Research
A SCIENCE PARTNER JOURNAL

RESEARCH ARTICLE

A Deep Generative Model with Multiscale Features Enabled Industrial Internet of Things for Intelligent Fault Diagnosis of Bearings

He-xuan Hu[1,2], Yicheng Cai[1,2], Qiang Hu[1,2], and Ye Zhang[1,2*]

[1]Key Laboratory of Water Big Data Technology of Ministry of Water Resources, Hohai University, Nanjing 211100, P. R. China. [2]College of Computer and Information, Hohai University, Nanjing 211100, P. R. China.

*Address correspondence to: yzh@hhu.edu.cn

Citation: Hu Hx, Cai Y, Hu Q, Zhang Y. A Deep Generative Model with Multiscale Features Enabled Industrial Internet of Things for Intelligent Fault Diagnosis of Bearings. *Research* 2023;6:Article 0176. https://doi.org/10.34133/research.0176

Submitted 15 December 2022
Accepted 29 May 2023
Published 7 July 2023

Copyright © 2023 He-xuan Hu et al. Exclusive licensee Science and Technology Review Publishing House. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY 4.0).

Effective condition monitoring and fault diagnosis of bearings can not only maximize the life of rolling bearings and prevent unexpected shutdowns caused by equipment failures but also eliminate unnecessary costs and waste caused by excessive maintenance. However, the existing deep-learning-based bearing fault diagnosis models have the following defects. First of all, these models have a large demand for fault data. Second, the previous models only consider that single-scale features are generally less effective in diagnosing bearing faults. Therefore, we designed a bearing fault data collection platform based on the Industrial Internet of Things, which is used to collect bearing status data from sensors in real time and feed it back into the diagnostic model. On the basis of this platform, we propose a bearing fault diagnosis model based on deep generative models with multiscale features (DGMMFs) to solve the above problems. The DGMMF model is a multiclassification model, which can directly output the abnormal type of the bearing. Specifically, the DGMMF model uses 4 different variational autoencoder models to augment the bearing data and integrates features of different scales. Compared with single-scale features, these multiscale features contain more information and can perform better. Finally, we conducted a large number of related experiments on the real bearing fault datasets and verified the effectiveness of the DGMMF model using multiple evaluation metrics. The DGMMF model has achieved the highest value under all metrics, among which the value of precision is 0.926, the value of recall is 0.924, the value of accuracy is 0.926, and the value of F1 score is 0.925.

Introduction

Bearings, as the key components supporting the rotating body in rotating machinery, are widely used in important fields such as transportation, energy, and chemical industry [1–3]. With the improvement of the level of information technology, mechanical equipment is developing toward intelligence and automation, and the requirements for the reliability of rolling bearings are getting higher and higher [4–6]. However, with the accumulation of working hours of mechanical equipment, less than 10% of the bearing failures reach the natural fatigue limit, and most of the rolling bearings fail because of complex and changeable working conditions, which brings immeasurable economic losses or casualties to the society [7–9]. Therefore, effective condition monitoring and fault diagnosis of bearings, timely and reliable assessment of equipment conditions, maintenance, and corrective measures can not only maximize the life of rolling bearings [10–14] but also eliminate unnecessary costs and waste caused by excessive maintenance.

The existing deep-learning-based bearing fault diagnosis models are all implemented on the basis of sufficient training data, and their prediction results are affected by the distribution of abnormal data and normal data [15–19]. In a real factory environment, because of the limitation of the quantity and quality of sensors and collectors, the fault data of various bearings collected by the data collection system will be insufficient. Specifically, when the amount of fault data is small, the prediction effect of the bearing fault diagnosis model based on deep learning will drop obviously. Therefore, how to construct a reliable fault diagnosis model under the condition of small samples is an urgent problem to be solved.

In addition, the existing deep-learning-based bearing fault diagnosis models all use multilayer neural networks to capture the features of fault data [20–24]. These models, which only consider single-scale features, are generally poor in the diagnosis of bearing faults. This is because features of different scales have different advantages. High-level features have strong abstraction ability but are not easy to understand. Low-level features are highly interpretable but have insufficient abstraction ability. Therefore, how to effectively fuse multiscale features is the key to improving the performance of the bearing fault diagnosis model.

In this work, we design a bearing fault data collection platform based on the Industrial Internet of Things (IIoT) [25,26], which is used to collect bearing status data from sensors in real time and feed it back into the diagnostic model. In addition,

on the basis of this platform, we propose a bearing fault diagnosis model based on deep generative models with multiscale features (DGMMFs). The DGMMF model augments the bearing data with 4 different variational autoencoder (VAE) models. Specifically, there are 4 kinds of bearing signal data in the dataset, which are normal data, abnormal inner ring, abnormal outer ring, and abnormal rolling. Then, the above 4 kinds of bearing signal data are input into 4 different VAE models for training respectively. Then, from the 4 trained VAE models, the normal data, abnormal inner ring, abnormal outer ring, and abnormal rolling data are respectively sampled and finally inserted into the training set to augment the data samples. The above steps realize the augmentation of normal data and fault data, making the distribution of various samples balanced, thus simplifying the training difficulty of the model.

In addition, the DGMMF model integrates features of different scales. Compared with single-scale features, these multiscale features contain richer information, thereby improving the fault diagnosis performance of the model. Specifically, the DGMMF model uses a multilayer one-dimensional convolutional neural network to extract the high-level features of the bearing data and a single-layer fully connected neural network to extract the low-level features of the bearing data, and then the high-level features and low-level features are fused to obtain multiscale features of bearing data. Finally, the DGMMF model uses the fully connected neural network and multiscale features to predict whether the bearing is faulty and its fault category. We validate the effectiveness of the DGMMF model on a real-world bearing dataset.

The main contributions of this work are as follows:

1. We designed a bearing fault data collection platform based on the IIoT, which is used to collect bearing status data from sensors in real time and feed it back to the bearing fault diagnosis model.
2. We propose a bearing fault diagnosis model DGMMF based on deep generative models with multiscale features. The model can integrate features of different scales and realize the augmentation of normal data and fault data, so that the distribution of various samples is balanced.
3. We conducted a large number of related experiments on the real bearing fault datasets and verified the effectiveness of the DGMMF model using multiple evaluation metrics.

The rest of the content is organized as follows. Related Work introduces the research progress of the bearing fault diagnosis model. Method introduces the bearing fault data collection platform and DGMMF model proposed in this work. Results and Discussion describes the experimental results of the DGMMF model on real bearing datasets. Conclusion summarizes the content of this work and discusses the direction of future improvement.

Related Work

In this section, we first introduce related work on traditional abnormal detection algorithms. Then, we will introduce the latest research progress of the bearing fault diagnosis model.

Abnormal detection

Tang et al. [27] proposed a deep-neural-network-based fault diagnosis model for rotating devices. They believe that shallow neural networks are flawed in performance and cannot meet the actual needs of smart devices. Therefore, they utilized more powerful convolutional neural networks, which are good at capturing nonlinear relationships between features. They verified the effectiveness of the model in the experiment.

Ranjith et al. [28] used the density-based spatial clustering of applications with noise (DBSCAN) algorithm to monitor traffic videos to detect abnormal driving and suspicious behavior. The DBSCAN [29] is a density-based clustering algorithm. DBSCAN can be used not only for clustering problems but also for abnormal detection problems. The basic idea is to find low-density abnormal samples in the sample space. The operation of DBSCAN can be decomposed into 4 parts. The first step is to select any data point from the datasets. In the second step, the selected data point is used as the core point, and all data points that can directly reach the core point are found using hyperparameters to form a cluster. The third step is to reselect if the selected data point is an edge point. The fourth step is to repeat the process of steps 2 and 3 until all points are processed.

