IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 23 June 2022, accepted 18 July 2022, date of publication 21 July 2022, date of current version 28 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3193244

RESEARCH ARTICLE

Rolling Bearing Fault Diagnosis Based on Improved GAN and 2-D Representation of Acoustic Emission Signals

MINH TUAN PHAM[1], JONG-MYON KIM[2], AND CHEOL HONG KIM[3]
[1]Branch of Qualcomm Vietnam Company Ltd., Hanoi 11606, Vietnam
[2]School of IT Convergence, University of Ulsan, Ulsan 44610, South Korea
[3]School of Computer Science and Engineering, Soongsil University, Seoul 06978, South Korea

Corresponding author: Cheol Hong Kim (cheolhong@ssu.ac.kr)

This work was supported by the National Research Foundation of Korea (NRF) Grant through the Korean Government [Ministry of Science and ICT (MIST)] under Grant 2021R1A2C1009031.

ABSTRACT Bearing fault diagnosis is essential in manufacturing systems to avoid problems such as downtime costs. Convolutional neural network (CNN) models have enabled a new generation of intelligent bearing fault diagnosis methods for smart manufacturing owing to their capability to extract features for 2-dimensional (2D) representations, such as signals represented in the time-frequency domain. Nevertheless, the cost and time required to collect sufficient training data tend to result in a lack of data and data imbalance in real fault diagnosis scenarios. This inevitable consequence leads to a high misclassification rate in conventional CNN models. In this study, to address this problem, we propose a novel effective generative adversarial network (GAN)-based method for rolling bearing fault diagnosis in early-stage and low rotational speeds based on data enhancement, which uses acoustic emission (AE) as a monitoring signal. In the proposed approach, generator, discriminator, and fault classifier models are trained simultaneously with the proposed strategy for updating parameters to avoid the gradient vanishing problem and outperform conventional methods. The fault classifier was developed based on CNN models which are compatible with 2-D signal representations represented by a constant-Q transform. The results of experiments conducted with unbalanced compound fault datasets verify the capabilities of the proposed method in various diagnosis scenarios compared with traditional methods, including SVM, CNN, and DCGAN models.

INDEX TERMS Acoustic emission, bearing fault diagnosis, convolutional neural network, GAN, unbalance data.

I. INTRODUCTION

Bearings play a critical role in rotating machinery and are widely used in complex industrial equipment and systems, such as engines, electrical motors, and wind turbines [1]. However, harsh working environments cause defects in bearings, which leads to a high probability of incidents, downtime, low production quality, and economic losses [2]. Thus, timely and accurate detection of rolling bearing failure is essential to maintain the safety and stability of machinery operations. In bearing fault diagnosis, the monitoring signals collected by sensors located throughout the machinery convey bearing health information, which is then analyzed to determine the bearing status. Conventionally, vibration is the most common monitoring signal, which can provide hidden information about local bearing damage such as wear, fatigue pitting, or spalling. The collected monitoring signals commonly comprise multicomponent data owing to noise submersion in real working environments. Thus, time-frequency domain methods have been widely utilized compared to single-domain analysis (time- and frequency-domains). They include empirical mode decomposition (EMD) [3], independent component analysis (ICA) [4], and synchronous wavelet transform (SWT) methods [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Yu-Huei Cheng.

Compared with conventional inspection methods based on vibration, acoustic emission (AE) techniques have been considered promising candidates for system integrity monitoring of rotating machines and bearings. Intuitively, the AE releases strain energy caused by the initiation and propagation of cracks [6]. This shows the advantage of detecting a range of damage mechanisms in their early stages before they become significant issues. Balderston *et al.* showed the limitation of vibration signals compared to AE in terms of low rotational speed where the energy impact is low [7]. Balderston used two types of AE signals, including burst-type emissions to detect element defects and continuous-type emissions for the state of lubrication starvation. Rogers utilized AE-based techniques to propose a monitoring method for bearings cranes attached to gas equipment [8]. Based on statistics, the kurtosis value was measured to detect the abnormal state and specific location of cracks. The authors also claimed that while vibration analysis is not appropriate in this application for low operating speeds, AE analysis was nonetheless informative in these cases. The AE signal was also found to be compatible with online condition monitoring rather than vibration. Another method proposed by Yoshioka and Fujiwara found that the detection range of AE parameters to be wider than that of vibration acceleration by conducting an experiment with a bearing lubricated by machine grease at a rotational speed of 660 RPM and a load 4.2kN [9]. In addition, Hawman and Galinaitis concurred with this conclusion and explained that the AE bearing fault signals are modulated at high frequency when it comes to AE signals [10]. However, alongside these notable advantages, some significant challenges remain to be addressed, such as the overlapping signals of individual monitored components (gears and bearings) and ambient noise. Thus, to improve the capabilities of AE bearing fault diagnosis technologies at the levels of individual components or machine systems to adapt to unforeseeable circumstances, novel analysis methods are required that are robust to various conditions and scenarios.

Regarding intelligent fault diagnosis, data-driven methods have been brought to the attention of many researchers in this field thanks to their high superiority of automation. Han *et al.* proposed a method with high generalization capability that ensure to work with unseen working conditions by regularizing both intrinsic and extrinsic generalization objectives [11]. In recent years, the advent of deep learning has advanced the field of fault diagnosis as the extension of data-driven approaches [12], [13]. In particular, convolutional neural networks (CNNs) have been widely utilized owing to their capability to extract the features accurately and automatically [14], [15]. Cao *et al.* proposed an unsupervised domain-share CNN to extract domain-invariant features for fault diagnosis in the conditions of varying rotational speeds [16]. Zhou *et al.* proposed a fault diagnosis framework based on Bayesian deep learning which help fault data to be understandable and deal with the problems of domain shifts [17]. Shao *et al.* proposed a method consisting of parameter transfer and infrared thermal images for fault diagnosis in variable rotating speeds [18]. Janssens *et al.* proposed CNNs applied to a 1-dimensional signal in the time domain, which automatically extracted and learned the features of several bearing fault types and lubrication degradation patterns [19]. Zhang *et al.* introduced a novel CNN-based framework with techniques to enhance the anti-noise and domain adaption ability, namely drop-out and ensemble learning [20]. Other methods have utilized the image processing capabilities of CNN models by combining CNNs with signal 2-D representation. Ding and He proposed an energy-fluctuating multiscale feature mining approach with wavelet packet energy images and CNNs for spindle bearings [21]. Wen *et al.* introduced a CNN method based on Lenet-5 and conducted experiments on various datasets of a bearing self-priming centrifugal pump and an axial piston hydraulic pump [22]. Abdeljaber *et al.* proposed a CNN-based method using 1D input data that combined feature extraction and classification into a single block for real-time diagnosis [23].

Supervised learning-based methods exhibit high accuracy, assuming that sufficient training data are available, and that data are balanced among categories. However, the data acquisition process is costly and time-consuming in real working environments. Therefore, the availability of sufficient training data for faulty cases are typically limited. Bearings operate primarily in the normal state; thus, the amount of data collected in this state outweighs that of faulty states. Many studies have been conducted to address the problem of performing diagnosis using an unbalanced dataset. Duan *et al.* combined a binary tree and support-vector data description optimized using particle swarm optimization to address the issues of unbalanced data [24]. Zhang *et al.* proposed a diagnostic approach based on empirical mode decomposition with a dynamic model built to generate fault data and address the problem of data imbalance [25]. In 2014, a novel approach for data enhancement, called generative adversarial networks (GAN), was introduced, which provided a new approach for various problems related to limited training samples [26]. Wang *et al.* proposed a method of constructing a discriminator for GAN models using a stacked noise reduction automatic encoder for gearbox fault diagnosis [27]. Han *et al.* used a CNN to establish a generator and discriminator to alleviate segmented datasets and improve model generation [28]. Nevertheless, previous GAN-based methods commonly separate the training of the model to generate supplementary samples and train the fault diagnosis model. This leads to difficulties in handling the accuracy of fault diagnosis tasks, which is the primary goal of an intelligent factory.

To address these problems, in this study, we propose a novel bearing fault diagnosis method that uses an improved GAN and 2-D spectrograms of AE signals in cases of unbalanced data and insufficient training samples. The proposed generator and discriminator loss functions are modified to avoid gradient vanishing problems and trained simultaneously with the fault classifier using an

appropriate scheme. In addition, from the perspective of high-dimensional input data, CNN can be further adapted as a bearing fault classifier to improve feature mining to achieve accurate diagnosis in various scenarios, such as variable rotational speeds, various crack sizes, various fault types, and noise environments. The main contributions of this work are summarized as follows.

