IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 15, 2022, accepted April 3, 2022, date of publication April 18, 2022, date of current version May 3, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3167761

Fault Diagnosis of Bearings With the Common-Domain Data

TAEYUN KIM AND JANGBOM CHAI
Machine Diagnostics Laboratory, Department of Mechanical Engineering, Ajou University, Suwon 16499, Republic of Korea
Corresponding author: Jangbom Chai (jbchai@ajou.ac.kr)

This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety (KoFONS) and in part by the Nuclear Safety and Security Commission (NSSC) of the Republic of Korea, under Grand 1805007.

ABSTRACT Rolling element bearings are one of the important components in rotating machines. Therefore, many studies on bearing diagnosis have been conducted with artificial intelligence (AI) to do maintenance on the machines on time. In general, AI successfully diagnoses the defects of bearing when it is trained with the sufficient data of a specific machine, but it hardly provides reasonable results when it is untrained or insufficiently trained. However, it is hard to obtain sufficient data even for a specific machine in practice. In this paper, a new method was developed to increase training data by transferring the cross-domain data into the common-domain data. Therefore, all the data from different kinds of machines with various bearings can be combined as a big training data. Bearings under consideration in this paper have different specifications and characteristics. In transferring into the common-domain, it is important to get rid of structural and environmental noise by signal processing, which makes it plausible to extract common features. With the common-domain data, one-dimensional convolutional neural network (1D-CNN) with feature domain adaptation is applied and successfully classifies the defects of each bearing. Moreover, 1D-CNN combined with support vector machine (SVM) can also classify defects successfully without feature domain adaptation, which makes it possible to train the model only with normal data of the machine in concern. To verify the proposed method, not only the bearing data from Case Western Reserve University and Paderborn University but also the bearing data with flow noise of Ajou University are used.

INDEX TERMS Bearing fault diagnosis, common-domain data, convolutional neural network, cross-domain fault diagnosis, domain adaptation, signal processing, support vector machine.

I. INTRODUCTION

Rolling element bearings play a significant role in electro-mechanical drive systems and motors. Failures of these machines are related to the bearings with a high probability [1], [2]. Therefore, the diagnosis of rolling element bearings is essential for the safe operation of machines. Several signals such as vibrations, acoustics, and currents have been recorded for the diagnosis of bearings. Fault characteristics can be extracted from the vibration signals of bearings. However, because the vibration signals are nonlinear and non-stationary, the development of a strategy for feature extraction from the signals is important [3]. Several techniques, such as envelope [4], wavelet packet decomposition with random forest (RF) [5], bispectrum analysis [6], double sparse dictionary with orthogonal matching pursuit and Deep Belief Network (DBN) [7], empirical mode decomposition (EMD) [8], and local mean decomposition(LMD) [9], have been studied and utilized to extract features and diagnose bearings exactly. However, considerable knowledge is required to select an appropriate signal processing method and techniques. For example, users must know that the discrete wavelet transform presents some issues such as shift-variance behavior [10]. Also, EMD and LMD present a mode-mixing problem [8], [9], [11]. It is also well known that the accuracy of diagnosis is highly dependent on the skills or experience of experts. Yuan *et al.* employed a continuous wavelet transform (CWT) to obtain characteristic signal information for input data, and their classification model provided remarkably high accuracy using the transformed signals [12]. However, the mother wavelet should be selected accurately to produce input data to obtain satisfactory results from the model when CWT is employed.

To address these challenges, numerous studies have been conducted to automatically extract the features of faults and diagnose bearings without expertise. Convolutional neural

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Shu.

networks (CNNs) are good alternatives to address such challenges. Huang *et al.* developed a deep CNN model with multi-scale cascade layer, which is effective in bearing diagnosis by using kernels of different sizes to extract more useful information [13]. Wen *et al.* transformed the signals of various systems into a two-dimensional image and extracted features using a CNN model based on LeNet-5. Their model demonstrated improved feature extraction capability and produced better results than other models [14]. Because CNN can be combined with a variety of other models or algorithms such as q-xy [15], studies have also involved diagnoses using CNN combined with various classifiers. Yuan *et al.* adopted ResNet-18 for feature extraction and a support vector machine (SVM) for classification [12]. Multi-CNN structures and multi-sensor data were used to prevent information losses and the CNN structures were combined with SVM to diagnose bearings in [16]. Han *et al.* conducted a study combining CNN with SVM to diagnose bearings, while adding specific conditions for moving from extraction stage to classification stage [17]. Multi RF models were used with CNN and continuous wavelet transform to classify bearing faults [18].

At times, it is necessary to determine the state of a system to be diagnosed using data acquired from other systems. This is called a cross-domain fault diagnosis. There exist two domains of cross-domain fault diagnosis. The source domain is adopted to train the models, and the data and labels are defined. The other domain is the target domain, which is the domain to be diagnosed, and the labels are either not available or partially available. As both domains have different distributions, a model trained using one system cannot be applied to other systems without appropriate methods because the boundaries for each domain are not the same. To address this challenge, numerous studies have been conducted, and strategies with CNN has particularly gained recognition in recent years. First, the distance between domains is calculated using several methods, and the difference between the two domains is reduced by minimizing the metrics. Lu *et al.* proposed a deep neural network model with a maximum mean discrepancy (MMD) and weighted regularization method to diagnose another system [19]. Guo *et al.* constructed a CNN model with 16 layers to classify the bearing states, and introduced an additional module for domain adaptation. The domain adaptation module comprises a domain classifier and distribution discrepancy metrics which is MMD [20]. Yang *et al.* adopted a multi-kernel MMD (MK-MMD) in multiple layers for domain adaptation [21]. In [22], the discrete wavelet transform was adopted to produce input data for the DBN, and the developed model contained a domain adaptation method. Correlation alignment [23], [24], Wasserstein distance [25], etc., have also been widely adopted in diagnosis of bearings. Another strategy for domain adaptation is the addition of a discriminator that prevents features from knowing the domains from which they have been extracted. Using this concept, several studies have successfully addressed cross-domain fault diagnosis, [26], [27]. Studies have also been conducted to compare the diagnostic results produced by several domain adaptation methods, and codes for the models have been provided in [28].

However, it is extremely difficult to conduct diverse experiments using a single system. Studies have been conducted to obtain various experimental results using data generation methods such as Monte Carlo [29]. If the data in various systems are considered as one, the domain of the training data can be expanded, and various characteristics and specifications can be applied to the models without a specific method. Kim *et al.* transformed the signals of different systems into the same pattern space using signal processing, and the models trained with processed data provided excellent diagnostic performance for other systems [30]. Using expertise information, Zheng *et al.* combined signals acquired from different systems and proposed generalized bearing diagnostic models [31]. In addition, it should be possible to classify labels of the target domain that are not obtained, or even if only normal data are available. In [19] and [31], the normal data of the target domain were used for training to solve this problem. However, it would be more efficient if the target domain could be diagnosed using a model trained with only data acquired from other systems.

Because data acquired from bearings with different specifications and characteristics are placed in the cross-domain, they cannot be combined as training sets. In this study, a method is proposed to increase the training data by transferring cross-domain data into common-domain data to diagnose ball-bearing systems with high accuracy. In particular, this study focuses on addressing the following challenges: (1) insufficient training data with limited types or levels of defects, and (2) fault classification when only normal data from target domain are available. Therefore, a signal-processing method [30] is applied to transfer the data into a common domain and reduce noise. For the second issue, an artificial intelligence (AI) models are designed to classify bearing defects with only normal data from target domain using common-domain data.

The remainder of this paper is organized as follows. In Section II, relevant information for a better understanding of this study is presented. In Section III, three different systems are explained, and the data from each system are analyzed. The domain adaptation methods and proposed method are explained in Section IV. Results from the domain adaptation and proposed methods with various models are presented in Section V. Finally, the conclusions of this study are presented in Section VI.

II. THEORETICAL BACKGROUND

A. CROSS-DOMAIN FAULT DIAGNOSIS

Cross-domain fault diagnosis refers to the issue of diagnosing states using a single classifier by reducing the discrepancy in the distribution resulting from different domains. For bearings, the domains vary depending on a variety of factors, such as the rotational speed, number of balls, and load. There exist two domains in cross-domain fault diagnosis. One

is the source domain, which has both data and labels and is adopted to train the model, and the other is the target domain, which contains only data without labels or holds only certain labeled data. As each domain has a different distribution, it is usually difficult to use one domain for the diagnosis of other domains. That is, both domains have different boundaries for classifying the states, and additional methods are required to reduce the difference between the domains. The two domain adaptation methods used are as follows:

1) MAXIMUM MEAN DISCREPANCY (MMD)

In MMD, the source and target domain data are embedded in the reproducing kernel Hilbert space with kernel function, and the distance between the means of the two distribution is calculated [32]. By reducing the distance metric, the difference between the two domains decreases. MK-MMD is a method employed to further reduce the mismatch using multiple kernels [33]. In this study, MK-MMD was employed for the diagnosis of unlabeled data.

