Expert Knowledge Transfer from CAE Models to CNN Models Using Enhanced Adversarial Domain Adaptation

Minseok Choi, Dongmin Lee, Mikyung Hwang, Dongwhi Yoo, and Hyunseok Oh

School of Mechanical Engineering, Gwangju Institute of Science and Technology, Gwangju, 61005, Republic of Korea

mschoi1104@gm.gist.ac.kr
dmlee4045@gm.gist.ac.kr
mkhwang@gm.gist.ac.kr
donghwi8712@gm.gist.ac.kr
hsoh@gist.ac.kr

ABSTRACT

It is a common belief that convolutional neural networks (CNN) are incapable of acquiring knowledge from domain experts for fault detection and diagnosis. To address the challenge, this paper proposes a knowledge-transfer scheme from computer-aided engineering (CAE) models to CNN models. Domain experts build the CAE models that emulate the faulty behavior of rotating machines by incorporating fault symptom and controlling the degree of fault severity. Fault data are hardly acquired from rotating machines in the field, while a sufficient number of fault data can be generated using the CAE models. Then, a domain adaption model is trained using synthetic data (i.e., normal and fault data) from the CAE models and real data (i.e., normal data only) from rotating machines. To evaluate the validity of the proposed method, a small-scale testbed is regarded as the target system that does not have any fault data. This study contributes to resolve the dearth of fault data from most safety-related engineering assets such as power plant steam turbines, wind turbines, and urban air mobility.

1. INTRODUCTION

The fourth industrial revolution is pushing for automated and intelligent mechanical facilities, which has led to a focus on developing fault diagnostics. Deep learning techniques, empowered by advancements in artificial intelligence and sensor technologies, have gained popularity for their ability to autonomously acquire knowledge about faults and outperform traditional methods. Deep learning has shown impressive results, but its success depends on two crucial factors: the availability of sufficient labeled data for generalized performance and the similarity between the trends observed in the training and test data as has been shown (Liu & Gryllias, 2022). However, in diagnosing mechanical systems, both of these circumstances are typically not met. It is difficult to obtain labeled data under various operating conditions. Moreover, the distribution of training data is often different from that of test data. As a result, there is a need for research on deep learning fault diagnostics that can address the lack of data and inconsistent distribution of the data in this context. Several studies addressed the lack of data and distribution mismatch in diagnosing mechanical systems by combining physical modeling with domain adaptation techniques. Wang et al. (2023) attempted to diagnose triplex pump using a domain adaptive network, DANN, and a physical model. Feng et al. (2023) used MMD and a physical model to detect degradation of the gear surface. However, these approaches can be implemented to another equipment with complex working conditions and external noise. The generalization ability of deep learning is still negatively impacted when the data distribution changes to a different system, which poses a challenge for practical applications.

The paper proposes a knowledge transfer scheme from CAE models to CNN models for fault diagnostics of mechanical systems. This scheme incorporates a new adversarial domain adaptation approach to address the distribution mismatch between simulated and actual systems, which will be presented in following sections.

2. PROPOSED SCHEME

This section starts with CAE modeling. Then, the adversarial learning strategy is described for fault diagnostics by combining synthetic data from CAE models with real data from actual machines. The outline of the proposed scheme is shown in Fig. 2.

2.1. CAE Model Description

In this study, a multibody dynamics (MBD) simulation model is developed to replicate a lab-scale testbed. A simulation database is created through model calibration and verification. The test bed components, including the motor, gearbox, coupling, bearing, rotating shaft, and disk, are depicted in Fig. 1. The assumption of rigid body is used to model the motor and gearbox as they are only used to transmit rotational force. The green jig supports the rotating body system, the 6205 bearing, and the rotating disk, all of which are modeled to the same dimension as the testbed.

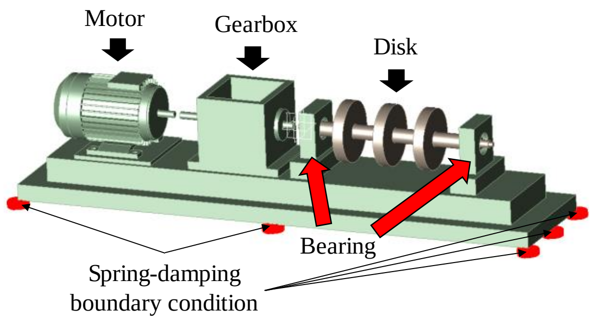

Fig. 1 Multibody dynamics model

A multibody dynamics model is built using fixed boundary conditions for jig-ground, inner ring-rotation shaft, and outer ring-jig to simulate the testbed. Contact conditions are used for ball-inner ring/outer ring and ball-cage contact parts. The rotation speed of the testbed is set to 900 rpm, and the acceleration sensor's sampling rate is 25,600 Hz. The model does not include spring and damping effects. Therefore, they are added as boundary conditions to the rotor system. To simulate a fault condition, an outer ring fault is seeded to bearing B. A groove is created at the location of the red box on the outer ring of the 6205 bearing.

