A Cloud-Edge Collaborative Intelligent Fault Diagnosis Method Based on LSTM-VAE Hybrid Model

Shaofei Lu, Xiaolin Tang, Yajun Zhu and Jingke She

May 4, 2021

A Cloud-Edge Collaborative Intelligent Fault Diagnosis Method Based on LSTM-VAE Hybrid Model

Shaofei Lu
College of Computer Science and Electronic Engineering
Hunan University
Changsha, China
sflu@hnu.edu.cn

Xiaolin Tang
College of Computer Science and Electronic Engineering
Hunan University
Changsha, China
xltanglx@hnu.edu.cn

Yajun Zhu
College of Computer Science and Electronic Engineering
Hunan University
Changsha, China
zhuyajun_summer@163.com

Jingke She
College of Computer Science and Electronic Engineering
Hunan University
Changsha, China
shejingke@hnu.edu.cn

***Abstract*—Fault diagnosis is of great significance for timely detection of safety hazards of machinery and ensures the normal operation of production. To address the problems of low accuracy and poor robustness of mechanical fault diagnosis methods in general, the paper proposes a cloud-edge collaborative intelligent fault diagnosis method based on the LSTM-VAE hybrid model. The method trains the LSTM-VAE hybrid model in the cloud by using the vibration signals of mechanical components at the early stage of operation, and then reconstructs the real-time vibration signals in the edge by using the trained LSTM-VAE, calculates the difference degree between the original signal and the reconstructed signal, compares them with the adaptive threshold, and combines the "3/5" strategy to achieve fault warning. The experimental results show that, compared with other fault diagnosis methods, the proposed method can accurately diagnose the fault of rolling bearings with different degradation modes, and significantly improve the fault warning time in slow degradation modes, with high timeliness and strong adaptability.**

Keywords—Cloud-Edge Collaborative, Long Short-Term Memory, Variational Auto-Encoder, Intelligent Fault Diagnosis, Deep Learning

I. INTRODUCTION

In recent years, the rapid development of information technology has promoted more and more technologies to become intelligent and networked [1], and machinery is made more and more complex. Once a minor failure occurs in the key components inside the machinery, it may lead to safety hazards in the whole machinery and then a safety accident. Therefore, it is of great practical significance to monitor and diagnose the operation status of key components of machinery in real-time.

Currently, mechanical fault diagnosis methods can be summarized into three categories: model-based methods, empirical knowledge-based methods, and data-driven methods [2]. The model-based diagnosis method explores the degradation process of machinery by constructing a mathematical model based on the failure mechanism of mechanical components [3]. For example, a model-based state estimation method for fault monitoring and identification of switching power converters was used by Poon [4] et al. However, the difficulty to fully understand the failure mechanism of damaged components makes it difficult to establish an accurate mathematical physical model, which limits the application of this method. The empirical knowledge-based fault diagnosis method transforms a large amount of expert experience and theory into a knowledge base for fault analysis by various reasoning means. Typical applications are in the form of expert systems and fuzzy logic. However, these methods often encounter bottlenecks in acquiring domain knowledge and converting it into rules, such as the knowledge base is easily limited by the expert's empirical knowledge; the rules of fuzzy logic are not easy to set, etc. In this case, it is recommended to adopt a data-driven fault diagnosis method. It uses signal processing technology and machine learning algorithms [5] or deep learning algorithms to deeply mine potential process information from the historical operation process data of key components of machinery, to perform fault diagnosis on machinery. Such as Dhamande [6] et al. extracted time-frequency features from vibration signals by Continuous Wavelet Transform and Discrete Wavelet Transform and accurately classified faults by combining artificial neural networks.

Since the dimensions of raw signals collected from machinery are generally large, and traditional machine learning algorithms have very limited ability to represent the data, it is difficult to learn effective features directly from the raw signals. As a branch in the field of machine learning, deep learning can create a multilayer network structure by continuously stacking neural networks to achieve deeper feature extraction, which solves the problem that traditional machine learning algorithms have insufficient ability to express features. Jia [7] et al. built a multilayer network based on DBN to mine features from spectral data and verified the effectiveness of the method in rolling bearing and gearbox datasets. Pan [8] et al. proposed an end-to-end fault diagnosis model CNN-LSTM for fault classification of bearings and the method achieved 99% accuracy in the test set.

