Anomaly Detection in Mechanical Vibration Using Combination of Signal Processing and Autoencoder

Ayaka Matsui¹, Shota Asahi², Satoshi Tamura¹, Satoru Hayamizu¹,
Ryosuke Isashi³, Akira Furukawa³ and Takayoshi Naitou³

¹ Graduate School of Natural Science and Technology, Gifu University

² Graduate School of Engineering, Gifu University

1-1 Yanagido, Gifu City, Gifu 501-1193, Japan

³ KYB Corporation, World Trade Center Bldg. 11F

2-4-1 Hamamatsu-cho, Minato-ku, Tokyo 105-6111, Japan

E-mail: ayaka@asr.info.gifu-u.ac.jp, tamura@asr.info.gifu-u.ac.jp

Abstract

The goal of this research is to detect abnormal behavior in vibration data of factory machinery and take measures against failure. In this study, an accelerometer is mounted on a mechanical device with a conveyor belt. Vibration data are collected to monitor the conveyor belt line. We propose a method of detecting vibration abnormalities by combining signal processing and an autoencoder (AE), one of the neural network models. In the learning phase, the vibration signal is converted to a spectrogram and used for the output of the neural network. Then, a random mask is applied to the horizontal direction of the spectrogram. This technique does not require a search for a valid frequency band. The masked spectrogram is used as the input for the neural network. A model that converts the masked spectrogram back to the original spectrogram is trained. This model is a type of AE, a deep convolutional encoder-decoder architecture. In the detection stage, the masked spectrogram is input to the model and a predicted image is obtained. The predicted images are evaluated with the weighted moving variance of the PSNR. In an experiment, a normal AE and two masked AEs were compared. Both the AE and the proposed method can detect faults, and it was shown that other faults can be detected by a masked AE.

1. Introduction

1.1 Background

Machines are an important component of the manufacturing floor. If a machine breaks down, it may cause significant losses, such as reduced production. However, determining whether a machine is operating abnormally depends on the intuition and experience of the operator, and it is difficult for human eyes to detect abnormalities. However, by detecting abnormalities from the vibrations generated by the equipment

and taking measures to prevent machine failures, the productivity of a factory should be able to be increased. In recent years, factories have required more advanced maintenance than ever before, and efforts are currently underway to collect various data from devices, including sensor data. These data are collected for use in statistical analysis and machine learning, and are used for a wide range of purposes, including product quality control and equipment failure prediction. In this study, vibration data are collected to monitor a conveyor belt line. An accelerometer is mounted on a mechanical device with a conveyor belt.

1.2 Related works

Anomaly detection is a task to detect anomalies in data by modeling training data. It has a wide range of applications, such as fraud detection, medical anomaly detection and video surveillance. A statistical method assumes a distribution of the data and detects outliers by an anomaly score based on statistical theory. Machine learning algorithms, such as clustering and Gaussian mixture models, are used for the modeling of normal data. For industrial laser cutting machines, a machine learning approach is used for cutting quality estimation [1]. For modeling the quality from sensor data, Gaussian mixture models, recurrent neural networks and convolutional neural networks are used in supervised settings.

Recently, various researchers have conducted studies on using deep learning to detect anomalies [2]. A denoising AE is a neural network model that is trained locally to denoise corrupted versions of the input [3]. In speech enhancement, a denoising AE is used for extracting robust features [4]. In computer vision, data augmentation by cutting out a region from an image [5] is used. For a spectrogram of sound, masking a region means cutting out time and frequency, therefore a data augmentation method by masking the spectrogram is proposed in [6]. We propose to combine an AE and signal

processing to detect abnormal vibrations. We use unsupervised training where the model can be trained using only normal vibrations.

1.3 Purpose of the research

The purpose of this research is to develop a method of detecting abnormal behavior from vibration by combining signal processing and an AE. The vibration data are converted to a spectrogram and used as image information. The spectrogram is masked for data augmentation. The masked spectrogram is input and a neural network model that predicts the original spectrogram is created. Deep learning technology has recently been developed and aims to incorporate this technology to create models that can detect anomalies with greater accuracy. Learning the neural network model requires the same amount of normal and abnormal data to identify whether the device is normal or abnormal. However, the abnormal data are more difficult to acquire than the normal data because they can only be acquired when the device is in an abnormal state. Therefore, the model is trained on the feature space of training data under the assumption that most or all of them are normal data. Both the normal and abnormal data are used for inference to determine the accuracy with which the proposed method detects abnormalities.