Yan et al. [30] used autoencoders to detect abnormal timing in video streams. The autoencoder is essentially used for representation learning tasks, that is, to compress the original data points into a low-dimensional vector and then restore the low-dimensional vector to the original data through the decoding operation [31]. The idea of using an autoencoder to handle abnormal detection tasks is as follows. Autoencoder trained on the basis of normal data can reconstruct and restore normal samples, but it cannot restore data points that are different from the normal distribution. Therefore, if a new sample is encoding and, after decoding, its error exceeds the error range of normal data after encoding and decoding, it is regarded as abnormal data.

Zhang et al. [32] used one-class support vector machine (SVM) to monitor intrusion behavior and abnormal traffic in the network. The algorithm idea of one-class SVM [33] is very simple. It is to find a hyperplane to circle the positive examples in the sample. The prediction is to use this hyperplane to make decisions. The samples in the circle are considered positive samples, and the samples outside the circle are negative samples. The idea of applying one-class SVM to abnormal detection tasks is as follows. We regard all abnormal samples as negative samples and all normal samples as positive samples. Then, through positive and negative samples, it is learned that it belongs to an unsupervised learning hyperplane to make a decision, and all samples that are not in the hyperplane are predicted as abnormal data.

Bearing fault diagnosis

Yang et al. [34] believed that the existing bearing fault diagnosis model based on deep learning directly analyzes the vibration information of the bearing, and such an operation lacks interpretability to a certain extent. Therefore, to increase the interpretability of the bearing fault diagnosis model, they combined convolutional neural network, recurrent neural network, and attention mechanism to realize automatic diagnosis of bearing signals. The reason why the model can be interpreted is that they use the attention mechanism to visualize the weight distribution of the input data, thereby increasing the interpretability.

Li et al. [35] believed that the existing bearing fault diagnosis model based on deep learning did not take into account the difference in distribution between training data and test data. This difference presents a domain shift problem, which leads to severe degradation in the diagnostic performance of the model. Therefore, for this problem, they proposed a domain adaptation method for bearing fault diagnosis. The main structure of the method is a convolutional neural network, and, on this basis, the representation of the source domain is adapted to the target domain by minimizing the average difference. The effectiveness of the domain adaptation method is verified by experiments.

Zhu et al. [36] believed that the fault features extracted by the existing bearing fault diagnosis model were not very effective, which led to the low performance of fault diagnosis. Aiming at this problem, they proposed a model based on feature extraction and fusion to diagnose bearing faults. The idea of the model is as follows. First, the frequency domain features of the bearing signal are transformed into a frequency domain matrix. Then, the features of the matrix are extracted by singular value decomposition and finally input into the SVM to obtain the predicted value.

Patel and Upadhyay [37] proposed a feature ranking and selection model based on Euclidean distance for bearing fault diagnosis. The model can select valuable features from the original feature set and can achieve high-precision fault diagnosis with a small number of features and calculation time. Experimental results on 2 real-world datasets demonstrate the superiority of the model, saving a lot of time for feature selection.

However, some previous research work defined bearing fault diagnosis as a binary classification problem, and it was impossible to infer the faulty bearing category, such as normal status, rolling fault, inner ring fault, and outer ring fault. In this work, we define bearing fault diagnosis as a multiclassification problem, that is, to infer the fault category of the bearing. We can compare various types of bearing faults together and can judge the fault location of the bearing at one time. This saves manpower and material resources to a great extent.

Method

In this section, we first introduce the architecture of the IIoT-based bearing fault data collection and diagnosis platform. How the VAE model generates fault data and the implementation details of the multiscale features of the DGMMF model is then introduced.

An IIoT-based bearing fault data collection and diagnosis platform

We use the IIoT technology to build an IIoT data collection and diagnosis platform for bearing fault. Figure 1 shows the architecture of the IIoT data platform. From left to right are the data collection layer, data analysis storage layer, and application service layer of the platform. The platform first uses the data collection layer to collect status data from sensors, actuators, and other field devices, including the collection of bearing running data. The role of the data analysis storage layer is to parse the collected data into a human-understandable format and store it in the database. The role of the application service layer is to develop various valuable applications for the status data stored in the database, including the diagnosis of bearing faults.

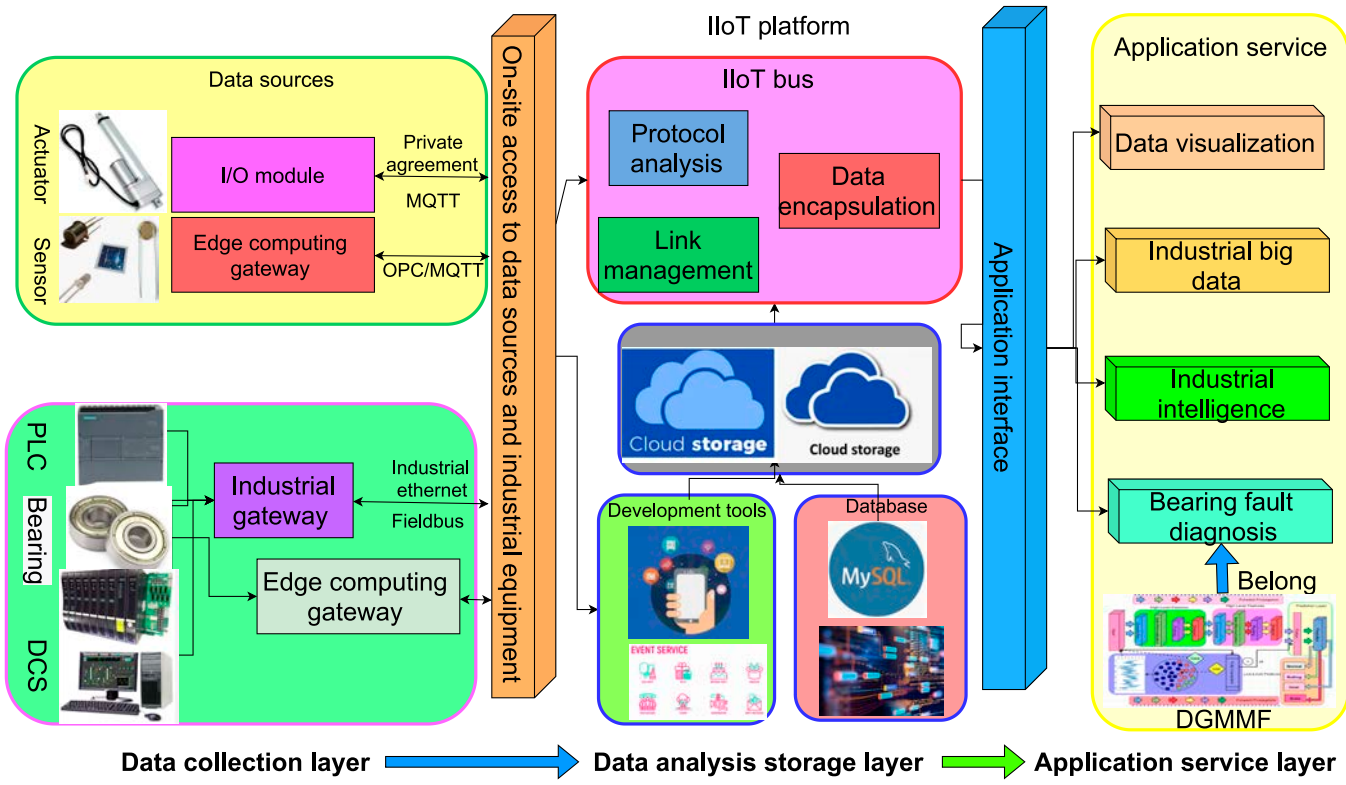


Fig. 1. The architecture of the IIoT-based bearing fault data collection and diagnosis platform.