1) A novel diagnosis based on an AE monitoring signal is proposed to perform fault prediction at an early stage and low rotational speeds, combining a time-frequency spectrogram representation by constant-Q transform [29] and a GAN-based method.
2) A new GAN generator, discriminator, and classifier are designed for bearing fault diagnosis under conditions of insufficient training samples and data imbalance, which alleviates the vanishing gradient problem.
3) We also establish a novel GAN training scheme that updates a fault classifier model during training of the generator and discriminator based on predetermined conditions to improve the performance of the classifier.

The remainder of this study is structured as follows.

Section II provides an overview of GAN models and the vanishing problem, Section III describes the proposed GAN scheme and architecture for bearing fault diagnosis, Section IV presents the proposed fault diagnosis process based on an improved GAN, Section V describes the datasets used for evaluations, as well as the experiments conducted and their results, as well as some analysis; finally, Section VI concludes the work.

II. RELATED PREREQUISITES

The proposed efficient method initially converts the raw AE signals into time-frequency spectrogram images by using CQT. Then, proposed efficient CNNs established by search is adopted to extract the insight features of those images. Finally, the conditions of bearing and fault types are classified by using fully connected layers. This section reviews previous works in this field of intelligent bearing fault diagnosis. Next, the background about techniques adopted in our proposed method is also introduced (CQT and efficient CNN design techniques) with some highlighted advantages in efficiency.

A. GENERATIVE ADVERSARIAL NETWORK

A general GAN architecture is depicted in Fig. 1, which consists of two main components, a generator G, and a discriminator D. The generator feeds the random parameter $z \in R^d$ called a latent variable (where d is the number of dimensions of variable z, and commonly, $z \sim N(0, 1)$). The fake sample is then generated by $x' = G(z)$ to fool the discriminator D into classifying it as true data. To summarize, there is a min-max game between D and G with the competition known as optimizing the loss function $L(D,G)$ with given real sample x:

$$\begin{aligned}
&\min_D \max_G L(D, G)\\
&= \min_D \max_G \{-E_{x\sim p_r(x)}[\log D(x)]\\
&\quad -E_{z\sim p_z(z)}[\log(1 - D(G(z)))]\}\\
&= \min_G \max_D \{E_{x\sim p_r(x)}[\log D(x)] + E_{x\sim p_g(x)}[\log(1 - D(x)]\}\\
&= \min_G \max_D \left\{ \int_x (p_r(x)\log D(x) + p_g(x)\log(1 - D(x)))dx \right\}
\end{aligned} \tag{1}$$

where E is the expectation, and p_r, p_z, p_g are probability distributions of x, z, x', respectively.

B. VANISHING GRADIENT PROBLEM IN GAN

The solution for this game is known as the Nash equilibrium, which is the point at which the GAN model converges, and neither D nor G can affect the result. From the loss function in (1), $\max_D L(D, G)$ can be expressed as [30]:

$$D^*(x) = \frac{p_r(x)}{p_r(x) + p_g(x)} \tag{2}$$

By the found D^*, the $\min_G L(D^*, G)$ can be then calculated as:

$$\min_G L(D^*, G) = 2D_{JS}(p_r||p_g) - 2\log 2 \tag{3}$$

where D_{JS} is Jensen–Shannon divergence.

Therefore, the gradient of L is linear to the gradient of D_{JS}, which vanishes when the probability distributions of the real and generated fake samples are significantly different in the early training stage. The reason for this is formulated by the following equation:

$$\nabla\theta_g L(D^*, G) = \nabla\theta_g(2D_{JS}(p_r||p_g)) \longrightarrow 0 \tag{4}$$

In contrast, a modified version of the loss functions of G and D is defined as [31]:

$$\begin{aligned}
&\min_D L(D, G)\\
&\quad = \min_D \left\{ \frac{1}{2}E_{x\sim p_r(x)}[(D(x) - 1)^2] \right.\\
&\qquad \left. + \frac{1}{2}E_{z\sim p_z(z)}[(D(G(z)) + 1)^2] \right\}\\
&\min_G L(D, G)\\
&\quad = \min_G \left\{ \frac{1}{2}E_{x\sim p_r(x)}[(D(x))^2] + \frac{1}{2}E_{z\sim p_z(z)}[D(G(z))^2] \right\}
\end{aligned} \tag{5}$$

The solution for $\min_D L(D, G)$ is found as:

$$D^*(x) = \frac{p_r(x) - p_g(x)}{p_r(x) + p_g(x)} \tag{6}$$

Then, $\min_G L(D^*, G)$ is transformed as follows:

$$\min_G L(D^*, G) = \chi^2_{Pearson}(p_r + p_g||2p_g) \tag{7}$$

The Pearson $\chi^2_{Pearson}$ divergence has $\nabla\theta_g\chi^2_{Pearson}$, which is not asymptotically zero, even though the distributions of

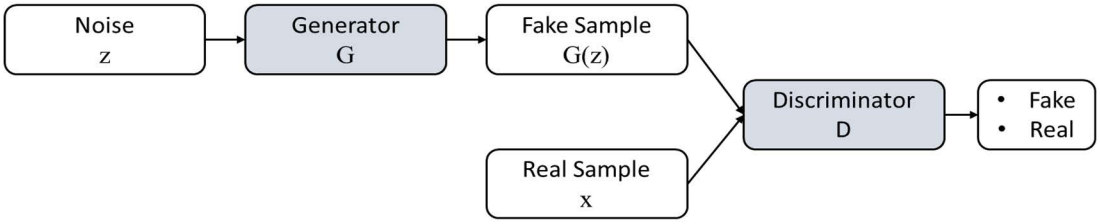

FIGURE 1. General GAN architecture.

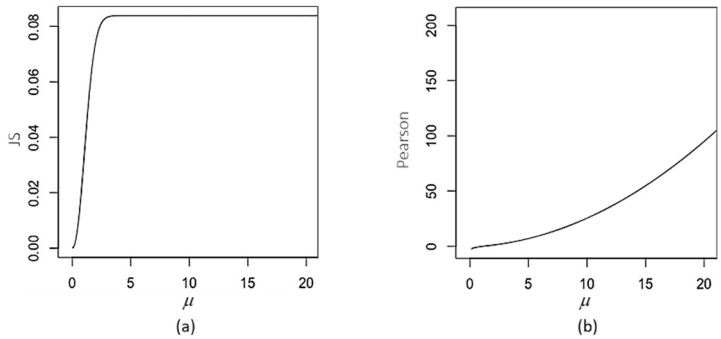

FIGURE 2. Graphs of two types of functions: (a) Jensen–Shannon divergence, (b) Pearson $\chi^2_{Pearson}$ divergence.

the real samples and generated fake samples are significantly different. Therefore, it helps to avoid vanishing problems during the early training stage. Fig. 2 illustrates the graphs of these two types of functions.

III. PROPOSED IMPROVED GAN FOR FAULT DIAGNOSIS

Data acquired from real industrial environments always exhibit the problem of imbalance because normal bearing conditions are the most common case. The cost and time requirements of data acquisition are also high, which leads to inefficiency in the conventional training strategy of intelligent fault diagnosis methods. In this section, we present the proposed improved GAN, which is designed to enhance fault diagnosis under the mentioned conditions, in which a generator, discriminator, and their loss functions are newly designed. Furthermore, a fault classifier based on a CNN model is also updated simultaneously by the scheme, as shown in Fig 3.

A. STRUCTURE OF THE PROPOSED GAN-BASED FAULT DIAGNOSIS NETWORK

To address gradient vanishing and control the GAN training process for rolling bearing fault datasets, we propose a GAN-based scheme consisting of three sub-models: a generator G, a discriminator D, and a classifier C. Instead of separating the fault classifier training stage from pseudo-sample generation, at every iteration, the generator G feeds on an input z' (a concatenation of random variable z and random pseudo-fault label y_{fake}) to generate a fake sample x' as output. Simultaneously, discriminator D is trained by fake samples and real samples (each data sample and its corresponding fault label are concatenated) to distinguish between fake and real samples. In addition, the classifier C is initially updated in a standard fashion by the available real samples, whose quantity is low. It is then trained by generating fake samples once the samples pass the predetermined conditions. Adaptation of the conditional GAN mechanism [32] produces labels for the generated images. In this manner, the classifier performs a supervised training process. Note that the generated images satisfying predetermined conditions are utilized to train the classifier immediately, which eliminates the step of saving newly generated samples.

In a common GAN, binary cross-entropy (BCE) is used as a generator and classifier loss function. However, given the abovementioned disadvantages of BCE-based loss causing gradient vanishing problems, we use the mean squared error loss (MSE) instead. According to (7), the loss function based on the L_2 – norm helps mitigate the gradient vanishing problem owing to the property of the Pearson $\chi^2_{Pearson}$ divergence.

The proposed GAN is used to solve a nonconvex optimization problem by iteratively solving each individual subproblem according to the gradient descent until an optimal solution is obtained.

