Size Estimation of Flaking in Rolling Bearings Using Deep Learning with Explainability

Osamu Yoshimatsu[1], Keiichirou Taguchi[2], Yoshihiro Sato[3], and Takehisa Yairi[4]

[1,2,3] *NSK Ltd., Kanagawa, 251-8501, Japan*
yoshimatsu-o@nsk.com
taguchi-kei@nsk.com
satou-yos@nsk.com

[1,4] *The University of Tokyo, Tokyo, 153-8904, Japan*
yairi@g.ecc.u-tokyo.ac.jp

Abstract

To improve the availability of rotating machines such as wind turbines, where rolling bearing replacement is costly and time-consuming, it is effective to estimate the damage progression of the rolling bearings. As one of the damage progressions, the size of flaking in rolling bearings is estimated by vibration analysis using rule-based methods. However, these rule-based methods require expert knowledge of rolling bearings. Therefore, an estimation model using deep learning was proposed and its performance was evaluated. Furthermore, it was verified that the proposed model had extracted the features of physical phenomena using Grad-CAM.

1. Introduction

Monitoring and diagnosing rotating machinery and performing maintenance at the appropriate stage is necessary to improve the operating efficiency of rotating machinery. One of the most critical maintenance targets is the fault of rolling bearings, a mechanical element loaded in rotating machinery parts. Failure of rolling bearings can take many forms, but the most common type of failure is 'flaking,' where the defect occurs when part of the raceway of the rolling bearing flakes off. Diagnosis of flaking involves detecting periodic shocks in bearing vibrations (Randall and Antoni (2011)).

One development of rolling bearing flaking diagnosis is estimating the remaining useful life by estimating the flaking size from vibrations. Flaking size progresses with continued operation after flaking, causing severe problems due to the machinery's rotational accuracy, vibration, and acoustics. Avoiding this severe damage by the diagnosis benefits ma-

chinery with high maintenance costs, such as wind turbines. A typical method for estimating the flaking size is to detect the vibration of the rolling elements of the bearing as they enter and exit the flaking and to calculate the interval between the two(Sawalhi and Randall (2011)).

However, estimating flaking size from vibration on a rule-based basis is time-consuming and financially costly, as it usually requires trial and error with high expertise. High expertise in installing appropriate vibration sensors and tuning the parameters of noise reduction methods in the vibration improves the detectability of feature vibrations for flaking size estimation. In the feature vibrations in flaking size estimation, the difficulty of detecting vibrations, particularly those of rolling elements entering the flaking, is mentioned in Smith, Hu, Randall, and Peng (2015). F. Zhang, Huang, Chu, and Cui (2020) focus on the problem of overlapping vibrations of multiple rolling elements with expanding flaking size, making it challenging to detect the feature vibrations.

We propose a flaking size estimation model based on deep learning and evaluate its performance to avoid diagnostic costs due to the high expertise. Deep learning models of rolling bearing vibration diagnosis have been studied to classify mainly flaking damaged parts. W. Zhang, Peng, Li, Chen, and Zhang (2017) proposed a WDCNN using 1d vibration acceleration waveforms as input and applying large kernel sizes to shallow layers. Lu et al. (2023) proposed a PICNN with the envelope spectrum of the vibration acceleration waveform as input, weighted along the impact vibration period in the case of bearing flaking. This study proposes a CNN-LSTM model with 1d vibration acceleration waveform and its integrated velocity corresponding waveform as parallel inputs. To verify the performance of the proposed flaking size estimation model, a dataset containing vibrations of various flaking sizes was made and used. The dataset measured the acceleration of

an artificial defect machined into the inner ring of a cylindrical roller bearing when it was allowed to progress under high load conditions for an extended period.