However, most of the abovementioned data-driven algorithm models require high computational performance and usually require edge devices with high parallel processing capability as well as memory throughput, which hinders the application of algorithm models in edge devices. With the development of big data [9], artificial intelligence [10], blockchain [11], edge computing [12], and cloud computing [13,14], the application of data-driven fault diagnosis methods to edge devices using a collaborative cloud-edge manner is being promoted by an increasing amount of researchers. For example, Zhang [15] et al. proposed an improved convolutional neural network model that is applied to detect high impedance faults in distribution networks using a migratory learning ground approach to the cloud-edge cooperative framework. Gai [16] et al. collaborated on mobile edge computing, cloud computing, and reinforcement learning to propose an intelligent model, SRL-RA, for resource allocation problems in a complicated networking environment.

Despite the remarkable achievements of data-driven fault diagnosis methods, the following problems still exist: most data-driven methods are supervised learning and rely on large-scale labeled datasets, which are, however, more difficult to obtain in real industrial scenarios [17]; the use of signal processing techniques to extract features relies on the expertise and may lose useful information in the original signal; the algorithmic models have low diagnostic accuracy and cannot detect potential failures of machinery in slow degradation mode promptly; the robustness of algorithmic models is poor in strong noise backgrounds.

To solve the above problems, this paper proposes a new fault diagnosis method. The main contributions are as follows:

- This paper proposes a new cloud-edge collaborative intelligent fault diagnosis method based on the hybrid model LSTM-VAE, which can work with only a small amount of monitoring data at the beginning of machinery operation without annotation the data set. Combined with the proposed adaptive threshold method and the "3/5" strategy, the proposed method can accurately detect potential failures of machinery with two general degradation modes and prevent further expansion of mechanical faults.
- The proposed method can learn deeper abstract features directly from the vibration signals collected by the machinery without relying on the expert's empirical knowledge, avoiding manual feature selection.
- The LSTM-VAE hybrid model is enhanced using the Dropout method, which improves the robustness of the model in a strong noise environment and prevents model overfitting, providing the possibility of its application to real industrial environments, which has important engineering applications.

The rest of the paper is organized as follows. Section II describes the proposed method. Section III introduces the experiments and analysis of their results. Finally, we summarize this paper and identify the future directions in Section IV.

II. Proposed Method

A. LSTM-VAE hybrid model

Long Short-Term Memory (LSTM) is a variant of Recurrent Neural Network (RNN) proposed by Hochreiter and Schmidhuber [18] et al. Compared to the general RNN, LSTM effectively solves the long-term dependency problem of RNN by introducing various gate units within the recurrent neurons. Fig. 1 illustrates a standard LSTM memory cell, which is well suited for processing medium to long time sequence data due to its unique design structure.

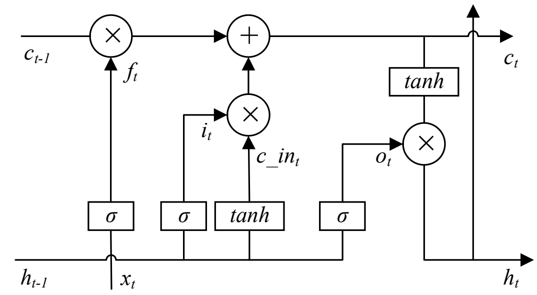

Fig. 1. LSTM memory cell

Variational Auto-Encoder (VAE) is an important generative model proposed by Kingma and Welling [19] et al. It is one of the most promising machine learning algorithms in unsupervised learning, which can self-learn the statistical distribution obeyed by the data and generate similar data. The VAE model consists of an encoder and a decoder, which can be implemented by various neural network models, such as Artificial Neural Network, Convolutional Neural Networks (CNN), RNN, etc. The structure principle of the VAE model is shown in Fig. 2, from which we can find that the data x is generated by the hidden variable z. $z \rightarrow \tilde{x}$ is the generative model $p_\theta(x \mid z)$, i.e., decoder, while $x \rightarrow z$ is the inferred model $q_\phi(z \mid x)$, i.e., encoder. VAE uses the inferred model $q_\phi(z \mid x)$ to approximate the true posterior probability distribution $p_\theta(x \mid z)$ and uses KL scatter to measure the similarity of the two distributions, and ϕ and θ are the parameters of the encoder and decoder, respectively.

