Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 107 (2022) 949–954

www.elsevier.com/locate/procedia

55th CIRP Conference on Manufacturing Systems

A novel model-independent data augmentation method for fault diagnosis in smart manufacturing

Pin Lyu[a,c], Hanbin Zhang[a], Wenbing Yu[b], Chao Liu[d,*]

[a]School of Electronic Information Engineering, Shanghai Dianji University, 300 Shuihua Road, Shanghai 201306, China
[b]School of Higher Vocational Technology, Shanghai Dianji University, 300 Shuihua Road, Shanghai 201306, China
[c]Department of Industrial and Systems Engineering, The Hong KongPolytechnic University, Hong Kong, China
[d] College of Engineering and Physical Sciences, Aston University, Birmingham B47ET, UK

* Corresponding author. Tel.: +852-9718-1997. E-mail address: c.liu16@aston.ac.uk

Abstract

With the rapid development of information technology, data-driven fault diagnosis has gained more and more attention because it provides a new way for enterprises to save costs. Considering that there are few abnormalities in equipment operation in actual industrial applications, it is still a challenge to implement data-driven fault diagnosis that requires a large amount of fault data. To tackle the challenge, this paper proposes a model-independent data augmentation method, which is a weighted combination of the two time series data augmentation methods, i.e. Gaussian noise and signal stretching. The experimental dataset is collected from an intelligent motor test platform. The fault diagnosis model based on support vector machine and feedforward neural network are applied to study the ability of the proposed data augmentation method in terms of model independence. Experimental results show that the proposed data augmentation methods can significantly improve the accuracy of fault diagnosis.

*Keywords:*data augmentation; fault diagnosis; time series; smart manufacturing

1. Introduction

The rapid development of the industrial Internet of Things, cyber-physical systems and big data continues to reshape modern manufacturing [1]. With the proliferation of data and computational intelligence, manufacturing processes are becoming more and more digital [2]. The new manufacturing model has increasingly higher requirements for equipment reliability and operational safety. Researches on data-driven fault diagnosis methods has been extensively developed [3,4]. This kind of data-driven method can use machine learning models, such as k-neighbor nearest [5,6], support vector machine (SVM) [7,8], feedforward neural network (FNN) [9,10], convolutional neural network (CNN) [11,12] and so on, to find out the complex relationship between the fault and its corresponding system response to determine the category of the fault by learning a large amount of fault data.

However, the frequency of faults in the actual intelligent manufacturing process is often very low, and the fault data obtained is far less than the normal data. If there is not sufficient fault data, the data-driven method may not be robust enough to achieve accurate diagnosis. To address this issue, many works of data augmentation have been done recently, of which the most representative methods can be roughly divided into three categories, digital twin-based methods, model-level methods, and data-level methods.

In general, digital twin-based methods try to produce a huge amount of training data by constructing the numerical model associated with a physical asset. Model-level methods attempt to design a new algorithm or an improved loss function, which

10.1016/j.procir.2022.05.090

biases the recognition process to the fault samples during the training stage. Data-level methods can be further divided into two categories. One aims to resampling fault samples or generating synthetic fault samples to augment the number of fault samples. The other is to use signal processing technologies to augment the whole dataset. Although these three methods have achieved certain performance for fault classification in case of lack of fault data, they are still some limitations to real-world applications. For example, it is very difficult to construct an accurate digital twin system based on physics-based models. The disadvantage of model-level methods is time-consuming to design a new algorithm and dependent different industrial domain of fault diagnosis. The deficiency of over-sampling fault data may insert a lot of noise samples, resulting in over generalization of the classification model. Most of existing methods of augmenting the time-series data by signal processing technologies are related to other research fields, of which is less discussed in fault diagnosis.

To overcome above-mentioned challenge, a model-independent time-series data augmentation method for fault diagnosis is proposed in this work to address the problem of insufficient fault data. Its core idea is to 1) use weighted the combination of Gaussian noise and signal stretching in a random conversion method to generate new samples with novel pattern features so that more representative fault features are extracted from the enhanced time series data; 2) Train a fault diagnosis model with the generated augmentation dataset.

