Development of Fault Diagnosis Model based on Semi-supervised Autoencoder

Yongjae Jeon[1], Kyumin Kim[2], YeLim Lee[3], Byeong Kwon Kang[4], and Sang Won Lee[5]

[1,2,3,4] *Department of Mechanical Engineering, Sungkyunkwan University, Suwon-si, Gyeonggi-do, 16419, Republic of Korea*
rhrhgudwp@g.skku.edu
aiden7@g.skku.edu
dldpfla1024@g.skku.edu
kevin0730@g.skku.edu

[5] *School of Mechanical Engineering, Sungkyunkwan University, Suwon-si, Gyeonggi-do, 16419, Republic of Korea*
sangwonl@skku.edu

ABSTRACT

The maintenance paradigm based on PHM (Prognostics and Health Management) technology, utilizing big data to predict process conditions through manufacturing intelligence, is rising. However, in most industries, there is lack of accurate labeling of sensor data, posing challenges in data utilization due to the significant cost of labeling tasks. Consequently, recent research has focused on semi-supervised learning methodologies, which are applicable in label-absent scenarios. Especially, there is a growing emphasis on semi-supervised autoencoder, which learns both labeled and unlabeled data simultaneously. Also, there is a demand for the development of fault diagnosis models for essential components, such as bearings in most mechanical systems. Vibrational data is actively being integrated with artificial intelligence for application in bearing fault diagnosis frameworks. Nonetheless, diagnosing the condition of bearings inside machine systems, especially within the machine tool spindle, remains challenging, as the labeling of collected data causes significant costs. Therefore, this paper aims to develop a fault diagnosis model for unlabeled bearings in machine tool spindle using a semi-supervised autoencoder. Initially, a monitoring system of bearing simulator that imitates a machine tool spindle bearing was constructed, and vibration signals from both normal and fault bearings were collected based on this system. Subsequently, a semi-supervised autoencoder model was developed to construct a fault diagnosis model using labeled simulator data and unlabeled machine tool spindle bearing data. To evaluate the model, additional data of normal and fault bearings in machine tool spindle were collected, and the performance of the model was compared with a conventional fault diagnosis model based on 1D-CNN.

1. INTRODUCTION

In manufacturing, machine system faults not only degrade the quality of the products but also lead to downtime, resulting in significant costs. Therefore, Prognostics and Health Management (PHM) techniques utilizing sensor data and artificial intelligence are widely used to monitor equipment and diagnose failures. In particular, deep learning methods, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), are used to enhance the reliability of failure diagnosis (Wen, Li, Gao & Zhang, 2017; Ince, Kiranyaz, Eren, Askar & Gabbouj, 2016; Cabrera, Guamán, Zhang, Cerrada, Sanchez, Cevallos, Long & Li, 2020; Abed, Sharma, Sutton & Motwani, 2015). However, acquiring a large amount of high-quality data for deep learning training is challenging in industries. While data are collected in various processes, almost data lack labels due to the high cost of labeling. Therefore, although data are abundant, distinguishing between normal and faulty states is difficult, posing challenges for developing failure diagnosis models. Particularly bearings, the core components of machine tools such as spindles, disturb the stable operation of the spindle when they break down because real-time confirmation and labeling of bearing failure are difficult.

Therefore, this paper proposes the utilization of data from a similar domain system and a semi-supervised autoencoder model to diagnose faults in unlabeled bearings of machining spindle bearings. First, a testbed capable of simulating the rotational motion of machining spindle bearings is constructed to create an environment for collecting enough labeled bearing data. Then a semi-supervised autoencoder structure, which can learn unlabeled real-world data and labeled testbed data, is used for fault diagnosis model. By learning unlabeled data, domain difference between the actual equipment and the testbed is reduced, allowing for the extraction of generalized features and thus the development of a fault diagnosis model applicable to actual equipment. To verify its effectiveness, the performance of the proposed model is compared with that of a conventional model trained only on testbed data.

This paper introduces the proposed model and its applications in Chapter 2. Chapter 3 describes the data collection process for machining spindle bearings and the testbed. Modeling is performed in Chapter 4, and a comparison of the performance between the conventional model and the proposed model is also conducted.