1) UPDATING DISCRIMINATOR D

The discriminator loss is defined as:

$$L_D(x, z) = MSE(D(x \oplus y_{real}), 1) + MSE(D(G(z') \oplus y_{fake}), 0) \tag{8}$$

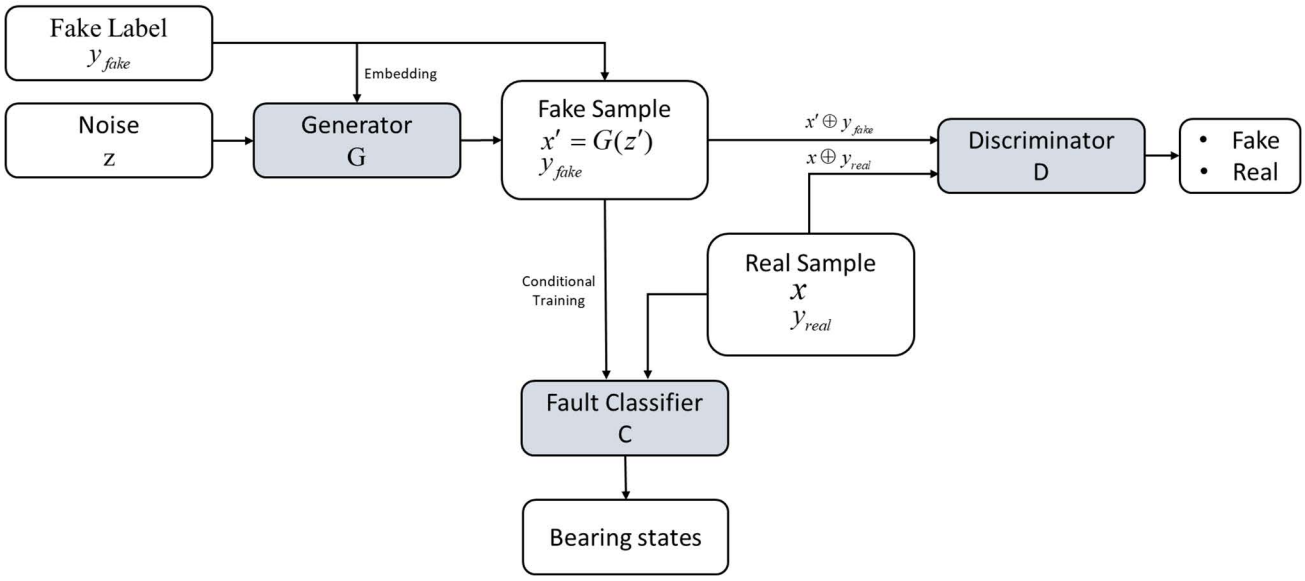

FIGURE 3. Proposed improved GAN architecture for bearing fault diagnosis.

where $z' = z \oplus Embedding(y_{fake})$, MSE is the mean squared error, L is the loss function, D is the discriminator, G is the generator, z is a random vector, x is the real data labeled y_{true}, y_{fake} is the generated label for fake data $x' = G(z')$ z(y_{true} and y_{fake} are fault-type labels), and $\oplus$ is the concatenation operator.

The defined discriminator loss indicates that the loss function includes two components. The former updates the discriminator on real data, which refers to real label 1, whereas the latter trains the discriminator on fake data, which refers to fake label 0.

2) UPDATING GENERATOR G

The loss of the generator is defined to indicate that the generated sample is real; therefore, its loss is simply calculated by the difference between the output of the discriminator fed by the fake sample and real label 1:

$$L_G(z) = MSE(D(G(z')), 1) \tag{9}$$

where z' is calculated from z and y_{fake}, which are randomly generated in each iteration. The embedding of y_{fake} acts as a supplemental feature indicating the belonging category of the generated sample, x'.

3) UPDATING CLASSIFIER C

Eventually, the loss of classifier is defined as:

$$L_C = CE(C(x), y_{real}) + \lambda \times CE(C(G(z')) > t, y_{fake}) \tag{10}$$

where C is the classifier, CE is the cross-entropy loss, t is the predetermined threshold of the prediction confidence, and λ is the adversarial weight (unsupervised loss weight).

Similar to the loss of the discriminator, L_C consists of two components updated by real samples and fake samples consecutively. The first component is supervised loss, which is calculated based on standard cross-entropy. The second is unsupervised loss, which is calculated once the probability of prediction is higher than a predetermined confidence threshold t. Owing to the low quality of the generated images in the early stage, the confidence threshold t is necessary to prevent a misleading update in the classifier. The value of the unsupervised loss is the cross-entropy between the predictions provided by the classifier for the given fake data and the corresponding fake labels. Moreover, unsupervised loss weight λ is used to control the influence of unsupervised learning, which we focus on less here because unsupervised learning plays a minor role as a supplement for supervised learning.

B. MODEL ARCHITECTURE

In this section, the proposed architectures for the classifier C, generator G, and discriminator D are introduced. The backpropagation algorithm updates all models with the support of Adam optimization during the training process, as suggested by [33]. First, the model of classifier C was constructed based on consecutive basic blocks which are designed based on CNN layers and shortcut or skip connections. The advantages of shortcut connections include capability to address feature degradation problems, providing better performance and efficiency (without requiring additional computational complexity and parameters) compared to other plain networks. Moreover, batch normalization is applied after every convolutional layer to perform a type of coordinated rescaling of the corresponding convolutional layer output.

Fig 4 illustrates the backbone of the proposed classifier architecture C, and Table 1 lists the configuration of classifier C in detail.

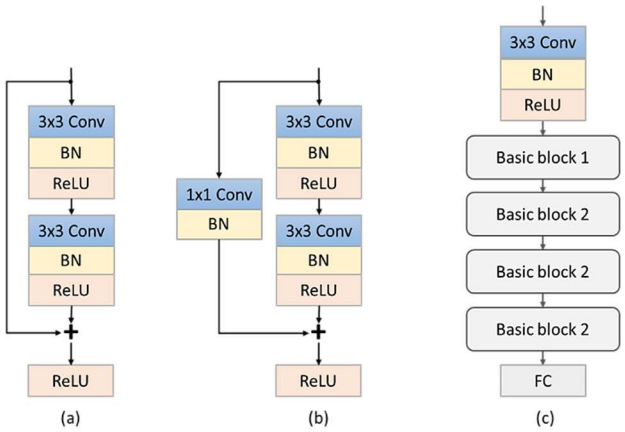

FIGURE 4. Proposed architecture based on basic blocks for classifier C: (a) basic block 1, (b) basic block 2, (c) the general model architecture where a×b Conv: convolutional layer with kernel size a×b, BN: Batchnorm operation, ReLU: rectified linear unit function, FC: fully connected layer.

TABLE 1. Configuration Of the architecture of classifier C.

Layer No.	Layer type	Output size (px)	Kernel size	Stride	Output channels
1	Input spectrogram	32×32	×	×	3
2	Conv	32×32	3×3	1	16
3	Basic block 1	32×32	×	×	16
4	Basic block 2	16×16	×	×	32
5	Basic block 2	8×8	×	×	64
6	Basic block 2	4×4	×	×	128
7	AvgPool	1×1	4×4	×	128
8	Classifier (FC)	×	×	×	8

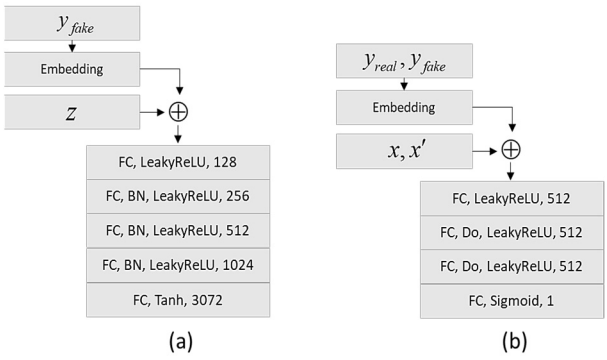

FIGURE 5. Proposed generator and discriminator architectures: (a) generator, (b) discriminator, where Do denotes drop-out operator, and the last number in each layer denotes the number of output neurons.

The generator and discriminator were built from common deep neural network (DNN) layers. LeakyReLU(0.2) (leaky rectified activation with factor 0.2), which maintains a negative value, is used as the activation instead of ReLU to work well. The final output of the generator is applied to the Tanh function to scale the output values to the range $[-1, 1]$, which has been shown to yield image quality [34]. The architectures of the generator G and discriminator D include consecutive layers, as shown in Fig. 5.

IV. GENERAL PROCEDURE OF THE PROPOSED METHOD

This section describes the entire proposed fault diagnosis process based on a GAN model, including AE data acquisition, signal preprocessing, the construction of the GAN-based model, as well as testing, and visualization of the results.