The rest of this paper is organized as follows. Section 2 briefly reviews the related work of data augmentation. Section 3 introduces the proposed method in this paper. Section 4 provides the results of the case study. Finally, the conclusion this work is summarized along with further researches in the future.

2. Related work

2.1. Digital twin-based methods for data augmentation

With significant advancement in information technologies, digital twin has gained increasing attention as it offers an enabling tool to continuously produce data reflected physical asset. Hence, digital twin is regarded as a means of data augmentation in fault diagnosis. For example, aiming at the problem of lack of fault data, Xia et al. [13] presented a fault diagnosis framework based on digital twins and deep transfer learning. Using limited measurement data, accurate fault diagnosis of the machine under the change of working conditions is realized. Mohamed et al. [14] developed a digital twin model of ball bearing to overcome the problem of insufficient experimental datasets in the machine learning algorithm-based diagnostic of rotating machines. Considering traditional machine learning algorithms requires a balanced dataset, Guo et al. [15] used digital twin technology to simulate many balanced datasets to train an improved Random Forest model to address the problem of less fault data. Although these approaches have been developed to overcome the practical issue of limited measured data, it is very difficult to construct a high-fidelity digital twin model of the physical assets.

2.2. Model-level methods for data augmentation

There are two kinds of model-level methods including unsupervised and semi-supervised machine learning models. Zhang et al. [16] employed an unsupervised learning model based on a flow model to solve the problem of sparse fault samples. To use condition-assisted sample generation for data augmentation, Luo et al. [17] proposed a condition-based deep convolution generative adversarial network model. Li et al. [18] introduced a fault diagnosis method based on semi-supervised learning and used three steps to achieve data augmentation. To expand the training dataset, Peng et al. [19] employed a method of weighted training samples to reduce the impact of class imbalance in the training stage of Bidirectional GRU (Gated Recurrent Unit) model. He et al. [20] proposed a new spatio-temporal multiscale neural network, which uses focal loss function to solve the imbalance problem of SCADA data. A cost-sensitive classifier learning method is designed to solve the problem of unbalanced distribution of fault instances. The superiority of this method was validated with the rotating machinery dataset collected from the refinery [21]. The above solutions are feasible in case verification, and most of them applies algorithms to solve the class imbalance. However, the high computational costs caused makes model-level methods not conducive to practical industrial application.

2.3. Data-level methods for data augmentation

Resampling technologies for data augmentation in data-level methods refer to random fault sample over-sampling. Synthetic minority over-sampling technique (SMOTE) [22] is one of the most famous technologies. It is an improved scheme based on random over-sampling. The purpose is to increase the sensitivity of the classifier to fault samples. However, its limitations are mainly reflected in the blindness of sampling.

On the contrary, signal processing technologies in data-level methods is to add a small amount of noise to the time-series data to enlarge the whole dataset. Liu et al. [23] used fully CNN and ResNet to compare time-series data augmentation methods such as jitter, arrangement, scaling, and time warping. In addition, the data augmentation technology of time-series data and its application in each time-series classification of neural networks were investigated [24]. Arthur et al. [25] used the time slicing method to achieve data augmentation and merged the original data and the enhanced data to form a CNN training dataset. Their experimental results show that using the time slicing method to enhance the data can improve the overall classification performance of CNN. To deal with the unevenly data distribution in the activities performed by the equipment, a warping window time-series data augmentation technology was introduced [26]. Wen et al. [27] reviewed the importance of data enhancement for time-series data from the excellent performance of deep learning in many time-series analyses tasks, and systematically analyzed different data enhancement methods for time-series data. The above most of time-series data enhancement techniques are used in biological information, health care, transportation, engineering construction and other fields have been researched. However, it was not widely discussed in the field of fault diagnosis.

