Advances in Artificial Intelligence, Big Data and Algorithms
G. Grigoras and P. Lorenz (Eds.)

doi:10.3233/FAIA230883

An End-To-End Fault Diagnosis Method for Emulsion Pump with Class-Imbalance

Yuehua LAI[a, b, 1], Ran LI[a, b], Mingliang LIU[a], Zaoyang WU[a] and Rongming CHEN[a]
[a] *Beijing Tianma Intelligent Control Technology Co., Ltd., Beijing, 101399, China*
[b] *China Coal Research Institute, Beijing, 100013, China*

Abstract. The condition monitoring data of emulsion pump follow the long-tail distribution. The amount of monitoring data for the normal condition is very large, while the amount of monitoring data for different fault conditions is very small, the problem of class-imbalance is prominent. The traditional intelligent fault diagnosis methods are proposed under the assumption of class balance, which the fault diagnosis model has the shortcoming of insufficient generalization ability when dealing with the class-imbalance problem. Thus, an end-to-end fault diagnosis method for emulsion pump with class-imbalance is proposed. conditional variational autoencoder is used to extract features and learn the state data distribution of emulsion pump, and the loss value of training samples is adjusted based on focal loss to balance the influence of different types of data on the model. Moreover, the end-to-end fault diagnosis model can be obtained based on the decoder model. Finally, the effectiveness of the proposed method is verified by simulation experiment data of emulsion pump faults. Compared with other methods under different types of imbalanced rates, the results show that the fault of emulsion pump can be accurately identified under the condition of only a small amount of fault data by the proposed method and the corresponding recognition accuracy is better than other methods.

Keywords. Fault diagnosis, class-imbalance, emulsion pump, conditional variational autoencoder, focus loss

1. Introduction

Emulsion pump is an important power equipment in the liquid supply system of coal mine. Once failure occurs, it will not only affect the efficiency of coal mining, but also cause serious economic losses and security risks [1-3]. Therefore, condition monitoring and fault diagnosis of emulsion pump are of great significance for the safe and stable operation of equipment [4].

In actual engineering, the status monitoring data of emulsion pump obey the long-tail distribution, and the sample size in normal condition is very large, while the sample size of failure is very small, resulting in the unbalance of data categories that must be paid attention to in engineering [5]. Fault diagnosis based on deep learning can complete automatic fault feature extraction and fault classification diagnosis. As an end-to-end working mode, it is widely used in various key parts of machinery. Li *et al.* used convolutional neural networks to extract fault features and generated fault samples based

[1] Corresponding Author: Yuehua LAI, Beijing Tianma Intelligent Control Technology Co., Ltd.; China Coal Research Institute; E-mail: laiyh@tdmarco.com

on generated adversarial networks to realize bearing fault diagnosis under unbalanced training data [6]. Guo *et al.* used Wasserstein distance conditional gradient punishment generation adversarial network to generate simulation samples to balance the data set and improve the accuracy of fault diagnosis [7]. Cao *et al.* used one-dimensional convolutional neural network to extract fault features, and used quasi-balance loss function instead of cross entropy loss function to offset the impact of data unbalanced distribution on the network, so as to improve the accuracy of bearing unbalance diagnosis [8]. Zhang *et al.* built a feature extraction network based on the DenseNet framework, added penalty coefficients to different types of samples in the loss function to achieve the weighted average of unbalanced sample errors, and proposed a fault diagnosis model for unbalanced samples of pressure reducing valves based on W-DenseNet [9].

According to the above research, the model of "data set expansion + model training" is mainly used in the construction of category-unbalance fault diagnosis model, and the training process is divided into two stages, which cannot achieve global optimization well. Using loss function to offset the impact of unbalanced distribution of data on the network can play a certain role, but it cannot model the data distribution well and improve the generalization ability of the diagnostic model. Therefore, in view of the above problems, this paper proposes an end-to-end fault diagnosis method for emulsion pump under class imbalance. This method uses conditional variational autoencoder (CVAE) [10] to extract features and learn data distribution about pump and valve state data. Based on focus loss, the loss values of normal samples are adjusted to balance the influences of different categories of data on the model. An end - to - end fault diagnosis model based on decoder is obtained. The experimental data of pump and valve fault simulation are used to verify the effectiveness of the method. The results show that the method can realize the accurate identification of pump and valve fault under the condition of only a few fault data. In addition, this method can also provide some reference for the diagnosis of other mechanical faults under category imbalance.