In this study, a novel GAN-based strategy for rolling bearing fault diagnosis with monitoring signals represented by 2-dimensional spectrogram images is proposed. The framework of the proposed method consists of bearing AE data acquisition, signal preprocessing, construction of the bearing fault diagnosis model, and testing and visualization of the results, as shown in Fig. 6. The general procedure is summarized as follows.

Step 1: Acquire AE monitoring signals that contain information on bearing health in various statuses (mostly in the normal state) under various working conditions using data acquisition systems.

Step 2: Segment the acquired AE signals by a fixed-length window (0.05 s) [34], and then convert each segment into a spectrogram image using a constant-Q transform (CQT), as shown in Fig. 7.

2-D spectrograms of time-frequency data have been shown to be advantageous when representing nonstationary signals. Among the methods used to convert the signals from time domain to time-frequency domain, constant-Q transform, a variant of wavelet transforms, is an efficient and effective method owing to their changeable resolutions in variable frequency bands. Indeed, the lower frequency bands containing useful information about bearing state have higher resolution compared to higher frequency bands where high-frequency noises occur. The CQT configuration is based on the following major factors: (1) window functions g_k, which are real-valued even functions. In the frequency domain, the Fourier transform of g_k is defined in the interval $[-F_s/2, F_s/2]$;(2) the sampling rate ω_s; (3) the number of bins per octave, and (4) the minimum and maximum frequencies, $\omega_{\min}$ and $\omega_{\max}$, respectively. After conversion into a spectrogram, all non-informative borders are removed before scaling to an image size of 32 × 32 by bilinear interpolation.

Step 3: Construct and train the models from the spectrogram dataset created using the proposed GAN strategy to update D, G, and C, as described in section III. After sufficient iterations, the training process ends once classifier C converges.

Step 4: Calculate the performance of the established model of classifier C using the testing set according to the metrics and visualize the features to ensure the interpretability of the method using t-SNE [35] and LIME [36].

With regards to metrics of evaluation, sensitivity was one of the essential metrics to compare diagnostic performances.

$$Sensitivity = \frac{N_{True\_Positive}}{N_{True\_Positive} + N_{False\_Negative}} \times 100(\%) \tag{11}$$

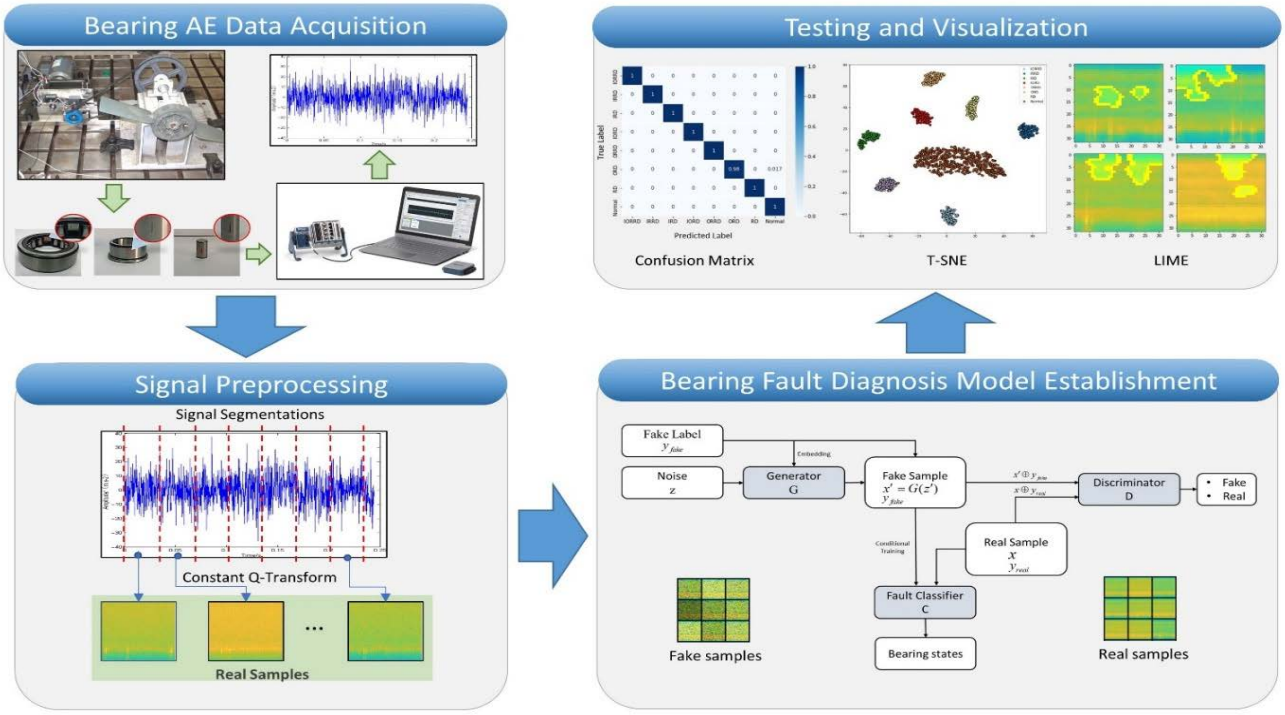

FIGURE 6. General procedure of the proposed method.

where $N_{True\_Positive}$ is the number of correctly predicted samples in a specific class and $N_{False\_Negative}$ is the number of incorrectly predicted samples in a specific class.

$$ACA = \frac{\sum \text{Sensitivity}}{\sum N_{Classes}} \tag{12}$$

Subsequently, the effect of z dimensions on the generated samples is also considered by using the QS score, which is based on the IS score [37]. A pre-trained classifier C using only a real dataset is used to measure the average distance between the label distribution $p(y|x')$ (less entropy to indicate high confidence) and marginal distribution $p(y)$ (high entropy to indicate high diversity in generated samples), where y is the label predicted by the pre-trained classifier C. Therefore, with a smaller QS score, a worse generator is established.

$$QS(G) = \exp(E_{x\sim p_g(x)} D_{KL}(p(y)||p(y|x))) \tag{13}$$

where D_{KL} is Kullback–Leibler divergence.

V. EXPERIMENTS AND ANALYSIS

This section introduces the datasets and describes the experiments conducted to evaluate and verify the effectiveness of the proposed method. The results were also compared and analyzed to demonstrate the performance of the proposed method under various scenarios.

A. DATASET

As shown in Fig. 8, the AE data were acquired from an experimental testbed consisting of a three-phase induction motor, a gearbox, and a load (an adjustable blade). The motion of the motor was transferred directly to the gearbox (damping ratio of 1.52:1), where two shafts were attached to the cylindrical bearings (FAG NJ206-E-TVP2). The non-drive-end shelf continued to transfer motion to the load via a belt and pulleys. Specifically, the monitoring bearing was attached to the non-drive end shaft.

TABLE 2. Bearing specification and FCFs.

Pitch diameter	46.5 (mm)
Rolling element diameter	9.0 (mm)
Contact angle	0°
Number of rolling elements	13
Outer defect frequency	5.241×f_r
Roller defect frequency	2.487×2×f_r
Inner defect frequency	7.758×f_r

For the data-acquisition system, a wideband AE sensor (PAC WSα) was mounted on the monitoring bearing. The signals from the AE sensor were captured by a PCI-2 based board at a sampling rate of 250 kHz. The frequency range of the AE sensor was 100–900 kHz, the peak sensitivity was −62 dB, and the directionality was ±1.5 dB. The range of the frequency response was between 1 kHz and 3 MHz, and the resonant frequency was 650 kHz. The specifications of the experimental bearing and its fault characteristic frequency (FCF) are shown in Table 2.

Various types of bearing faults were created using diamond cutting of different sizes at different positions on the bearing (Fig. 9). We initially prepared two primary datasets, A and B, from the data acquired from these seven types of faults and

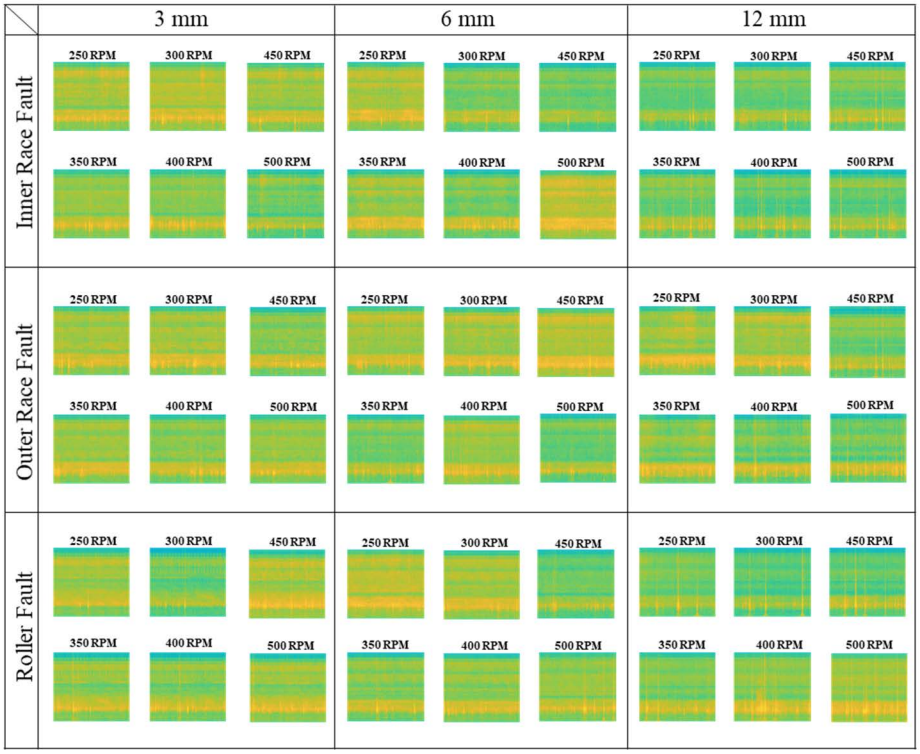

FIGURE 7. Spectrograms of fault types at various crack sizes (3mm to 12mm) and variable rotational speeds (250 RPM to 500 RPM) created by constant-Q transform.

