Fault Diagnosis of Tunnel Fan Motor Bearings Based on Three-dimensional Vibration Joint Prediction and Auto-Encoder Mechanism

Fu Wang
Research Institute of Intelligent Transportation,
Zhejiang Institute of Mechanical & Electrical Engineering CO., LTD., Hangzhou, China
603329336@qq.com

Hang Shen
Research Institute of Intelligent Transportation,
Zhejiang Institute of Mechanical & Electrical Engineering CO., LTD., Hangzhou, China

Qiaonan Wu
Traffic Engineering Department,
Zhejiang Institute of Mechanical & Electrical Engineering CO., LTD., Hangzhou, China

*Zhuoren Zhang
Research Institute of Intelligent Transportation,
Zhejiang Institute of Mechanical & Electrical Engineering CO., LTD., Hangzhou, China
*Corresponding Author: zhangzhuoren1@gmail.com

Zhaohui Lin
Traffic Engineering Department,
Zhejiang Institute of Mechanical & Electrical Engineering CO., LTD., Hangzhou, China

Bao Li
Research Institute of Intelligent Transportation,
Zhejiang Institute of Mechanical & Electrical Engineering CO., LTD., Hangzhou, China

Abstract—**Fault diagnosis of tunnel fan motor bearings is crucial to ensure the safe operation of transportation infrastructure. Existing bearing fault diagnosis methods have problems such as limited data dimension and insufficient model learning ability. To address this problem, this paper proposes a diagnostic model that integrates three-dimensional vibration joint prediction and Auto-Encoder mechanism. The model is tested on the open source dataset released by MUET. The experimental results show that the model performs significant advantages under complex working conditions and is superior to other models. Specifically, compared with the multilayer perceptron classifier, the precision, recall and F1-score are improved by 27.94%, 28.19% and 28.06% in absolute percentage, respectively.**

Keywords--Bearing Fault Diagnosis; Three-dimensional vibration joint prediction; Auto-Encoder; Deep learning

I. INTRODUCTION

Fans are the core ventilation equipment of tunnel traffic engineering, and their core components, rolling bearings, play a vital role in the operation of fans. However, rolling bearings are subjected to the combined stress of alternating loads and mechanical vibrations under harsh working conditions of high temperature, high humidity and long-term continuous operation, and are prone to various failures[1]. Studies have shown that tunnel fan motor bearing failures account for more than 40% of the overall failures. Once a failure occurs, it will directly lead to the paralysis of the ventilation system and even cause safety accidents such as excessive CO concentration. Therefore, studying efficient and accurate bearing fault diagnosis methods is of great practical significance for ensuring the safe operation of tunnel traffic lifeline.

Existing research mainly extracts fault features based on one-dimensional vibration signals. Early methods usually use time domain statistical features, such as root mean square, kurtosis, or frequency domain analysis, such as Fourier transform, wavelet packet decomposition, combined with shallow learning models such as support vector machines and random forests for diagnosis. These methods rely on expert experience for complex feature engineering and have difficulty capturing deep nonlinear relationships in vibration signals. With the continuous advancement of deep learning technology, researchers have begun to use convolutional neural networks (CNN) and long short-term memory networks (LSTM) to process vibration signals and improve diagnostic performance through automatic feature learning. However, existing methods still have the following key limitations:

1. Data dimension limitation: Existing studies usually only model and process one-dimensional signals, resulting in the loss of vibration feature information and difficulty in fully reflecting fault characteristics.

2. Dependence on feature construction: Although deep learning reduces the reliance on manual features, existing methods still need to construct different features as model inputs, such as Fourier transform or wavelet transform. This method of manually mining features requires experienced engineers to construct features and increases computational complexity.

To overcome the above limitations, this paper proposes a diagnostic model that integrates three-dimensional vibration joint prediction and Auto-Encoder mechanism, and makes the following contributions:

1. Most existing studies use one-dimensional vibration data for fault diagnosis. This paper innovatively uses three-dimensional vibration data for joint prediction.