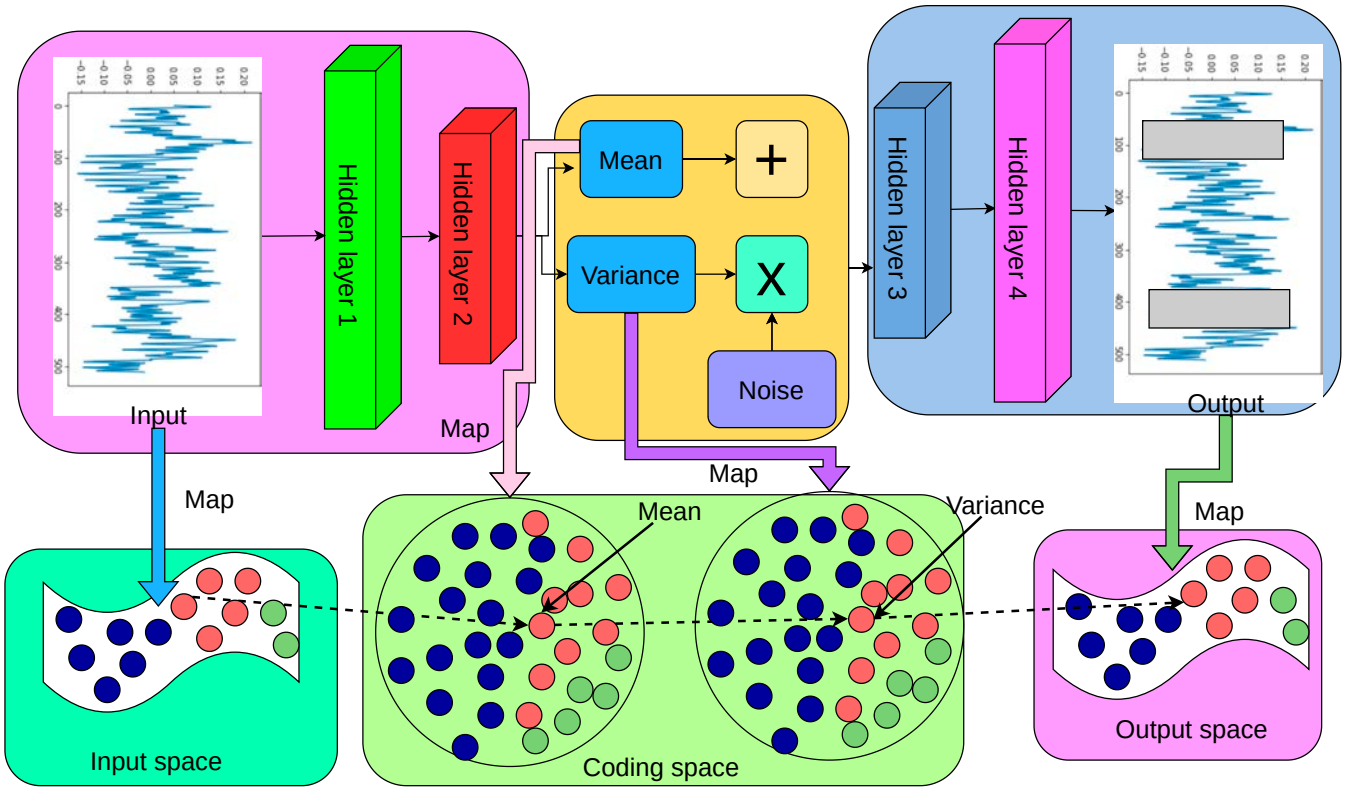


Fig. 2. The frame diagram of the VAE model to augment fault data.

Augmenting fault data with deep generative model

In a real factory environment, because of the limitation of the quantity and quality of sensors and collectors, the fault data of various bearings collected by the data collection system will be insufficient. Specifically, when the amount of fault data of various types of bearings is small, the prediction effect of the bearing fault diagnosis model based on deep learning will drop obviously. Hence, we use VAE model to augment the bearing signal data. This is actually a preprocessing operation on the dataset, which can be used to improve the performance of the model. The VAE model needs to augment the data with the signals of the identified bearing fault categories. The augmented data are combined with the original data to train the subsequent DGMMF model. In the experimental part of this work, the public bearing dataset is used. This dataset has known fault categories in advance, so VAE can be used to augment the data samples and then used for training the DGMMF model. In the real factory environment, the factory needs to manually collect the bearing fault category data determined in the machinery and then use the VAE model to augment the collected data to train the DGMMF model. Then, the factory uses the trained DGMMF model to carry out long-term intelligent fault monitoring and diagnosis for the mechanical bearings.

In this work, the DGMMF model is augmented with data from 4 different VAE models. Figure 2 is the frame diagram of the VAE model to augment fault data. First, the workflow of VAE is shown in Eqs. 1 to 4.

$$H_1 = HL_1(X) \tag{1}$$

$$M, V = HL_2(H_1) \tag{2}$$

$$H_2 = HL_3(M, V) \tag{3}$$

$$\hat{X} = HL_4(H_2) \tag{4}$$

Among them, HL_1, HL_2, HL_3, and HL_4 are hidden layers, respectively. X represents the time series signal of the bearing. M and V represent the mean and variance, respectively. A trained VAE model can be learned by minimizing the original data and reconstructing the data. Then, the VAE model is sampled to generate bearing data through Eqs. 5 and 6 and then added to the training set to augment the data samples.

$$Gen = VAE(S) \tag{5}$$

$$Input = Gen + Data \tag{6}$$

Among them, *Data* is the original data, and *Gen* is the generated data. *Input* represents the training set that is finally used to train the model. There are 4 kinds of bearing signal data in the dataset, which are normal data, abnormal inner ring, abnormal outer ring, and abnormal rolling. Then, the above 4 kinds of bearing signal data are input into 4 different VAE models for training. The above steps realize the augmentation of normal data and fault data, making the distribution of various samples balanced, thus simplifying the training difficulty of the model.

A DGMMF-based bearing fault diagnosis model

Relying on the IIoT data collection platform, we propose a bearing fault diagnosis model named DGMMF. Figure 3 is the