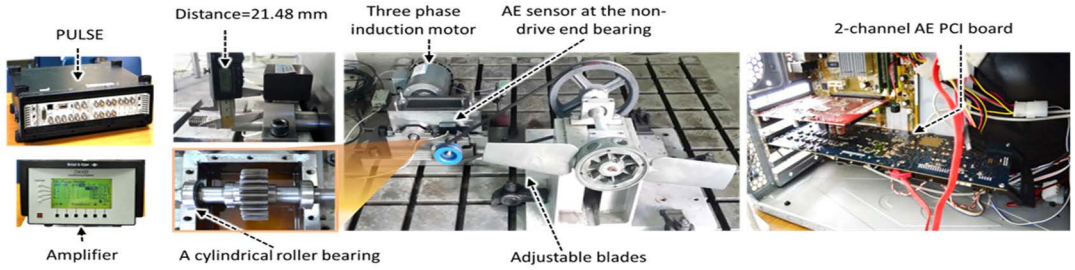

FIGURE 8. Data acquisition system and experimental testbed.

FIGURE 9. Rolling bearing fault types: (a) outer race defect (ORD), (b) inner race defect (IRD), (c) roller defect (RD), (d) inner and outer race defect (IORD), (e) outer race and roller defect (ORRD), (f) inner race and roller (IRRD), (g) inner race, outer race, and roller (IORRD).

from the normal state. The samples in training, validation, and testing were independent. In both datasets, the training samples were from signals acquired at rotational speeds of 300, 400, and 500 RPM. Samples for validation and testing were obtained from signals acquired at rotational speeds of 250 rpm, 350 rpm, and 450 RPM. Although most traditional fault diagnosis methods do not analyze variable rotational speeds using the prepared datasets, the proposed method may be observed under different rotational speeds. Dataset A consists of eight classes (seven fault types of crack size 3 mm, and normal state), and dataset B consists of 22 classes (seven fault types, three crack sizes (3 mm, 6 mm, and 12 mm) for each type of fault, and normal state).

Table 3 shows the imbalance and the low number of samples of faulty states compared to normal state samples

TABLE 3. Primary datasets A and B.

Dataset	Subset	Crack size			Speed (RPM)	No. of Classes	No. of Normal Samples Each Speed	No. of Fault Samples Each Speed	No. of Total Samples Each
		Length (mm)	Width (mm)	Depth (mm)					
Dataset A	Training	3	0.6	0.3	250, 350, 450	8	240	80	2400
	Validation				300, 400, 500	8	80	80	1920
	Testing				300, 400, 500	8	300	300	7200
Dataset B	Training	3	0.6	0.3	250, 350, 450	22	240	80	5760
		6	0.6	0.5					
		12	0.6	0.5					
	Validation	3	0.6	0.3	300, 400, 500	22	80	80	5280
		6	0.6	0.5					
		12	0.6	0.5					
	Testing	3	0.6	0.3	300, 400, 500	22	300	300	19800
		6	0.6	0.5					
		12	0.6	0.5					

TABLE 4. Down-sampling datasets created from datasets A and B.

A1 = A×20%	A2 = A×40%	A3 = A×60%	A4 = A×80%
B1 = B×20%	B2 = B×40%	B3 = B×60%	B4 = B×80%

because most of the bearing normally works in reality. From primary datasets A and B, four down-sampling datasets of each (A1 to A4 and B1 to B4) were created with the number of samples, as shown in Table 4 (the ratio of samples between classes was kept unchanged). Despite of the steady ratio of imbalance, the extremely low number of samples in fault classes (dataset A1 and B1) are also used to verify the extreme cases of imbalance because any increase of normal-state samples cannot enhance the performace of the model.

B. METHODS FOR COMPARISON

To demonstrate the effectiveness of the proposed GAN method, the experimental results were compared with several state-of-the-art algorithms used for intelligent fault diagnosis. All experiments were conducted using a computer system consisting of 1 GPU GeForce RTX 2080 Ti with a training batch size of 64. To minimize specificity and chance, experimental results were calculated on average after ten repetitions.

a) *Support-vector machine (SVM)* [38]: The typical method for supervised classification problem by finding the separating hyperplane for two categories of samples with the largest margin. Typically, this method is only compatible with low noise and a low number of features.

b) *CNN-based Lenet-5* [22]: one of the earliest and simplest CNN models, a type of feed-forward neural network adopted for image data.

c) *Using only classifier C*: the proposed CNN-based model C was used to train directly without a GAN model to compare the effect of GAN methods on low-sample datasets.

d) *Deep Convolutional GAN (DCGAN)-based method* [33]: A popular GAN method for image data in which the generator and discriminator are built based on CNN layers and with the same classifier C, is used to compare the results of the GAN-based method.

C. CONVERGENCE ABILITY OF THE PROPOSED GAN-BASED METHOD

The experiment was conducted using Dataset A4. The change curves of the loss function values are shown in Fig.10. With regard to G and D, the data values experienced significant fluctuations during the iterations and tended to be more stable in the later period. Classifier C exhibited overfitting at an early stage owing to the lack of data diversity. After approximately 20 epochs, both the training and validation curves converged with high accuracy.

In addition, we measured the training time after some epochs with the corresponding achieved validation accuracy, as shown in Tables 5 and 6, when training with datasets A1 and B1, respectively. In comparison, with dataset A1, the model training process using only classifier C and DCGAN converged after approximately 50 epochs (11.7 min) and 100 epochs (23.3 min), respectively. With dataset B1, the model using only classifier C and DCGAN required 80 epochs (25.1 min) and 190 epochs (60.16 min), respectively, to converge.

Fig. 11 illustrates the images generated in a batch by the generator of two GAN-based methods on a 5-epoch basis. Intuitively, samples generated by the proposed GAN method represent features better and closer to the real samples than those generated by the DCGAN method.

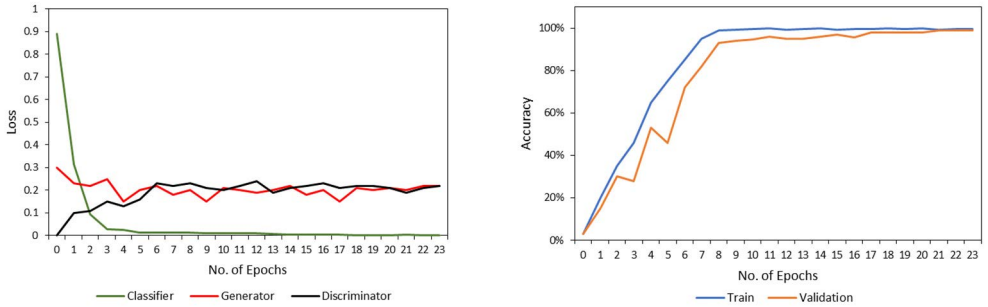

FIGURE 10. Loss curves and accuracy curves of the proposed GAN-based method during the training stage.

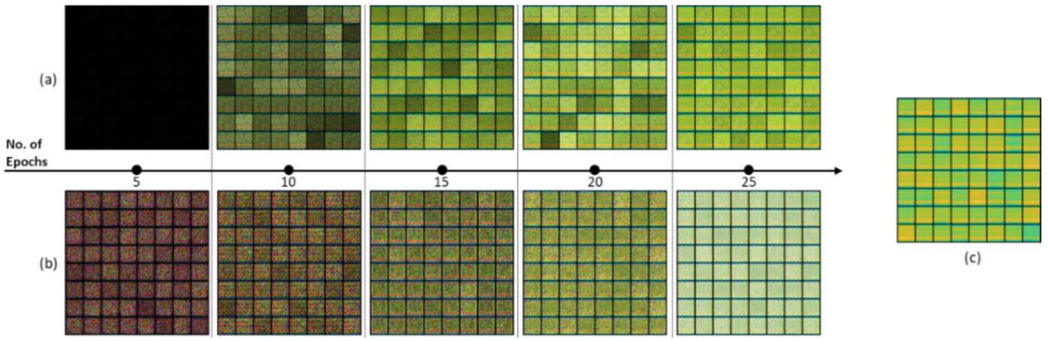

FIGURE 11. Spectrograms in a batch of 64 samples during training process: (a) spectrogram batch generated by the proposed GAN-based method, (b) spectrogram batch generated by DCGAN, (c) a batch of real sample spectrograms.