2. Unlike existing studies that require manual feature design, this paper adopts an end-to-end diagnostic method to learn fault features directly from raw data without manual design, simplifying the process and improving efficiency.

3. The accuracy of the bearing fault diagnosis model proposed by this paper is significantly better than that of traditional methods.

This paper is organized as follows: Section I provides an overview, which introduces the background of tunnel fan motor bearing fault diagnosis and summarizes the contributions of this paper. Section II reviews the related work in the field of bearing fault diagnosis. Section III focuses on the methodology and details the operation principle of the proposed model. Section IV introduces the experimental procedure, compares the proposed model with several models and analyzes the experimental results. Section V concludes the paper.

II. RELATED WORK

At present, the status of bearing fault diagnosis classification mainly includes the following three aspects. First, research institutions and enterprises in various countries have begun to carry out research and application of rolling bearing fault diagnosis classification, such as the Department of Mechanical Engineering at Stanford University in the United States, Siemens in Germany, and NTN in Japan, and have achieved certain research results and application practices. Second, machine learning has been widely used in the field of bearing fault diagnosis classification. Traditional machine learning fault diagnosis methods generally cover three main steps: data collection, feature extraction, and fault classification. In the data collection step, raw data that can reflect the health status of the bearing is obtained through a variety of sensors such as vibration sensors, current sensors, temperature sensors, and acoustic emission sensors. In the feature extraction step, features are manually selected and extracted based on predefined formulas designed by expert knowledge. In the fault classification step, some intelligent methods are usually used to establish the mapping relationship between features and faults, including decision trees, support vector machines[2], artificial neural networks[3], etc. In addition, the deep learning-driven fault diagnosis model can autonomously extract fault features directly from the input data without prior knowledge, and judge the health status of the bearing based on this, including convolutional neural networks[4], recursive neural networks[5], restricted Boltzmann machines, etc. Yu[6] et al. developed a rolling bearing fault diagnosis technology that integrates empirical mode decomposition (EMD) energy entropy and neural network. The combination of EMD and neural network provides an effective means to realize intelligent diagnosis of rolling bearing faults. Chen et al.[7] proposed a fault diagnosis method based on cyclic spectral coherence (CS Coh) two-dimensional graph representation and convolutional neural network. These technologies can analyze rolling bearing faults from different angles and improve the accuracy and reliability of diagnosis.

However, the current rolling bearing fault diagnosis and classification still faces some problems and challenges, such as the difficulty of data collection and processing, complex and changeable signals, and low classification accuracy. Therefore, how to improve the accuracy and efficiency of rolling bearing fault diagnosis and classification is still a problem that needs to be solved in current research and application. This paper proposes a tunnel fan motor bearing fault diagnosis method based on three-dimensional vibration joint prediction and Auto-Encoder mechanism, which makes the fault characteristics more accurately expressed and input into the fault diagnosis model. By comparing with other algorithms in experiments, it is proved that this method has a higher accuracy rate.

III. METHODOLOGY

Before describing the model structure, we first describe the flow of data in the model. Firstly, the original three-dimensional time-domain vibration data will be removed invalid data and then aligned according to time. Missing data is filled utilizing linear interpolation. Then the min-max method usually utilized to re-scale the data to the range of [0, 1]. The encoder part of the Auto-Encoder is then used to reconstruct the data features to obtain the latent representation of the data. Next, the latent representation will be sent to 3 blocks with convolutional neural network for feature extraction. The extracted features will be eventually mapped to the dimensions of the number of categories through a linear layer. Finally, softmax will be used to represent the probability of each predicted category.

The Auto-Encoder[8] is a unsupervised model, which is trained to reconstruct the input itself. It learns the low-dimensional latent representation of the original data by compressing and reconstructing the data. Its structure is shown in the figure 1.