2) DOMAIN ADVERSARIAL NEURAL NETWORK (DANN)

A DANN performs domain adaption by adding other neural network layers with a gradient reversal layer. Thus, the domain from which the features are extracted remains unknown [34].

B. SUPPORT VECTOR MACHINE

An SVM is a supervised learning model, which determines the hyperplane that maximizes a margin. The margin is the distance between the decision boundary and points that are closest to the boundary. Linearly separable data are illustrated in Figure 1 [35]. When data are not linearly separable, they are mapped to a higher dimension via kernel functions, and then a linearly separating hyperplane can be observed. Soft-margin SVM allows certain instances that cross the margin plane, unlike the hard-margin SVM [35].

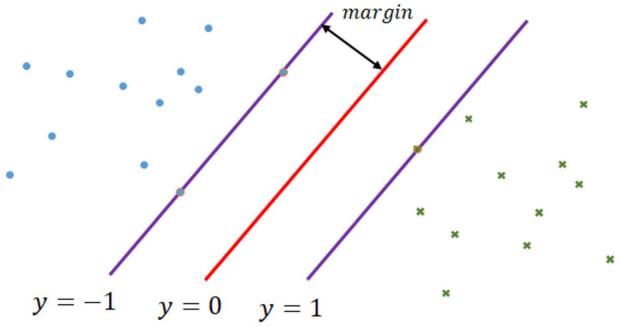

FIGURE 1. Hard margin support vector machine. The margin represents the vertical distance between the decision boundary (y = 0) and data points (support vectors). The support vectors are the set of points at y = 1 and y = −1 and are marked with orange circles.

C. CONVOLUTIONAL NEURAL NETWORK

CNN has been employed in several fields, such as the classification of music [36], text [37], and images [38]. As can be inferred from the examples, there exist various types of input data, such as pictures [38] and sound waves [39]. Before the CNN was actively studied, machine learning methods were executed by extracting hand-crafted features [40]. Because a CNN comprises feature extractors and classifiers, feature extraction is conducted automatically with classification. This method presents the advantage of automatically extracting features and reducing information loss. In general, convolution and pooling layers are sequentially stacked in the extractor. The kernels (filters) in the convolution layer move by set values, and the dot product of the input data and the weights in the filters go through activation functions to create new feature maps. Pooling layers reduce the dimensions of the feature maps by calculating and representing the maximum or average values of a particular range. A classifier is added following feature extraction. The output layer of the classifier comprises as many nodes as the number of classes and commonly adopts the Softmax activation function [41].

D. SIGNAL PROCESSING

A study in which cross-domain fault diagnosis was performed after the signal processing of bearing data collected from different domains into a common space has been reported [30]. The CWRU data were moved to the same pattern space using the Paderborn University (PU) data in [30], whereas the PU and Pump data were moved to the same pattern space using the CWRU data in this study. Various filters, such as low-pass and minimum-phase filters were utilized. For the CWRU data, a minimum-phase filter was applied to the resampled signal, whereas other data were transformed using the CWRU normal data, and then, a minimum-phase filter was applied. A total of 1,024 points from the minimum-phase signals were extracted for the input data. Figure 2 illustrates the flowchart of making signals utilized in this study. The extracted signals were stored as one-dimensional signals and were input to the classification model. To obtain the similar length data as the raw signal, the data were extracted by moving the window from the resampled signal (3 kHz). The window size was 9,000, and the moving value was set differently for each system.

III. DATA DESCRIPTION AND ANALYSIS

A. PUMP DATA

Vibration datasets from three ball-bearing systems were utilized to train and test the AI models. The first datasets were acquired while running a centrifugal pump at Ajou University, as shown in Figure 3(a) and 3(b), hereinafter referred to as the Pump data. A defect in the pump bearing was introduced using a drill, as illustrated in Figure 4. While the pump with a defective bearing operated at the best efficiency point, the vibration from the bearing housing of the pump was measured using accelerometers with sampling rates of 12,000 Hz. The rotational frequency of the pump was 59 Hz,

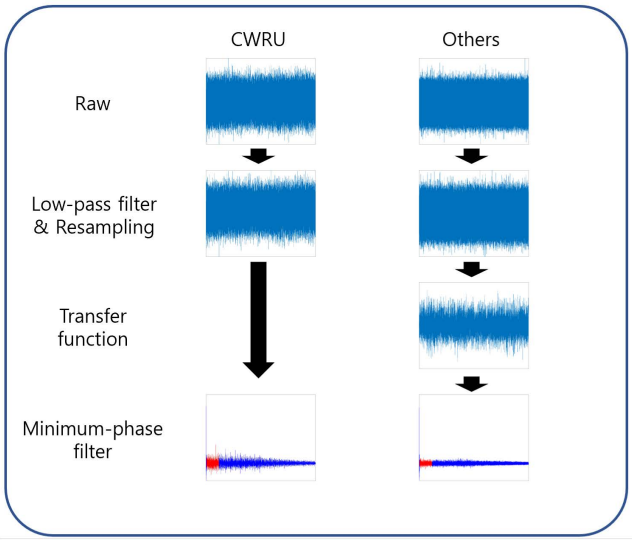

FIGURE 2. Flowchart of signal processing for the Case Western Reserve University (CWRU) data and others.

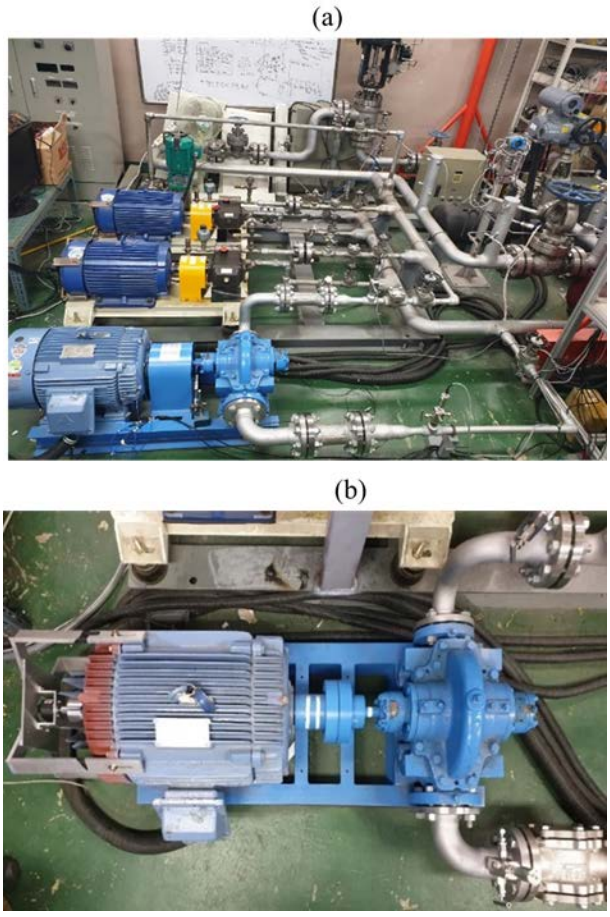

FIGURE 3. (a) Flow loop for the experimental data and (b) A centrifugal pump for the experimental data.

and the working fluid was water. To check whether the characteristics of faults could be observed in the signals, the signals of the Pump data were analyzed in both time domain and

FIGURE 4. Faulted ball bearing of centrifugal pump in Ajou University using a drill, and the fault is located on the inner raceway.

TABLE 1. Fault frequencies (Hz) of the pump data.

Rotational Freq.	Outer Raceway Freq.	Inner Raceway Freq.	Ball Freq.	Fundamental Freq.
59	153.67	259.33	107.77	37.05

frequncy domain first. The characteristic frequencies were calculated and are presented in Table 1 using the following equations [42]:

$$BPFO = \frac{nf_r}{2}(1 - \frac{d}{D}cos\phi), \quad (1)$$

$$BPFI = \frac{nf_r}{2}(1 + \frac{d}{D}cos\phi), \quad (2)$$

$$BSF = \frac{D}{2d}[1 - (\frac{d}{D}cos\phi)^2], \quad (3)$$

$$FTF = \frac{f_r}{2}(1 - \frac{d}{D}cos\phi). \quad (4)$$

where $BPFO$ is the ball pass frequency of the outer race, $BPFI$ is the ball pass frequency of the inner race, BSF is the ball spin frequency, FTF is the fundamental train frequency, f_r is the shaft speed, n is the number of rolling elements, d is the diameter of the ball, D is the diameter of the pitch, and ϕ is the angle of the load from the radial plane.