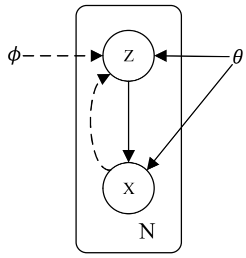

Fig. 2. VAE model structure principle

In the field of mechanical fault diagnosis, the vibration signals are often used as the main object of research on machinery fault diagnosis methods because they contain rich condition information of machinery [20]. The vibration signal belongs to a kind of time-series data, and there is a temporal correlation between its samples. And under the complex working conditions, the vibration signal contains more noise and presents the characteristics of nonlinear, nonperiodic, and nonstationary. Given the advantages of the LSTM model's excellent expressiveness to time series data and the VAE model's robustness to noise, this paper proposes a new data-driven hybrid model LSTM-VAE for fault diagnosis.

The LSTM-VAE hybrid model combined the advantages of both LSTM and VAE models. On the one hand, by introducing LSTM to establish time-series dependence on vibration signals, the long and short-term time-series features of vibration signals are extracted. On the other hand, it combines VAE to model the time series and map the vibration signal and the correlation between dimensions into the hidden space to achieve deeper feature extraction. Meanwhile, VAE learns the distribution of the vibration signal of the mechanical health state by modeling the vibration signal of the mechanical health state through variational inference. Because the

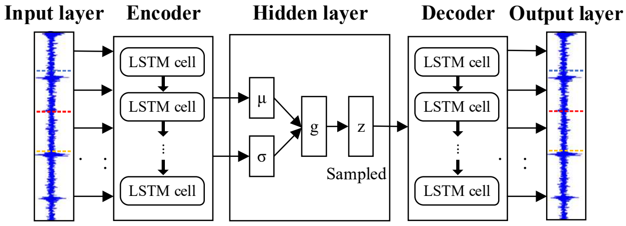

Fig. 3. LSTM-VAE hybrid model structure

probability distributions learned in the hidden space are continuous, VAE is robust to noise caused by operating conditions, equipment differences, and other factors.

The structure of the LSTM-VAE hybrid model is shown in Fig. 3, which mainly consists of an input layer, an encoder, a hidden layer, a decoder, and an output layer. Among them, the input layer is responsible for the segmentation of the received signal. The encoder is composed of multiple LSTM cells and takes a 3D sequence as input. Like all encoders in the VAE architecture, it outputs a 2D sequence for approximating the mean μ and standard deviation σ of the potential distribution. The hidden layer is responsible for sampling from the 2D potential distribution and outputting the sampled compressed features to the decoder. The decoder decodes the compressed features to generate a 3D sequence. Finally, the output layer concatenates the generated sequence back to reconstruct the original time series.

B. *Adaptive threshold method and "3/5" strategy*

LSTM-VAE is still essentially a codec, which encodes and decodes the input data, and then outputs the reconstructed result. In this paper, Mean Square Error (MSE) is used as a measure of the difference between the input data and the reconstructed output results, and its calculation formula is as follows. When the MSE is greater than the warning threshold, the machinery can be considered to be malfunctioning and an alarm is triggered.

$$\mathrm{MSE}(Y,\hat{Y}) = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}^i - y^i)^2 \tag{1}$$

At present, the warning thresholds for most fault diagnosis methods are pre-set into the system by the operator. However, in the actual industrial environment, even for the same type of machinery, the warning thresholds may vary depending on the operating conditions, materials, manufacturing, or processing technology level, so an empirically given warning threshold cannot be applied to all machines. Given the excellent generative capability of the LSTM-VAE hybrid model, the warning thresholds can be self-learning by partial normal state data of the machinery as well as newly generated normal state data. More data means that the learned warning thresholds are more accurate and reliable, which enables the system to alarm at the early stage of fault occurrence with strong self-adaptability. The calculation process of the early warning threshold is shown in Fig. 4. Firstly, the trained LSTM-VAE is used to generate more healthy state data X. The new data X is then input to the LSTM-VAE, and the reconstructed data X' is obtained after LSTM-VAE encoding and decoding. The MSE between X and X' is calculated and noted as loss2. Finally, the mean μ and standard deviation σ are calculated after the combination of loss1 and loss2, where loss1 is the MSE generated by the trained LSTM. According to the 3σ criterion, since the probability of a normal distribution taking values outside $(\mu - 3\sigma, \mu + 3\sigma]$ is less than 0.3%, which is

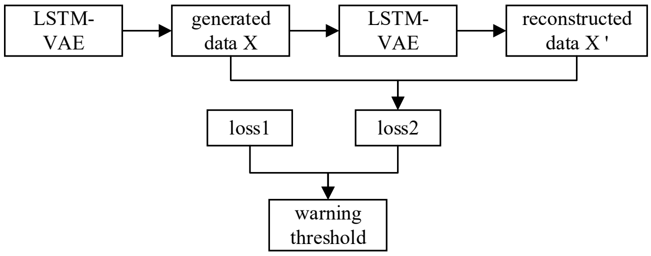

Fig. 4. Calculation of warning threshold

a small probability event, the warning threshold can be set to $\mu + 3\sigma$.