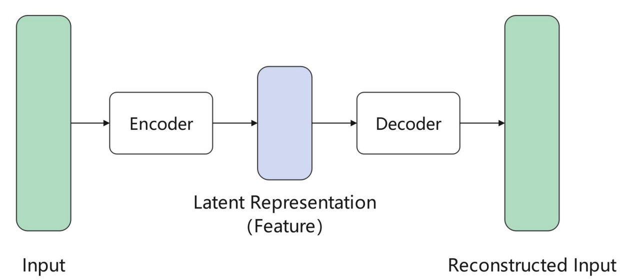

Figure 1. Auto-Encoder

The Auto-Encoder consists of two parts: the encoder and the decoder. The encoder extracts the features of the original data by mapping the high-dimensional input data to a low-dimensional latent space. The decoder reconstructs the original data from the low-dimensional latent representation, restoring the original input as much as possible. Its formula can be expressed as (1) and (2),

$$z = \mathrm{Encoder}(x) \quad (1)$$

$$\tilde{x} = \mathrm{Decoder}(z) \quad (2)$$

where z represents latent representation, $\tilde{x}$ represents the reconstructed input data. Usually, the Auto-Encoder is trained using loss functions such as MSE.

Convolutional neural network [9] is a deep learning model designed for processing grid-like data. It is mainly composed of convolutional layers, activation functions, pooling layers and fully connected layers.

Assume that the input size of X is $H_{in} \times W_{in} \times C_{in}$, the size of the convolution kernel K is $K_h \times K_w \times C_{in} \times C_{out}$, step size is S, padding is P. Then the convolutional layer can be expressed as (3),

$$F_{i,j,c_{out}} = \sum_{c_{in}=1}^{C_{in}} \sum_{m=0}^{K_{h-1}} \sum_{n=0}^{K_{w-1}} X_{i\times S+m, j\times S+n} \cdot K_{m,n,c_{in},c_{out}} + b_{c_{out}} \quad (3)$$

Where $b_{c_{out}}$ denotes bias.

The activation function usually uses ReLU, and the formula can be expressed as (4),

$$f(x) = \max(0, x) \quad (4)$$

The pooling layer usually use the Max-Pooling operation, assuming that the pooling window is $P_h \times P_w$, pooling step size is S_p, the formulas can be written as (5),

$$Y_{i,j} = \max_{m=0}^{P_{h-1}} \max_{n=0}^{P_{w-1}} X_{i\times S_p+m, j\times S_p+n} \quad (5)$$

where $Y_{i,j}$ represents each position of the output feature map.

The fully connected layer usually performs a linear transformation, and its formula is

$$y = WX + b \quad (6)$$

where W is the weight matrix and b is the bias matrix.

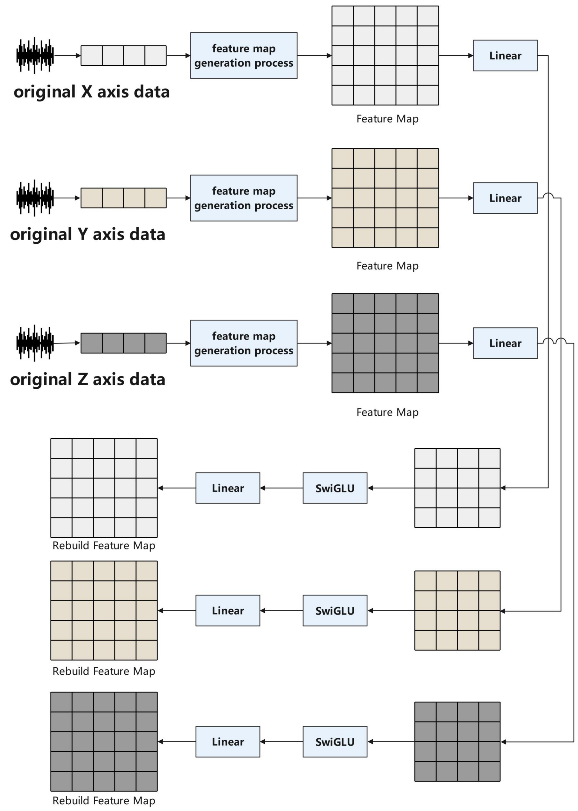

Figure 2. Auto-Encoder Structure in this paper

The model proposed by this paper utilizes Auto-Encoder to train the latent representation of the original three-dimensional time-domain vibration data. In the subsequent diagnostic task, only the latent representation output by the encoder of the autoencoder will be used. Its structure is shown in the figure 2.