2. Feature Extraction

A masked spectrogram is used as the feature value. A spectrogram is a diagram that shows how a spectrum changes over time. Each datum is represented in three dimensions: time, frequency and power. The spectrogram used in this study has time on the horizontal axis and frequency on the vertical axis, and uses color to represent the power in each frequency band. To clarify the difference in power, we consider the logarithm. Next, the spectrogram is masked horizontally. The masking position is changed to a random position at each learning epoch. Spectrogram values are converted to values between 0 and 1 and the range of mask values is set to 0.5. Figure 1 shows the original spectrogram (A) and masked spectrograms (B, C).

Figure 1: A: Original spectrogram, B: Spectrogram masked by 1/32 width in vertical direction, C: Spectrogram masked by 1/8 width in vertical direction

3. Proposed Method

Figure 2 outlines the proposed method. Spectrogram images are created from vibration data and used as features. When creating a spectrogram image, the spectrogram is masked in the horizontal direction. In the training phase, a model that converts the masked spectrogram back to its original form is trained. This model is a type of AE and has a deep convolutional encoder–decoder architecture. We named this model the masked Autoencoder (masked AE). In the anomaly detection phase, the masked spectrogram of the vibration is input to the model and the original spectrogram is predicted. The deviation between the predicted and original spectrograms is quantified to detect anomalies. For normal vibrations, the output produces a spectrogram similar to the original spectrogram of the periodic peak, and for abnormal vibrations, the output has some differences that represent the condition of the device and the machine line of the conveyor belt.

Figure 2: Overview of the proposed method

3.1 Procedure

The procedure for the proposed method is as follows. Note that the numbers used here correspond to those in Figure 2.

1. Create a spectrogram image from the vibration waveform.
2. Apply a horizontal mask to the spectrogram image created in step 1.
3. Perform the masked AE model learning with the masked spectrogram image as the input and the original spectrogram image as the output.
4. Perform steps 1 and 2 on the new vibration waveform data, and input the masked image created in step 2 into the model built in step 3.

- Evaluate the abnormality by comparing the predicted image with the correct image.

3.2 Evaluation method

In this paper, we use the weighted moving variance obtained by the peak signal-to-noise ratio (PSNR) as the abnormality. The PSNR is calculated from the original and predicted spectrograms. The weighted moving variance is calculated by rearranging the PSNR in a sequential order. The window size of the weighting is 161 frames and the shift is 1 frame. Then, 3% of the top and 3% of the bottom of each frame are removed to remove outliers. The variance is obtained from the PSNR in the frame.

4. Experiment

We compare the masked AE with the normal AE. Two different mask widths are used for the masked AE.

4.1 Data

Table 1 shows the conditions for collecting acceleration data. The conveyor broke once during data collection and the conveyor was rapidly roughly repaired. Then, the chain was replaced and cut. After replacing the chain, the condition of the chain became unstable and the overall length tended to increase. Therefore, after a short time, the chain was expected to cut. In this experiment, the data after adjusting the chain were set as normal data, and the data before the failure were set as abnormal data. Figures 3 and 4 show graphs in which the normal and abnormal data are plotted, respectively. Some of the regular data after chain adjustment were used to train the data. All the data were used for test data.

Table 1: Conditions for collecting acceleration data

Sampling frequency	12.8 kHz
Data length	10 s

Figure 3: Normal data

Figure 4: Abnormal data

4.2 Experimental conditions

When learning a regular AE, the original spectrogram is used as both the input and output images. We prepare masks with two widths and investigate the effects of changing the mask width. A mask with a thickness of 1/32 or 1/8 is applied to the input image. Table 2 indicates the parameters of the spectrogram image and Table 3 indicates the learning parameters. The number of epochs is set to 301 for the normal AE and the masked AE with a mask width of 1/32, and to 1001 for the masked AE with a mask width of 1/8. Figure 5 shows the model structure.