The Pump data were analyzed as illustrated in Figure 5. Vibration signals of defective bearings, such as raceway faults, generally demonstrate impulse-like symptoms, and their intervals represent frequencies of the fault. However, it is difficult to identify impulse-like signals from a defective bearing, and only the rotational frequency and its harmonics can be observed in the Pump data. It seems that there is no noticeable difference between the defected and normal data not only in the time domain but also in the frequency domain. This is because the characteristic signals are masked by noise. Therefore, it is necessary to first reduce the noise resulting from the Pump data to clarify fault signatures. The dashed red line in the second column represents the fault frequency.

B. CASE WESTERN RESERVE UNIVERSITY DATA

The next dataset is from the Bearing Data Center of CWRU, and the testbed of the CWRU data is illustrated in

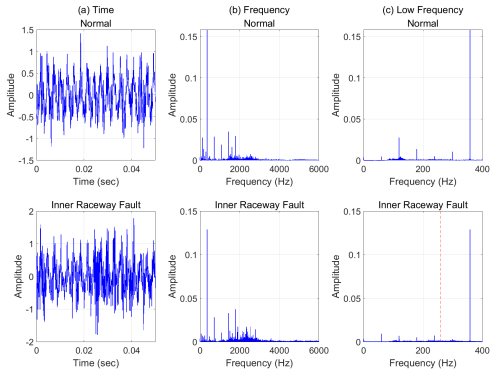

FIGURE 5. Pump data in the (a) time, (b) frequency, and (c) low frequency domains. The red dashed lines represent characteristic frequency.

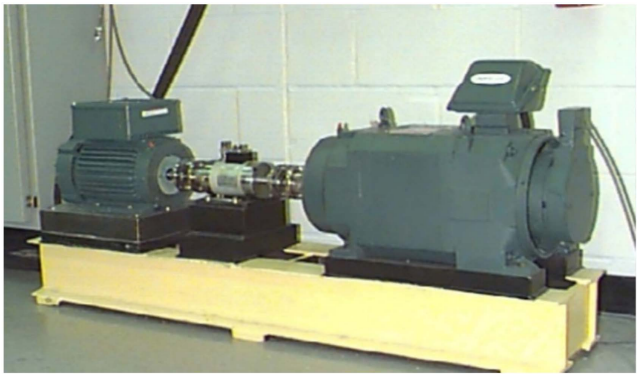

FIGURE 6. Testbed of the CWRU data (electric motor, torque transducer, encoder and dynamometer, to control electronics) [43].

Figure 6 [43]. The CWRU data have frequently been adopted in studies on bearing diagnosis. The datasets comprise normal, inner raceway, outer raceway, and ball fault data. Drive-end fault data are measured at 12,000 Hz and 48,000 Hz, whereas normal baseline data and fan-end fault data are measured at a sampling rate of only 12,000 Hz. Herein, bearing loads of 0, 1, 2, and 3 hp were applied, and data were obtained by mounting defective bearings with defect diameters of 0.007, 0.014, and 0.021-in, respectively. In addition, the drive-end inner raceway fault and drive-end ball fault data also included 0.028-in defect version, which were recorded at 12,000 Hz. Experiments for the outer raceway were conducted in three different positions in some cases. The details can be found in [43].

As with the Pump data, the fault frequencies of the CWRU data were also calculated, and fault characteristics could be found in time domain and frequency domain signals. The calculation results are presented in Table 2. An example of analyzing the CWRU datasets is illustrated in Figure 7. As illustrated in the figure, the fault characteristics are indicated in the time and frequency domains, although the data are not processed. In this study, normal and abnormal (0.007-in inner raceway fault) data on the drive end measured at sampling rates of 12 kHz and applied loads of 0 to 3 hp were used. Regardless of the horsepower, data in the same state were adopted as one.

TABLE 2. Fault frequencies (Hz) of the CWRU data.

Load level (hp)	0	1	2	3
Rotational Freq.	29.93	29.53	29.17	28.73
Outer Raceway Freq.	107.30	105.87	104.56	103.00
Inner Raceway Freq.	162.09	159.93	157.97	155.60
Ball Freq.	141.09	139.20	137.48	135.43
Fundamental Freq.	11.92	11.76	11.62	11.44

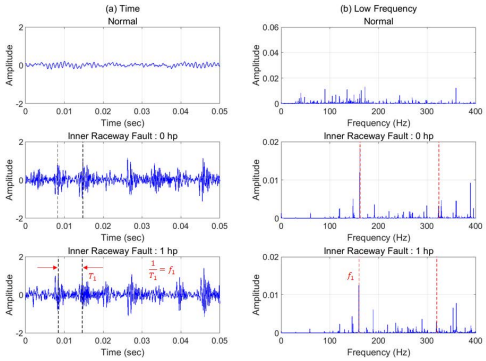

FIGURE 7. Example of analyzing the CWRU data with load 0 hp and 1 hp and a 0.007-in defect in the (a) time and (b) low frequency domains. The red dashed lines represent the characteristic frequencies (inner raceway frequency) and their harmonic.

C. PADERBORN UNIVERSITY DATA

The last datasets, abbreviated as PU data, are from the Konstruktions–Und Antriebstechnik (KAt) datacenter at PU [44]. Figure 8 illustrates the testbed of the PU datasets. The PU data contain healthy bearing, artificially defected, and real defected data with various conditions, such as rotational speed(900 and 1500 rpm), load torque(0.1 and 0.7 Nm) and radial force(400 and 1000 N). Data were measured at 64 kHz. The details can be found in [44]. Similar to the other two datasets, the frequencies of the PU data were also calculated for the defects, and the results are presented in Table 3. For the inner raceway fault signal, it is difficult to determine periodic intervals in the time domain compared with that in

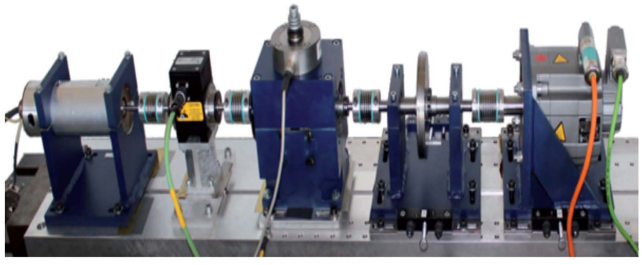

FIGURE 8. Testbed of the PU data: test motor, measuring shaft, bearing module, flywheel, and load motor (from left to right) [44].

TABLE 3. Fault frequencies (Hz) of the PU data.

Rotational Freq.	Outer Raceway Freq.	Inner Raceway Freq.	Ball Freq.	Fundamental Freq.
25	76.35	123.64	49.91	9.54

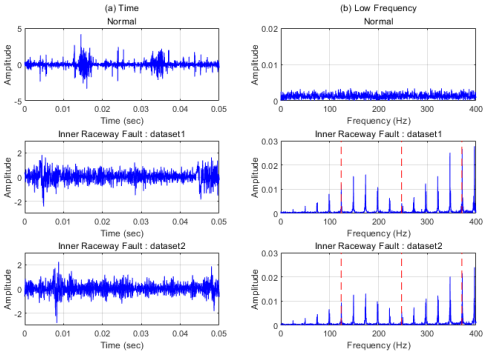

FIGURE 9. Example of analyzing the PU data in the (a) time and (b) low frequency domain. The red dashed lines represent the characteristic frequency (inner raceway frequency) and their harmonics.

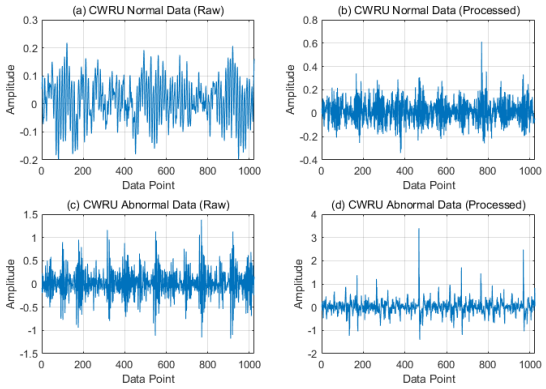

FIGURE 10. Example of processed CWRU data: (a) CWRU normal data (raw), (b) CWRU normal data (processed), (c) CWRU abnormal data (raw), and (d) CWRU abnormal data (processed).

the CWRU data. However, the characteristic frequency and its harmonics can be detected in frequency domain, as illustrated in Figure 9 (b). In addition, the normal PU data are more complex than CWRU normal data. That is, although the fault characteristics are included in the signal, the characteristics of faults may not be distinguished from noise if signal processing is not performed properly. To develop the proposed method, normal data (K001) and inner raceway fault data (KI16), which were measured with a rotating speed of 1,500 rpm, load torque of 0.7 Nm, and radial force of 1,000 N were adopted. Seven out of the twenty datasets were used for each state.