Firstly, the original window data is generated into a feature map through the feature map generation process, and then the feature map is compressed into a low-dimension latent space using a linear layer. After that, the SwiGLU activation function is performed for activation. Finally, a linear layer is used to restore the features to the dimension of the original feature map. The loss function uses the mean square error, the formula is (7),

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \tilde{y}_i) \quad (7)$$

SwiGLU intergrates the Swish activation function and the gated linear unit to a combined function. SwiGLU replaces the Sigmoid function in the GLU with the Swish function, the formula is (8),

$$\mathrm{SwiGLU}(x) = \mathrm{Swish}(W_2 x + b_2) \otimes (W_1 x + b_1) \quad (8)$$

where $\otimes$ denotes Hadamard product.

The model we proposed innovatively combines the Auto-Encoder mechanism and performs joint prediction based on the 3-dimension vibration data. The structure of the model is shown in the figure 3.

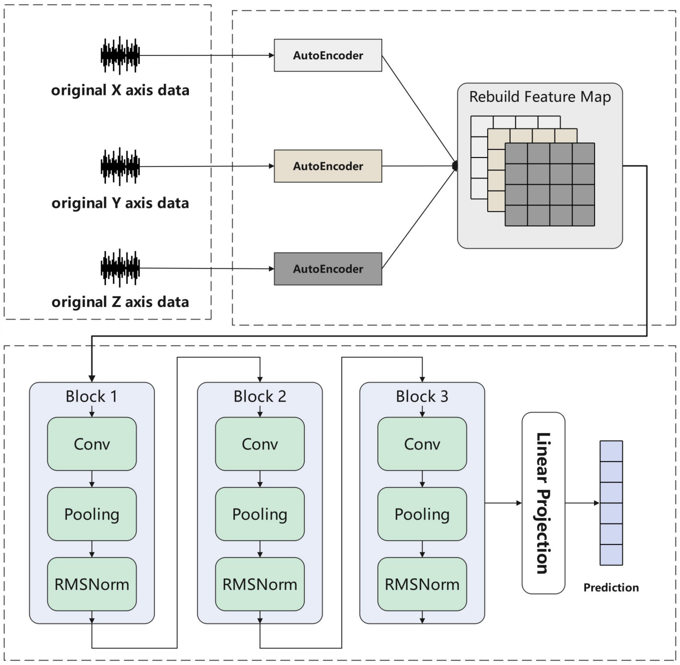

Figure 3. Proposed Model

Firstly, the data of the X-axis, Y-axis, and Z-axis are extracted to represent the latent space through the Auto-Encoder model. The extracted features are concatenated into feature maps of three channels and sent to Block1. Block1 uses a convolution kernel with an input channel number of 3, an output channel number of 10, and 3×3 size to convolve the feature map. Then the convolution result is pooled using the Max-Pooling operation. After that, the data is regularized using the RMSNorm method. The formula for RMSNorm is as (9),

$$\mathrm{RMS}(x) = \sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2} \tag{9}$$

$$y = \frac{x}{\sqrt{\mathrm{RMS}(x)+\epsilon}} \cdot \gamma \tag{10}$$

d refers to the input dimension of x and γ is a learnable scaling parameter. ϵ prevent the denominator from being 0.

Then the data is sent to Block2 and Block3 in order. Block2 and Block3 have the same structure except that the convolution kernel size is different from that of Block1. The convolution kernel of Block2 has an input channel number of 10, an output channel number of 10, and 5×5 size. The convolution kernel of Block3 has an input channel number of 10, an output channel number of 3, and 3×3 size. After flattening the output of Block3, using a fully connected layer to map the flat result to the dimension of the diagnostic category, and finally the result is converted into the predicted probability of each category using softmax function.