We also verified the explainability of the proposed model in estimating the flaking size. As an explanation, we verified in Grad-CAM whether the vibration features are related to the estimation results in a way equivalent to the rule-based method with high expertise. Explanability has been studied in deep learning vibration diagnosis of rolling bearings, mainly in classifying flaking-damaged parts, by confirming whether periodic shocks, equivalent to rule-based methods, are related to the classification results. Li, Zhang, and Ding (2019) added an Attention mechanism to a model for classifying flaking-damaged parts using 1d vibration acceleration waveforms as input. They confirmed that Attention is higher for periodic shock vibration. Chen, Liu, He, Liu, and Zhang (2022) proposed GS-CAM, which provides a richer visualization of the relationship between periodic shocks in the deep learning model and the classifying results. In this study, Grad-CAM is used to visualize that the results of the proposed model's flaking size estimation are related to the vibration of the rolling element as it enters and exits the flaking, similar to the rule-based model.

2. PROPOSED METHOD

2.1. CNN-LSTM Flaking Size Estimation Model

Figure 1 gives an overview of the proposed model. The proposed model consists of an acceleration feature extractor(FEA), a velocity feature extractor(FEV), a composite feature extractor(FEC), an LSTM layer, and a regressor. The CNN layer of all feature extractors consists of a Convlution layer, Average pooling layer, and SE-block, with Swish as the activation function. The parallel inputs to the model are the 1d vibration acceleration waveform and the velocity equivalent waveform integrated from it, respectively. The parallel inputs are used to extract the feature vibrations required for flaking size estimation and to verify the explainability using Grad-CAM. The feature vibrations required for rule-based flaking size estimation correspond to the velocity-equivalent signal for the low-frequency range of the entry vibration of the rolling element into the delamination and the acceleration signal for the broadband range of the exit vibration.

2.2. Explanability Verification Method

Grad-CAM was applied to the proposed trained model to verify whether the vibration of the rolling element entering and exiting the bearing flaking size estimation model contributes to the estimation results. Grad-CAM is a method for visualizing the parts of the input data that contribute to the output results in deep learning models and is mainly used for image models. In this study, Grad-CAM was applied to the final layer of the velocity and acceleration feature extractors (FEA
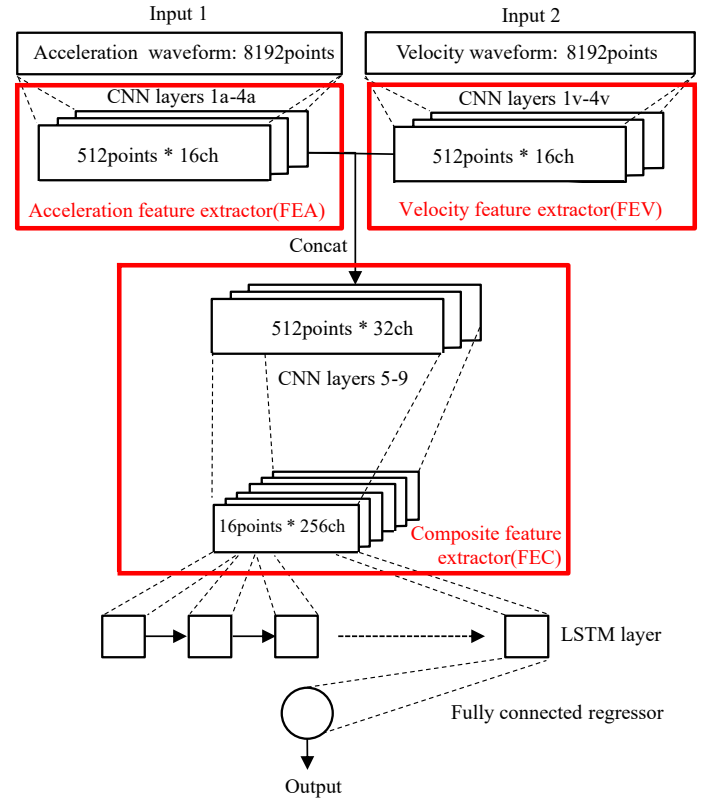

Figure 1. Overview of proposed flaking size estimation CNN-LSTM model.

and FEV respective fourth layer outputs) to verify the contribution of feature vibrations similar to rule-based flaking size estimation.

