$$

where y_i and $\hat{y}_i$ are the actual and predicted values, respectively.

The model is trained in a representation learning-based framework in which the input sequence serves simultaneously as the target. Training proceeds by minimizing the mean squared error between the input and its reconstruction, thereby forcing the latent representation to encode salient temporal and structural information from the data. Parameter optimization is performed using Adam optimizer and Huber loss. Table 1 summarizes the parameters of the representation-learning model.



a) b)

Figure 8. Manually labeled segments on spectral RMS of insufficient lubrication

a). Last stage of dataset 1; b). Dataset 2

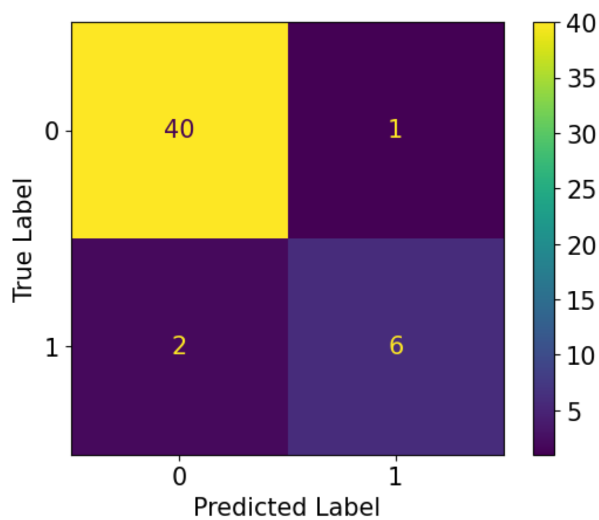


Figure 9. Confusion matrix of model classification performance

4. Experiment results and discussion

The proposed approach is validated by two practical experiments. The first experiment used the IMS bearing dataset (Lee, Qiu, Yu, Lin, & Rexnord Technical Services, 2007), which consists of run-to-failure tests under controlled operating conditions. The test rig included four bearings mounted on a shaft rotating at a constant speed of 2000 rpm under a radial load of 6000 lbs (~2721.55 kg). Vibration data were recorded as 1 s snapshots at a sampling rate of 20 kHz, resulting large-scale multichannel time-series data. The experiment comprised three datasets involving four failed bearings: bearing 3 and bearing 4 from dataset 1, bearing 1 from dataset 2, and bearing 3 from dataset 3.

The second experiment was conducted on a machining-center spindle under insufficient-lubrication conditions. Four accelerometers were installed to measure bearing vibration signals along the X and Y axes at the front and rear bearing positions. Spindle operated at a constant speed of 20000 rpm, and the signals were sampled at 25.6 kHz. This experiment was also performed under run-to-failure conditions. Two datasets were collected from the front bearings, and failure occurred at the end of each test when the spindle broke down.

4.1. Training procedure

In the IMS experiment, four bearing-failure cases from three datasets were analyzed for early fault detection. The frequency range of interest extended from three times the rotating frequency (approx. 100 Hz) to 10000 Hz. This range was divided into eight equal frequency segments (j =3). The suspicious segment was segment Seg_{10} (Figure. 3), corresponding to 2475 - 3712.5 Hz, as its spectral RMS exhibited significant variation over time.

In the second experiment, the desired frequency range was also extended from three times of rotating speed (approx. 1000 Hz) to 10000 Hz. Similarly, this range was divided into eight equal segments (j = 3), and the selected suspicious segment was again Seg_{10}, corresponding to 3250 – 4375 Hz.

The resulting spectral RMS sequence were then processed using a sliding-window approach with a window length of L = 60. The spectral RMS data obtained from four failed bearings were divided into training set (left side of the dashed line in Figure. 5) and test set (right side of the dashed line in Figure. 5). The training set consisted of data collected during the early stage, when the bearing remained in healthy condition. In contrast, the test set corresponded to later stages, during which fault progression and eventual failure occur. Similarly, the spectral RMS values from dataset 1 were partitioned into training and test sets (Figure. 6) according to the same criterion. This partitioning strategy ensured that the model learned only the statistical representation of normal bearing behavior, allowing deviations observed during testing phase could be interpreted as potential fault-induced anomalies.