2. METHODS

2.1. Autoencoder

Autoencoder is one of the unsupervised learning algorithms, that has an Encoder neural network that reduces the dimension of input data, and a Decoder neural network that reconstructs the input data from the reduced dimensions. Both networks are connected through the latent space, to reconstruct the original data from the latent space where features of the data are preserved. Typically, the Mean Squared Error (MSE) is used as the loss function to minimize the difference between the input data and the reconstructed data. Hinton and Salakhutdinov (2006) confirmed Autoencoder is available for data dimensionality reduction, and Kingma and Welling (2013) proposed the Variational Autoencoder structure, which combined with probabilistic models, for data generation applications. Additionally, Sakurada and Yairi (2014) confirmed its usability in anomaly detection.

2.2. Semi-supervised Autoencoder

Semi-supervised learning is a method of training model where one part of the input data is labeled (Reddy, Viswanath & Reddy, 2018). This approach is typically used in situations where labeled and unlabeled data are mixed. Semi-supervised learning uses labeled data to train the model and unlabeled data to improve the model's generalization performance. This offers the advantage that utilizing data efficiently and building models with a shortage of labeled data.

Semi-supervised Autoencoder is the conventional Autoencoder structure with a separate Fully-Connected layer(FC layer) to enable learning from labeled data. For the task of classification, a Classifier can be added to the Encoder and Decoder structure as shown in Figure 1. In this structure, labeled data is used to train the Classifier to perform well in classification from the reduced dimensions by the Encoder. Also, with the unlabeled data the Autoencoder is trained to extract generalized features. Additionally, the loss function is defined as a joint loss combined with the cross-entropy for classification and the Mean Squared Error(MSE) for reconstruction, as shown in Eq. (1). Each weight is one of the hyperparameters that need to be optimized during the training process. As a result, by using both labeled and unlabeled data in training, an Encoder that extracts generalized features and a Classifier that has high classification performance can be obtained.

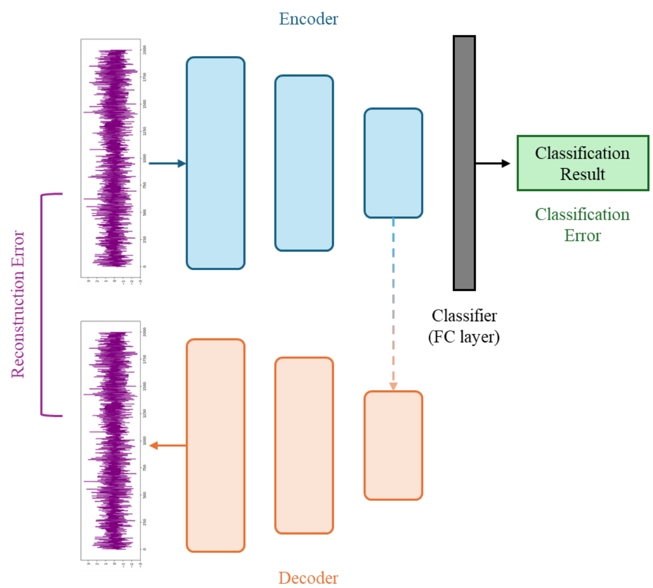


Figure 1. Semi-supervised Autoencoder structure

$$L_J = w_R L_R + w_C L_C \quad (1)$$

2.3. Proposed Method

In this paper, labeled bearing simulator(source) data, which is similar to an unlabeled machining spindle bearing(target) data, is used for fault diagnosis using the Semi-supervised Autoencoder, as shown in Figure 2. Initially, both labeled source data and unlabeled target data are used to train a feature extractor(Encoder) and reconstructor(Decoder). This allows the feature extractor to extract generalized features that reduce domain difference between target and source. Furthermore, the labeled source data is additionally used to train the feature extractor and classifier, making the classifier determine a decision boundary that can classify normal and fault conditions from generalized features. Through this structure and method, it can diagnose normal and fault conditions of the unlabeled target data.