2.2. Domain Adaptation for Fault Diagnosis

This section presents a 1D CNN-based adversarial domain adaptation model that consist of a feature extractor $F(\cdot)$, label classifier $LC(\cdot)$, domain classifier $DC(\cdot)$, and decoder $D(\cdot)$, as illustrated in Figure 7.

The input data is prepared by transforming the raw time-series vibration signal using the absolute value of Hilbert transform. A single batch of training data $X = \{X^S, X^T\}$ is created by connecting the source and target data after applying the Hilbert transform.

The feature extractor includes four feature extraction modules and one convolution (Conv) layer. Each feature extraction module comprises a Conv layer, the ReLU activation function, Mixstyle layer, and maxpooling layer. The goal is to transform high-dimensional vibration signals from source and target domains into lower-dimensional vibration features. The Mixstyle layer combines feature vectors from both domains to learn a generic expression that is not limited to a specific domain as has been shown Zhou, Yang, Qiao, and Xiang (2021).

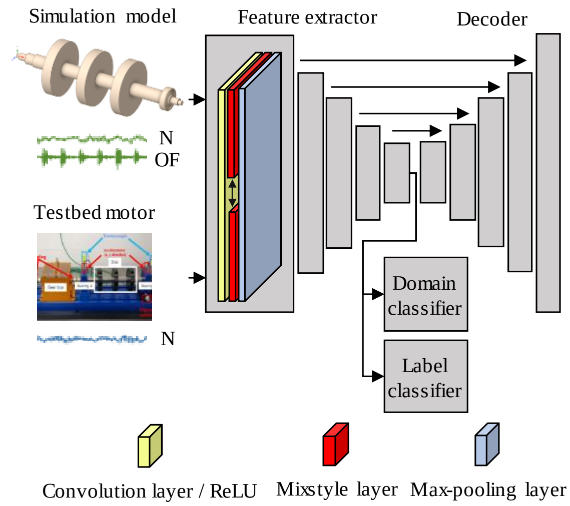

Fig. 2 Overview of the proposed scheme

The Mixstyle layer blends the feature statistics of the source and target domains using a following formula:

$$\gamma_{mix} = \lambda\sigma(f_i^S) + (1-\lambda)\sigma(f_i^T) \tag{1}$$

$$\beta_{mix} = \lambda\mu(f_i^S) + (1-\lambda)\mu(f_i^T) \tag{2}$$

where f_i^S is the source domain feature vector of ith Conv layer; f_i^T is the target domain feature vector of ith Conv layer; and λ are instance-wise weights sampled from the Beta distribution. The ith feature vectors of the source and target domains are each style-normalized into blended features:

$$Mixstyle(f_i^S) = \gamma_{mix}\frac{f_i^S - \mu(f_i^S)}{\sigma(f_i^S)} + \beta_{mix} \tag{3}$$

$$Mixstyle(f_i^T) = \gamma_{mix}\frac{f_i^T - \mu(f_i^T)}{\sigma(f_i^T)} + \beta_{mix} \tag{4}$$

The label classifier, LC, is responsible for diagnosing the state of the system by receiving the low-dimensional feature vector f_5 from the feature extractor F. It consists of two fully connected layers (FC) and outputs a class probability vector $p = \{p_S, p_T\}$, which indicates the probability of the input data belonging to each of the possible classes. This vector is computed using the following equation:

$$p = LC(f_5) \tag{5}$$

The deep learning model uses the loss function LC to diagnose the source domain with high accuracy:

$$L_C = -\sum_{i=1}^{N} y_i \log p_i \tag{6}$$

where N is the number of the data; y_i is the encoded probability of the ground truth label of i^{th} data; and p_i is the estimated probability of the i^{th} data.