In addition, to improve classification accuracy, some researches integrated signal processing technologies in data-level methods with models. For example, Wu et al. [28] applied successively adding Gaussian noise, random scale, random stretching, and random cropping methods to augment the original data in the time-series data enhancement module of an agnostic framework. To improve the accuracy of fault diagnosis, this research mainly emphasizes how data enhancement module collaborates with other three modules to eliminate class imbalances. Therefore, their data augmentation method is still related to models. In another study, five kinds of data augmentation methods such as Gaussian noise, masking noise, signal translation, amplitude offset and signal stretching are respectively used to expand CWRU and IMS dataset. They evaluated the proposed deep learning model with the five enhanced data [29]. Although their data augmentation method improved the performance of the proposed deep learning model, they found that their data augmentation method has a clear relationship with the residual block used in the deep network, which indicates if the time-series data augmentation technology is used alone, its effect is also closely related to the complexity of the model.

To address above-mentioned limitations, this article weighted the combination of Gaussian noise and signal stretching in the random conversion method to generate new samples with novel pattern features so that more representative fault features are extracted from the enhanced time series data using either traditional machine learning models or complex deep learning models. The main contribution of this paper is the proposed data augmentation method is model-independent.

3. Methodology

Fig. 1 shows the flowchart of fault diagnosis with the proposed data augmentation method. In this flowchart, the vibration sensor is mounted on a motor to collect vibration data through an intelligent motor test platform. The original data is

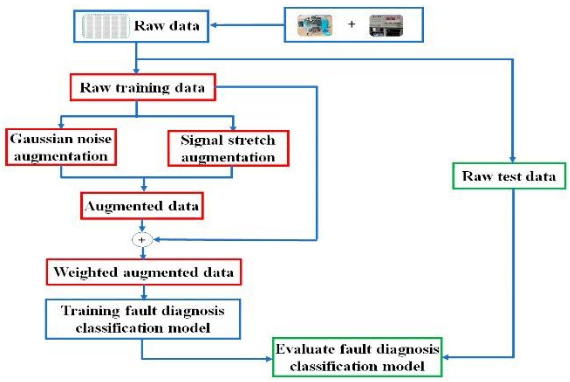

Fig. 1. The method for fault diagnosis with data augmentation

first divided into training set and test set. The original training data set is used for data augmentation, and the weighted combination of the original data set and the two augmentation training data is used to train the machine learning model. The trained model is evaluated with the original test data.

3.1. Model-independent data augmentation

For convenience, taking the vibration signal generated by the motor as an example, the structure of the original time-series data is defined as the following three rows and n columns matrix $\mathbf{M}$, where m_{a_i}, m_{h_i} and m_{v_i} are the vibration data in the axial, horizontal and vertical directions of the motor driving end respectively, i denotes the index of the data and n is the length of the data.

$$\mathbf{M}=\begin{bmatrix} m_{a_i} \\ m_{h_i} \\ m_{v_i} \end{bmatrix}=\begin{bmatrix} m_{a_1} & m_{a_2} & \cdots & m_{a_n} \\ m_{h_1} & m_{h_2} & \cdots & m_{h_n} \\ m_{v_1} & m_{v_2} & \cdots & m_{v_n} \end{bmatrix} \tag{1}$$

A new sample matrix $\mathbf{M}^{gu}$ is obtained by adding Gaussian noise $\varepsilon_i \sim N(\mu, \sigma^2)$ to each sample of the original time series-data $\mathbf{M}$.

$$\mathbf{M}^{gu}=\mathbf{M}+\varepsilon_i=\begin{bmatrix} m_{a_1}+\varepsilon_{a_1} & m_{a_2}+\varepsilon_{a_2} & \cdots & m_{a_n}+\varepsilon_{a_n} \\ m_{h_1}+\varepsilon_{h_1} & m_{h_2}+\varepsilon_{h_2} & \cdots & m_{h_n}+\varepsilon_{h_n} \\ m_{v_1}+\varepsilon_{v_1} & m_{v_2}+\varepsilon_{v_2} & \cdots & m_{v_n}+\varepsilon_{v_n} \end{bmatrix} \tag{2}$$