As observed above, the CWRU and PU data have distinct characteristic frequencies, but the shapes of the spectra differ according to the system. In addition, energies that are not related to faults also appear. Because the shape of the signal

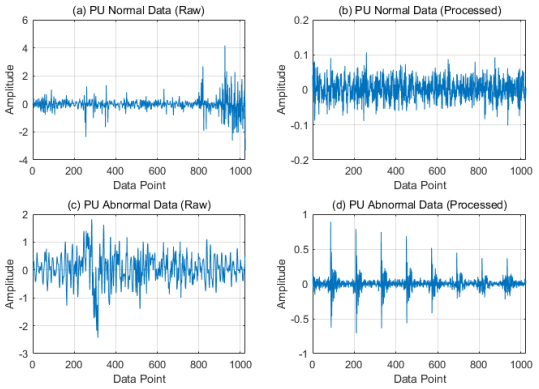

FIGURE 11. Example of processed PU data: (a) PU normal data (raw), (b) PU normal data (processed), (c) PU abnormal data (raw), and (d) PU abnormal data (processed).

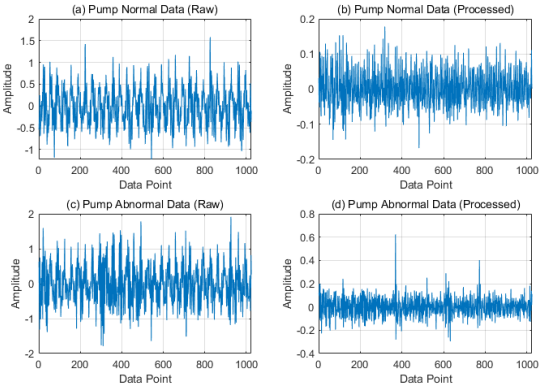

FIGURE 12. Example of processed Pump data: (a) Pump normal data (raw), (b) Pump normal data (processed), (c) Pump abnormal data (raw), and (d) Pump abnormal data (processed).

varies for several reasons other than defects, the AI model can be trained with other characteristics. Therefore, the models that are trained using one system may not accurately extract the features of faults for other bearings. However, because experiments with various conditions cannot be treated as a single system without proper methods, the models cannot train several characteristics with raw data from different systems. For the Pump data from Ajou University, no apparent peaks can be observed in the time domain owing to flow noise. These phenomena are the same as those in the frequency domain. The reduction in noise will help the model find fault characteristics better by making the characteristics of the fault distinct. To address the above challenges, a preprocessing method developed in [30] was applied to the three different systems. Because a model trained with one system cannot extract defect characteristics from other systems accurately, the model should reflect the various characteristics to improve the ability to extract features. Transforming and combining data with different conditions will expand the domain of the training data and allow the identification of defect characteristics in more diverse systems.

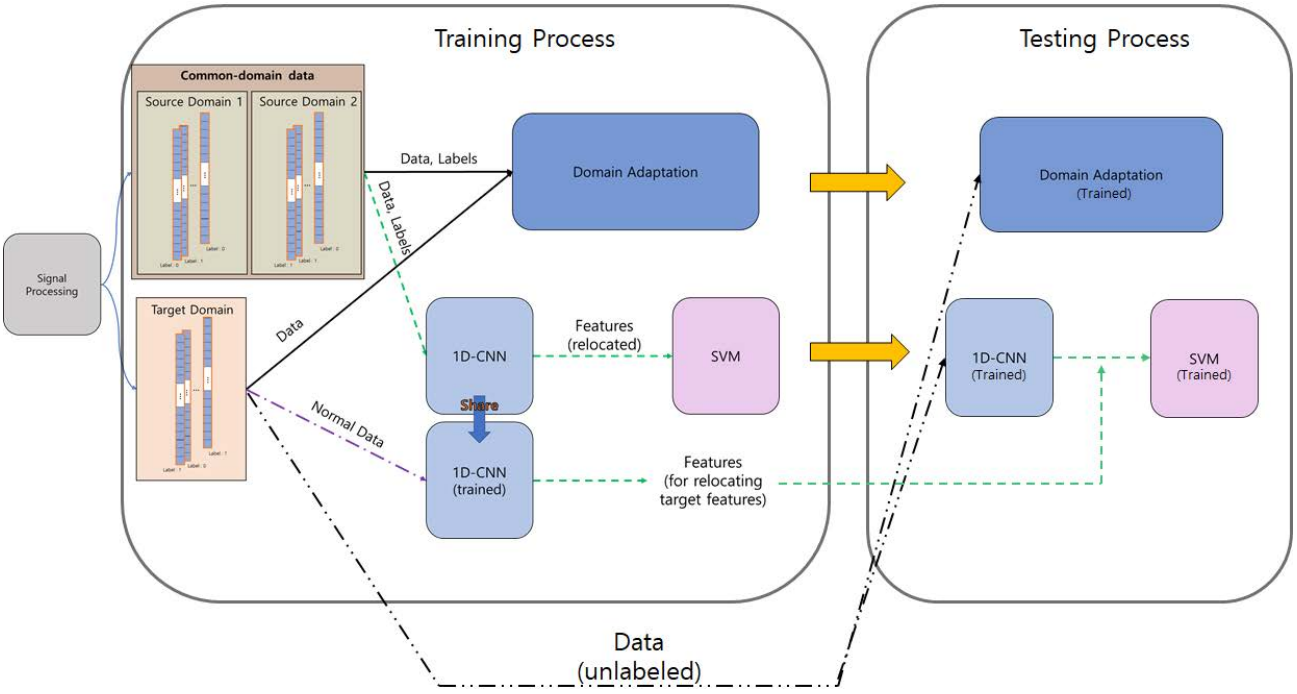

FIGURE 13. Processes of domain adaptation and AI models with common domain data.

IV. METHODOLOGY

A. COMMON-DOMAIN DATA

It may be possible to develop an AI model that can diagnose any ball bearing if sufficient data are available, including normal states and various types of abnormal states for ball bearings. In other words, AI models for a specific ball bearing can be developed and validated when the data are sufficient to train a model. However, these data are generally insufficient. In addition, diagnosis cannot be performed successfully if the noise is so dominant that the defect signatures are hardly observable in the measured signals. Therefore, signal processing is essential to get rid of noise and to transform into the common pattern space to perform cross-domain fault diagnosis and make generalize the AI models. Accordingly, the method developed in [30] was employed, and the results of signal processing to eliminate noise in the raw data and transform the features into a common pattern space are presented. The results of the processed data (CWRU, PU, and Pump) are illustrated in Figure 10, 11, and 12, respectively. Each figure illustrates both the raw and processed signals of the normal and abnormal data. It is shown that raw data have different patterns according to the systems while the processed data have similar patterns. In the processed data, the changes of fault characteristics are more clearly recognizable even in noisy situation such as flow noise condition as shown in Figure 12. MATLAB was utilized for the signal processing.

Common-domain data are defined as cross-domain data extracted from different systems and transferred into a common domain. To examine the effect of signal processing, the models were trained and tested using the combined raw data. Also, target domain data were transferred into the same domain when common-domain data were used in training the models.

B. MODELS

Several models were tested and analyzed to verify the usefulness of transforming signals from different systems into the common-domain signals. Figure 13 illustrates a flowchart of the proposed method and processes followed by the domain adaptation models. The methods are explained as follows.

1) DOMAIN ADAPTATION

To validate the effectiveness of common-domain data and address the challenge presented by a diagnosis without labels of the target domain, the domain adaptation models designed in [30] were adopted and are presented in Table 4. First, the signals were processed before combining the data of the source domains. Common-domain data and labels are used for training CNN and target data are used for domain adaptation without labels. During the testing process, the models received and classified data that were not adopted for training. A CNN with MK-MMD and DANN were employed to check the ability of the common-domain data. The CNN was trained for 20 epochs, and the domain adaptation methods were trained for the remaining 100 epochs. In order to perform domain adaptation, training is performed by adding a new loss function obtained by multiplying a trade-off term to a loss function for classification. An Adam optimizer was adopted for training, and the batch size was set to 64. Pump data were input into the model without labels, whereas the CWRU