IV. EXPERIMENTS

A. Experimental Data and Settings

The bearing vibration data used in this paper comes from the open source data set released by MUET[10]. The data set constitute of a total of 7 types data, with the diameter from 0.7mm to 1.7mm and fault-free data. In addition to the healthy data, each fault data is divided into two types: inner fault and outer fault, and each inner fault and outer fault is also tested using 100w, 200w and 300w loads.

The fault classification and number of sampling points of the data set are shown in the table 1, the number of sampling points without failure is 235630.

Table 1. Sampling Points

	0.7mm	0.9mm	1.1mm	1.3mm	1.5mm	1.7mm
Inner Fault	382234	422066	420391	404280	410993	404911
Outer Fault	326168	403876	426461	399726	392474	404108

To ensure the quality of the data, we preprocessed the data as follows:

Remove invalid data (more than 3 standard deviations).

Fill missing data using linear interpolation.

Normalize the data to the range of [0, 1].

Slide the data on each axis with a window of size 5000 to obtain window data.

The sizes of the training set, validation set, and test set are divided into 70%, 15% and 15%.

The experiment was performed on the server, which configuration is as follows:

CPU: Intel Core i9; GPU: RTX 4090 24GB; RAM: 128GB; PyTorch: 2.6; Python: 3.10

The parameters of the Auto-Encoder model are set as follows:

The matrix size of linear layer 1: 100×50, the matrix size of linear layer 2: 50×100 , Learning rate: 1e-3, Optimizer: Adam, Training Epochs: 20(with early stopping), Dropout: 0.5

The parameters of the diagnose model are set as follows:

The initial learning rate is set to 1e-4, the number of warm-up epochs is set to 20, and the learning rate from 1^{st} to 20^{th} epoch is increased to 2e-3 based on the momentum mechanism. Then until the last epoch, the learning rate is gradually reduced to 2e-4 based on the momentum mechanism.

Optimizer: Adam, Batch Size: 32, Training Epochs: 100 (with early stopping), Dropout: 0.3, CNN Kernal Size: [3, 10, 3, 3], [10, 10, 5, 5], [10, 3, 3, 3].

B. Experimental results and analysis

The loss function is used to measure the difference between two probability distributions. This article utilizes the cross entropy loss function, which is widely used in classification tasks. Its formula is as (11):

$$\mathrm{CrossEntropyLoss} = -\frac{1}{n}\sum_{m=1}^{n}\sum_{i=1}^{class} y_i^m * \log(\widehat{y_i^m}) \tag{11}$$

y_i^m represents the true label value of the m^{th} sample, $\widehat{y_i^m}$ represents the probability of the corresponding category predicted by the m^{th} sample, and n represents the total number of samples.

In addition, this paper utilizes precision, recall, and F1-score to test the performance of the proposed model. The formula for precision is (12):

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{12}$$

The formula for recall is (13),

$$\text{Recall} = \frac{\text{TP}}{\text{TP+TN}} \tag{13}$$

The formula for the F1 value is (14):

$$\text{F1} = 2 \times \frac{\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \tag{14}$$

The results of the model proposed by this paper is compared with MLP Classifier, SVM, CNN, BiLSTM, FFT+CNN and FFT+BiLSTM, and the test results can be read from the Table 2.

Table 2. Test Results

Model	Precision	Recall	F1-Score
MLP Classifier	70.72%	70.04%	70.38%
SVM	75.63%	73.38%	74.49%
CNN	79.17%	77.03%	78.08%
BiLSTM	80.76%	79.64%	80.19%
FFT+CNN	94.19%	94.01%	94.10%
FFT+BiLSTM	95.12%	94.02%	94.56%
ours	**98.66%**	**98.23%**	**98.44%**

The experimental results show that the multilayer perceptron classifier has the worst performance, whilst the model proposed by this paper performs the best. The precision, recall and F1-score of the model proposed by this paper reached 98.66%, 98.23% and 98.44 respectively. The absolute percentages are 27.94%, 28.19% and 28.06% higher than the worst multilayer perceptron classifier respectively. At the same time, the absolute percentages are also 3.54%, 4.21% and 3.88% higher than the FFT+BiLSTM model.