According to reference [30], when the standard deviation σ of Gaussian noise model is 0.15, it can ensure that the label of original time-series data remains unchanged after adding noise. Therefore, this paper sets the standard deviation σ to 0.15. The new data $\mathbf{M}^{gu}$ and the original data $\mathbf{M}$ are concatenated vertically to obtain Gaussian enhanced data $\mathbf{M}^{gu}_{new}$.

$$\mathbf{M}^{gu}_{new}=\begin{bmatrix} \mathbf{M} \\ \mathbf{M}^{gu} \end{bmatrix} \tag{3}$$

Signal stretching has the function of amplifying vibration data in the time domain and changing the speed or duration of time series signals. It can compensate for signal differences that may be generated due to the placement of sensors in different positions of the device. Signal stretching can maintain the overall shape of the time-series data, and at the same time can achieve a slight change in the sampling frequency. The length l of the stretched original time-series data is determined by the stretch factor $\alpha<1$. The stretch factor α is the ratio of the length of the original time-series data sample, defined as $l=n(1+\alpha)$, and the stretched original time-series data $\mathbf{M}'$ is as follows.

$$\mathbf{M}^{'}=\begin{bmatrix} m^{'}_{a_i} \\ m^{'}_{h_i} \\ m^{'}_{v_i} \end{bmatrix}=\begin{bmatrix} m^{'}_{a_1} & m^{'}_{a_2} & \cdots & m^{'}_{a_l} \\ m^{'}_{h_1} & m^{'}_{h_2} & \cdots & m^{'}_{h_l} \\ m^{'}_{v_1} & m^{'}_{v_2} & \cdots & m^{'}_{v_l} \end{bmatrix} \quad (4)$$

where $m^{'}_{a_i}$, $m^{'}_{h_i}$ and $m^{'}_{v_i}$ are the results of the three samples interpolation of the original time-series data $\mathbf{M}$. Then, the stretched data $\mathbf{M}'$ is shrunk to reduce its size to the size of the original actual sample length *n*. The newly generated stretch data $\mathbf{M}^{''}$ can be obtained by selecting *n* column data from a new starting point *t* in $\mathbf{M}'$ and $t=(l-n)/2$.

$$\mathbf{M}^{''}=\begin{bmatrix} m^{'}_{a_t} & m^{'}_{a_{t+1}} & \cdots & m^{'}_{a_{t+n}} \\ m^{'}_{h_t} & m^{'}_{h_{t+1}} & \cdots & m^{'}_{h_{t+n}} \\ m^{'}_{v_t} & m^{'}_{v_{t+1}} & \cdots & m^{'}_{v_{t+n}} \end{bmatrix} \quad (5)$$

The new sample $\mathbf{M}^{''}$ and the original data $\mathbf{M}$ are concatenated in the vertical direction to obtain the enhanced data $\mathbf{M}^{st}_{new}$ after stretching.

$$\mathbf{M}^{st}_{new}=\begin{bmatrix} \mathbf{M} \\ \mathbf{M}^{''} \end{bmatrix} \quad (6)$$

According to the literature [30], when the stretch factor α is set to 0.1, the original time-series data can be stretched while maintaining the label, so this paper sets the stretch factor α to 0.1. Finally, the parameters $\lambda \in [0, 1]$ are used to weight the data enhanced $\mathbf{M}^{gu}_{new}$ and $\mathbf{M}^{st}_{new}$ by Gaussian noise and signal stretching as follows and concatenated with the original data $\mathbf{M}$ in the vertical direction to obtain the final enhanced data $\mathbf{A}$.

$$\mathbf{A}=\begin{bmatrix} \lambda \mathbf{M}^{gu}_{new} \\ (1-\lambda)\,\mathbf{M}^{st}_{new} \\ \mathbf{M} \end{bmatrix} \quad (7)$$

According to formula (7), the data volume of the enhanced dataset $\mathbf{A}$ is 5 times of the original data volume. The parameter λ adjusts both the original time-series data in $\mathbf{M}^{gu}_{new}$ and original time-series data in $\mathbf{M}^{st}_{new}$. In addition, the ratio of the added noise level and the stretching amplitude is further controlled by parameter λ. Therefore, the enhanced randomness of the time-series signal expressed by the weight parameter can effectively reflect the non-stationarity of the fault motor vibration signal under different working conditions.