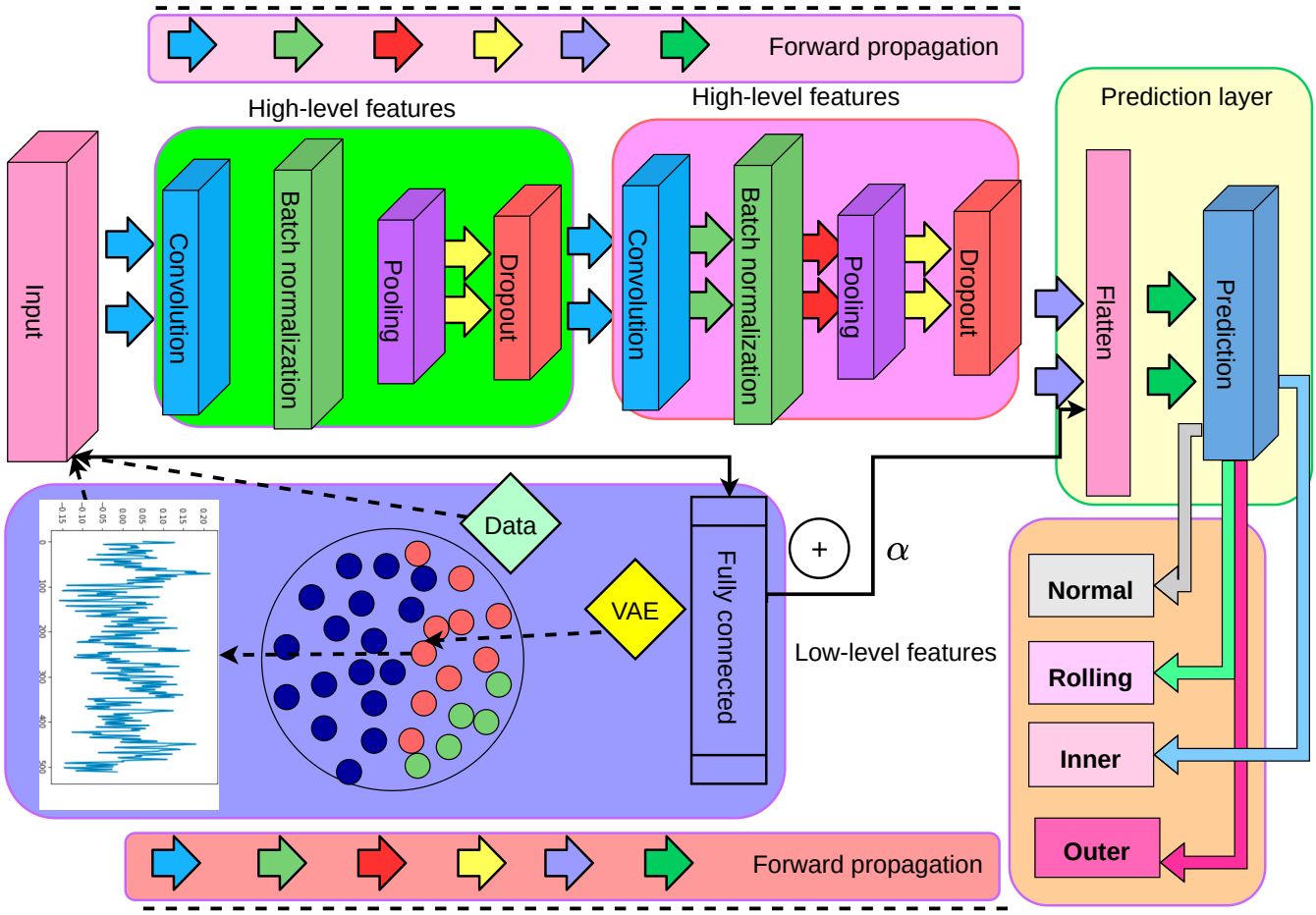


Fig. 3. The frame diagram of the DGMMF model.

frame diagram of the DGMMF model. Previously diagnostic models that only consider single-scale features are generally poor in the diagnosis of bearing faults. This is because features of different scales have different advantages, and the abstraction ability of high-level features is strong, but it is not easy to understand. Low-level features have strong interpretability but insufficient abstraction ability. Therefore, the DGMMF model integrates features of different scales. Compared with single-scale features, these multiscale features contain richer information, thereby improving the fault diagnosis performance of the model. Specifically, as shown in Eq. 7, the DGMMF model uses a multilayer one-dimensional convolutional neural network to extract high-level features of bearing data.

$$HF = Con1D\left(X'\right) \tag{7}$$

$$Con1D = [C, B, P, D] \tag{8}$$

Among them, HF represents the high-level feature, and X' represents the time series signal of the bearing that needs to be diagnosed. $Con1D$ represents one-dimensional convolution operation, generally using multiple one-dimensional convolutions to extract features. The composition of one-dimensional convolution operation is shown in Eq. 8. Among them, C, B, P, and D are respectively represented as convolution operation, batch normalization, pooling operation, and dropout operation. The DGMMF model then utilizes Eq. 9 to extract low-level features.

$$LF = FC\left(X'\right) \tag{9}$$

where LF represents the low-level features of the bearing and FC represents the fully connected network. Then, the DGMMF model uses Eqs. 10 and 11 to fuse low-level features and high-level features to form multiscale features MSF.

$$FHF = Flatten(HF) \tag{10}$$

$$MSF = FHF + \alpha LF \tag{11}$$

where FHF is the straightened high-level feature and $Flatten$ represents the straightening operation. Finally, the DGMMF model inputs MSF into the predictor shown in Eq. 12 and outputs the fault category Y of the bearing.

$$Y = Predictor(MSF) \tag{12}$$

where $Predictor$ generally consists of a fully connected network. The value of Y is one of normal, rolling, inner ring, and outer ring. The pseudocode of the DGMMF model is shown in Algorithm 1.

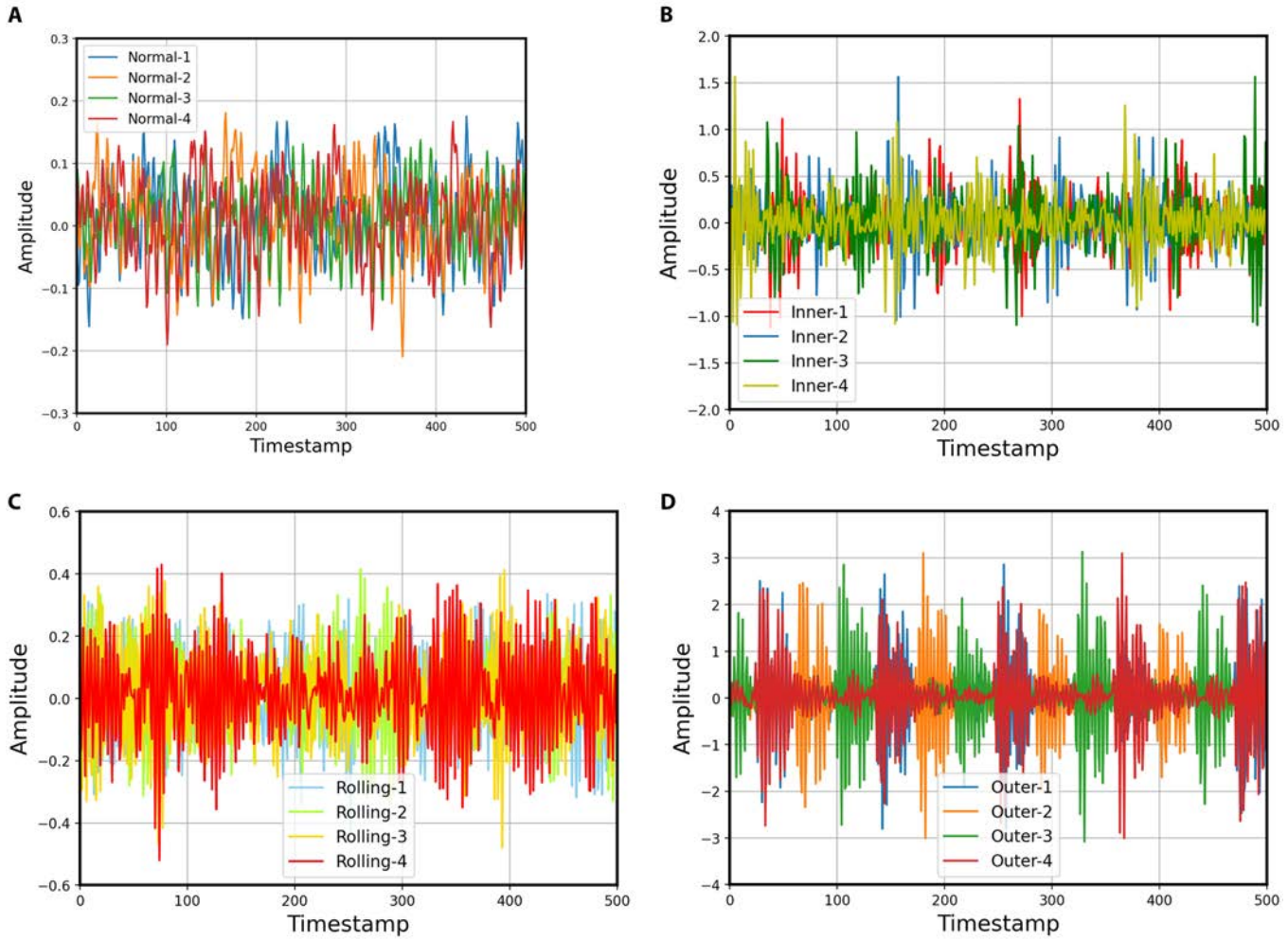


Fig. 4. The data distribution of normal and fault data. (A) Normal data, (B) inner ring fault data, (C) rolling fault data, and (D) outer ring fault data.