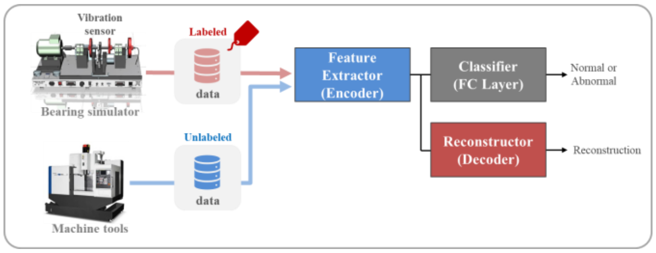


Figure 2. Diagram of proposed method

3. Experiment and Data Acquisition

3.1. Machining spindle bearing

The machining spindle bearing, which is the target domain for fault diagnosis, has a 6204 ball bearing inside the spindle, as illustrated in Figure 3. An accelerometer is attached to the spindle for data acquisition. While the spindle was rotating at 2,000 RPM, 200 data were collected every 0.1 seconds with 20,000Hz sampling frequency without overlapping, so each data had 2,000 points. However, this data was collected without labels, lacking information about normal or fault conditions. Therefore, to develop a fault diagnosis model, it is necessary to use a source domain that is similar to the machining spindle bearing, but with information on the condition.

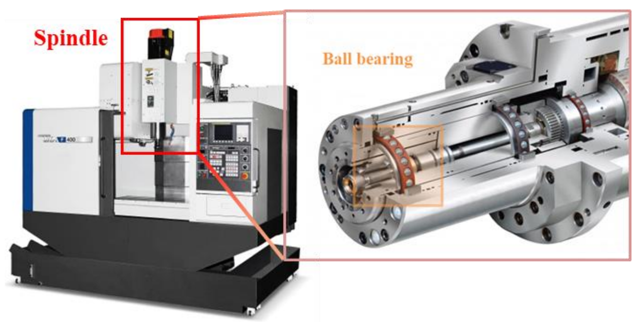


Figure 3. Image of machining spindle bearing

3.2. Bearing Simulator

The bearing simulator is a source domain to imitate the operation of a machining spindle bearing, as shown in Figure 4. It connects the motor on the left and the bearing on the right by the shaft to allow rotation. The attached bearing is the 6204 ball bearing used in the machining spindle, and an accelerometer is installed in the Y direction to collect data during rotation. The way of data acquisition was identical to those for the machining spindle bearing, and 400 data are each collected using both normal and fault bearings.

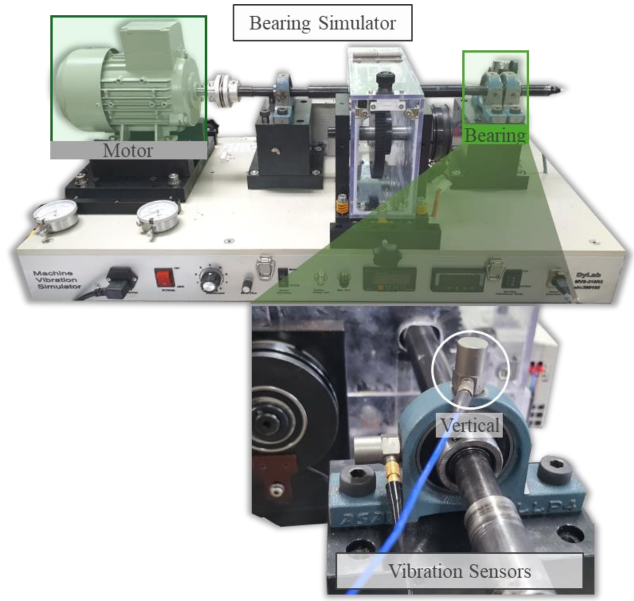


Figure 4. Image of bearing simulator

4. Modeling

4.1. Model Architecture

The architecture of the proposed model is explained in Table 1. The Feature Extractor and Reconstructor were composed of 1D convolutional layers and transposed 1D convolutional layers, and the hidden layer of the classifier was set as one, as shown in Figure 5. A 1D convolutional layer is a network that replaces vertical and horizontal convolution with unidirectional convolution to apply convolution operation in vector data. Through this layer, time-series data can be used in training without converting into a matrix.