2.3. Various Flaking Size Dataset

To train and test the proposed model, we used a vibration dataset from a test rig using a cylindrical roller bearing NU2228EM with an artificial defect machined on the inner ring to simulate flaking. This data set contains vibration data for various flaking sizes during the long-term operation of the test rig containing the bearing with the machined defects. The test rig was stopped several times during the long-term operation when flaking size were measured, and vibration data were acquired under several conditions with loads and rotational speeds. Table 1 describes the operating and measurement conditions of the test.

2.4. Verification

To validate the performance of the proposed model in estimating the separation size, we performed a five-fold cross-validation in which the dataset was split into five parts, and the train-test data was combined 4:1 a total of five times. In addition, the train-test data were split under conditions where

Table 1. Operating and measurement conditions of test for the various flaking size dataset

Bearing number	NU2228EM
Bearing type	cylindrical roller
Rotational speed	3 speeds(1200, 1500, 1750min^{-1})
Load	7 radial loads(29.3k - 205kN)
Test pieces num.	2
All condition num.	220
Sensor direction	2 directions(Radial, Axial)
sampling frequency	96kHz
sampling time	30s

Figure 2. Comparison of estimates and true values for each operating condition in five-fold cross validation tests.

the real-time of the rolling elements entering and exiting the flaking was not matched. The vibration data under each operating condition was split into frames of 8,192 points each, and an equal number of frames were randomly selected from the files for each operating condition for a total of 61,772 frames used for training. The output labels of the model during the training were set to the following equation.

$$Y = -\ln(\frac{P_{flaking}}{P_{frame}}) \qquad (1)$$

Y is the ture output value, $P_{flaking}$ is the number of data points equivalent to the grand-truth flaking size, and P_{frame} is the number of data points for the entire frame.

Table 2 shows the training conditions. All frames divided from the entire file under each operating condition were inferred in sequence in the inference using the test data. The average value of the entire output was used as the result of the flaking size estimation.

Table 2. Training conditions

Epoch num.	50
Learning rate	2.5e-4 to 1.0e-3(WarmUp)
Train frames num.	61,776
Mini-batch num.	32

3. RESULTS

3.1. Flaking Size Estimation Performance

The results of the five-fold cross-validation show the estimated and actual flaing lengths for each operating condition of the test data, as shown in Figure 2. In this figure, the estimated and actual flaking sizes are expressed as a ratio to the pitch of the rolling element spacing in order to compare the results under all operating conditions. The difference between the estimated and grand-truth values is mostly within 0.15 pitch of the rolling element pitch, except for the operating condi-
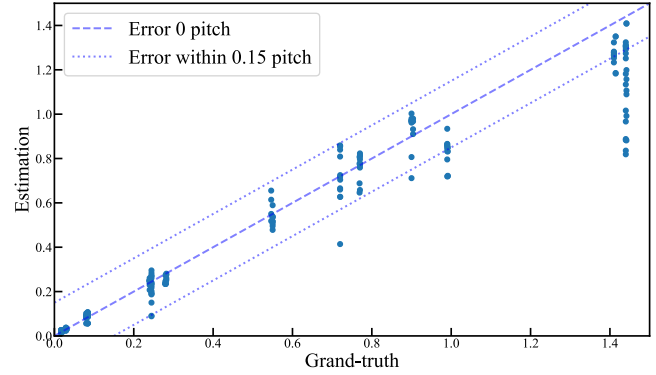
tions where the flaking size greatly exceeds one pitch of the rolling element pitch. The accuracy of the estimation results for test data with a flaking size exceeding 1.4 pitches tends to be low, which can be assumed to be due to the test conditions being the closest to extrapolation. However, in flaking size estimation, the effect of low extrapolation performance is limited because estimation accuracy of less than one pitch is essential for bearing remaining useful life estimation. These results verified the construction of a deep learning model that can accurately estimate the flaking size from rolling bearing vibration.