Table 2: Spectrogram image parameters

Maximum frequency	6.4 kHz
Data length	7,875 ms (100,800 points)
Frame length	70 ms (896 points)
Frameshift length	35 ms (448 points)
Window function	hamming

Table 3: Model learning conditions

Learning data	2,839
Input size	$224 \times 224 \times 1$
Batch size	112
Epochs (normal AE)	301
Epochs (1/32 masked AE)	301
Epochs (1/8 masked AE)	1,001
Optimizer	Adam
Loss function	$MSE * (\tanh(y_{true} * 4 + 0.5))$
Framework	Keras (2.2.4) using TensorFlow backend

Figure 5: Model structure: filter size of convolution: 16*2, 32,*2, 64*3, 128*3, 128*2, 64*3, 32*2, 16*2 (from the left)

4.3 Results and discussion

Figure 6 shows the abnormality with the normal AE, Figure 7 shows the abnormality with the masked AE with the mask width of 1/32 and Figure 8 shows the abnormality with the

masked AE with the mask width of 1/8. A red horizontal line indicates a threshold on the 90% of whole abnormality. From Figures 6, 7, and 8, the normal AE and the two masked AEs all show a high abnormality above the threshold before failure. After the threshold is exceeded, failures can be predicted to occur soon. In the abnormality graph of the normal AE, the values gradually increase before failure. This indicates that the condition of the conveyor gradually deteriorated before the failure. Also, the abnormality from failure to chain replacement is higher than the normal data. This indicates that the condition of the conveyor is not completely restored by the rough repair. In the abnormality graph of the masked AEs, the abnormalities increase temporarily after changing and adjusting the chain. This indicates that the chain does not mesh well for a short time after it is replaced and cut. Comparing the abnormality graphs of the masked AEs, it can be seen that the larger the mask width, the greater the pre-failure anomalies and the lower the change in the value owing to unknown errors (green line). We consider that masking makes it more difficult to reconstruct spectrograms and makes more susceptible to smaller accidents.

Figure 6: Abnormality with normal AE: The four vertical lines indicate days when the conveyor failed and was roughly repaired, when the chain was replaced, when the chain was adjusted and when the value changed owing to unknown errors (from the left). The red area (left of leftmost line) indicates the data that seem to be abnormal before the failure. The yellow area (right of the 3rd line from left) indicates the data that seem to be normal after chain adjustment. The green area (part of the normal data) indicates the data used for learning

Figure 7: Abnormality with masked AE with mask width of 1/32

Figure 8: Abnormality with masked AE with mask width of 1/8

5. Conclusions

In this paper, we proposed a method of detecting anomalies by using vibration data for feature extraction, and a method of detecting conveyor belt failures by using a machine learning architecture. The vibration data are converted to a spectrogram and used as image features. A horizontal mask is applied to the spectrogram. The proposed method has the advantage in that the mask is applied randomly and there is no need to search for valid frequency bands. An experiment was performed to compare a normal AE and two masked AEs, all of which showed promising performance for predicting failure. We also showed that another anomaly could be detected by masking. As a future task, we intend to collect more data to see whether we can detect more failure patterns. In addition, we set the threshold to only 90% of whole abnormality in this paper. It is necessary to set a value that can detect any failure.

References

- [1] G. Santoline, P. Rota, D. Gandolfi and P. Bosetti: Cut quality estimation in industrial laser cutting machines: A machine learning approach, CVPR Workshop, 2019.
- [2] R. Chalapathy and S. Chawla: Deep learning for anomaly detection: A survey, arXiv:1901.03407v2, 2019.
- [3] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio and P. A. Manzagol: Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion, Journal of Machine Learning Research, Vol.11, pp.3371-3408, 2016.
- [4] X. Lu, Y. Tsao, S. Matsuda and C. Hori: Speech enhancement based on deep denoising autoencoder, Interspeech, pp.436-440, 2013.
- [5] T. DeVries and G. W. Taylor: Improved regularization of convolutional neural networks with cutout, arXiv: 1708.04552, 2017.
- [6] D. S. Parks et al.: Spec augment: A simple data augmentation method for automatic speech recognition, arXiv:1904.08779, 2019.