2. The Proposed Method

2.1. Data Distribution Learning Based on CVAE

CVAE introduces the label category of data into VAE model so that the model can generate the category sample according to the preset category information. The sample label is regarded as the condition information, and the input of the encoder of CVAE is the combination of the condition information and the data sample, and the input of the decoder is the combination of the condition information and the hidden variable $\boldsymbol{z}$. As the category information is joined in VAE model, the expression of encoder can be described as $q_\theta(\boldsymbol{z}|\boldsymbol{x}, y)$ and the expression of decoder can be described as $p(\boldsymbol{x}|\boldsymbol{z}, y)$. As with the VAE, the loss function of the CVAE that includes Sample reconstruction error and KL divergence error is shown as follows:

$$L_{\mathrm{CVAE}}(\boldsymbol{x}, y) = E_{z\sim q_\phi(z|x,y)}\left[\log p(\boldsymbol{x} \mid z, y)\right] - KL\left[q(z \mid \boldsymbol{x}, y) \,\|\, p(z|y)\right] \tag{1}$$

According to the classical CVAE model, we can control the categories of generated samples. Therefore, augmentation of the data that has small amount can be conducted based on CVAE, and the generalization ability of the trained model can be improved after data augmentation. However, the training mode of fault diagnosis model is divided into two separate steps: data augmentation and model training, the global optimization of the model cannot be guaranteed. Therefore, the classical CVAE model is further modified to train by changing the input positions of sample $\boldsymbol{x}$ and label y in this paper. Thus, the obtained decoder, which is also the fault diagnosis model, can identify the class of samples according to the hidden variables that follow normal distribution and the input samples. The structure schematic of the modified CVAE is shown in Figure 1.

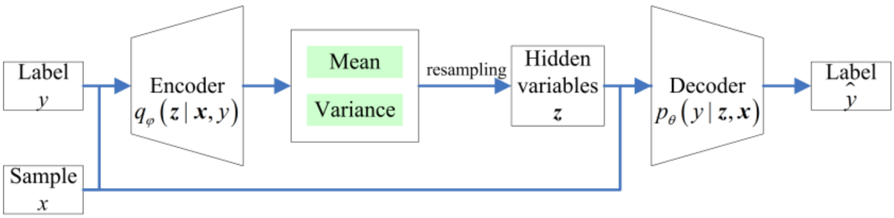

Figure 1. Structure schematic of the modified CVAE.

The loss function of modified CVAE is shown as follows:

$$\tilde{L}_{\mathrm{CVAE}} = -\mathrm{KL}\left(p_\theta(z \mid \boldsymbol{x}, y) \| q_\varphi(z \mid \boldsymbol{x})\right) + \mathrm{E}_{q_\varphi(z \mid \boldsymbol{x}, y)} \log p_\theta(y \mid \boldsymbol{x}, z) \tag{2}$$

2.2. Loss Weight Adjustment Based on Focus Loss

Focus loss is proposed based on cross-entropy loss, and its definition is given as follows:

$$L_{\mathrm{FL}}(p_t) = -\alpha_t (1 - p_t)^\gamma \log(p_t) \tag{3}$$

Compared with cross-entropy loss, focus loss function increases the focus parameter γ (γ>0), which can reduce the loss of easily classified samples and strengthen the model's learning of difficult classified samples. The value of focus parameter γ is suggested to be 2 in [11]. At the same time, the focus loss function also introduces a balance parameter α to deal with the imbalance of the proportion of positive and negative samples, which usually represents the frequency of the corresponding category in the training set.

2.3. Imbalanced Fault Diagnosis Method for Emulsion Pump

Under the framework of CVAE, multi-layer perceptron (MLP) is used to extract sensitive features to fully describe the distribution of state data in encoder. At the same time, MLP is also used to construct the decoder of CVAE. In addition, the loss value of the sample is adjusted based on the focus loss function to balance the influence of different types of data on the model. Finally, the fault diagnosis model can be obtained based on the decoder model. The specific steps of the proposed method are as follows:

Step 1: The vibration acceleration sensors are used to collect the vertical vibration signals of the pump valve, and the vibration signals are processed by adding windows to form signal samples.

Step 2: The vibration signal is transformed by FFT to obtain the signal spectrum, which will be further normalized and used as the input of the CVAE model.

Step 3: The CVAE model is constructed by MLP, and focal loss function is used to balance the influence of different types of data on the model.

Step 4: The decoder of CVAE is taken as the fault diagnosis model, and the sample and hidden variable parameters are input to realize the recognition of sample categories.

The flow diagram of the proposed method is shown in Figure 2.