3.2. Explanability

Figure 3 shows an example of grad-cam results when a test data frame is an input to the trained proposed model. The test data in this example has a flaking size of 0.77 pitch and a rotational speed of $1,200min^{-1}$, which means that the interval between feature vibrations in the input data is approximately 500 points. The importance of the final layer output of the velocity feature extractor is high for the vibration of the rolling body entering the flaking. In contrast, the importance of the final layer output of the acceleration feature extractor is high for the impact vibration of the rolling body exiting the flaking. The results show that the same vibration features as in the rule-based flaking size estimation method with high expertise are related to the estimation results in the proposed model.

4. CONCLUSION

We proposed a CNN-LSTM model for estimating the flaking size of rolling bearings from vibration and verified its performance using a dataset with progressed artificial defects. The verification results showed that the size estimation was highly accurate except for the case of long flaking size. Grad-CAM was also performed to verify the explainability of the proposed model, and the essential vibration features in the rule-based model were also essential in estimating the proposed model. This estimation following physical phenomena
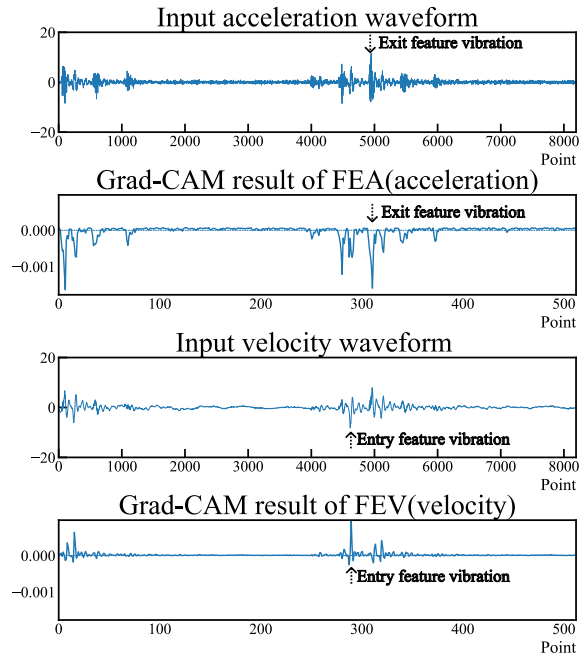
Figure 3. Example of Grad-CAM results for a trained model.

is the explainability of the proposed model. On the other hand, the proposed model has the same problem of training data availability as the conventional deep learning model for rolling bearing diagnostics, so a method for training models with higher generalization performance using transfer learning and other method is needed.

REFERENCES

Chen, B., Liu, T., He, C., Liu, Z., & Zhang, L. (2022, June). Fault diagnosis for limited annotation signals and strong noise based on interpretable attention mechanism. *IEEE Sens. J.*, *22*(12), 11865–11880.

Li, X., Zhang, W., & Ding, Q. (2019, August). Understanding and improving deep learning-based rolling bearing fault diagnosis with attention mechanism. *Signal Processing*, *161*, 136–154.

Lu, H., Pavan Nemani, V., Barzegar, V., Allen, C., Hu, C., Laflamme, S., . . . Zimmerman, A. T. (2023, May). A physics-informed feature weighting method for bearing fault diagnostics. *Mech. Syst. Signal Process.*, *191*, 110171.

Randall, R. B., & Antoni, J. (2011, February). Rolling element bearing diagnostics—a tutorial. *Mech. Syst. Signal Process.*, *25*(2), 485–520.

Sawalhi, N., & Randall, R. B. (2011, April). Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size. *Mech. Syst. Signal Process.*, *25*(3), 846–870.

Smith, W. A., Hu, C., Randall, R. B., & Peng, Z. (2015). Vibration-Based spall size tracking in rolling element bearings. In *Proceedings of the 9th IFToMM international conference on rotor dynamics* (pp. 587–597). Springer International Publishing.

Zhang, F., Huang, J., Chu, F., & Cui, L. (2020, December). Mechanism and method for the full-scale quantitative diagnosis of ball bearings with an inner race fault. *J. Sound Vib.*, *488*, 115641.

Zhang, W., Peng, G., Li, C., Chen, Y., & Zhang, Z. (2017, February). A new deep learning model for fault diagnosis with good Anti-Noise and domain adaptation ability on raw vibration signals. *Sensors*, *17*(2).