Real industrial systems are not ideal models. They are susceptible to strong noise or variable operating conditions. And there is a high probability that a disturbance will cause the vibration signal to jitter violently and return to smoothness quickly, resulting in false alarms in normal systems. To reduce the occurrence of such false alarms, this paper proposes a "3/5" strategy based on the idea of sliding windows. The specific idea of the "3/5" strategy is to set up a sliding window of size 5, and the system will store the results of each diagnosis in the window and check the situation in the window. If there are 3 faults in the 5 diagnostic results in the window, the machinery is considered to be truly faulty and an alarm is triggered immediately. Otherwise, remove the diagnostic result in the first window and store the next diagnostic result in the last window.

C. *Dropout method*

The LSTM-VAE hybrid model is robust to noise though. However, when the model has more parameters, insufficient training data, and more noise, the model is prone to overfitting. In this paper, the Dropout method proposed by Hinton [21] et al. was introduced to augment the LSTM-VAE to prevent overfitting of the model during the training process. The Dropout method is based on the principle that a certain percentage of neurons are randomly deactivated in each training batch, thus reducing the interaction between neurons in each layer and enhancing the generalization of the model.

D. *Cloud-Edge collaborative intelligent fault diagnosis method*

The specific implementation of the cloud-edge collaborative intelligent fault diagnosis method is: firstly, the sensing data of the machinery is monitored at the edge and uploaded to the cloud for storage and analysis, and then the AI fault diagnosis model is trained in the cloud and downlinked to the edge for machinery fault diagnosis, thus realizing the collaboration between the cloud and the edge. In this paper, the proposed LSTM-VAE hybrid model is the AI fault diagnosis model.

As shown in Fig. 5, the overall architecture of proposed method is divided into two phases: offline training and online testing. The offline training phase is conducted in the cloud, mainly to train the LSTM-VAE hybrid model and calculate the warning threshold. And then migrate the trained LSTM-VAE and warning threshold to the edge devices. The online testing phase is performed at the edge, and its main task is to use the trained LSTM-VAE hybrid model to encode and decode the monitored real-time vibration signals, output the reconstructed vibration signals, calculate the losses of the original vibration signals and the reconstructed vibration signals, compare them with the warning thresholds, and

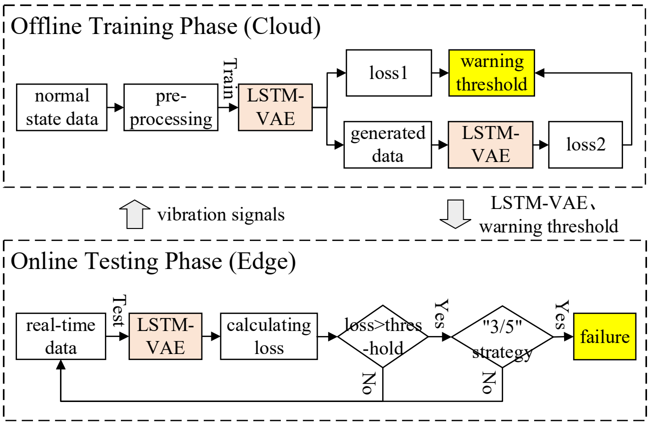

Fig. 5. Overall framework of the proposed method

combine them with the "3/5" strategy to achieve an online diagnosis of machinery.

1) Offline Training Phase (Cloud)

Throughout the life cycle of machinery, without a fault diagnosis system, we do not know when it is potentially failing, but it is considered normal for some time after it has started to be used. And for some reason, it fails later in its use. Therefore, in general, we can obtain some of the normal state data at the start of the machine operation.

First, we do some pre-processing on these normal state data. For example, the number of training samples is increased by data augmentation. The convergence speed and accuracy of the model are improved by normalizing the data, and the gradient explosion can be prevented to some extent. The pre-processed data are then fed into the LSTM-VAE hybrid model so that the LSTM-VAE can learn the probability distribution of potential features of these normal state data. When training the LSTM-VAE, the LSTM-VAE will reconstruct the original normal state data. And through calculation, we can get the loss between the reconstructed and the original data. And this loss is recorded as loss1. After that, the warning threshold is calculated by combining the abovementioned adaptive thresholding method. Finally, the trained LSTM-VAE and the warning threshold are downlinked to the edge side for fault diagnosis of the monitored machinery.