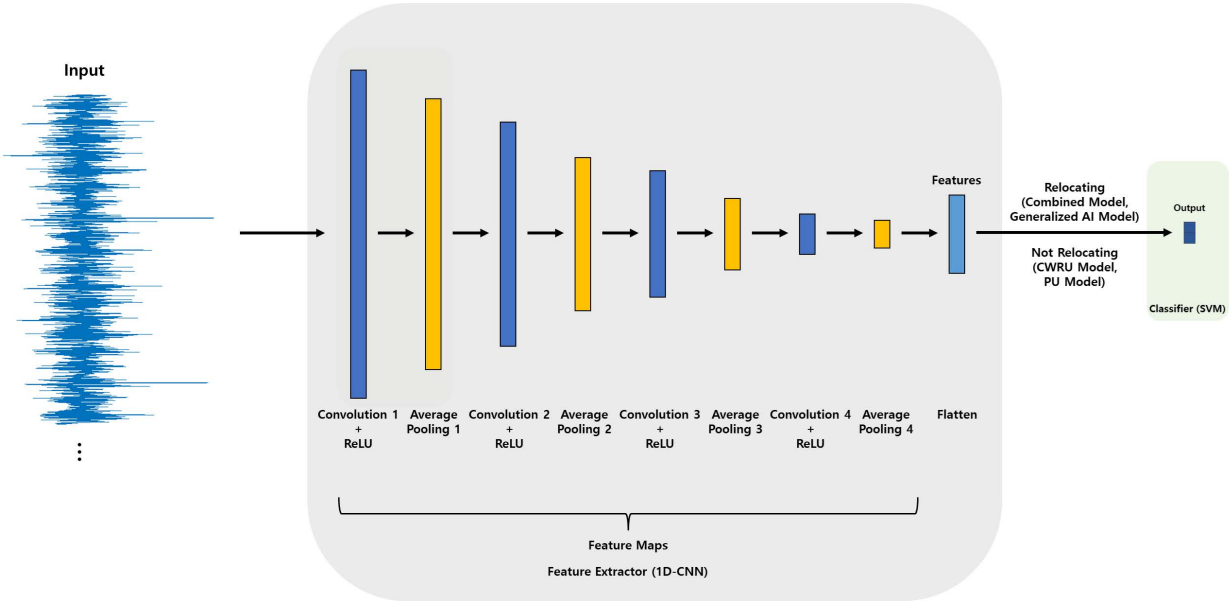

FIGURE 14. Example of structure for AI models.

TABLE 4. Structure for domain adaptation models [30].

Role	Layer	Parameters
-	Input	-
Extractor	Convolution 1	Kernel Size = 20, Stride = 1, Channel = 32
	Batch Normalization 1	-
	ReLU 1	-
	Average Pooling 1	Kernel Size = 2, Stride = 2
	Convolution 2	Kernel Size = 5, Stride = 1, Channel = 64
	Batch Normalization 2	-
	ReLU 2	-
	Average Pooling 2	Kernel Size = 2, Stride = 2
	Convolution 3	Kernel Size = 3, Stride = 1, Channel = 128
	Batch Normalization 3	-
	ReLU 3	-
	Adaptive Average Pooling	Output = 4
	Fully Connected	Out Features = 256
	ReLU 4	-
Classifier	Fully Connected	Output = 2
Discriminator	Fully Connected 1	Out Features = 512
	ReLU 1	-
	Fully Connected 2	Out Features = 1024
	ReLU 2	-
	Fully Connected 3	Out Features = 1
	Sigmoid	-

and PU data were adopted with labels. Normal and abnormal values were labeled as 0 and 1, respectively. The codes for domain adaptation published and demonstrated in [28] were modified and used with PyTorch and Python.

2) GENERALIZED AI MODEL AND COMBINED AI MODEL

A 1D-CNN (extractor) with SVM (classifier) was designed as illustrated in Figure 14 and utilized to address the challenge in which only normal data from the target domain are available. This model is similar to the one utilized in the domain adaptation models but has a simplified structure with no specific layers. The architecture of the CNN is one-dimensional, as illustrated in Figure 14, to make the input data traceable. The CNN was trained using common-domain data and labels in generalized AI models. Features extracted from the 1D-CNN were relocated before training the SVM, which was utilized for the classifier in the combined and generalized AI models. Relocation was performed using the normal data of each system. For example, the features of CWRU were rearranged using the average values of features extracted from the normal state of the CWRU training data. The target features were relocated in the same manner as the source domain, for which normal data in the target domain were fed into a trained extractor and transformed into features for relocating. Only this data were known to be normal, and the status for the rest of the data were unlabeled.

Herein, three types of models (AI Models 1–3) were designed and evaluated. The structure of AI Model 1 adopted two convolution and pooling layers, respectively. The kernel size of the first convolution layer was 20, and the other was set to 5. The pooling kernels were equally set to 2. In AI Model 2, a convolution layer with a kernel size of 3 and a pooling layer were added. The kernel of pooling layer 3 was 2. For AI Model 3, the kernel sizes of the convolution layers were adopted as 20, 10, 5, and 3, respectively. The size of the pooling layer was 2, as in previous models. Figure 14 illustrates the structure of AI Model 3, which is the longest among the AI models considered. The number of features extracted from each AI model was 249, 123, and 59, before fully-connected layer (size 20), respectively. The strides for the convolution layers were 1, and the pooling was 2. Adam optimizer and cross-entropy loss were, and the activation

TABLE 5. Descriptions of AI models.

Model Name	Training Set	Label		Feature Extractor	Classifier
Combined AI Model (combined raw data)	CWRU Normal	0*	0**	1D-CNN	SVM
	PU Normal	1*	0**		
	CWRU Abnormal	2*	1**		
	PU Abnormal	3*	1**		
Generalized AI Model (common-domain data)	CWRU Normal	0*	0**	1D-CNN	SVM
	PU Normal	1*	0**		
	CWRU Abnormal	2*	1**		
	PU Abnormal	3*	1**		

* : Labels for feature extraction in the CNN
** : Labels for classification in the SVM

TABLE 6. Descriptions of the CWRU and PU models.

Model Name	Training Set	Labels	Feature Extractor	Classifier
CWRU Model (raw data)	CWRU Normal	0*,**	1D-CNN	SVM
	CWRU Abnormal	1*,**		
PU Model (raw data)	PU Normal	0*,**	1D-CNN	SVM
	PU Abnormal	1*,**		

* : Labels for feature extraction in the CNN
** : Labels for classification in the SVM

function for the convolution layers was a rectified linear unit (ReLU). Early stopping using the callbacks function in Keras was adopted in training CNN. The training data were divided into a specific ratio and adopted as the training, validation, and test sets, respectively. The length of the input data for the CNN was 1,024. The generalized AI models were trained with common-domain data, whereas the combined AI models were trained with combined raw data. In the combined and generalized AI models, the datasets were labeled as presented in Table 5.

3) CWRU MODEL AND PU MODEL

The CWRU and PU models are AI models trained with the CWRU and PU data without signal processing, respectively. The models were designed to examine the characteristics of features extracted from each system before examining the results of the proposed method. The training processes were identical to those for the generalized and combined AI models, except that the features were not relocated. In addition, a structure with two convolution and pooling layers was utilized for both models. Trained models were evaluated using data that were not utilized in the training. For example, the CWRU model was trained with the CWRU and PU data, and the Pump data were evaluated. Similar to the proposed method, features were extracted from the CNN structure, and SVM was utilized as the classifier. Table 6 presents the manner in which each model is trained. Keras was adopted to train the CNN-based models, whereas MATLAB was adopted to train the SVM.

TABLE 7. Results of domain adaptation models.

Model	Number of Train Set	Number of Test Set (Pump)	Test Number	Test Accuracy (%) (Classification of Pump data)	Running Time (s)
CNN + MK-MMD (combined raw data)	4480	512	1	2.93	268.90
			2	3.71	276.28
			3	3.71	276.74
CNN + MK-MMD (common-domain data)	4480	512	1	90.62	307.62
			2	91.21	310.63
			3	97.66	447.60
CNN + DANN (combined raw data)	4480	512	1	0	269.82
			2	0	267.53
			3	0	461.81
CNN + DANN (common-domain data)	4480	512	1	100	297.46
			2	100	304.46
			3	100	299.63

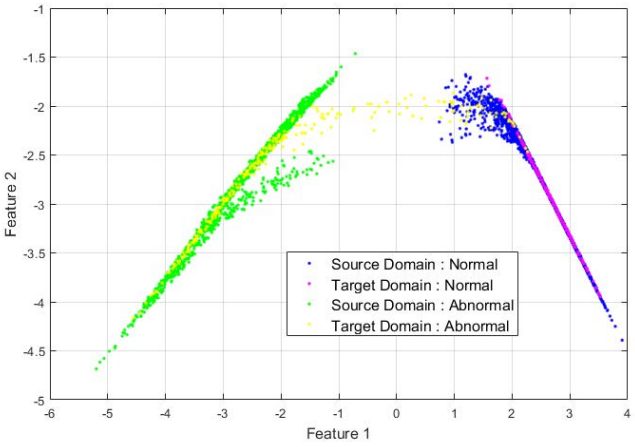

FIGURE 15. Comparison of features extracted from the CNN model with MK-MMD using common-domain data.