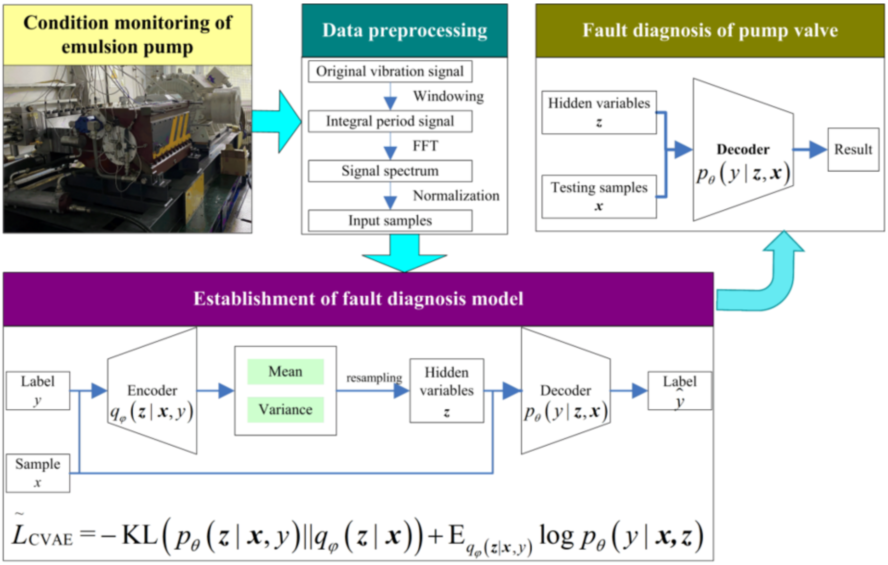

Figure 2. The flow diagram of imbalanced fault diagnosis method for pump valve.

3. Experimental Verification

3.1. Test Bench of Emulsion Pump

The test bench of BRW630/40 emulsion pump is shown in Figure 3. Five vibration acceleration sensors are arranged above the pump valve to collect vertical vibration signals of 5 groups of suction and discharge pump valves. Meanwhile, a Hall sensor is installed on the end cover of the crankcase to obtain the phase information of the crankshaft.

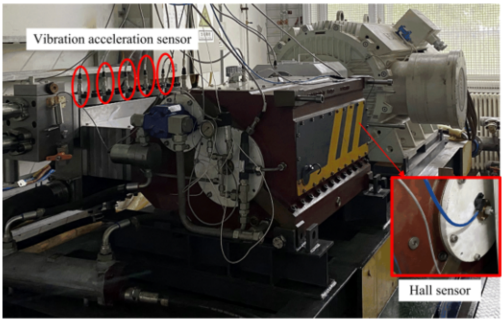

Figure 3. Test bench of BRW630/40 emulsion pump.

After replacing the faulty spool of the normal group 5 suction and discharge spool individually or together, the fault data of the pump valve is collected. Set the vibration sampling frequency as 10.24kHz, and obtain vibration signals under different states through the monitoring system.

3.2. Fault Diagnosis of Emulsion Pump with Class-Imbalance

The applicability and effectiveness of the proposed method is verified by using the measured data set of pump and valve status. All the samples in the data set correspond to the vibration signals of a single complete suction and drainage cycle. The data set includes four different states of the pump and valve, namely normal state, suction valve failure, discharge valve failure and suction and discharge valve failure. Pump and valve data sets are divided into training sets and test sets. Data sample sizes corresponding to different pump and valve states under different data sets are shown in Table 1.

Table 1. Data sets divided under different states of pump valve.

Status of pump valve	Training set	Test set
Normal condition	5000	500
Drain valve fault	50	500
Suction valve fault	50	500
Suction and drain valve fault	50	500

CVAE model mainly consists of an encoder and a decoder. The encoder based on MLP network used a hidden layer of 512-256 nodes to extract sensitive features to describe the distribution of data, and the rectified linear unit (ReLU) function was selected as the activation function by the neurons. The features of the encoder are connected to two networks consisting of 64 neurons to learn the mean and variance of the features, respectively. Since both the encoder and decoder will use normalized spectrum of vibration signal as input, the input dimensions for the encoder and decoder are 1028 and 1088, respectively. At the same time, a decoder was built using four layers of MLP and the number of neurons in the network was set to 1088-512-256-4. In order to improve the sparsity of the model and prevent overfitting of the model, the dropout method is used in the full connection layer.

In addition, the focus loss function was used to adjust the loss value of the sample, and the parameters were selected as $\gamma = 2$ and $\alpha= 0.25$ to balance the influence of different types of data on the model [11]. After model parameters were initialized, Adam optimizer was selected to train the model, the learning rate was set to 1e-4, and the number of samples used for single model training was set to 64.