Algorithm 1. DGMMF.
Input: Bearing's time series signal, X
Output: Predicted category of the bearing's time series signal, Y

Augmenting fault data with deep generative model
$H_1 = HL_1(X)$
$M, V = HL_2(H_1)$
$H_2 = HL_3(M, V)$
$\hat{X} = HL_4(H_2)$
$Gen = VAE(S)$
$Input = Gen + Data$
A DGMMF-based bearing fault diagnosis model
$HF = Con1D(X')$
$Con1D = [C, B, P, D]$
$LF = FC(X')$
$FHF = Flatten(HF)$
$MSF = FHF + \alpha LF$
$Y = Predictor(MSF)$
Return Y

Results and Discussion

In this section, we first introduce the bearing fault datasets used, followed by the relevant setup of the experiments. Then, we introduce the influence of the weight parameter α in Eq. 11 on the DGMMF model, the influence of the weight decay of the optimizer, and learning rate on the DGMMF model. Finally, we will introduce the comparison of experimental results between the DGMMF model and other models.

Datasets

We use a publicly available bearing failure dataset for this experiment. The diameters of the bearings are 0.007, 0.014, and 0.021 mm, respectively. The status of each bearing is normal status, rolling fault, inner ring fault, and outer ring fault, respectively. Examples of each status are shown in Fig. 4. Figure 4A to D respectively shows 4 normal bearing signals, 4 signals of rolling fault, 4 signals of inner ring fault, and 4 signals of outer ring fault.

Experimental setup

We slide the window to intercept and sample the time series data. Seventy percent of the data are used as training samples, and 30% of the data are used as test samples. To numerically display the diagnostic ability of the model, we use evaluation metrics such as precision, recall, accuracy, and F1 score to measure the performance of various models.

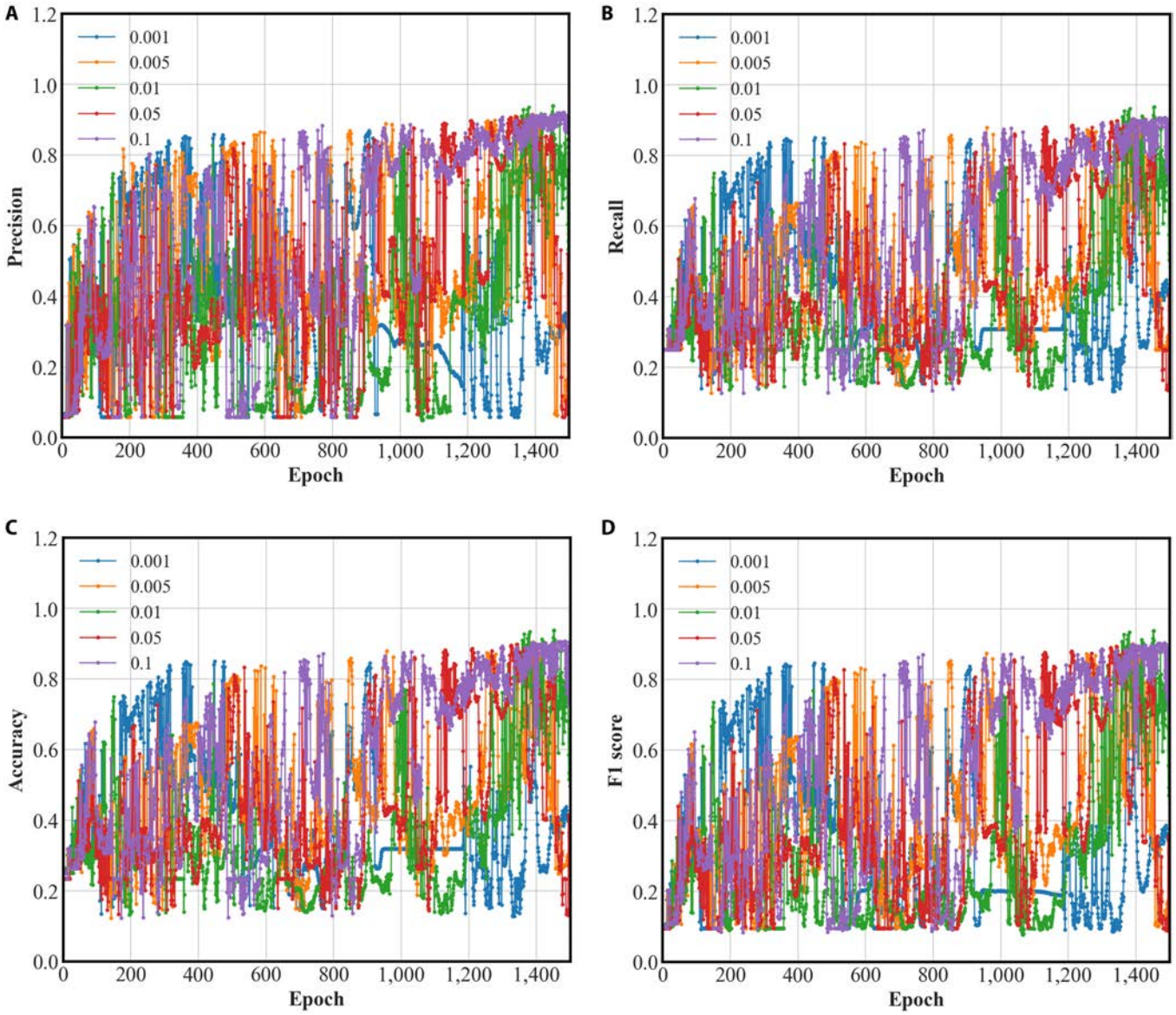


Fig. 5. The influence of weight parameter α on DGMMF model. (A) Precision, (B) recall, (C) accuracy, and (D) F1 score.

The influence of weight parameter α on DGMMF model

The weight parameter α controls the proportion of low-level features in multiscale features. The larger the value of the weight parameter α, the higher the proportion of low-level features in multiscale features. The smaller the value of the weight parameter α, the lower the proportion of low-level features in multiscale features will be. Therefore, it can affect the amount of information contained in the multiscale features, thereby affecting the diagnostic ability of the subsequent DGMMF model. Therefore, we run the following experiments to compare the diagnostic ability of the DGMMF model under different weight parameters α. We manually adjust the size of the weight parameter α to obtain different evaluation metrics. The variation range of α is $[0.001, 0.005, 0.01, 0.05, 0.1]$. The experimental results are shown in Fig. 5.

The abscissa in Fig. 5 is the number of iterations of the DGMMF model, and the ordinate is the score of the DGMMF model under different evaluation metrics. The DGMMF model has different diagnostic capabilities under different weight parameters α. The above experimental results show that in most cases, the low-level features in the multiscale features are helpful to the diagnostic ability of the DGMMF model. The weight of low-level features in multiscale features needs to be carefully designed. Otherwise, inappropriate weight values will decrease the diagnostic ability of the DGMMF model.