Table 1. Description of proposed model architecture

Network	Layer type
Feature Extractor (Encoder)	Conv. layer 1
	Conv. layer 2
	Max Pooling layer
	Conv. layer 3
	Conv. layer 4
Reconstructor (Decoder)	Transposed Conv. layer 1
	Transposed Conv. layer 2
	Up sampling
	Transposed Conv. layer 3
	Transposed Conv. layer 4
Classifier (FC)	Hidden layer
	Output layer

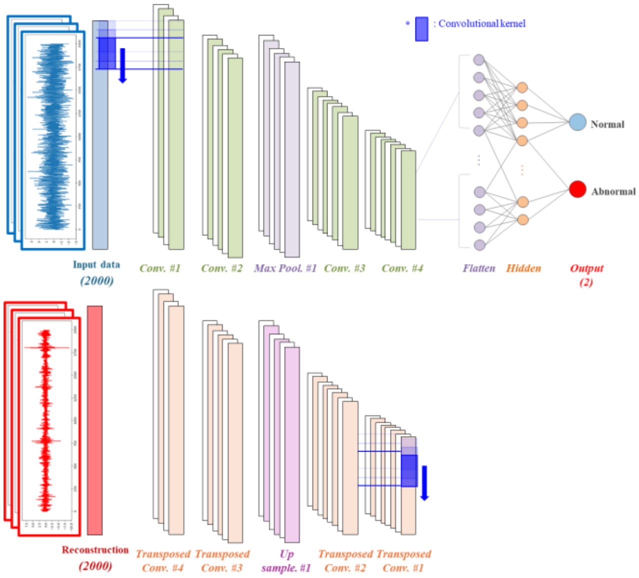


Figure 5. Proposed model architecture

4.2. Model Training

While training the model, the hyperparameters were set as follows. The number of filters in all convolutional layers was 16, the neuron of the hidden layer was set to 16, kernel size was set to 7, and stride was set to 5. ReLU was used for the activation function for all layers, except the output layer, which was set to softmax. The optimizer was Adam, with a learning rate of 0.0001, the weights for classification error (w_C) and reconstruction error (w_R) in the joint loss were set to 0.5 arbitrarily. A total of 700 iterations were done for training with a mix of 200 unlabeled target domain data and 800 labeled source domain data.

To evaluate the performance of the proposed model, a conventional fault diagnosis model was additionally trained. It had only a feature extractor and classifier as shown in Figure 6, with the same model structure and hyperparameters as the proposed model. Since this model can only be trained with labeled data, it is trained with only 800 labeled source domain data.

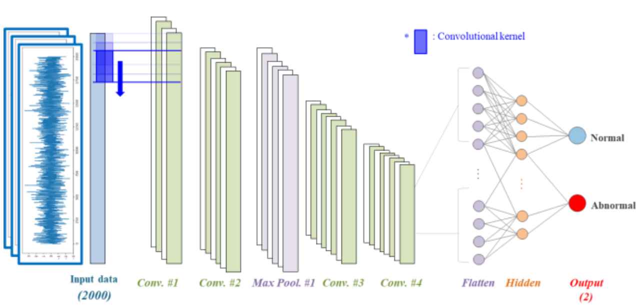
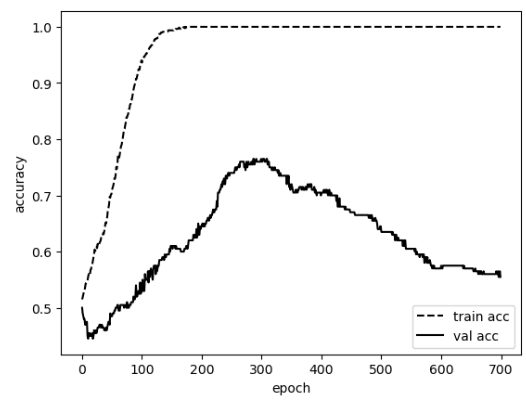