The study utilizes CDAN to learn feature vectors that are invariant to domain shift as has been shown Long, Cao, Wang, and Jordan (2018). The domain classifier DC comprises one gradient reversal layer (GRL) and two fully connected (FC) layers to achieve this.

$$w(p) = 1 + e^{\sum_{c=1}^{C} p_c \log p_c} \tag{7}$$

$$L_D = \sum_{i=1}^{N} w(p_i^S)\log(h_i^S) + w(p_i^T)\log(1-h_i^T) \tag{8}$$

where N is the number of the data; C is the number of classes; and p_c is the probability of prediction a sample to class c. The entropy weight $w(p)$ quantifies the level of uncertainty in the predictions of the label classifier LC.

The decoder D uses the compressed feature vector f_5 obtained from the encoding process to reconstruct the input data X, improving its representation. It consists of four upsampling modules and one convolutional layer that are connected sequentially. Each upsampling module includes a transposed convolution layer and a convolutional layer. Skip connections are used to reconstruct both local and global information in the signal, connecting the Conv layer of the upsampling module with the feature vector obtained from the feature extractor F intermediate layer and the output vector of the corresponding upsampling module. The loss function is used to determine the difference between the reconstructed signal X' and the input signal X.

$$L_R = \sum_{i=1}^{N} \lVert X_i' - X_i \rVert \tag{9}$$

where N is the number of the data.

The proposed domain adaptation model consists of three losses: the classification loss L_C, an adversarial domain loss L_D, and a reconstruction loss L_R.

$$L = L_C + \lambda_1 L_D + \lambda_2 L_R \tag{10}$$

where λ_1 is the weight constant that controls the strength of domain loss and λ_2 is the weight constant that controls the strength of reconstruction loss.

3. Dataset

The bearing failure dataset from Smart Diagnosis and Design Optimization (SDDO) was employed as testbed data. As shown in Fig. 3, the SDDO dataset was collected from a rotating body testbed apparatus consisting of a motor, gearbox, disk, and a 6205 bearing. Bearing B in Fig. 1 had an outer ring defect of 1.0 mm, and an accelerometer was used to evaluate the y-axis acceleration of bearing B. The accelerometer sampled data at a rate of 25.6 kHz. Data were collected under 900 RPM operating settings to ensure data diversity.

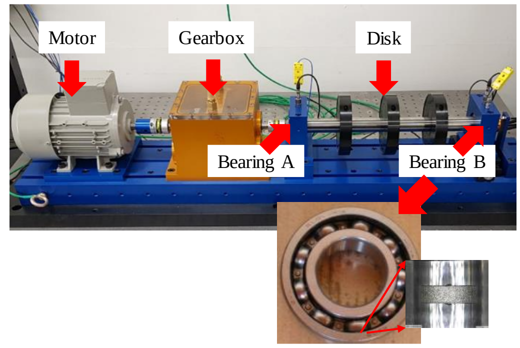

Fig. 3 SDDO testbed

The simulation model generated acceleration data that produced BPFO characteristic frequency components when analyzed in the frequency domain.

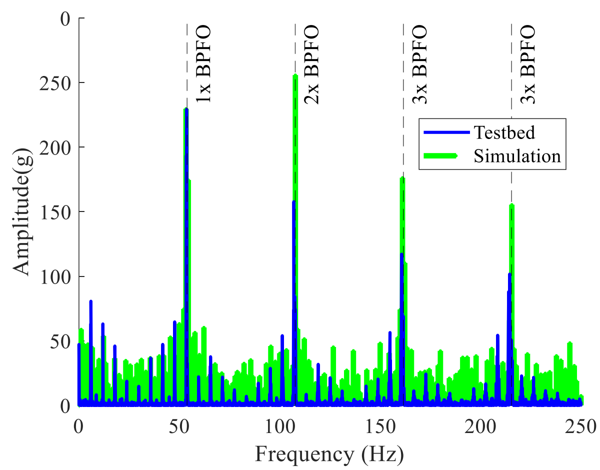

Fig. 4 Simulation and testbed frequency spectrum after model calibration

To verify the model, real 1.0 mm outer ring defect data obtained from a rotational speed test bed was used. The density of the spring, damper, and rotor system was chosen as a calibration variable, and model calibration was performed using the 1x BPFO characteristic component. The calibration result showed an error of 0.20% for the 1x BPFO component, as shown in Fig. 4.