After training, the anomaly threshold was defined as 2.3×10^{-4}, corresponding to 95th percentile of reconstruction errors obtained from normal training data. Any reconstruction error exceeding this value in test phase is classified as anomalous. Based on this decision boundary, anomalous spectral RMS values were identified as dots in Figure. 5 and 6. The detected anomaly indices for the spectral RMS sequences in the testing phase of the four bearing cases were 1787, 1557, 643, 5967 respectively (Table. 2). These results demonstrate earlier fault detection on the first three bearing cases compared with the approach previously reported by Ahmad, Styp-Rekowski, Nedelkoski, & Kao, (2020). Furthermore, the proposed model successfully detected the degradation process associated with insufficient lubrication at approximately the 39th minute of operation in the lubrication-deficient dataset.

Table 2. Comparison of the initial points that identifies as anomaly in the previous and proposed researches

Dataset	Initial detected sample numbers	
	Previous method	Proposed method
Set 1 bearing 3	2039	1787
Set 1 bearing 4	1704	1557
Set 2 bearing 1	667	643
Set 3 bearing 3	5241	5967

4.2. Model evaluation

The spectral RMS values of the second insufficient-lubricant dataset are presented in Figure. 7. This dataset was then applied as unseen data to the trained representation-learning model together with the previously determined threshold. As illustrated in Figure. 7, the model was capable of detecting the degradation process at an early stage of operation, at approximately the 83rd minute. The dots represent the anomalous of machining spindle.

To further quantify the model's performance, the insufficient-lubrication datasets were divided into multiple sub-datasets, each containing 128 data points. Each sub-dataset was manually labeled according to the following criteria: Normal = 0 and Abnormal = 1, as illustrated in Figure 8.

The final stage of Dataset 1 consisted of 9 labeled segments, including 5 abnormal segments, whereas Dataset 2 contained 40 labeled segments, of which 3 were labeled as abnormal.

Figure. 9 demonstrates the confusion matrix of the model classification results. Based on this results, evaluation metrics of the model, including precision, recall, F1-score and accuracy were calculated and summarized in Table. 3.

Although the obtained evaluation metrics indicate that the model achieved a reasonable level of classification performance, the relatively limited amount of testing data restricts the reliability and generalizability of the evaluation results. Therefore, additional experiments using larger and more diverse datasets are required to further validate the robustness and practical applicability of the proposed approach.

Table 3. Evaluation metrics

Precision	0	0.95
	1	0.86
Recall	0	0.98
	1	0.75
F1-score	0	0.96
	1	0.8
Accuracy		0.94

5. CONCLUSION

This study proposed an anomaly-detection framework for monitoring machining-spindle. The framework integrated a feature-extraction approach with a multi-head attention-based representation learning. Firstly, the feature-extraction method obtained the spectral root mean square of vibration signals in the selected frequency range. Secondly, the representation learning model was trained to reconstruct the spectral RMS sequences under normal operating conditions, and a statistical threshold was defined as 95th percentile of the reconstruction error distribution.

Experimental results on two practical case studies showed the effectiveness of the proposed framework. The model successfully detected failure conditions in all four faulty bearings for NASA IMS data and identified abnormal conditions prior to machine stoppage in spindle data collected under insufficient-lubrication conditions.

Nevertheless, the proposed framework currently relies solely on the spectral RMS feature. Additional informative features, such as kurtosis, power spectral density, crest factor, and other statistical indicators, could be extracted and fused into a unified health index to provide a more comprehensive representation of bearing degradation Moreover, the model remains purely data-driven. To improve prediction accuracy and generalization, future work may incorporate physics-based knowledge approaches such as physics-informed neural networks. In addition, representation-learning models may occasionally reconstruct abnormal samples too accurately, which can reduce anomaly-detection performance. To address this limitation, memory-augmented architectures provide a more effective alternative by constraining memory usage and improving anomaly detection.