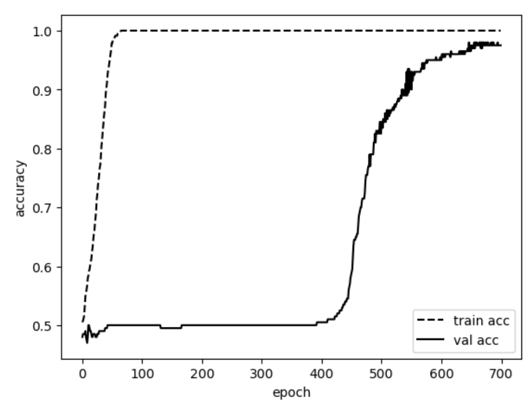
Figure 6. Conventional model architecture

4.3. Result

Model performance comparison was done using additional validation data, collected by machining spindle bearing. Normal and fault data were collected by operating the machining spindle at 2,000 RPM, with normal bearings and bearings damaged by impacts. Under the same setting as the training data acquisition, 100 data for each normal and fault condition were collected. Using these data, the accuracy of the proposed model and conventional model were evaluated to compare the performance of the models. The final accuracy was 76.5% for the conventional model and 97.5% for the proposed model which the proposed model has higher performance. Figure 7 shows the accuracy of both train and validation data at each epoch for both models. Both models reached 100% accuracy on the train data. However, the validation accuracy of the conventional model initially shows an increase, but the accuracy decreases in the end part because of the overfitting problem caused by domain differences. In contrast, the proposed model's validation accuracy did not increase significantly at the beginning but reached high accuracy at the end. This happens due to the delay in optimizing the autoencoder's error compared to the classifier, which can be improved by changing the weight ratio.



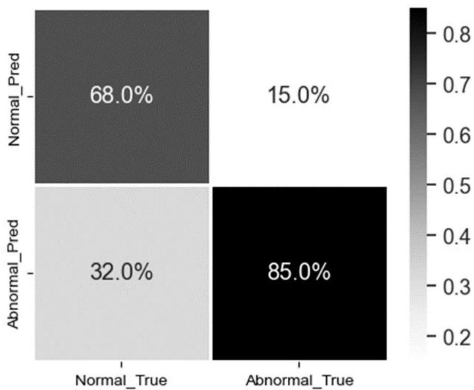
(a) Conventional model



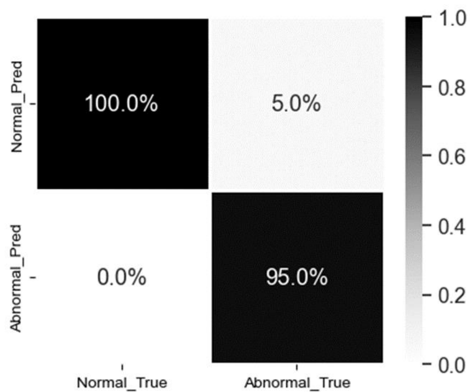
(b) Proposed model

Figure 7. Training and validation accuracy at each epoch

Figure 8 is a confusion matrix that compares the accuracy of the proposed and conventional models with validation data. Only 5% of the fault data are misdiagnosed by the proposed model, whereas 15% of fault data are misdiagnosed by the conventional model. Also, the conventional model misdiagnoses 32% of normal data, resulting in low performance.