V. CONCLUSIONS

Intelligent fault diagnosis of tunnel fan motor bearings is a key technology to ensure the safe operation and maintenance of transportation infrastructure. The diagnostic model of three-dimensional vibration joint prediction and Auto-Encoder mechanism proposed in this paper effectively breaks through the bottleneck of traditional methods in terms of limited data dimension and insufficient feature representation ability. By combining the low-dimensional latent feature extraction capability of the Auto-Encoder and fusing the joint prediction of three-dimensional vibration signals, the efficient fusion and accurate prediction of multi-dimensional vibration information under complex working conditions are achieved. Experiments based on the Mehran University open source dataset show that this model has significant advantages in various scenarios. Experimental comparison further verifies the superiority of the model. Its precision of 98.66%, recall of 98.23% and F1-score of 98.44% are significantly ahead of existing methods, and are 27.94%, 28.19% and 28.06% higher than the worst-performing multi-layer perceptron classifier, and 3.54%, 4.21% and 3.88% higher than the suboptimal FFT+BiLSTM model, respectively. Although the results show significant improvements over existing methods, future research still faces some challenges, including modeling of environmental interference factors and improving cross-device migration diagnostic capabilities. This study provides a new technical path for intelligent diagnosis of rotating bearing faults, which has practical value in promoting the intelligent management of transportation infrastructure health.

REFERENCES

[1] Bin Liu, Changfeng Yan, Yaofeng Liu, “ISEANet: An interpretable subdomain enhanced adaptive network for unsupervised cross-domain fault diagnosis of rolling bearing,” Advanced Engineering Informatics, vol. 62, 2024.

[2] Guo H . Fault Diagnosis of Motor Bearings Based on Ensemble Empirical Mode Decomposition and Minimum Support Vector Machine[J].Journal of Engineering System,2024,2(3):

[3] Najim S H ,Alsalaet K J .Cross-domain diagnosis of roller bearing faults based on the envelope analysis adaptive features and artificial neural networks[J].Journal of Vibration and Control,2024,30(13-14):3168-3183.

[4] Ji Y ,Gao J ,Shao X , et al.Multi-scale quadratic convolutional neural network for bearing fault diagnosis based on multi-sensor data fusion [J]. Nonlinear Dynamics,2025, (prepublish):1-22.

[5] Ruxue B, Zong M, Quansheng X, et al.Fractional Fourier and time domain recurrence plot fusion combining convolutional neural network for bearing fault diagnosis under variable working conditions[J].Reliability Engineering and System Safety,2023,232

[6] YU Y, JUNSHENG C, “A roller bearing fault diagnosis method based on EMD energy entropy and ANN,” Journal of Sound and Vibration, 2006, 294(1/2):269-277.

[7] CHEN Z, MAURICIO A, LI W, et al, “A deep learning method for bearing fault diagnosis based on Cyclic Spectral Coherence and Convolutional Neural Networks,” Mechanical Systems and Signal Processing, 2020, 140:106683.

[8] Wang Y, Yao H, Zhao S. Auto-encoder based dimensionality reduction[J]. Neurocomputing, 2016, 184: 232-242.

[9] O'shea K, Nash R. An introduction to convolutional neural networks[J]. arXiv preprint arXiv:1511.08458, 2015.

[10] Kumar, Dileep; Mehran, Sanaullah; Shaikh, Muhammad Zakir; Hussain, Majid; Chowdhry, Bhawani Shankar; Hussain, Tanweer (2022), “Triaxial Bearing Vibration Dataset of Induction Motor under Varying Load Conditions”, Mendeley Data, V2, doi: 10.17632/fm6xzxnf36.2