TABLE 5. Training efficiency of the proposed method with dataset A1.

No. Of Epochs	Training time	Validation time	Total time (minutes)	Validation Accuracy
5	1.5	0.7	2.2	45.47%
10	3.1	1.5	4.6	89.38%
15	4.6	2.3	6.9	92.81%
20	6.2	3.2	9.4	99.84%

TABLE 6. Training efficiency of the proposed method with dataset B1.

No. Of Epochs	Training time	Validation time	Total time (minutes)	Validation Accuracy
20	11.2	1.1	12.3	62.68%
40	22.4	2.0	24.4	89.13%
60	33.7	3.1	36.6	92.51%
80	44.6	4.3	48.9	94.54%

D. DIFFERENT INPUT DIMENSION AND VARIANCE

To study the effects of noise on the proposed GAN performance, we fixed the architectures of generator G and discriminator D of the proposed GAN, varied the dimension of noise z, and observed the QS value on classifier C trained using Dataset A. The results in Fig. 12 show that the QS score is low in the range of 15 to 25 but does not change significantly for the other. Because the generator G does not feed z directly as input (concatenated with a fake label vector), the low dimension of z may still ensure the creation of high-quality fake samples. We choose the dimension of noise z as 10, which is sufficient for a good performance.

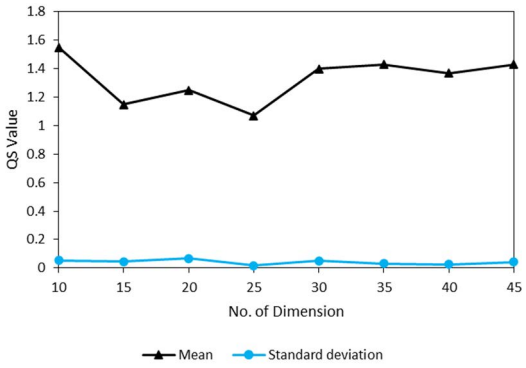

FIGURE 12. QS scores achieve by varying the dimension of z.

E. COMPARISON ACCURACY

To further demonstrate the performance of the proposed method, we conducted a number of experiments on variants of dataset A using the methods introduced in Section B. As may be observed in Table 7, the ACA of the SVM exhibited a failure in diagnosis even with the number of samples in dataset A, owing to the complicacy of fault features. Among the CNN-based methods, two methods based on the GAN strategy show the highest accuracy, especially in the low-sample dataset A1. The proposed method outperformed the others by reaching 99.79% with dataset A1 and 100% with the other variants of dataset A. Moreover, between

TABLE 7. Performance comparison of approaches in various datasets of fault type diagnosis.

Method	Average Classification Accuracy (ACA)					Parameter	MACs
	Dataset A1	Dataset A2	Dataset A3	Dataset A4	Dataset A		
SVM	12.54 % (100 epochs)	15.52% (100 epochs)	16.55% (100 epochs)	17.42% (100 epochs)	18.92% (100 epochs)	24584	24584
CNN based Lenet-5	82.29% (182 epochs)	86.67% (156 epochs)	88.12% (100 epochs)	89.00% (85 epochs)	90.62% (90 epochs)	61836	658068
CNN only	93.96% (55 epochs)	96.64% (47 epochs)	96.91% (40 epochs)	98.15% (45 epochs)	98.84% (40 epochs)	308568	16356360
DCGAN	96.04% (122 epochs)	97.71% (110 epochs)	98.75% (68 epochs)	99.38% (132 epochs)	99.58% (108 epochs)	308568	16356360
Proposed GAN	99.79% (20 epochs)	100% (5 epochs)	100% (5 epochs)	100% (5 epochs)	100% (5 epochs)	308568	16356360

TABLE 8. Performance Of the proposed method on variants of dataset B.

B1	B2	B3	B4	B
94.54% (80 epochs)	95.46% (65 epochs)	96.34% (65 epochs)	96.88% (60 epochs)	97.09% (60 epochs)

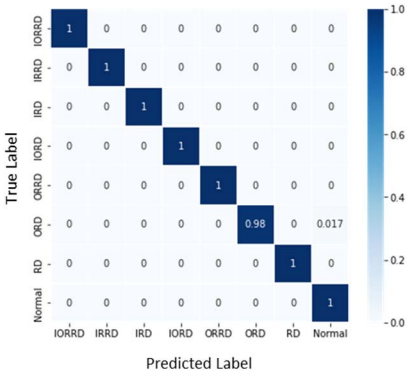

FIGURE 13. Confusion matrix of the result with dataset A1 by using the proposed method.

the two GAN-based methods, the convergence ability of the proposed method was significantly improved. However, because the last three models in Table 7 use only classifier C in the inference stage, the computational complexity (number of MACs) and the number of parameters were the same. Therefore, the GAN strategy did not affect the computational overhead.

To comprehensively evaluate the performance of the proposed GAN method, Fig. 13 depicts the confusion matrix for the diagnosis task for Dataset A1. Whereas each row of the confusion matrix represents a specific predicted bearing state, each column shows the corresponding true bearing state. Among the classes' accuracies, the value for ORD reached 98.33%, whereas the model achieved 100% accuracy for all the others.

To verify the ability of the proposed GAN-based method to perform the complicated task of diagnosis, which not only classifies the fault type but also the level of degradation, we conducted experiments using variants of Dataset B (The class name is an abbreviation of fault type and the length of the corresponding crack size.). There is a similar pattern of accuracy in the variants of Dataset B; the accuracy decreases according to the decline in the number of samples. The lowest accuracy was 94.54% for dataset B1, as listed in Table 8. In addition, the confusion matrix in Fig. 14 illustrates that there were two classes of faults that were difficult to recognize (IRRD, IORD) in an early stage when the crack size was small (3 mm length), whereas the others achieved 100% accuracy.

F. FEATURE VISUALIZATION AND FEATURE INTERPRETABILITY

The result of Dataset A1 was randomly selected to analyze the feature extraction capability of the proposed method. In an effort to enhance the interpretability of the proposed deep learning-based method, which is normally treated as a black box machine learning model, we introduce results obtained by adapting the local surrogate model, namely, the local interpretable model-agnostic explanation (LIME). LIME generates a variant of an image by segmenting the image into superpixels and turning these pixels off or on. Thus, it can calculate and weigh the effect of a superpixel according to the corresponding proximity to the instance of interest. The results in Fig. 15 show the most important areas of the input samples of different bearing states, which are highlighted inside the yellow borders. LIME allows explanations to be depicted directly on image samples. Moreover, T-distributed stochastic neighbor embedding visualization (t-SNE) was performed using the feature distribution illustrated in Fig. 16. It may be observed that the samples tend to separate after each block layer. After the last block, each fault-type class is well clustered. These results demonstrate the powerful classification of the proposed approach.

G. FAULT DIAGNOSIS UNDER NOISE ENVIRONMENTS

Experiments were conducted to validate the stability of the proposed method in noisy environments using dataset A1. We emulated the real noise environment by adding white