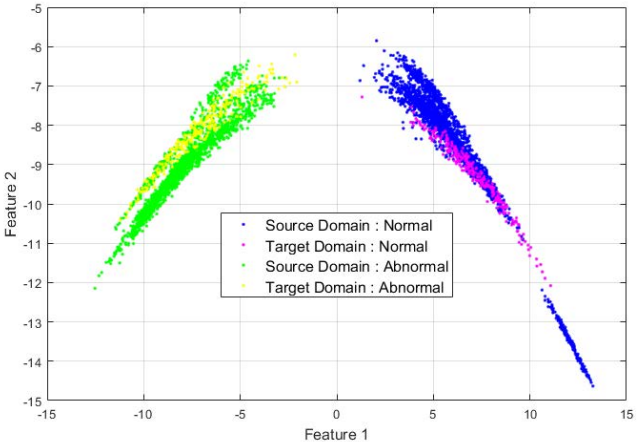

FIGURE 16. Comparison of features extracted from the CNN model with DANN using common-domain data.

V. RESULTS

A. RESULTS OF DOMAIN ADAPTATION MODELS

The accuracies of the domain adaptation methods presented in Table 7 are the results produced for the last epoch of each model. Combined raw and common-domain data were input into each model and compared. The two models could not accurately predict the states of the Pump data when trained with combined raw data. However, the use of common-domain data not only made it possible to accurately

TABLE 8. Results of the CWRU and PU models.

Model Name	System Name	Number of		Test Accuracy (%)		Running Time (s)
		Train Set	Test Set	Avg.	Std.	Avg.
CWRU Model (raw)	CWRU	1279	427	100	0	60.70
	PU	-	3500	50	0	
	Pump	-	2848	50	0	
PU Model (raw)	CWRU	-	2133	25.89	44.85	73.85
	PU	2100	700	99.95	0.08	
	Pump	-	2848	52.15	5.09	

extract fault characteristics, but it also classified the states of processed Pump data with high accuracy. The accuracies were over 90 percent regardless of the method employed when common-domain data were adopted.

Figures 15 and 16 illustrate one of the three experiments for each model using principal component analysis (PCA). Via the distributions, it can be confirmed that the results of the domain adaptation models are valid, and common-domain data are effective in diagnosing noisy systems, even though the labels of the target domain are not available.

B. RESULTS PRODUCED BY CWRU AND PU MODELS

Table 8 presents the data sizes for each process and classification results of the CWRU and PU models. The results are presented using the average and standard deviation of the classification results with an average running time of three times. The CWRU model is an AI model that was trained using raw CWRU data. A total of 1,279 datasets were utilized for training, the model was verified with 427 datasets in training CNN, and tests were conducted with 427 different datasets. The model classified the CWRU test sets perfectly, whereas it failed to classify the states of other bearings. Figure 17 shows the results for one of the three iterations using confusion matrix. The PU and Pump datasets were not used in training.

To analyze the CWRU model, features of the three datasets extracted by the CWRU model are drawn in a two-dimensional space using PCA, as illustrated in Figure 18. Because the criterion of the normal state for the CWRU data is different from that for the PU data, the PU data are misclassified with the criterion of the CWRU model. The model also fails to classify the Pump data. However, it can be observed that the fault characteristics are extracted for each system because the features extracted from the normal and abnormal data are distinguishable depending on the states.

The PU model was developed by training the AI model using raw PU data. The PU model was trained and evaluated with CWRU and Pump datasets, as presented in Table 8. Similar to the CWRU model, the results are represented by the average and standard deviation of classification with an average running time of three times. The labels were the same as those utilized in the CWRU model. A total of 2,100

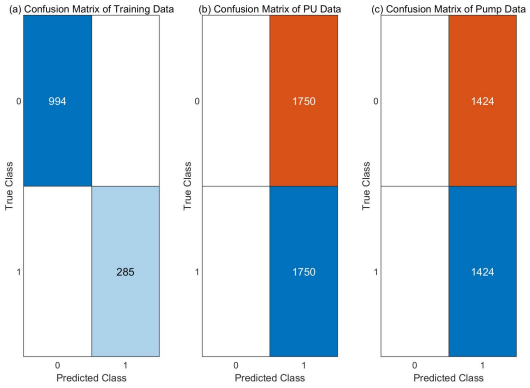

FIGURE 17. Confusion matrices of the CWRU model: (a) confusion matrix of the training set. (b) confusion matrix of the PU data (not adopted in training). (c) confusion matrix of the Pump data (not adopted in training).

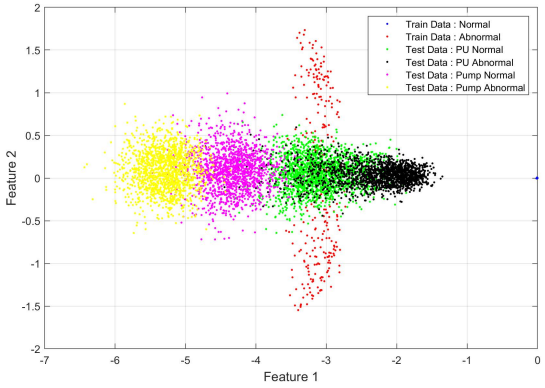

FIGURE 18. Comparison of features extracted from the CWRU model.

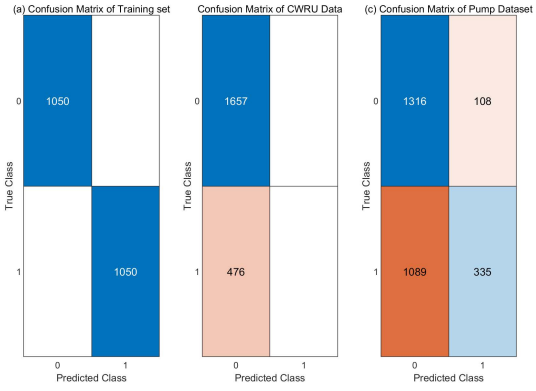

FIGURE 19. Confusion matrices of the PU model: (a) confusion matrix of the training set, (b) confusion matrix of the CWRU data (not adopted in training), (c) confusion matrix of the Pump data (not adopted in training).

datasets were utilized for training, and 700 different sets were utilized for validation and testing. When the PU model was evaluated with the PU test sets, it provided an accuracy of approximately 100 percent. However, the PU model failed to classify the CWRU and Pump data, similar to the CWRU

TABLE 9. Test results of the combined AI models.

Model Name	Model Number	Test Number	System Name	Size of		Test Accuracy (%)		Running Time (s)
				Training data	Test data	Normal	Abnormal	
Combined AI Model	1	1	CWRU	994/285	346/109	100	100	93.33
			PU	1050/1050	351/321	0	100	
			Pump	712*/-	712/1424	0	100	
		2	CWRU	994/285	328/94	100	100	79.83
			PU	1050/1050	361/344	0	100	
			Pump	712*/-	712/1424	0	100	
		3	CWRU	994/285	319/98	100	100	89.94
			PU	1050/1050	355/355	0	100	
			Pump	712*/-	712/1424	0	100	
	2	1	CWRU	994/285	344/91	100	100	100.07
			PU	1050/1050	337/355	0	100	
			Pump	712*/-	712/1424	0	100	
		2	CWRU	994/285	313/99	100	100	127.99
			PU	1050/1050	358/357	34.36	100	
			Pump	712*/-	712/1424	0	100	
		3	CWRU	994/285	319/99	100	100	110.08
			PU	1050/1050	342/367	0	100	
			Pump	712*/-	712/1424	0	100	
	3	1	CWRU	994/285	323/108	100	100	111.05
			PU	1050/1050	327/369	18.04	100	
			Pump	712*/-	712/1424	1.12	100	
		2	CWRU	994/285	328/90	100	100	108.30
			PU	1050/1050	359/350	10.03	100	
			Pump	712*/-	712/1424	57.02	70.08	
		3	CWRU	994/285	330/107	100	100	100.59
			PU	1050/1050	345/345	0	100	
			Pump	712*/-	712/1424	0	100	

*: Used in relocating

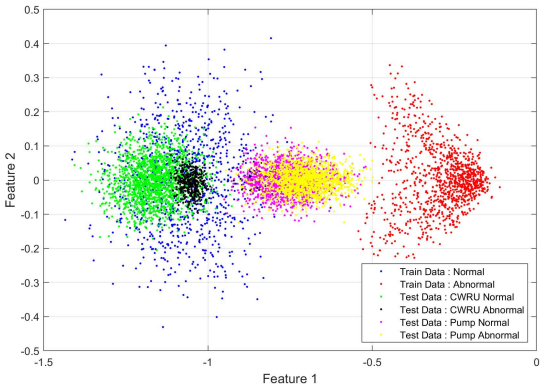

FIGURE 20. Comparison of features extracted from the PU model.