2) Online Testing Phase (Edge)

In the online testing phase, use Algorithm 1 to determine whether the current mechanical component is faulty. The vibration signal monitored in real-time is fed into the trained LSTM-VAE hybrid model. LSTM-VAE will reconstruct this vibration signal and calculate the reconstructed loss value. If this loss is less than or equal to the warning threshold, the machinery is not faulty and does not need any treatment and continues to monitor the machinery; if this loss is greater than the warning threshold, the point may be a potential failure point. Combine this with the abovementioned "3/5" strategy. If the "3/5" strategy is not met, it means that the machinery is not faulty. At this point, the system continues to monitor the machinery and no additional processing is required. If the "3/5" strategy is met, the machinery does fail, and the system immediately issues a control command to stop the equipment and notify maintenance personnel of the failure. Since the time-consuming process of training the LSTM-VAE hybrid model is already done in the cloud, the online diagnosis of Algorithm 1 at the edge is efficient and can meet the real-time performance.

TABLE I. ALGORITHMS 1

Determining whether the current mechanical component is faulty
1: Input: real-time vibration signals of mechanical components $\{x_1, x_2, \ldots, x_n\}$
2: Output: true or false
3: calculate the loss of vibration signals $\{x_1, x_2, \ldots, x_n\}$ by the trained LSTM-VAE hybrid model
4: if the loss is greater than the warning threshold:
5: if meet the "3/5" strategy:
6: return true
7: end if
8: end if
9: return false

III. EXPERIMENT

A. Data Description

The experimental data come from the publicly available PRONOSTIA [22] platform, which takes only a few hours to complete the bearing degradation. During the test, the bearing speed was set to 1800 r/min. And a radial force of 4kN was applied to the test bearing to accelerate the bearing degradation. The sampling frequency is 25.6 kHz, and each sample contains 2560 data points, i.e., 0.1 s, and the sampling is repeated every 10 s. And the test was stopped when the vibration amplitude of the bearing exceeded 20g.

B. Data Preprocessing

To have more samples for training and testing algorithmic models, when dealing with vibration signals, they are often sliced into small segments of the same length, and then the newly sliced segments are used as the training and testing sets for the model. The process of slicing the vibration signal into multiple segments is shown in Fig. 6. As the sampling frequency of the bearings is 25.6kHZ, a small segment is cut into 256 sizes to facilitate the processing of the data, i.e. one sample contains 256 data points.

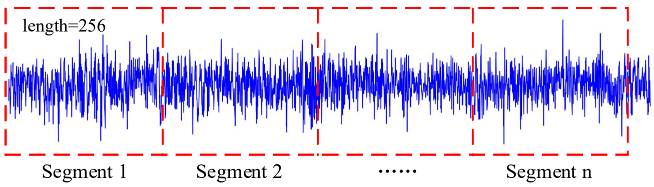

Fig. 6. Signal segmentation

C. Selection of training set and test set

The degradation process of rolling bearings varies even under the same operating conditions, but most of them follow a slow degradation mode and a fast degradation mode [23]. Two bearings were randomly selected from the total data set for the analysis and their vibration signals throughout their life cycle are shown in Fig. 7. The root mean squared (RMS) values were extracted from the vibration signals and are shown in Fig. 8. RMS will increase with the degree of bearing failure and can reflect the overall bearing degradation trend [17]. As can be seen from Fig. 8, the RMS of bearing 1 is gradually increasing, which indicates that it is a slow degradation mode. For machinery in the slow degradation mode, most fault diagnosis methods suffer from untimely warnings, which will lead to a slow increase in the degree of failure, thus rendering the machinery inoperable. In contrast, the RMS of bearing 2 is rapidly increasing towards the end of the bearing life, which indicates a rapid degradation mode. For

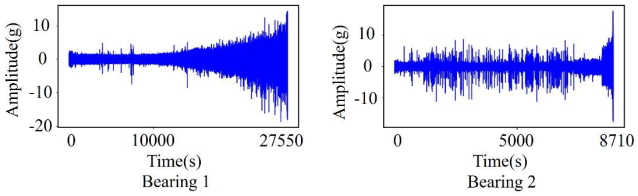

Fig. 7. Vibration signal of bearing

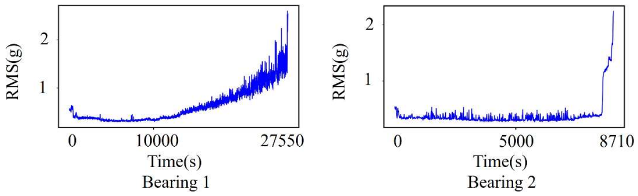

Fig. 8. RMS of bearing

machines in the rapid degradation model, accurate and timely warning must be required, otherwise, the machinery will not work properly in the lightest case, or even a safety accident will occur.