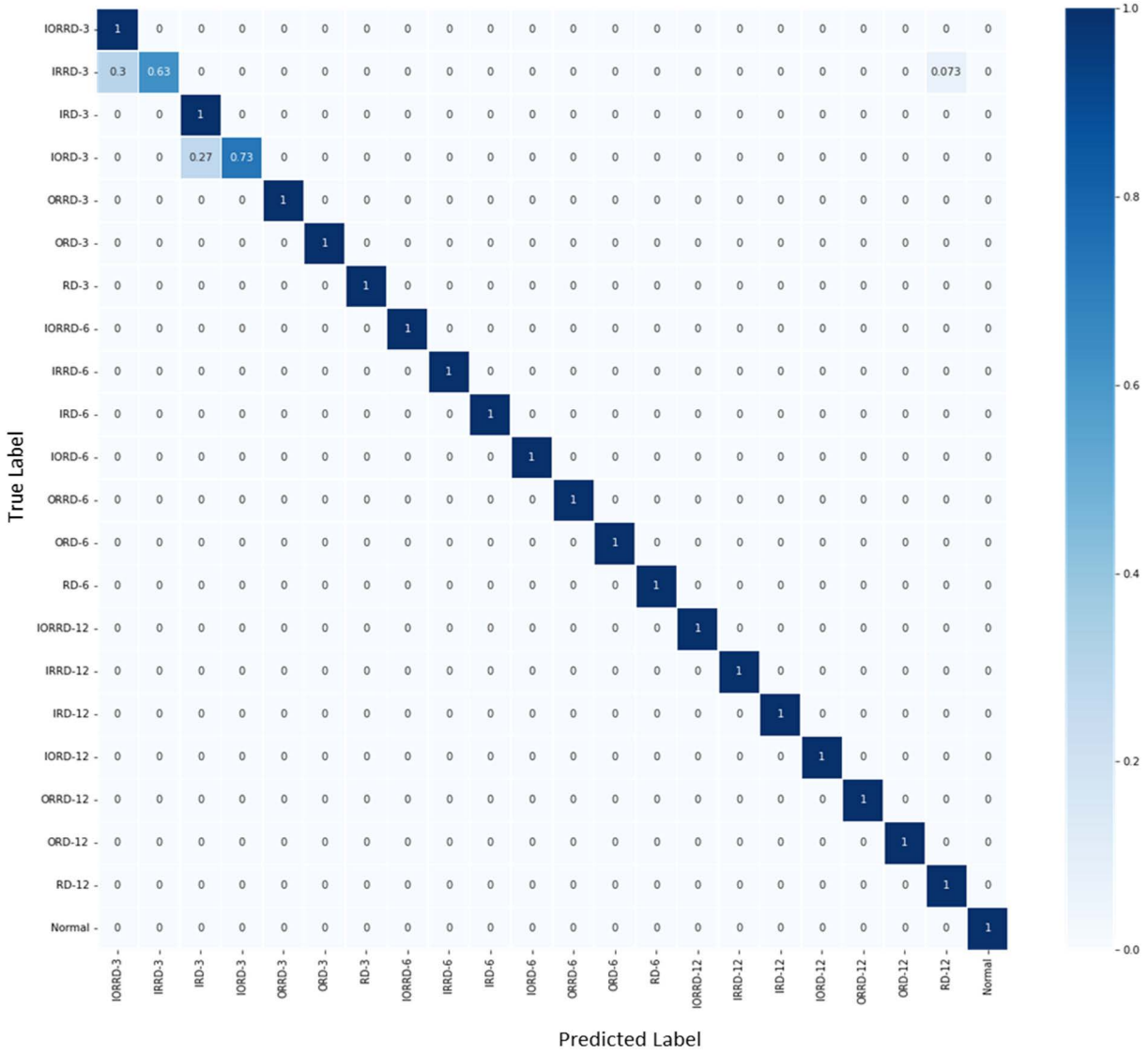

FIGURE 14. Confusion matrix of the result with dataset B by using the proposed method.

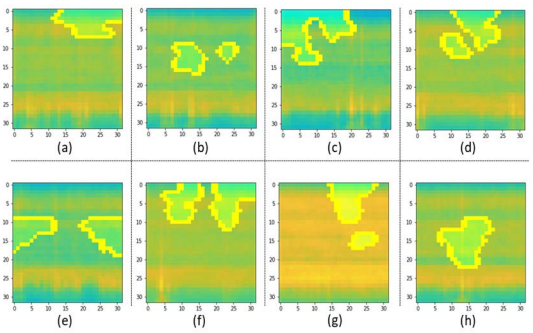

FIGURE 15. Model interpretability with LIME for classifier C in various fault samples: (a) IRD, (b) ORD, (c) RD, (d) IORD, (e) IRRD, (f) ORRD, (g) IORRD, (h) Normal state.

noise to the monitoring signals at various levels. In reality, when AE signals are collected, the threshold voltage for signals is often set to shield the environmental noise; thus, once the energy of the AE signals and the noise are of the same order of magnitude, the signals will be unable to trigger the threshold voltage, resulting in the failure to collect effective signals [39].

TABLE 9. Diagnosis ACA under various Snr with dataset A1.

SNR	No noise	4 dB	3dB	2dB	1dB
ACA	99.79%	96.71%	96.34%	93.89%	92.20%

Therefore, we consider the positive signal-to-noise ratio (SNR) values with the SNR calculated as follows:

$$SNR_{dB} = 20\log(P_{signal}/P_{noise}) \tag{14}$$

The results in Table 9 and Fig. 17 show that the diagnostic performance of the proposed methods gradually decreased

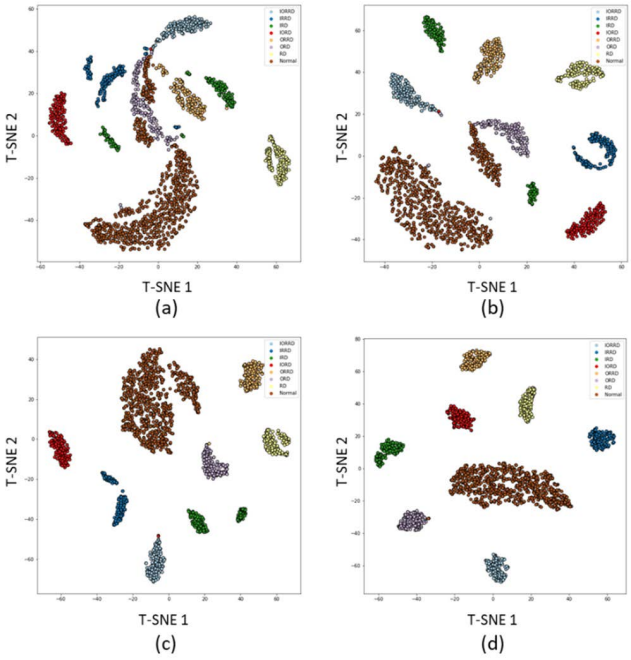

FIGURE 16. Feature visualization by t-SNE by using classifier C trained by the proposed method: (a) the first basic block, (b) the second basic block, (c) the third basic block, (d) the last basic block.

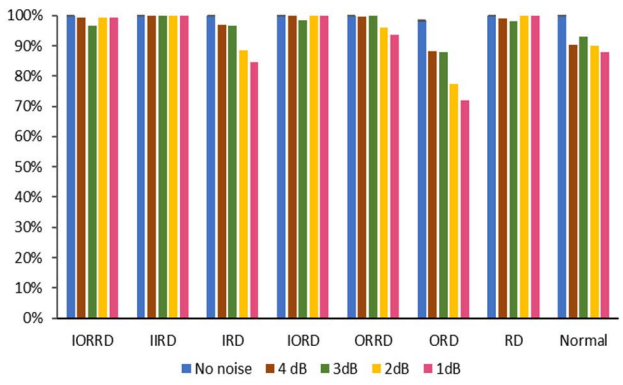

FIGURE 17. Fault classification result of each under various SNR with dataset A1.

with increasing noise power, but even at a high noise level, the accuracy was higher than 92.2%. The lack of training samples had a negative impact on these cases.

VI. CONCLUSION

In this study, we proposed a GAN-based method with AE monitoring signals represented in a 2-D spectrogram for bearing fault diagnosis under various conditions, including variable rotational speed, various crack sizes, and fault types. The proposed architectures and training scheme enhance the performance of the CNN-based classifier compared with low-sample, imbalanced, and noisy datasets. In the proposed framework, real monitoring segments are initially converted into spectrograms by CQT, which are the form of data used for the training process. The results of experiments, t-SNE, and LIME visualization show that the form of signal representation is appropriate to reveal hidden fault features in different bearing statuses and under variable working conditions. Furthermore, the effectiveness of the GAN-based method on 10 variants of experimental datasets outperformed that of conventional deep learning methods (SVM and CNN) and previous DCGAN methods in the absence of training data and data imbalance. The improved GAN-based method achieved high accuracy and rapid convergence. This provides a promising new approach for intelligent factories to adopt in real working environments. In future work, we intend to expand the scope of experiments on different machinery systems and develop a more powerful architecture to verify the feasibility of using a limited number of samples to achieve high accuracy and high efficiency in other applications.

REFERENCES

[1] B. Yang, Y. Lei, F. Jia, and S. Xing, "An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings," *Mech. Syst. Signal Process.*, vol. 122, pp. 692–706, May 2019, doi: 10.1016/j.ymssp.2018.12.051.

[2] L. Cui, X. Wang, H. Wang, and J. Ma, "Research on remaining useful life prediction of rolling element bearings based on time-varying Kalman filter," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2858–2867, Jun. 2020, doi: 10.1109/TIM.2019.2924509.

[3] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A*, vol. 454, no. 1971, pp. 903–995, 1998, doi: 10.1098/rspa.1998.0193.