FIGURE 21. Confusion matrices of combined AI Model 1: (a) confusion matrix of the training set, (b) confusion matrix of the CWRU data (not adopted in training), (c) confusion matrix of the PU data (not adopted in training) and (d) confusion matrix of the Pump data (not adopted in training and calculating relocating features).

model. One of the results produced by the PU model is illustrated in Figure 19, using confusion matrices.

In Figure 20, each feature of the three datasets extracted from the PU model is compared in the same way as for the CWRU model. The features of the training sets are distributed more widely than those of the CWRU model. Thus, if the two datasets are trained together, a more general boundary may be set, and various characteristics may be applied to the model. However, because the signals are obtained from entirely different systems, they must be processed to be able to treat them as the same domain and expand the domains of the training data. The necessity of signal processing and signal combination is confirmed using combined and generalized AI models.

C. RESULTS PRODUCED BY COMBINED AI MODELS

Combined AI models, which are AI models trained with combined raw data, were developed to check the validity of the signal processing and combination. Generalizing the AI

TABLE 10. Test results of the generalized AI models.

Model Name	Model Number	Test Number	System Name	Size of		Test Accuracy (%)		Running Time (s)
				Training data	Test data	Normal	Abnormal	
generalized AI Model	1	1	CWRU	994/285	325/90	99.69	100	114.93
			PU	1050/1050	355/357	100	100	
			Pump	712*/-	712/1424	100	100	
		2	CWRU	994/285	333/96	99.70	100	109.65
			PU	1050/1050	356/342	100	100	
			Pump	712*/-	712/1424	100	97.47	
		3	CWRU	994/285	324/88	99.69	100	104.27
			PU	1050/1050	363/352	100	100	
			Pump	712*/-	712/1424	100	99.37	
	2	1	CWRU	994/285	330/94	100	100	80.86
			PU	1050/1050	346/357	100	100	
			Pump	712*/-	712/1424	100	99.86	
		2	CWRU	994/285	313/94	100	100	125.89
			PU	1050/1050	364/356	100	100	
			Pump	712*/-	712/1424	99.72	99.30	
		3	CWRU	994/285	330/94	99.39	100	150.38
			PU	1050/1050	335/368	100	100	
			Pump	712*/-	712/1424	99.58	99.37	
	3	1	CWRU	994/285	301/96	100	0	92.49
			PU	1050/1050	382/348	100	100	
			Pump	712*/-	712/1424	100	0	
		2	CWRU	994/285	338/99	100	0	124.69
			PU	1050/1050	345/345	100	99.71	
			Pump	712*/-	712/1424	100	0	
		3	CWRU	994/285	344/94	100	5.32	80.56
			PU	1050/1050	346/343	100	100	
			Pump	712*/-	712/1424	100	0	

*: Used in relocating

model without signal processing was difficult, even though the data were combined. In Table 9, the results are summarized for the combined AI models with data sizes for both training and testing. As indicated for most models and trials, the classification of CWRU is obtained with an accuracy of up to 100 percent, but classifications for the PU and Pump data are not performed exactly. Three tests were performed for each model, and one of the results for combined AI Model 1 is illustrated in Figure 21 using confusion matrices.

D. RESULTS PRODUCED BY GENERALIZED AI MODELS

Generalized AI models were trained using common-domain data, and the model was utilized to verify that both training and test data were classified even in a situation wherein only normal data of the target domain were available. Testing data were also processed. The feature extraction network was trained using common-domain data to extract the common fault characteristics. The generalized AI model demonstrated an accuracy of up to approximately 100 percent for the test data comprising processed CWRU and PU data when Models 1 and 2 were adopted. In addition, Pump data that were not adopted for training the CNN and SVM were classified with an accuracy of over 97 percent when the size of the features was 249 and 123. Therefore, the generalized AI

FIGURE 22. Confusion matrices of generalized AI Model 1: (a) confusion matrix of the training sets, (b) confusion matrix of the CWRU datasets (not adopted in training), (c) confusion matrix of the PU datasets (not adopted in training), and (d) confusion matrix of the Pump datasets (not adopted in training and calculating relocating features).

model can be concluded to demonstrate a reliable performance in diagnosing the three bearing systems. In addition, the model accurately classifies the states of Pump data even though normal Pump data are only adopted in relocating target features and are not adopted in training both the CNN and SVM. However, when the number of features was 59,

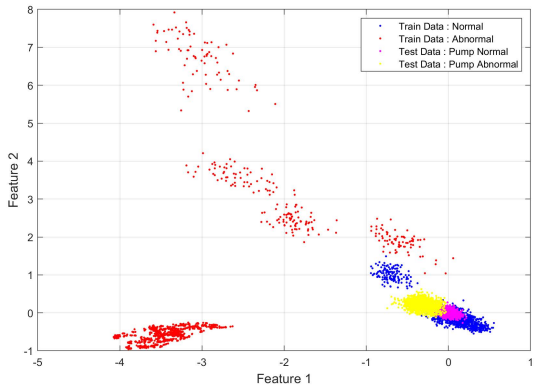

FIGURE 23. Comparison of features extracted from generalized AI Model 1, which utilizes type 1 structure and common-domain data.

the model could not be generalized. That is, generalized AI Model 3 is not suitable for diagnosing the three bearing systems. All the results are presented in Table 10.

Finally, the features extracted from generalized AI Model 1 are compared, as illustrated in Figure 23. The features extracted using the processed data are all located in the same feature space, and the inner faults of the three systems can be classified using the same decision boundary.

VI. CONCLUSION

Developing an AI model to diagnose machines with only available data, such as normal operating data, is the key to its practical usage in industries because it is impractical to have sufficient training data for all the systems concerned. Feature domain adaptation methods have been employed to address these challenges. However, such methods fail to extract the features of defects without eliminating structural and operational noise. Therefore, herein, signal processing was performed in advance to transfer each data into the common domain, which made it possible to increase the amount of training data. In addition, the machine was successfully diagnosed using only normal data by transferring the data into the common-domain. Thus, in this study, a novel method is proposed to increase the training data for extracting general fault characteristics by transferring data from different systems into a common-domain and combining them. The following is a summary of the results.

1) Two different datasets (CWRU data and PU data) were processed and combined to build big training data, referred to as common-domain data. When using domain adaptation models with common-domain data, Pump data were classified with an accuracy of over 90 percent. However, combined raw data in the cross domain failed to train the models to extract fault features.

2) Features are extracted with 1D-CNN from the common-domain data and, SVMs are developed with the features which are relocated in the feature domain using the normal data of all the systems. Via this process, all the features are placed in the common-domain, and all the systems can be classified with common decision boundary. Models with different number of layers were tested and validated. Pump data were classified with over 97 percent accuracy regardless of states when AI Model 1 and 2 were adopted. The generalized AI models exhibited good accuracy similar to the domain adaptation models. However, they required less information.

Timely-maintenance of machines is extremely important for the safe operation of a plant. However, it is often impossible to predict the lifetime because the entire history of the machine is unknown. Therefore, a study on aging diagnosis using the common-domain method can be useful because information on the history of a specific machine can be obtained.

REFERENCES

[1] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005.

[2] M. Djeddi, P. Granjon, and B. Leprettre, "Bearing fault diagnosis in induction machine based on current analysis using high-resolution technique," in *Proc. IEEE Int. Symp. Diag. Electr. Mach., Power Electron. Drives*, Sep. 2007, pp. 23–28.

[3] Q. Hu, X.-S. Si, A.-S. Qin, Y.-R. Lv, and Q.-H. Zhang, "Machinery fault diagnosis scheme using redefined dimensionless indicators and mRMR feature selection," *IEEE Access*, vol. 8, pp. 40313–40326, 2020.

[4] V. Kannan, H. Li, and D. V. Dao, "Demodulation band optimization in envelope analysis for fault diagnosis of rolling element bearings using a real-coded genetic algorithm," *IEEE Access*, vol. 7, pp. 168828–168838, 2019.

[5] Z. Wang, Q. Zhang, J. Xiong, M. Xiao, G. Sun, and J. He, "Fault diagnosis of a rolling bearing using wavelet packet denoising and random forests," *IEEE Sensors J.*, vol. 17, no. 17, pp. 5581–5588, Sep. 2017.

[6] G. Wang, F. Gu, I. Rehab, A. Ball, and L. Li, "A sparse modulation signal bispectrum analysis method for rolling element bearing diagnosis," *Math. Problems Eng.*, vol. 2018, Jan. 2018, Art. no. 2954094.

[7] J. Guo and P. Zheng, "A method of rolling bearing fault diagnose based on double sparse dictionary and deep belief network," *IEEE Access*, vol. 8, pp. 116239–116253, 2020.