The influence of weight decay on the DGMMF model

The purpose of weight decay in the optimizer is to prevent overfitting of the DGMMF model. In the loss function, weight decay is a coefficient placed in front of the regular term. The regular term generally indicates the complexity of the model, so the role of weight decay is to adjust the influence of model complexity on the loss function. Therefore, we run the following experiments to compare the diagnostic ability of the DGMMF model

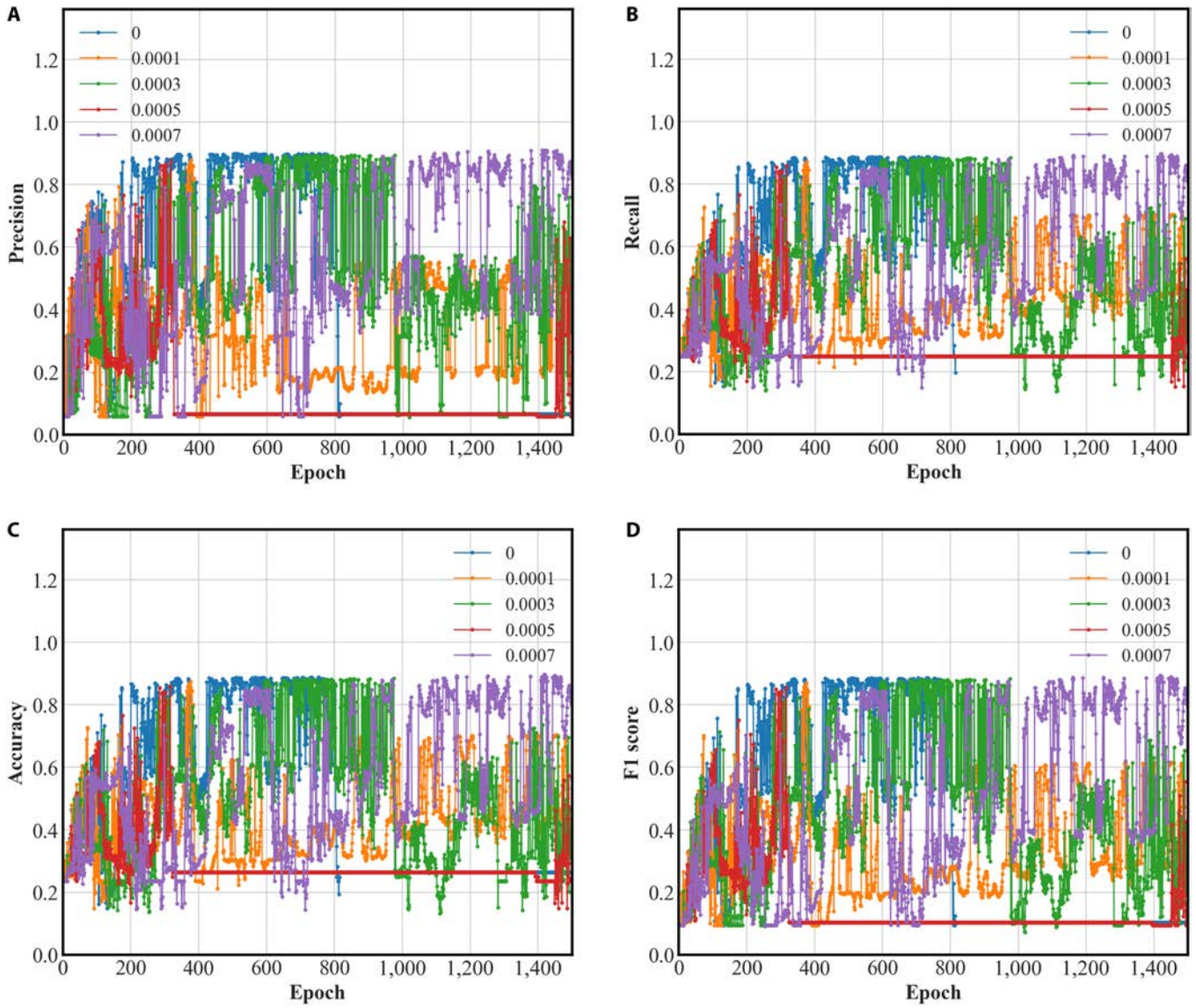


Fig. 6. The influence of weight decay on the DGMMF model. (A) Precision, (B) recall, (C) accuracy, and (D) F1 score.

under different weight decay. We manually adjust the size of weight decay to obtain different evaluation metrics. The range of weight decay is [0, 0.0001, 0.0003, 0.0005, 0.0007]. The experimental results are shown in Fig. 6.

The abscissa of Fig. 6 is the number of iterations of the DGMMF model, and the ordinate is the score of the DGMMF model under different evaluation metrics. The diagnostic performance of the DGMMF model is different under different weight decay values. We can clearly find that when the value of weight decay is 0.0007, the experimental results of the DGMMF model under the 4 evaluation metrics of precision, recall, accuracy, and F1 score are the best. It can be found that the DGMMF model with a weight decay value of 0.0007 is better than the DGMMF model with a weight decay value of 0, which shows that the weight decay can improve the diagnostic performance of the DGMMF model.

The above experimental results show that an appropriate weight decay value can adjust the influence of the complexity of the DGMMF model on the loss function, thereby avoiding the occurrence of overfitting. Therefore, we need to use the validation set to find the appropriate weight decay value, so as to further improve the diagnostic ability of the DGMMF model. We change the value of weight decay in the interval [0, 0.0001, 0.0003, 0.0005, 0.0007]. We choose the optimal weight decay value based on its performance on the validation set.

The influence of learning rate on the DGMMF model

We run the following experiments to observe the effect of learning rate on the DGMMF model. The specific settings are as follows, the change interval of the learning rate is [0.001, 0.01, 0.05, 0.1, 0.5]. The evaluation metrics used have mentioned above.

Figure 7 shows the performance of the DGMMF model at different learning rates. We found that when the learning rate is too large, the DGMMF model fails to train. When the learning rate is 0.001, DGMMF is trained successfully. Therefore, we need to choose an appropriate learning rate according to the complexity of the model.

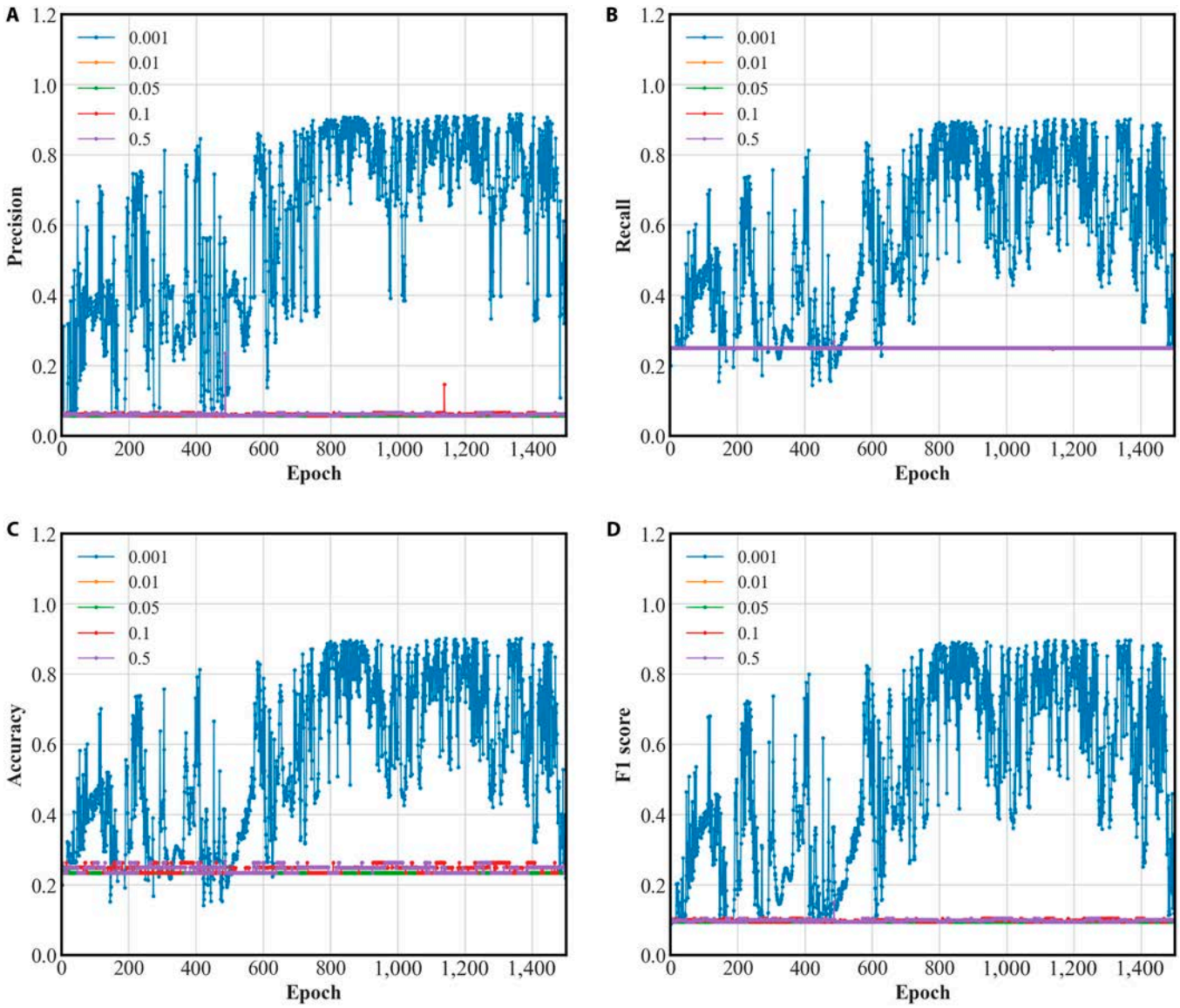


Fig. 7. The influence of learning rate on the DGMMF model. (A) Precision, (B) recall, (C) accuracy, and (D) F1 score.