Training LSTM-VAE requires some bearing data in a normal state, and the RMS of bearing 1 and bearing 2 did not increase at the beginning of the run, which indicates that they were in a healthy state for a while when they started to be used. For safety reasons, the experiment selected a small number of data sets with smoother RMS from bearing 1 and bearing 2 as the training set, and the rest as the test set. The results of the delineation are shown in TABLE II.

TABLE II. TRAINING SET AND TEST SET DIVISION

Dataset / Bearings	Total dataset	Training set	Testing set
Bearing 1	0~27550s	0~10000s	10001~27550s
Bearing 2	0~8710s	0~5000s	5001~8710s

D. *Build LSTM-VAE hybrid model*

The structure of the LSTM-VAE hybrid model has been explained in detail in Fig. 3. The input data of the model has three dimensions. The first dimension is the number of training samples n; the second dimension is the time step, i.e., the size of the segment of the cut, which is 256 from above; the third dimension is the default value, because only the vibration signal is used as a feature, and a vibration signal is a one-dimensional form of temporal data, so the size of the third dimension is 1. Therefore, the dimension of the input data is n × 256 × 1. The encoder of LSTM-VAE is an LSTM network with 32 LSTM memory units, its input size is n × 256 × 1, and its output size is n × 32. The hidden layer part of LSTM-VAE is a neural network with 16 ordinary neurons, its input size is fixed at n × 2, and its output size is n × 16. The hidden layer decodes the data by sampling randomly from the sampling layer and feeding the sampled data back to the decoder of LSTM-VAE. The decoder of LSTM-VAE is also an LSTM network with 32 LSTM memory units, its input size is n × 16, and its output size is n × 256 × 1. The proposed model is developed based on Python and implemented in the open repository Keras, and the ReLU function is used as the activation function. RMSProp algorithm (Root Mean Square Prop, RMSProp) is used as the optimizer. TABLE III. shows all the parameter settings of the LSTM-VAE hybrid model. Besides, All experiments were executed on a computer equipped with a 4 GB GPU (GTX 1050Ti).

TABLE III. ALL PARAMETERS OF THE LSTM-VAE HYBRID MODEL

Number	Name	Parameters of each layer	Other Parameters
1	Input layer	shape=[n,256,1]	epoch=50
2	Encoder	lstm cell=32	activation='ReLU'
3	Hidden layers	latent_dim=16	optimizer='RMSProp'
4	Decoders	lstm cell=32	batch_size=64
5	Output layer	shape=[n,256,1]	dropout=0.2

E. *Experimental comparison*

With the above analysis, the fault diagnosis was performed on bearing 1 and bearing 2 using the method proposed in this paper, and the results are marked with red dashed lines in Fig. 9. For comparison, another fault diagnosis method proposed by Ginart [24] et al. was also experimented on bearing 1 and bearing 2, and the results are marked with green dashed lines in Fig. 9. The experimental results showed that for bearing 2 in the fast degradation mode, both methods found its potential failure point at the 8280ths. For bearing 1 in the slow degradation mode, it is difficult to identify its potential failure point. Because the performance of bearing 1 is slowly decaying and the health index of the bearing oscillates between normal and abnormal. But this problem is solved in the proposed method. Because of the "3/5" strategy, the potential failure point of bearing 1 is identified at 14150s, while the comparison method is at 17330s, so the proposed method significantly advances the warning time of the machinery, effectively preventing further expansion of the failure and providing more reliable protection for the machinery.