4. Result and Discussion

The proposed method includes both normal and fault data by incorporating simulation data into the source domain, and domain adaptation is achieved by considering only normal data from the test bed as the target domain. The diagnostic performance of various models, including DNN, CNN, DANN, CDAN, and the proposed method, was evaluated using defect and normal data from the test bed as shown in Fig. 5. While DNNs and CNNs showed limited ability to classify normal and outer race defects, DANN and CDAN exhibited slightly improved diagnostic performance. In contrast, the proposed method achieved a significantly improved accuracy of 94.5%, despite not being trained on the test bed defects. Therefore, the proposed model can potentially diagnose defective components in the test bed by learning from a combination of the defective components from the simulation and the normal components from the test bed.

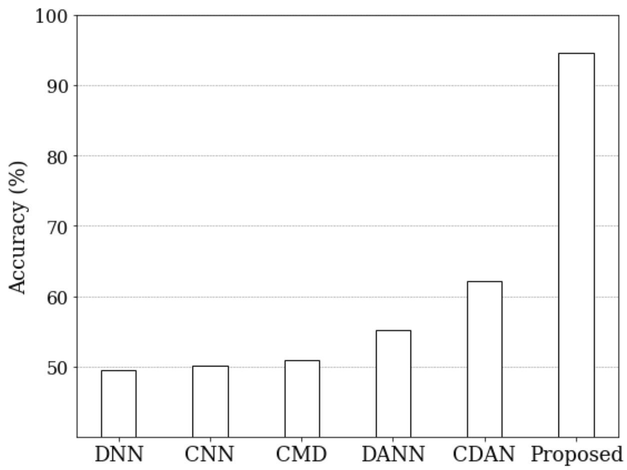

Fig. 5 Transfer result between simulation and testbed

5. CONCLUSION

This paper proposed a expert knowledge transfer scheme for bearing fault diagnosis. A simulation model is built to emulate the healthy and faulty behavior of the rotating machine. A sufficient number of faulty data can be obtained. Then, the domain adaptation model is trained using synthetic data from the simulation model and real normal state data collected from rotating machines. The proposed scheme addresses the distribution mismatch between simulated and actual machine and can diagnose the target machine even without access to its defect data. The performance of the domain adaption method was compared with existing methods using test bed equipment and a CAE model to demonstrate the effectiveness of the proposed method. The results show that the proposed scheme can be implemented to practical applications.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A2C1008143). This work was supported by GIST Research Project grant funded by the GIST in 2023.

REFERENCES

Liu, C. & Gryllias, K. (2022). Simulation-driven domain adaptation for rolling element bearing fault diagnosis. *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 5760-5770.

Wang, J., Zhang, Z., Liu, Z., Han, B., Bao, H., & Ji, S. (2023). Digital twin aided adversarial transfer learning method for domain adaptation fault diagnosis. *Reliability Engineering & System Safety*, vol. 234, 109152.

Feng, K., Ji, J. C., Zhang, Y., Ni, Q., Liu, Z., & Beer, M. (2023). Digital twin-driven intelligent assessment of gear surface degradation. *Mechanical Systems and Signal Processing*, vol. 186, 109896.

Long, M., Cao, Z., Wang, J., & Jordan, M. I. (2018). Conditional adversarial domain adaptation. *Proceedings of the Advances in Neural Information Processing Systems*, Montreal, Canada, pp. 1645-1655.

Zhou, K., Yang, Y., Qiao, Y., & Xiang, T. (2021). Domain generalization with MixStyle. *Proceedings of the International Conference on Learning Representations*, May 3-7, Vienna, Austria.

Minseok Choi received the B.S. and M.S. degrees in mechanical engineering from Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, in 2021 and 2022 respectively, where he is currently pursuing the Ph.D. degree in mechanical engineering. His current research topics include prognostics and health management for rotating machines.

Donghwi Yoo received the B.S. degree in mechanical engineering from Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, in 2021, where he is currently pursuing the M.S. degree in mechanical engineering. His current research topics include prognostics and health management for rotating machines.

Hyunseok Oh received the B.S. degree in mechanical engineering from Korea University, Seoul, South Korea, in 2004, the M.S. degree in mechanical engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2006, and the Ph.D. degree in mechanical engineering from the University of Maryland, College Park, MD, USA, in 2012. He is an Associate Professor with the School of the Mechanical Engineering, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea. His research interests include fault diagnostics, industrial artificial intelligence, and model verification and validation.