3.2. Data Preparation

The data preparation step of the proposed method includes data segmentation and feature extraction from collected data. A single data point of the motor vibration signal cannot provide information on whether the motor is faulty because it only represents the instantaneous state of the bearing, like the snapshot image. In contrast, motor vibration consists of continuous vibration distributed over a slice of time, like an image or video sequence. Therefore, the data stream containing a single data point needs to be divided into data windows, that is, continuous time-series points.

This paper considers a 0.1 second fixed length data window with 50% overlap between adjacent windows. After the time-series data is divided into windows, a set of time-domain statistical features are extracted from each window. These features represent the signal pattern in the corresponding window and are finally used as the input of classification algorithm. In this paper, nine features are extracted from each window, which are maximum, minimum, mean value, standard deviation, root mean square value, skewness, kurtosis, crest factor, and form factor. These features are used as input to train the classification model.

3.3. Training and evaluation of classification

Since the supervised learning model can provide better fault classification performance, FNN and SVM are considered for fault diagnosis with augmented data. The FNN has 1 hidden layer with a total of 1000 neurons. The dropout is set to 0.5, and the softmax classifier is used. SVM uses radial basis kernel function and uses grid search to find the optimal parameters C and gmma. Their values are 50 and 0.05, respectively. After each model is trained, it is evaluated with test data. The test data is obtained from the original dataset before data augmentation. The performance of the model is evaluated by *accuracy*, *precision*, *recall* and *F1-score*. A confusion matrix is also used to analyze the effect of the model on each fault category.

4. Case study and results

In this case, three vibration sensors are mounted on a three-phase asynchronous normal motor with a power of 7.5KW, a rated voltage of 380V, a speed of 2900r/min, 0 load, and a connection method of Δ. The vibration signal of the motor is collected in three channels, that is, the axial, horizontal, and the vertical of the drive end of the motor. The actual sampling frequency is 10HKz. The acquisition time is about 104 seconds. Each channel collects 1,048,576 collection points. Taking 1024 collection points as a sample fragment. So, 1024 normal samples can be obtained for each channel, and a total of 3072 normal samples can be obtained. Then, replace the normal motor with the same type but with faulty bearings (inner race fault and outer race fault) for data collection. In order to simulate the case of a small number of faulty samples, the collection time is about 35 seconds, and the three channels are used to obtain the data. The number of fault samples is 1024. The normal sample data and the fault sample data are mixed to obtain the original dataset, which is a matrix with 4096 rows and 1024 columns.

The original data is divided into training set and test set at a ratio of 1:1. The original training data is augmented using the proposed method, and 5-fold enhanced training data is generated (2-fold Gaussian enhancement, 2-fold stretching enhancement, 1-fold the original data). Table 1 summarizes the number of samples in the original dataset and enhanced dataset in the case. It can be seen from Table 1 that when the data enhancement method proposed in this paper is adopted, the sample data of each category in the enhanced dataset has reached a balanced state. For example, the number of outer race faults in the data collection stage is 586 samples, 294 are used for training, 294 are used for testing, and a total of 1470 training samples are obtained after 4-fold enhancement. Next, use 5-fold training data to train the FNN and SVM.

Table 1. Number of samples of raw and augmented data.

Health condition of bearing	Raw data	Raw testing data	Raw training data	Augmented training data	Combined training data(5-fold)
normal	3027	1514	**1514**		
Outer race fault	586	294	294	1176	**1470**
Inner race fault	438	219	219	876	**1095**

We first observe the changes of the accuracy of the fault diagnosis model after the FNN and SVM are trained with the combined training data by changing different weight parameters λ. Fig. 2 shows the changes of the diagnosis accuracy of the two fault diagnosis models with different weights. The fault diagnosis accuracy of SVM and FNN respectively reached 93.45% and 98.73% on the enhanced data when λ is equal to 0.007.