Experimental results compared to other methods

We compare the experimental results of the DGMMF model with the following model.

- Att-MBiGRU [34]: Multilayer bidirectional gated recurrent units with attention mechanism (Att-MBiGRU) combined convolutional neural network, recurrent neural network, and attention mechanism to realize automatic diagnosis of bearing signals.
- SAE [38]: The stacked autoencoder (SAE) diagnose bearing faults through multiple encoding and decoding operations.
- NN [39]: The neural network (NN) uses neural units and activation functions to diagnose bearing faults.

Table. The experimental results of the DGMMF model and benchmarks.

Methods	Precision	Recall	Accuracy	F1 score
NN	0.876	0.862	0.863	0.860
SAE	0.885	0.869	0.870	0.867
Att-MBiGRU	0.907	0.906	0.907	0.905
DGMMF	**0.926**	**0.924**	**0.926**	**0.925**

Table and Fig. 8 show the final experimental results of the above comparison models. Observing the above results, we can find that the DGMMF model has achieved the highest value under all metrics, among which the value of precision is 0.926, the value of recall is 0.924, the value of accuracy is 0.926, and the value of F1 score is 0.925. The second model is the Att-MBiGRU model, and its scores under the 4 evaluation metrics are 0.907, 0.906, 0.907, and 0.905, respectively. The last model is the NN model, and its scores under the 4 evaluation metrics are 0.876, 0.862, 0.863, and 0.860, respectively. Compared with

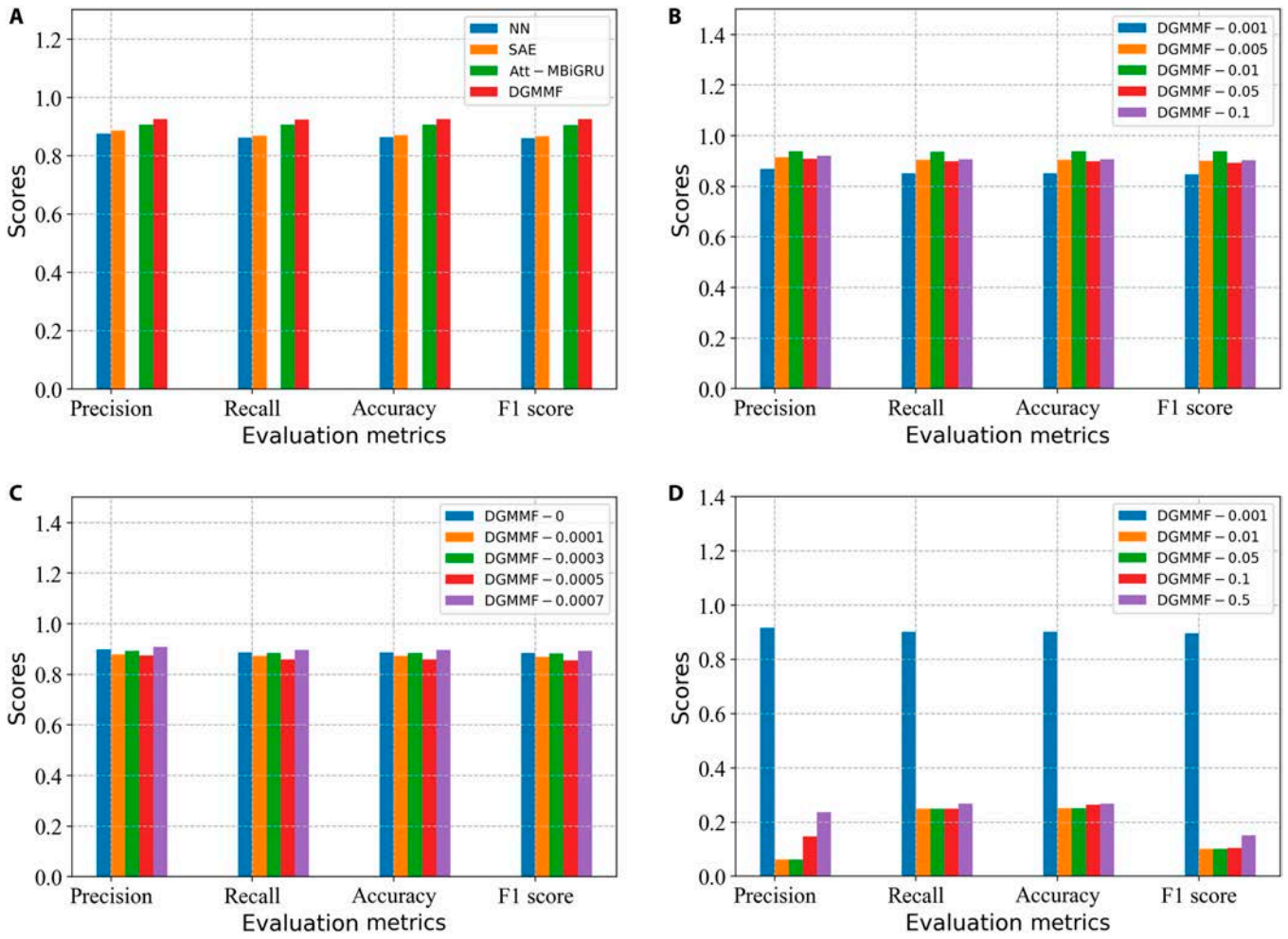


Fig. 8. The experimental results of the DGMMF model. (A) The experimental results of the DGMMF model and benchmarks. (B) The experimental results of the DGMMF model with different α (maximum value in 1,500 iterations). (C) The experimental results of the DGMMF model with different weight decay (maximum value in 1,500 iterations). (D) The experimental results of the DGMMF model with different learning rate (maximum value in 1,500 iterations).

the Att-MBiGRU model, the DGMMF model has an improvement of 2%, 1.9%, 2%, and 2.2% under the 4 evaluation metrics. Compared with the NN model, the DGMMF model has an improvement of 5.7%, 7.1%, 7.3%, and 7.5% under the 4 evaluation metrics.

The above experimental results illustrate the effectiveness of the DGMMF model. The reason why the DGMMF model can achieve better diagnosis results is that the use of VAE as a generative model realizes the augmentation of normal data and fault data, which makes the distribution of various samples balanced, thus simplifying the training difficulty of the model. In addition, the DGMMF model integrates features of different scales. Compared with single-scale features, these multiscale features contain richer information and can better express features, thereby improving the fault diagnosis performance of the model.

Conclusion

In this work, we design a bearing fault data collection platform based on the IIoT, which is used to collect bearing status data from sensors in real time and feed it back to the bearing fault diagnosis model. We propose a bearing fault diagnosis model DGMMF based on deep generative models with multiscale features. The model can integrate features of different scales and realize the augmentation of normal data and fault data, so that the distribution of various samples is balanced. Finally, we conducted a large number of related experiments on the real bearing fault datasets, which verified the effectiveness of the DGMMF model using multiple evaluation metrics. However, the DGMMF model has not yet been run in a real factory environment. Therefore, in future work, we will look for cooperation opportunities to verify the effectiveness of the DGMMF model in a real factory environment.