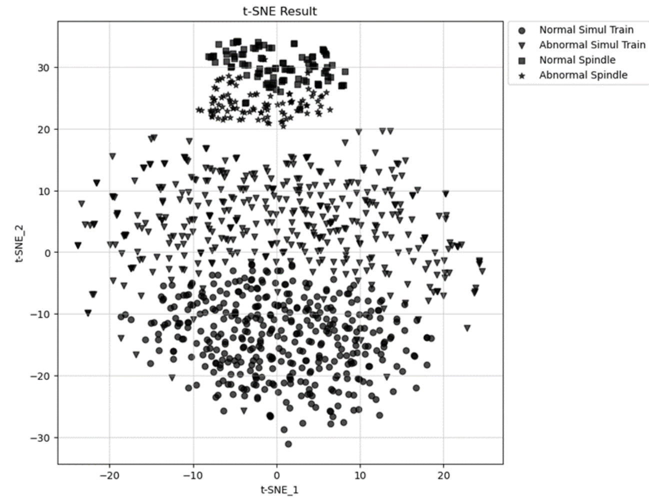
(a) Conventional model



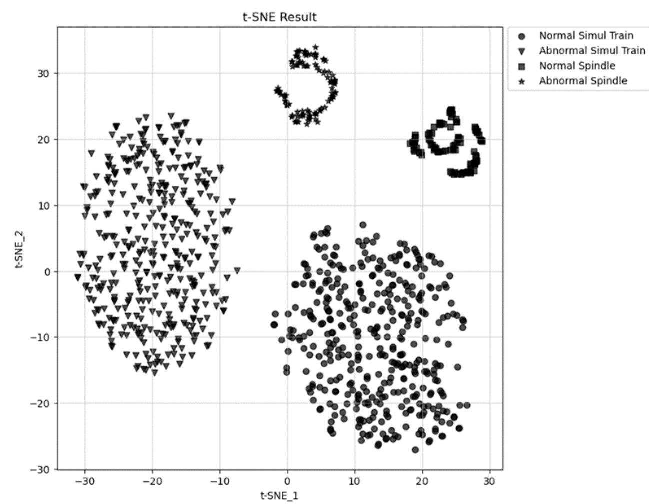
(b) Proposed model

Figure 8. Confusion matrix of final accuracy

Figure 9 visualizes features extracted by the feature extractor from the source domain data used for training and the target domain data used for validation, using t-distributed Stochastic Neighbor Embedding (t-SNE). Unlike the conventional model, which makes it hard to define a decision boundary due to the domain differences, the proposed model can make a clear decision boundary, showing reduced domain differences.



(a) Conventional model



(b) Proposed model

Figure 9. Result of t-SNE by feature

5. Conclusion

This paper proposes a fault diagnosis methodology with multi-domain data using a semi-supervised autoencoder to solve the problem of developing fault diagnosis models with the lack of label data in the real-world industry. A testbed, similar to the domain of actual equipment, was constructed to train the proposed model with sufficient labeled data and unlabeled equipment data. From this approach, a feature extractor that extracts generalized features by reducing the influence of domain information, and a classifier that can diagnose conditions based on generalized features were developed. The model was validated with additional machining spindle bearing data, resulting in the development of a high-performance fault diagnosis model. This approach enables the practical utilization of unlabeled data collected from industrial machines. Furthermore, it has been demonstrated that a high-performance fault diagnosis model can be developed with unlabeled data. This can be applied to many manufacturing, contributing to the reduction of labeling costs across various industries.

However, this study conducted a fault diagnosis model from the perspective of deep learning without considering physical method. In case of bearings, the amplitude of certain fault frequency increase depending on fault characteristics of bearings. The reason for not applying this method is that the value at the fault frequency does not appear clearly, because of the various signals of the machine tool. Therefore, just monitoring the value in the fault frequency is not appropriate for bearing fault diagnosis of the machine tool. However, by considering both data-driven model and physical method, a hybrid model with higher performance can be developed.

And this study did not include the uncertainty caused by various operating conditions of bearings, as it focused on a diagnostic model under fixed operating conditions. Also, the comparison with other models, that can use the unlabeled data, was insufficient. Therefore, the development of a fault diagnosis model that can robustly operate under changing bearing operating conditions and validation of the effectiveness of the proposed methodology through comparison with other semi-supervised learning, synthetic data generation techniques, and domain adaptation technologies are planned for future works.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea (No. 2022R1A2C3012900) and the Ministry of Trade, Industrial and Energy (No. 20012807) grant funded by the Korea government.