[8] X. Ye, Y. Hu, J. Shen, R. Feng, and G. Zhai, "An improved empirical mode decomposition based on adaptive weighted rational quartic spline for rolling bearing fault diagnosis," *IEEE Access*, vol. 8, pp. 123813–123827, 2020.

[9] Y. Li, S. Si, Z. Liu, and X. Liang, "Review of local mean decomposition and its application in fault diagnosis of rotating machinery," *J. Syst. Eng. Electron.*, vol. 30, pp. 799–814, Aug. 2019.

[10] M. Kobayashi and K. Nakano, "Two problems of wavelet packet transform," in *Proc. 10th Int. Conf. Inf. Technol., New Generat.*, Las Vegas, NV, USA, Apr. 2013, pp. 153–159.

[11] Y. Gao, G. Ge, Z. Sheng, and E. Sang, "Analysis and solution to the mode mixing phenomenon in EMD," in *Proc. Congr. Image Signal Process.*, 2008, pp. 223–227.

[12] L. Yuan, D. Lian, X. Kang, Y. Chen, and K. Zhai, "Rolling bearing fault diagnosis based on convolutional neural network and support vector machine," *IEEE Access*, vol. 8, pp. 137395–137406, 2020.

[13] W. Huang, J. Cheng, Y. Yang, and G. Guo, "An improved deep convolutional neural network with multi-scale information for bearing fault diagnosis," *Neurocomputing*, vol. 359, pp. 77–92, Sep. 2019.

[14] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.

[15] D. Prenga, "General features of the q-XY opinion model," *J. Hum., Earth, Future*, vol. 1, no. 2, pp. 87–96, Jun. 2020.

[16] Y. Liu, X. Yan, C.-A. Zhang, and W. Liu, "An ensemble convolutional neural networks for bearing fault diagnosis using multi-sensor data," *Sensors*, vol. 19, no. 23, p. 5300, Dec. 2019.

[17] T. Han, L. Zhang, Z. Yin, and A. C. C. Tan, "Rolling bearing fault diagnosis with combined convolutional neural networks and support vector machine," *Measurement*, vol. 177, Jun. 2021, Art. no. 109022.
[18] G. Xu, M. Liu, Z. Jiang, D. Söffker, and W. Shen, "Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning," *Sensors*, vol. 19, no. 5, p. 1088, 2019.
[19] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang, "Deep model based domain adaptation for fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2296–2305, Mar. 2017.
[20] L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, "Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data," *IEEE Trans. Ind. Electron.*, vol. 66, no. 9, pp. 7316–7325, Sep. 2019.
[21] B. Yang, Q. Li, L. Chen, C. Shen, and S. Natarajan, "Bearing fault diagnosis based on multilayer domain adaptation," *Shock Vib.*, vol. 2020, pp. 1–11, Sep. 2020.
[22] Y. Zou, Y. Zhang, and H. Mao, "Fault diagnosis on the bearing of traction motor in high-speed trains based on deep learning," *Alexandria Eng. J.*, vol. 60, no. 1, pp. 1209–1219, Feb. 2021.
[23] Z. Yang, X. Wang, and R. Yang, "Transfer learning based rolling bearing fault diagnosis," in *Proc. IEEE 10th Data Driven Control Learn. Syst. Conf. (DDCLS)*, May 2021, pp. 354–359.
[24] X. Li, H. Jiang, S. Liu, J. Zhang, and J. Xu, "A unified framework incorporating predictive generative denoising autoencoder and deep coral network for rolling bearing fault diagnosis with unbalanced data," *Measurement*, vol. 178, Jun. 2021, Art. no. 109345.
[25] M. Zhang, D. Wang, W. Lu, J. Yang, Z. Li, and B. Liang, "A deep transfer model with Wasserstein distance guided multi-adversarial networks for bearing fault diagnosis under different working conditions," *IEEE Access*, vol. 7, pp. 65303–65318, 2019.
[26] W. Mao, Y. Liu, L. Ding, A. Safian, and X. Liang, "A new structured domain adversarial neural network for transfer fault diagnosis of rolling bearings under different working conditions," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–13, 2021.
[27] T. Han, C. Liu, W. Yang, and D. Jiang, "A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults," *Knowl.-Based Syst.*, vol. 165, pp. 474–487, Feb. 2019.
[28] Z. Zhao, Q. Zhang, X. Yu, C. Sun, S. Wang, R. Yan, and X. Chen, "Applications of unsupervised deep transfer learning to intelligent fault diagnosis: A survey and comparative study," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–28, 2021.
[29] J. Furch and Q. H. Nguyen, "Lifetime test of tracked vehicle torsion bars using Monte Carlo method," *Emerg. Sci. J.*, vol. 4, no. 5, pp. 376–389, Oct. 2020.
[30] T. Kim and J. Chai, "Pre-processing method to improve cross-domain fault diagnosis for bearing," *Sensors*, vol. 21, no. 15, p. 4970, Jul. 2021.
[31] H. Zheng, Y. Yang, J. Yin, Y. Li, R. Wang, and M. Xu, "Deep domain generalization combining a priori diagnosis knowledge toward cross-domain fault diagnosis of rolling bearing," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.
[32] A. Gretton, B. Sriperumbudur, D. Sejdinovic, H. Strathmann, S. Balakrishnan, M. Pontil, and K. Fukumizu, "Optimal kernel choice for large-scale two-sample tests," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1214–1222.
[33] M. Long, Y. Cao, Z. Cao, J. Wang, and M. I. Jordan, "Transferable representation learning with deep adaptation networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 12, pp. 3071–3085, Dec. 2019.
[34] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, pp. 1–35, Jan. 2016.
[35] M. C. Bishop, "Sparse kernel machines," in *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006, pp. 325–356.
[36] A. K. Sharma, G. Aggarwal, S. Bhardwaj, P. Chakrabarti, T. Chakrabarti, J. H. Abawajy, S. Bhattacharyya, R. Mishra, A. Das, and H. Mahdin, "Classification of Indian classical music with time-series matching deep learning approach," *IEEE Access*, vol. 9, pp. 102041–102052, 2021.
[37] D. Alsaleh and S. Larabi-Marie-Sainte, "Arabic text classification using convolutional neural network and genetic algorithms," *IEEE Access*, vol. 9, pp. 91670–91685, 2021.
[38] I. Hababeh, I. Mahameed, A. A. Abdelhadi, and A. Barghash, "Utilizing convolutional neural networks for image classification and securing mobility of people with physical and mental disabilities in cloud systems," *IEEE Access*, vol. 8, pp. 163730–163745, 2020.
[39] S. Tiwari, A. Jain, A. K. Sharma, and K. M. Almustafa, "Phonocardiogram signal based multi-class cardiac diagnostic decision support system," *IEEE Access*, vol. 9, pp. 110710–110722, 2021.
[40] A. K. Sharma, A. Panwar, P. Chakrabarti, and S. Vishwakarma, "Categorization of ICMR Using feature extraction strategy and MIR with ensemble learning," *Proc. Comput. Sci.*, vol. 57, pp. 686–694, Jan. 2015.
[41] I. Goodfellow, Y. Bengio, and A. Courville, "Convolutional networks," *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, pp. 321–361.
[42] T. A. Harris and M. N. Kotzalas, "Vibration, noise, and condition monitoring," in *Essential Concepts of Bearing Technology*, 5th ed. Boca Raton, FL, USA: CRC Press, 2006.
[43] Case Western Reserve University Bearing Data Center. Accessed: Apr. 10, 2022. [Online]. Available: https://engineering.case.edu/bearingdatacenter
[44] C. Lessmeier, J. K. Kimotho, D. Zimmer, and W. Sextro, "Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification," in *Proc. Eur. Conf. Prognostics Health Manage. Soc.*, 2016, pp. 5–8.

TAEYUN KIM received the B.S. and M.S. degrees in mechanical engineering from Ajou University, Suwon, South Korea, in 2015 and 2017, respectively, where he is currently pursuing the Ph.D. degree with the Department of Mechanical Engineering. His research interests include machine learning, deep learning, signal processing, and vibration analyses.

JANGBOM CHAI received the B.S. and M.S. degrees in mechanical engineering from Seoul National University, South Korea, in 1984 and 1986, respectively, and the Ph.D. degree in mechanical engineering from MIT, in 1993. He studied machine diagnostics at the Acoustics/Vibration Laboratory, MIT. He is currently a Professor with the Department of Mechanical Engineering, Ajou University. His research interests include prognostics and health management based on artificial intelligence in the power plant industry. He received the O. Hugo Schuck Best Award from the American Automatic Control Council and awards from the Korean Nuclear Society and the Korean Society of Pressure Vessel and Piping.