Next, to verify the effect of data augmentation proposed in this paper, we compare the accuracy, precision, recall and F1-score of the model on the raw training data and combined training data corresponding to the optimal λ, as shown in Fig.3 and Fig. 4 From the two figures, we can see that compared with only using the raw training data, the combined training data can significantly improve the performance of the fault diagnosis model. For example, the accuracy, precision, recall and F1-score of SVM range from 67.54% to 93.45%, 78.56% to 91.92%, 65.43% to 92.32% and 71.39% to 92.10% respectively. From the performance coordinates, we can also observe that the improvement of SVM is higher than that of FNN. This shows that the data augmentation method proposed in this paper has nothing to do with the complexity of the model.

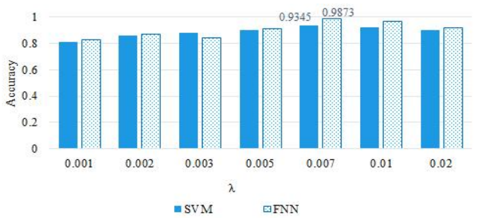

Fig. 2. Accuracy of model on augmentation training data with different

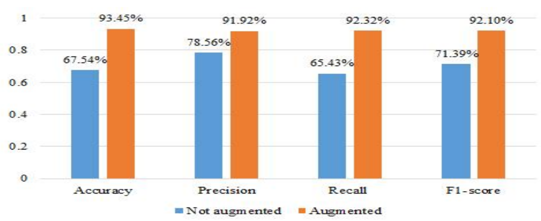

Fig. 3. Performance measures of SVM

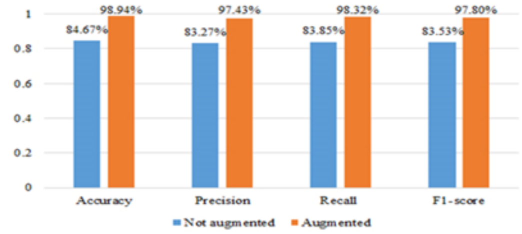

Fig. 4. Performance measures of FNN

Fig. 5 and Fig. 6 are the confusion matrix of the SVM model on the test set. The confusion matrix explores the classification of the trained models in different categories on the test set before and after data augmentation. As can be seen from Fig. 5, the classification result error rate of the model trained on the original training data is high on the test set. In addition, SVM can easily misclassify the inner race fault as a health condition. On the other hand, the errors in the test set of the model trained on the combined training data are significantly reduced, and the three different health conditions of the bearing can be identified with 94.5%, 93.6% and 92.8% respectively.

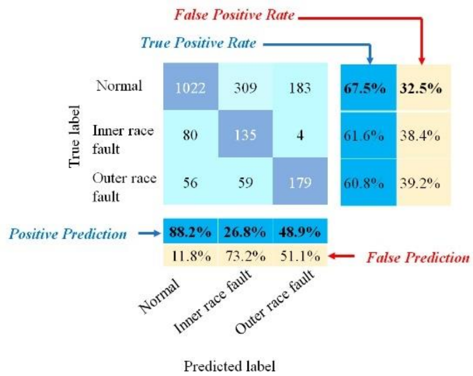

Fig. 5. SVM with no augmentation

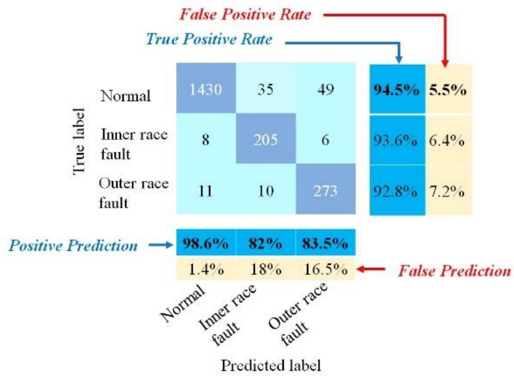

Fig. 6. SVM with augmentation

Finally, we compared the proposed methods with the Gaussian noise, masking noise, signal translation, amplitude offset and signal stretching used in literature [29] based on FNN. The experimental results are shown in Table 2. From Table 2, it can be observed that the proposed method is superior to other methods. Among the individual noise augmentation method, signal stretching and Gaussian noise are best. This indicates the weighted combination of Gaussian noise and signal stretching is a promising method in real-world industrial application.