NOMENCLATURE

L_J Joint loss
w_R Reconstruction weight
L_R Reconstruction loss
w_C Classification weight
L_C Classification loss

REFERENCES

Wen, L., Li, X., Gao, L., & Zhang, Y. (2017). A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Transactions on Industrial Electronics*, 65(7), 5990-5998. doi: 10.1109/TIE.2017.2774777

Ince, T., Kiranyaz, S., Eren, L., Askar, M., & Gabbouj, M. (2016). Real-time motor fault detection by 1-D convolutional neural networks. *IEEE Transactions on Industrial Electronics*, 63(11), 7067-7075. doi: 10.1109/TIE.2016.2582729

Cabrera, D., Guamán, A., Zhang, S., Cerrada, M., Sanchez, R. V., Cevallos, J., Long, J. & Li, C. (2020). Bayesian approach and time series dimensionality reduction to LSTM-based model-building for fault diagnosis of a reciprocating compressor. *Neurocomputing*, 380, 51-66. doi: https://doi.org/10.1016/j.neucom.2019.11.006

Abed, W., Sharma, S., Sutton, R., & Motwani, A. (2015). A robust bearing fault detection and diagnosis technique for brushless DC motors under non-stationary operating conditions. Journal of Control, *Automation and Electrical Systems*, 26, 241-254. doi: https://doi.org/10.1007/s40313-015-0173-7

Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., Walle, R. V. & Hoecke, S. V. (2016). Convolutional neural network based fault detection for rotating machinery. *Journal of Sound and Vibration*, 377, 331-345. doi: https://doi.org/10.1016/j.jsv.2016.05.027

Lu, C., Wang, Z., & Zhou, B. (2017). Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification. *Advanced Engineering Informatics*, 32, 139-151. doi: https://doi.org/10.1016/j.aei.2017.02.005

Jiang, H., Li, X., Shao, H., & Zhao, K. (2018). Intelligent fault diagnosis of rolling bearings using an improved deep recurrent neural network. *Measurement Science and Technology*, 29(6), 065107. doi: 10.1088/1361-6501/aab945

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507. doi: 10.1126/science.1127647

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*. doi: https://doi.org/10.48550/arXiv.1312.6114

Sakurada, M., & Yairi, T. (2014). Anomaly detection using autoencoders with nonlinear dimensionality reduction. *In Proceedings of the MLSDA 2014 2nd workshop on machine learning for sensory data analysis* (pp. 4-11). doi: https://doi.org/10.1145/2689746.2689747

Reddy, Y. C. A. P., Viswanath, P., & Reddy, B. E. (2018). Semi-supervised learning: A brief review. *International Journal of Engineering &Technology*, 7(1.8) (pp. 81-85). doi: https://doi.org/10.14419/ijet.v7i1.8.9977

BIOGRAPHIES

Yongjae Jeon is now Ph.D. candidate in the Sustainable Design and Manufacturing Laboratory, Department of Mechanical Engineering, Sungkyunkwan University. He obtained his bachelor's degree (System Management Engineering) from Sungkyunkwan University in 2020. His research interest is AI-based prognostics and health management for smart manufacturing.

Kyumin Kim is now master candidate in the Sustainable Design and Manufacturing Laboratory, Department of Mechanical Engineering, Sungkyunkwan University. He obtained his bachelor's degree (Mechanical Engineering) from Sungkyunkwan University in 2023. His research interest is AI-based prognostics and health management for smart manufacturing.

Yelim Lee is now master candidate in the Sustainable Design and Manufacturing Laboratory, Department of Mechanical Engineering, Sungkyunkwan University. She obtained her bachelor's degree (Mechanical Engineering) from Sungkyunkwan University in 2023. Her research interest is AI-based prognostics and health management for smart manufacturing.

Byeong Kwon Kang has a master's degree in the Sustainable Design and Manufacturing Laboratory, Department of Mechanical Engineering, Sungkyunkwan University. He obtained his bachelor's degree (Mechanical Engineering) from Sungkyunkwan University in 2022 and completed his master's degree in 2024. His research interest is AI-based prognostics and health management for smart manufacturing.

Sang Won Lee is now professor in the school of Mechanical Engineering, Sungkyunkwan University. He obtained his bachelor's degree in 1995 and master's degree in 1997 (Mechanical Design and Production Engineering) from Seoul National University. He obtained his Ph.D. degree (Mechanical Engineering) from University of Michigan in 2004. His research interest includes prognostics and health management (PHM), cyber-physical system (CPS), additive manufacturing, and data-driven design.