REFERENCES

Ahmad, S., Styp-Rekowski, K., Nedelkoski, S., & Kao, O. (2020). Autoencoder-based Condition Monitoring and Anomaly Detection Method for Rotating Machines. Proceedings - *2020 IEEE International Conference on Big Data*, Big Data 2020, 4093–4102. doi:10.1109/BigData50022.2020.9378015

Gong, F., Ma, P., Zhang, H., Wang, C., Li, X., & Wu, Y. (2025). Rolling bearings remaining useful life estimation using digital twin and physics-informed methods with uncertainty quantification. *Engineering Applications of Artificial Intelligence*, 154(May), 111070. doi:10.1016/j.engappai.2025.111070

Huang, C. G., Huang, H. Z., Li, Y. F., & Peng, W. (2021). A novel deep convolutional neural network-bootstrap integrated method for RUL prediction of rolling bearing. *Journal of Manufacturing Systems*, 61, 757–772. doi:10.1016/j.jmsy.2021.03.012

Kim, S., Seo, Y. H., & Park, J. (2024). Transformer-based novel framework for remaining useful life prediction of lubricant in operational rolling bearings. *Reliability Engineering and System Safety*, 251(March), 110377. doi:10.1016/j.ress.2024.110377

Lee, J., Qiu, H., Yu, G., Lin, J., & Rexnord Technical Services. (2007). 'Bearing Data Set', *NASA Ames Prognostics Data Repository*. IMS, University of Cincinnati.

Li, H., Zhang, Z., Li, T., & Si, X. (2024). A review on physics-informed data-driven remaining useful life prediction: Challenges and opportunities. *Mechanical Systems and Signal Processing*, 209(January), 111120. doi:10.1016/j.ymssp.2024.111120

Mei, J., Jia, J., Zeng, R., Zhou, B., & Zhao, H. (2016). A multi-order FRFT self-adaptive filter based on segmental frequency fitting and early fault diagnosis in gears. *Measurement: Journal of the International Measurement Confederation*, 91, 532–540. doi:10.1016/j.measurement.2016.05.092

Shang, Y., Tang, X., Zhao, G., Jiang, P., & Ran Lin, T. (2022). A remaining life prediction of rolling element bearings based on a bidirectional gate recurrent unit and convolution neural network. *Measurement: Journal of the International Measurement Confederation*, 202(December 2021), 111893. doi:10.1016/j.measurement.2022.111893

Shen, S., Lu, H., Sadoughi, M., Hu, C., Nemani, V., Thelen, A., … Kenny, S. (2021). A physics-informed deep learning approach for bearing fault detection. *Engineering Applications of Artificial Intelligence*, 103, 104295. doi:10.1016/j.engappai.2021.104295

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention Is All You Need. *Advances in neural information processing systems*, 30. Retrieved from https://arxiv.org/pdf/1706.03762

BIOGRAPHIES

Jooho Hwang is currently a Principal Researcher at the Korea Institute of Machinery and Materials (KIMM), Daejeon, South Korea, in the Dept. of Ultra-Precision Machines and Systems. He received his B.S. degree from Inha University, South Korea, in 1993, and his M.S. degree from the Korea Advanced Institute of Science and Technology (KAIST), South Korea, in 1995. He obtained his Ph.D. degree from KAIST in 2006. He joined KIMM as a Researcher in 1995, became a Senior Researcher in 2000, and has been serving as a Principal Researcher since 2009. His research interests include machine tools metrology, precision machine design, and precision machine control. His recent research focuses on diagnostic and predictive maintenance technologies for machine tools

Dung Minh Nguyen was born in Vinh Phuc, Vietnam, on October 19th, 1996. He is currently a Ph.D student at University of Science and Technology, Daejeon, South Korea. He received his B.S degree in Transportation Mechanical Engineering from Hanoi University of Science and Technology, Hanoi, Vietnam in 2019 and Master degree at mechanical engineering from University of Science and Technology, Daejeon, South Korea in 2024. His current research includes predictive maintenance in machining-center spindle and applied machine learning.

Sung Joon Kim received the B.S. degree in mechanical engineering from Sungkyunkwan University, Suwon, South Korea, in 2015, and the Ph.D. degree in mechanical engineering from the same university in 2021. His doctoral research focused on slip detection systems for robotic skin. From 2021 to 2024, he worked as a Manager at LS Electric. He is currently a Senior Researcher at the Korea Institute of Machinery and Materials (KIMM), Daejeon, South Korea. His research interests include AI-based condition monitoring, predictive maintenance, deep learning for fault diagnosis, and collaborative robotics.