Table 2. Performance comparison of different augmentation methods

Methods	Accuracy
Proposed method	98.84%
Gaussian noise	84.98%
Masking noise	83.56%
Signal translation	74.85%
Amplitude offset	83.23%
Signal stretching	85.76%

5. Conclusion and future work

A model-independent time-series data augmentation method is proposed to expand the fault samples in the training set to address the problem of fault data in actual industrial scenarios. The specific implementation technology adopts the weight combination of Gaussian noise and signal stretching. The proposed method is verified on a bearing dataset collected from our developed intelligent motor test platform. The experimental results show that the performance of fault diagnosis model is significantly improved after using the proposed augmentation method. This shows that the proposed data augmentation method has a certain potential to solve the problem of lack of fault samples in industrial application scenarios. Our future research will further explore the effect of fault diagnosis model based on deep learning on proposed data augmentation method and the application of other time series data augmentation technologies in the field of fault diagnosis.

Acknowledgements

This research work was partially supported by the National Natural Science Foundation of China (No. 52105534) and Shanghai Science and technology program (Project No. 22010500900).

References

[1] Chao Liu, Pai Zheng, Xun Xu. Digitalisation and servitisation of machine tools in the era of Industry 4.0: a review. International Journal of Production Research, 2021,DOI: 10.1080/00207543.2021.1969462.

[2] Jinjiang Wang, Lunkuan Ye, Robert X. Gao, Chen Li & Laibin Zhang. Digital Twin for rotating machinery fault diagnosis in smart manufacturing. International Journal of Production Research, 2018,57(12):3920-3934.

[3] Jinjiang Wang, Yulin Ma, Laibin Zhang, Robert X. Gao, Dazhong Wu. Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems, 2018,48:144–156.

[4] Duy-Tang Hoang, Hee-Jun Kang. A survey on Deep Learning based bearing fault diagnosis. Neurocomputing,2019, 335:327–335.

[5] A. Dasgupta, S. Debnath, A. Das. Transmission line fault detection and classification using cross-correlation and k-nearest neighbor, International Journal Knowledge-Based Intelligent Engineering System, 2015,19 (3):183–189.

[6] Z. Zhou, C. Wen, C. Yang. Fault isolation based on k-nearest neighbor rule for industrial processes, IEEE Trans. Ind. Electron.2016, 63 (4): 2578–2586.

[7] X. Zhang, B. Wang, X. Chen. Intelligent fault diagnosis of roller bearings with multivariable ensemble-based incremental support vector machine, Knowledge-Based System, 2015,89: 56–85.

[8] J. Zheng, H. Pan, J. Cheng. Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines, Mechanical System Signal Process.2017, 85:746–759.

[9] M.-Y. Cho, T.-F. Lee, S.-W. Gau, C.-N. Shih. Power transformer fault diagnosis using support vector machines and artificial neural networks with clonal selection algorithms optimization. in: International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, Springer, 2006,179–186.

[10] M. Kordestani, M.F. Samadi, M. Saif, K. Khorasani. A new fault diagnosis of multifunctional spoiler system using integrated artificial neural network and discrete wavelet transform methods, IEEE Sensors Journal, 2018,18 (12): 4990–5001.

[11] Wen, Long; Li, Xinyu; Gao, Liang; Zhang, Yuyan. A new convolutional neural network-based data-driven fault diagnosis method. IEEE transactions on industrial electronics,2018, 65 (7), 5990-5998.