According to the flow chart of the proposed method shown in Figure 2, the measured data set was used for training and verification. In order to illustrate the feature extraction capability of the model, the t-distributed stochastic neighbor embedding (t-SNE) algorithm was used to reduce the dimension of the second layer feature of the decoder, and three-dimensional visualization is carried out for the first three principal components of the feature, as shown in Figure 4.

The generalization ability of fault diagnosis model is quantitatively measured by using test data set of pump and valve. For the identification of test sets, the results are presented in the confusion matrix, as shown in Figure 5. As can be seen from the results in the figure, the recognition accuracy of the fault diagnosis model for normal state can reach 100%, and the recognition accuracy of different fault states can reach more than 95%. The average accuracy of the fault diagnosis model is 97.05%, indicating that the method can accurately diagnose the emulsion pump faults under unbalanced conditions.

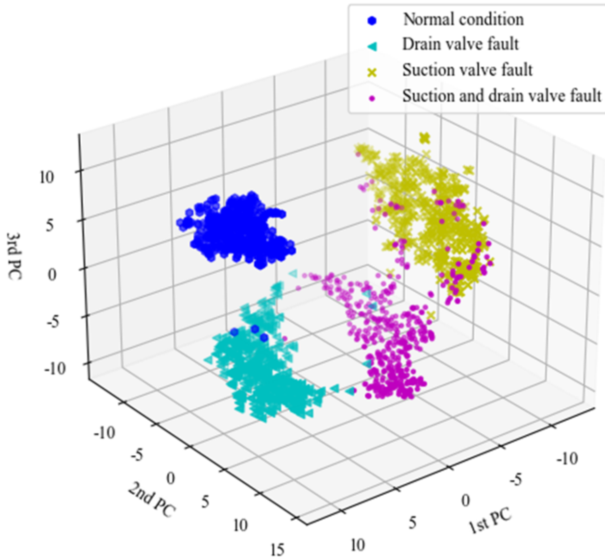

Figure 4. The 3D visualization of features based on t-SNE.

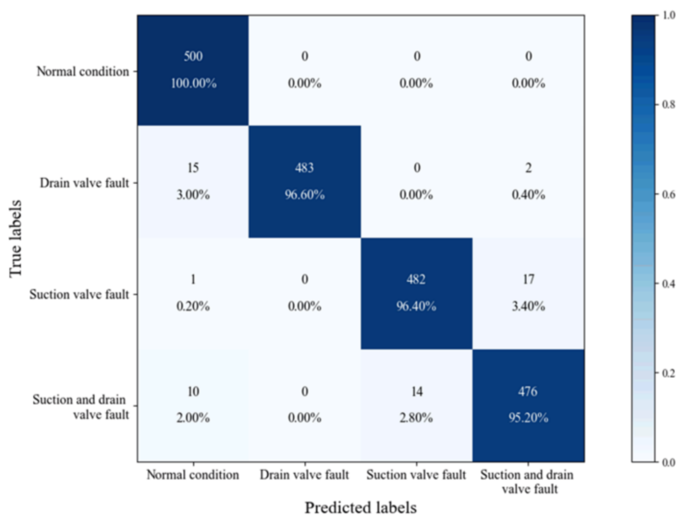

Figure 5. Confusion matrix of test set results.

3.3. Contrastive Analysis

In order to verify the performance of the fault diagnosis model, the proposed method is compared with the following methods.

1. SMOTE+DT: Based on the multi-domain feature set [12], use SMOTE over-sampling to increase the number of failure samples. Based on the balanced data set after data enhancement, a fault diagnosis model is established by using decision tree algorithm.
2. SMOTE+SVM: Based on the same balance data set as the first method, select SVM to build the fault diagnosis model with radial basis function as kernel.
3. ADASYN+SVM: Based on multi-domain feature set, ADASYN algorithm is used for over-sampling to increase the number of fault samples. Based on the balanced data set and support vector machine algorithm after data enhancement, a fault diagnosis model is established.
4. Downsampling +RF: The downsampling technique is adopted for normal samples to keep a balance between normal samples and fault samples. Based on Random Forest (RF) of integrated learning, a fault detection classifier is established.
5. CVAE Oversampling +MLP: The network structure of CVAE model is constructed by MLP network, and the network nodes are set as 1028-512-256-64(x2) and 10888-512-256-1024. Aiming at the problems of class unbalance, the CVAE model is adopted to oversample the fault samples so that the data categories reach a balanced state, and then the fault diagnosis model is trained in combination with MLP.