[4] J. E. Garcia-Bracamonte, J. M. Ramirez-Cortes, J. de Jesus Rangel-Magdaleno, P. Gomez-Gil, H. Peregrina-Barreto, and V. Alarcon-Aquino, "An approach on MCSA-based fault detection using independent component analysis and neural networks," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 5, pp. 1353–1361, May 2019, doi: 10.1109/TIM.2019.2900143.

[5] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mech. Syst. Signal Process.*, vol. 35, nos. 1–2, pp. 108–126, Feb. 2013, doi: 10.1016/j.ymssp.2012.09.015.

[6] R. Wells, M. A. Hamstad, and A. K. Mukherjee, "On the origin of the first peak of acoustic emission in 7075 aluminium alloy," *J. Mater. Sci.*, vol. 18, no. 4, pp. 1015–1020, Apr. 1983, doi: 10.1007/BF00551968.

[7] H. L. Balderston, "The detection of incipient failure in bearings," *Ultrasonics*, vol. 8, no. 1, p. 70, Jan. 1970, doi: 10.1016/0041-624X(70)90973-X.

[8] L. M. Rogers, "The application of vibration signature analysis and acoustic emission source location to on-line condition monitoring of anti-friction bearings," *Tribol. Int.*, vol. 12, no. 2, pp. 51–58, Apr. 1979.

[9] T. Yoshioka and T. Fujiwara, "A new acoustic emission source locating system for the study of rolling contact fatigue," *Wear*, vol. 81, no. 1, pp. 183–186, Sep. 1982, doi: 10.1016/0043-1648(82)90314-3.

[10] M. W. Hawman and W. S. Galinaitis, "Acoustic emission monitoring of rolling element bearings," in *Proc. IEEE Ultrason. Symp.*, Oct. 1988, pp. 885–889, doi: 10.1109/ULTSYM.1988.49503.

[11] T. Han, Y.-F. Li, and M. Qian, "A hybrid generalization network for intelligent fault diagnosis of rotating machinery under unseen working conditions," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021, doi: 10.1109/TIM.2021.3088489.

[12] Z. He, H. Shao, X. Zhang, J. Cheng, and Y. Yang, "Improved deep transfer auto-encoder for fault diagnosis of gearbox under variable working conditions with small training samples," *IEEE Access*, vol. 7, pp. 115368–115377, 2019, doi: 10.1109/ACCESS.2019.2936243.

[13] X. Jiang, X. Cheng, J. Shi, W. Huang, C. Shen, and Z. Zhu, "A new l_0-norm embedded MED method for roller element bearing fault diagnosis at early stage of damage," *Measurement*, vol. 127, pp. 414–424, Oct. 2018, doi: 10.1016/j.measurement.2018.06.016.

[14] S. Shao, S. McAleer, R. Yan, and P. Baldi, "Highly accurate machine fault diagnosis using deep transfer learning," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2446–2455, Apr. 2019, doi: 10.1109/TII.2018.2864759.

[15] H. Wang, J. Xu, R. Yan, C. Sun, and X. Chen, "Intelligent bearing fault diagnosis using multi-head attention-based CNN," *Proc. Manuf.*, vol. 49, pp. 112–118, Jan. 2020, doi: 10.1016/j.promfg.2020.07.005.
[16] H. Cao, H. Shao, X. Zhong, Q. Deng, X. Yang, and J. Xuan, "Unsupervised domain-share CNN for machine fault transfer diagnosis from steady speeds to time-varying speeds," *J. Manuf. Syst.*, vol. 62, pp. 186–198, Jan. 2022, doi: 10.1016/j.jmsy.2021.11.016.
[17] T. Zhou, T. Han, and E. L. Droguett, "Towards trustworthy machine fault diagnosis: A probabilistic Bayesian deep learning framework," *Rel. Eng. Syst. Saf.*, vol. 224, Aug. 2022, Art. no. 108525, doi: 10.1016/j.ress.2022.108525.
[18] H. Shao, W. Li, M. Xia, Y. Zhang, C. Shen, D. Williams, A. Kennedy, and C. W. de Silva, "Fault diagnosis of a rotor-bearing system under variable rotating speeds using two-stage parameter transfer and infrared thermal images," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021, doi: 10.1109/TIM.2021.3111977.
[19] O. Janssens, "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib.*, vol. 377, pp. 331–345, Sep. 2016, doi: 10.1016/j.jsv.2016.05.027.
[20] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, Feb. 2018, doi: 10.1016/j.ymssp.2017.06.022.
[21] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 8, pp. 1926–1935, Aug. 2017, doi: 10.1109/TIM.2017.2674738.
[22] 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, doi: 10.1109/TIE.2017.2774777.
[23] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks," *J. Sound Vib.*, vol. 388, pp. 154–170, Feb. 2017, doi: 10.1016/j.jsv.2016.10.043.
[24] L. Duan, M. Xie, T. Bai, and J. Wang, "A new support vector data description method for machinery fault diagnosis with unbalanced datasets," *Expert Syst. Appl.*, vol. 64, pp. 239–246, Dec. 2016, doi: 10.1016/j.eswa.2016.07.039.
[25] J. Zhang, W. Ma, J. Lin, L. Ma, and X. Jia, "Fault diagnosis approach for rotating machinery based on dynamic model and computational intelligence," *Measurement*, vol. 59, pp. 73–87, Jan. 2015, doi: 10.1016/j.measurement.2014.09.045.
[26] I. Goodfellow, "NIPS 2016 tutorial: Generative adversarial networks," 2016, *arXiv:1701.00160*.
[27] Z. Wang, J. Wang, and Y. Wang, "An intelligent diagnosis scheme based on generative adversarial learning deep neural networks and its application to planetary gearbox fault pattern recognition," *Neurocomputing*, vol. 310, pp. 213–222, Oct. 2018, doi: 10.1016/j.neucom.2018.05.024.
[28] 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, doi: 10.1016/j.knosys.2018.12.019.
[29] N. Holighaus, M. Dörfler, G. A. Velasco, and T. Grill, "A framework for invertible, real-time constant-Q transforms," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, no. 4, pp. 775–785, Apr. 2013, doi: 10.1109/TASL.2012.2234114.
[30] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014, *arXiv:1406.2661*.
[31] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, "Least squares generative adversarial networks," 2016, *arXiv:1611.04076*.
[32] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, *arXiv:1411.1784*.
[33] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.
[34] M. T. Pham, J.-M. Kim, and C. H. Kim, "Efficient fault diagnosis of rolling bearings using neural network architecture search and sharing weights," *IEEE Access*, vol. 9, pp. 98800–98811, 2021, doi: 10.1109/ACCESS.2021.3096036.
[35] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.
[36] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?': Explaining the predictions of any classifier," 2016, *arXiv:1602.04938*.
[37] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training GANs," 2016, *arXiv:1606.03498*.
[38] S. E. Pandarakone, Y. Mizuno, and H. Nakamura, "Evaluating the progression and orientation of scratches on outer-raceway bearing using a pattern recognition method," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1307–1314, Feb. 2019, doi: 10.1109/TIE.2018.2833025.
[39] X. Wang, X. Liu, T. He, D. Xiao, and Y. Shan, "Structural damage acoustic emission information enhancement through acoustic black hole mechanism," *Measurement*, vol. 190, Feb. 2022, Art. no. 110673, doi: 10.1016/j.measurement.2021.110673.

MINH TUAN PHAM received the Engineering degree in automation and control from the Hanoi University of Science and Technology, Hanoi, Vietnam, in 2019, and the M.S. degree from the Department of System Engineering, Graduate School of ICT Convergence, Chonnam National University, Gwangju, South Korea. He worked as a Researcher with Soongsil University, Seoul, South Korea. His research interests include embedded systems, fault detection, deep learning, digital signal processing, vibration, and acoustic emission.

JONG-MYON KIM received the B.S. degree in electrical engineering from Myongji University, Yongin-si, South Korea, in 1995, the M.S. degree in electrical and computer engineering from the University of Florida, Gainesville, in 2000, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, in 2005. He was a Senior Research Staff with the Chip Solution Center, Samsung Advanced Institute of Technology, from 2005 to 2007. Since 2007, he has been with the IT Convergence Department, University of Ulsan, Ulsan, South Korea, where he is currently a Professor. His research interests include embedded systems, deep learning, fault diagnosis and prognostics, and parallel processing.

CHEOL HONG KIM received the B.S. degree in computer engineering from Seoul National University, Seoul, South Korea, in 1998, the M.S. degree, in 2000, and the Ph.D. degree in electrical and computer engineering from Seoul National University, in 2006. He worked as a Senior Engineer with Samsung Electronics, South Korea, from 2005 to 2007. He also worked as a Professor with Chonnam National University, South Korea, from 2007 to 2020. He is currently working as a Professor with the School of Computer Science and Engineering, Soongsil University, Seoul. His research interests include computer systems, embedded systems, deep learning, fault diagnosis, and low power systems.