[12] T. Ince, S. Kiranyaz, L. Eren, M. Askar and M. Gabbouj. Real-time motor fault detection by 1-D convolutional neural networks. IEEE Transactions on Industrial Electronics,2016, 63(11),7067-7075.

[13] Min Xia, Haidong Shao, Darren Williams, Siliang Lu, Lei Shu, Clarence W. de Silva. Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning. Reliability Engineering and System Safety, 2021,215:107938.

[14] Mohamed Habib Farhat, Xavier Chiementin, Fakher Chaari, Fabrice Bolaers. Digital twin-driven machine learning:ball bearings fault severity classification, Measurement Science and Technology, 2021,32,044006.

[15] Kai Guo, Xiang Wan, Lilan Liu, Zenggui Gao and Muchen Yang. Fault Diagnosis of Intelligent Production Line Based on Digital Twin and Improved Random Forest. Applied Science,2021,11,7733.

[16] Liangwei Zhang, Jing Lin, Haidong Shao, Zhicong Zhang, Xiaohui Yan, Jianyu Long. End-to-end unsupervised fault detection using a flow-based model. Reliability Engineering and System Safety, 2021,215: 107805.

[17] Jia Luo, Jinying Huang, Hongmei Li. A case study of conditional deep convolutional generative adversarial networks in machine fault diagnosis. Journal of Intelligent Manufacturing, 2021,32:407–425.

[18] Xiang Li, Xu Li, Hui Ma. Deep representation clustering-based fault diagnosis method with unsupervised data applied to rotating machinery. Mechanical Systems and Signal Processing,2020, 143:106825.

[19] Peng Peng, Wenjia Zhang, Yi Zhang, Yanyan Xu, Hongwei Wang, Heming Zhang. Cost sensitive active learning using bidirectional gated recurrent neural networks for imbalanced fault diagnosis. Neurocomputing,2020,407: 232–245.

[20] Qun He, Yanhua Pang, Guoqian Jiang, Ping Xie. A Spatio-Temporal Multiscale Neural Network Approach for Wind Turbine Fault Diagnosis With Imbalanced SCADA Data. IEEE transactions on industrial informatics, 2021,17(10):6875-6884.

[21] Qifa Xu, Shixiang Lu, Weiyin Jia, Cuixia Jiang. Imbalanced fault diagnosis of rotating machinery via multi-domain feature extraction and cost-sensitive learning. Journal of Intelligent Manufacturing,2020,31:1467–1481.

[22] N. v. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 2002,16,321-357.

[23] Liu B, Zhang Z, Cui R. Efficient Time Series Augmentation Methods. In: CISP-BMEI,2020.

[24] Brian Kenji Iwana, Seiichi Uchida. An empirical survey of data augmentation for time series classification with neural networks. PLoS ONE 16(7): e0254841. https://doi.org/10.1371/journal.pone.0254841,2021.

[25] Arthur Le Guennec, Simon Malinowski, Romain Tavenard. Data Augmentation for Time Series Classification using Convolutional Neural Networks. ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, Riva Del Garda, Italy. halshs-01357973,2016.

[26] K.M. Rashida, J. Louisa. Window-Warping: A Time Series Data Augmentation of IMU Data for Construction Equipment Activity Identification. Advanced Engineering Informatics, 2019,42: 100944.

[27] Wen Q, Sun L, Song X, Gao J, Wang X, Xu H. Time series data augmentation for deep learning: A survey.arXiv preprint arXiv:200212478. 2020.

[28] Jingyao Wu, Zhibin Zhao, Chuang Sun, Ruqiang Yan, Xuefeng Chen. Reliability Engineering and System Safety, 2021,216:107934.

[29] Xiang Li, Wei Zhang, Qian Ding, JianQiao Sun. Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation. Journal of Intelligent Manufacturing, 2020,31:433–452.

[30] Mooseop Kim, Chi Yoon Jeong. Label-preserving data augmentation for mobile sensor data. Multidimensional Systems and Signal Processing, 2021, 32:115–129.