The degree of class imbalance will definitely affect the generalization ability of fault diagnosis model. Class unbalance rate is defined as the ratio of the number of normal samples to the number of wrong samples, as shown in formula (13).

$$ratio = \frac{n_n}{n_f} \tag{4}$$

Where, n_n and n_f respectively represent the number of normal samples and fault samples in the training data set.

The number of samples in the fixed normal state is 5000, and the number of samples in each fault state gradually increases, so that data sets under different types of unbalance rates can be obtained. Different training models can be obtained from training data sets of different types of unbalance rates. The fault diagnosis accuracy of the above methods under different types of unbalance rates is shown in Figure 6.

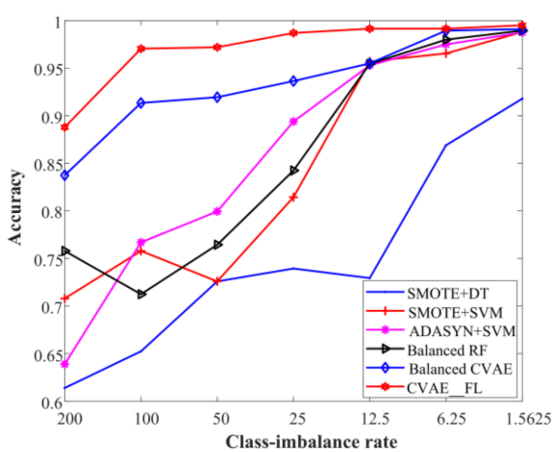

Figure 6. Accuracies of different fault diagnosis methods under different class-imbalance rates.

As can be seen from the figure, the effect of class unbalance rate on different methods is different. When the class unbalance rate tends to the extreme, the emulsion pump end-to-end fault diagnosis method has the best performance, especially when the class unbalance rate reaches 200.

4. Conclusion

In order to solve the problem of insufficient generalization ability of pump valve fault diagnosis model under class imbalance, an end-to-end fault diagnosis method for pump valve under class imbalance was proposed. In this method, conditional variational autoencoder variants are used for feature extraction and data distribution learning. The loss value of training samples is adjusted based on focus loss to balance the influence of different types of data on the model. The decoder obtained by training is the fault diagnosis model. The effectiveness of the method is verified by simulation experiment data of pump and valve fault.

References

[1] R. Li, W. Wei, H. Liu, et al., Comparative evaluation of wear behavior of tribo-pairs in reciprocating pumps with multiple materials under different conditions, *Journal of Theoretical and Applied Mechanics* **61** (2023), 77–87.

[2] R. Li, W. Wei, H. Liu, et al., Experimental and numerical study on dynamic and flow characteristics of the reciprocating pump valve, *Processes* **10**(2022), 1328.

[3] R. Li, D. Wang, W. Wei, et al., Analysis of the Movement Characteristics of the Pump Valve of the Mine Emulsion Pump Based on the Internet of Things and Cellular Automata, *Mobile Information Systems* **3**(2021), 1–8.

[4] R. Niu, H. Ding, R. Shi , et al., Fault diagnosis of emulsion pump based on depth residual network, *Coal Mine Machinery* **42**(2021), 177–180.

[5] J. Pan, Q. Ding, A. Jiang, et al., Fault diagnosis of unbalanced data of chillers based on xgboost, *Journal of Mechanical Strength* **43**(2021), 27–33.

[6] Z. Li, H. Yin, J. Zuo, et al., Bearing fault diagnosis based on generative adversarial network on imbalanced data, *Journal of Chinese Computer Systems* **42**(2021), 46–51.

[7] J. Guo, M. Wang, L. Sun, et al., New method of fault diagnosis for rolling bearing imbalance data set based on generative adversarial network, *Computer Integrated Manufacturing Systems* **28**(2022), 2825–2835.

[8] J. Cao, Z. He, P. Yu, et al., Bearing fault diagnosis method under unbalanced data distribution, *Journal of Jilin University(Engineering and Technology Edition)* **52**(2022), 2523–2531.

[9] H. Zhang, Y. Sheng, Z. Huang, et al., W-DenseNet-based fault diagnosis model of pressure-reducing valve withunbalanced samples, *Control and Decision* **37**(2022), 1513–1520.

[10] L. Girin, S. Leglaive, X. Bie, et al., Dynamical variational autoencoders: a comprehensive review, *Foundations and Trends in Machine Learning* **15**(2021), 1–175.

[11] T. Lin, P. Goyal, R. Girshick, et al., Focal loss for dense object detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **42**(2020), 318–327.

[12] X. Yan and M. Jia, 2018 A novel optimized SVM classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing, *Neurocomputing* **313**(2018), 47–64.