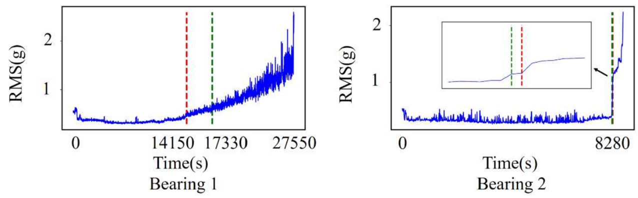

Fig. 9. Bearing fault diagnosis results

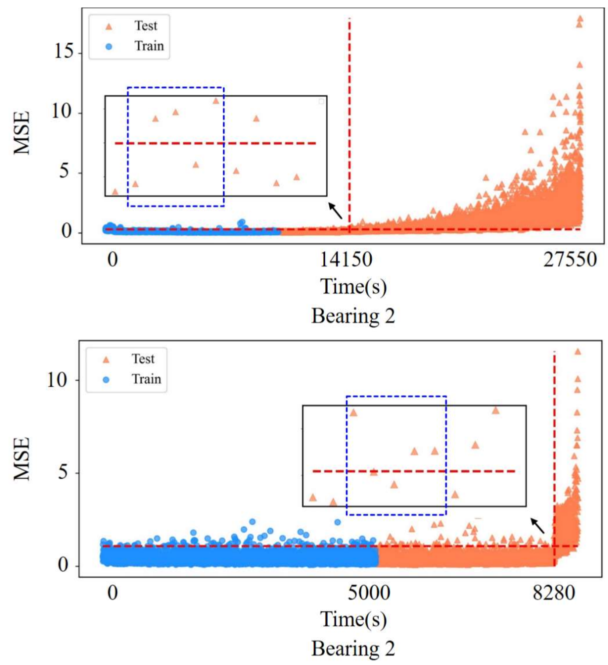

Fig. 10. RMS near the potential failure point of bearing

From Fig. 7, it is easy to find that the vibration signal of bearing 1 has occasional jitter at the beginning of the operation, and the vibration signal of bearing 2 is even more unstable and contains a lot of noise. The experimental results prove that the proposed method has good anti-interference capability and the "3/5" strategy improves the diagnostic accuracy to a certain extent. Fig. 10 shows some details of the MSE near the potential failure points of bearing 1 and bearing 2, where the vertical red dashed line identifies the location of the potential failure point and the horizontal red dashed line is the adaptive warning threshold. It can be found that the 5 diagnoses within the blue dashed box satisfy the "3/5" strategy. However, the previous diagnoses did not satisfy the "3/5" strategy due to thepresence of random noise in the vibration signal, so they were not identified as potential faults.

IV. CONCLUSIONS

In this paper, a cloud-edge collaborative intelligent fault diagnosis method based on the LSTM-VAE hybrid model is proposed for machinery, and the effectiveness of the method is proved through experiments. The use of this method can timely and accurately detect the abnormal state of the key components of the machinery when they are working so that the machinery can be produced safely within the controllable range. This is of great significance to ensure the safe and stable operation of the system and improve the production efficiency of the enterprise. However, in the actual industrial environment, there are still many limitations in mechanical fault diagnosis. For example, under the limited and dynamically changing network resources, computing resources, and storage resources, how to perform fault diagnosis on machinery to ensure the reliability and timeliness of diagnosis needs further research.

ACKNOWLEDGMENT

This work was supported in part by the Industrial Internet Innovation and Development Project of China under Grant TC19084DY, and in part by the Special Funds for Construction of Innovative Provinces in Hunan Province of China under Grant 2020SK2066 and Grant 2020GK2016.

REFERENCES

[1] W. Liang, J. Long, K.-C. Li, J. Xu, N. Ma, and X. Lei, "A Fast Defogging Image Recognition Algorithm Based on Bilateral Hybrid Filtering," ACM Trans. Multimedia Comput. Commun. Appl., vol. 17, no. 2, p. Article 42, 2021.

[2] T. Zhang and J. Dai, "Mechanical Fault Diagnosis Methods Based on Convolutional Neural Network: a Review," Journal of Physics: Conference Series, vol. 1750, no. 1: IOP Publishing, p. 012048, 2021.

[3] K. Javed, R. Gouriveau, and N. Zerhouni, "State of the art and taxonomy of prognostics approaches, trends of prognostics applications and open issues towards maturity at different technology readiness levels," Mechanical Systems and Signal Processing Signal Processing, vol. 94, pp. 214-236, 15 Sept. 2017.

[4] J. Poon, P. Jain, I. C. Konstantakopoulos, C. Spanos, S. K. Panda and S. R. Sanders, "Model-Based Fault Detection and Identification for Switching Power Converters," in IEEE Transactions on Power Electronics, vol. 32, no. 2, pp. 1419-1430, Feb. 2017.