Acknowledgments

Author Contributions: Hexuan Hu constructed the model framework, completed some experiments, and double-checked this paper carefully before submission. Yicheng Cai performed the research, analyzed the data, and wrote this paper. Qiang Hu contributed to the drawing illustrations and polishing the language. Ye Zhang contributed to refining the ideas, carrying

out additional analyses and finalizing this paper. **Competing interests:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

References

1. Werfel F, Floegel-Delor U, Rothfeld R, Riedel T, Goebel B, Wippich D, Schirrmeister P. Superconductor bearings, flywheels and transportation. *Supercond Sci Technol*. 2012;25(1):Article 014007.
2. Mukoyama S, Nakao K, Sakamoto H, Matsuoka T, Nagashima K, Ogata M, Yamashita T, Miyazaki Y, Miyazaki K, Maeda T, et al. Development of superconducting magnetic bearing for 300 kw flywheel energy storage system. *IEEE Trans Appl Supercond*. 2017;27(4):1–4.
3. Zhou W, Habetler TG, Harley RG, Bearing condition monitoring methods for electric machines: A general review. Paper presented at: 2007 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives; 2007 Dec 04; Cracow, Poland. p. 3–6.
4. He Y, Zhang X. Approximate entropy analysis of the acoustic emission from defects in rolling element bearings. *J Vib Acoust*. 2012;134(6).
5. He D, Li R, Zhu J, Zade M. Data mining based full ceramic bearing fault diagnostic system using AE sensors. *IEEE Trans Neural Netw*. 2011;22(12):2022–2031.
6. Ulatowski A, Bazzi AM. A combinational-logic method for electric vehicle drivetrain fault diagnosis. *IEEE Trans Ind Appl*. 2016;52(2):1796–1807.
7. Li B, Chow M-Y, Tipsuwan Y, Hung JC. Neural-network-based motor rolling bearing fault diagnosis. *IEEE Trans Ind Electron*. 2000;47(5):1060–1069.
8. Lou X, Loparo KA. Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mech Syst Signal Process*. 2004;18(5):1077–1095.
9. Guo X, Chen L, Shen C. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Measurement*. 2016;93:490–502.
10. Immovilli F, Bellini A, Rubini R, Tassoni C. Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison. *IEEE Trans Ind Appl*. 2010;46(4):1350–1359.
11. Li X, Jiang H, Wang R, Niu M. Rolling bearing fault diagnosis using optimal ensemble deep transfer network. *Knowl Based Syst*. 2021;213:106695.
12. Jin X, Zhao M, Chow TW, Pecht M. Motor bearing fault diagnosis using trace ratio linear discriminant analysis. *IEEE Trans Ind Electron*. 2014;61(5):2441–2451.
13. Wang H, Liu Z, Peng D, Qin Y. Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis. *IEEE Trans Industr Inform*. 2019;16(9):5735–5745.
14. Li X, Zhang W, Ding Q. Understanding and improving deep learning-based rolling bearing fault diagnosis with attention mechanism. *Signal Process*. 2019;161:136–154.
15. Hoang D-T, Kang H-J. A survey on deep learning based bearing fault diagnosis. *Neurocomputing*. 2019;335:327–335.
16. Sohaib M, Kim C-H, Kim J-M. A hybrid feature model and deep-learning-based bearing fault diagnosis. *Sensors*. 2017;17(12):2876.
17. Neupane D, Seok J. Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review. *IEEE Access*. 2020;8:93155–93178.
18. Sun J, Yan C, Wen J. Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning. *IEEE Trans Instrum Meas*. 2017;67(1):185–195.
19. Chen Z, Mauricio A, Li W, Gryllias K. A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks. *Mech Syst Signal Process*. 2020;140:Article 106683.
20. He M, He D. Deep learning based approach for bearing fault diagnosis. *IEEE Trans Ind Appl*. 2017;53(3):3057–3065.
21. Wang H, Xu J, Yan R, Gao RX. A new intelligent bearing fault diagnosis method using SDP representation and SE-CNN. *IEEE Trans Instrum Meas*. 2019;69(5):2377–2389.
22. Zhang A, Li S, Cui Y, Yang W, Dong R, Hu J. Limited data rolling bearing fault diagnosis with few-shot learning. *IEEE Access*. 2019;7:110895–110904.
23. Guo S, Zhang B, Yang T, Lyu D, Gao W. Multitask convolutional neural network with information fusion for bearing fault diagnosis and localization. *IEEE Trans Ind Electron*. 2019;67(9):8005–8015.
24. Wu Z, Jiang H, Zhao K, Li X. An adaptive deep transfer learning method for bearing fault diagnosis. *Measurement*. 2020;151:Article 107227.
25. Chi Y, Dong Y, Wang ZJ, Yu FR, Leung VC. Knowledge-based fault diagnosis in industrial internet of things: A survey. *IEEE Internet Things J*. 2022;9(15):12886–12900.
26. Wang X, Lu S, Huang W, Wang Q, Zhang S, Xia M. Efficient data reduction at the edge of industrial internet of things for pmsm bearing fault diagnosis. *IEEE Trans Instrum Meas*. 2021;70:1–12.
27. Tang S, Yuan S, Zhu Y. Convolutional neural network in intelligent fault diagnosis toward rotatory machinery. *IEEE Access*. 2020;8:86510–86519.
28. Ranjith R, Athanesious JJ, Vaidehi V, Anomaly detection using dbscan clustering technique for traffic video surveillance. Paper presented at: 2015 Seventh International Conference on Advanced Computing (ICoAC); 2015 Dec 15–17; Chennai, India, pp. 1–6.
29. Schubert E, Sander J, Ester M, Kriegel HP, Xu X. DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Trans Database Syst*. 2017;42(3):Article 19.
30. Yan S, Smith JS, Lu W, Zhang B. Abnormal event detection from videos using a two-stream recurrent variational autoencoder. *IEEE Trans Cogn Dev Syst*. 2018;12(1):30–42.
31. Liou C-Y, Cheng W-C, Liou J-W, Liou D-R. Autoencoder for words. *Neurocomputing*. 2014;139:84–96.
32. Zhang M, Xu B, Gong J, An anomaly detection model based on one-class SVM to detect network intrusions. Paper presented at: 2015 11th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN); 2015 Dec 16–18; Shenzhen, China. p. 102–107.
33. Erfani SM, Rajasegarar S, Karunasekera S, Leckie C. High-dimensional and large-scale anomaly detection using a

linear one-class SVM with deep learning. *Pattern Recogn*. 2016;58:121–134.
34. Yang Z-b, Zhang J-p, Zhao Z-b, Zhai Z, Chen X-f. Interpreting network knowledge with attention mechanism for bearing fault diagnosis. *Appl Soft Comput*. 2020;97:Article 106829.
35. Li X, Zhang W, Ding Q, Sun J-Q. Multi-layer domain adaptation method for rolling bearing fault diagnosis. *Signal Process*. 2019;157:180–197.
36. Zhu H, He Z, Wei J, Wang J, Zhou H. Bearing fault feature extraction and fault diagnosis method based on feature fusion. *Sensors*. 2021;21(7):2524.
37. Patel SP, Upadhyay SH. Euclidean distance based feature ranking and subset selection for bearing fault diagnosis. *Expert Syst Appl*. 2020;154:Article 113400.
38. Zhou P, Han J, Cheng G, Zhang B. Learning compact and discriminative stacked autoencoder for hyperspectral image classification. *IEEE Trans Geosci Remote Sens*. 2019;57(7):4823–4833.
39. Li Z, Liu F, Yang W, Peng S, Zhou J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans Neural Netw Learn Syst*. 2021;33(12): 6999–7019.