[5] W. Liang, L. Xiao, K. Zhang, M. Tang, D. He, and K. C. Li, "Data Fusion Approach for Collaborative Anomaly Intrusion Detection in Blockchain-based Systems," in IEEE Internet of Things Journal, pp. 1-1, 2021.

[6] L. S. Dhamande and M. B. Chaudhari, "Bearing Fault Diagnosis Based on Statistical Feature Extraction in Time and Frequency Domain and Neural Network," International Journal of Vehicle Structures Systems, vol. 8, no. 4, p. 229, 2016.

[7] F. Jia, Y. G. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," Mechanical Systems and Signal Processing, vol. 72-73, pp. 303-315, 1 May. 2016.

[8] H. H. Pan, X. X. He, S. Tang, and F. M. Meng, "An Improved Bearing Fault Diagnosis Method using One-Dimensional CNN and LSTM," Strojniski Vestnik-Journal of Mechanical Engineering, vol. 64, no. 7-8, pp. 443-452, 2018.

[9] W. Dai, L. Qiu, A. Wu and M. Qiu, "Cloud Infrastructure Resource Allocation for Big Data Applications," in IEEE Transactions on Big Data, vol. 4, no. 3, pp. 313-324, 1 Sept. 2018.

[10] S. Lu, Q. Zeng, H. Wu, J. Su, X. Liu and H. Tang, "A New Power Load Forecasting Model (SIndRNN): Independently Recurrent Neural Network Based on Softmax Kernel Function," 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 2019, pp. 2305-2310.

[11] W. Liang, D. Zhang, X. Lei, M. Tang, K. Li, and A. Zomaya, "Circuit Copyright Blockchain: Blockchain-based Homomorphic Encryption for IP Circuit Protection," in IEEE Transactions on Emerging Topics in Computing, pp. 1-1, 2020.

[12] S. Lu, Q. Xia, X. Tang, X. Zhang, Y. Lu and J. She, "A Reliable Data Compression Scheme in Sensor-Cloud Systems Based on Edge Computing," in IEEE Access, vol. 9, pp. 49007-49015, 2021.

[13] M. Qiu, Z. Ming, J. Wang, L. T. Yang and Y. Xiang, "Enabling Cloud Computing in Emergency Management Systems," in IEEE Cloud Computing, vol. 1, no. 4, pp. 60-67, Nov. 2014.

[14] K. Gai, M. Qiu, H. Zhao and X. Sun, "Resource Management in Sustainable Cyber-Physical Systems Using Heterogeneous Cloud Computing," in IEEE Transactions on Sustainable Computing, vol. 3, no. 2, pp. 60-72, 1 April-June 2018.

[15] Y. Zhang, X. Wang, J. He, Y. Xu, F. Zhang and Y. Luo, "A Transfer Learning-Based High Impedance Fault Detection Method Under a Cloud-Edge Collaboration Framework," in IEEE Access, vol. 8, pp. 165099-165110, 2020.

[16] K. Gai and M. Qiu, "Reinforcement Learning-based Content-Centric Services in Mobile Sensing," in IEEE Network, vol. 32, no. 4, pp. 34-39, July/August 2018.

[17] Y. G. Lei, N. P. Li, L. Guo, N. B. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," Mechanical Systems and Signal Processing, vol. 104, pp. 799-834, 1 May. 2018.

[18] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-80, Nov 15 1997.

[19] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," arXiv preprint arXiv, 2013.

[20] P. P. Kharche and S. V. Kshirsagar, "Review of fault detection in rolling element bearing," International Journal of Innovative Research in Advanced Engineering, vol. 1, no. 5, pp. 169-174, 2014.

[21] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing coadaptation of feature detectors," arXiv preprint arXiv, 2012.

[22] P. Nectoux et al., "PRONOSTIA: An experimental platform for bearings accelerated degradation tests," in Proc. Int. Conf. Prognost. Health Manage., Denver, CO, USA, 2012, pp. 1-8.

[23] B. Yang, R. Liu and E. Zio, "Remaining Useful Life Prediction Based on a Double-Convolutional Neural Network Architecture," in IEEE Transactions on Industrial Electronics, vol. 66, no. 12, pp. 9521-9530, Dec. 2019.

[24] A. Ginart, I. Barlas, J. Goldin and J. L. Dorrity, "Automated Feature Selection for Embeddable Prognostic and Health Monitoring (PHM) Architectures," 2006 IEEE Autotestcon, 2006, pp. 195-201.