

Research on fault diagnosis method of bearing based on image transformation of time series

Chuanchao Xu
Institute of Engineering
Thermophysics & University of
Chinese Academy of Sciences,
CAS
xuchuanhao@iet.cn
Beijing, China

Yue Liu
Institute of Engineering
Thermophysics & University of
Chinese Academy of Sciences,
CAS
liuyue16@foxmail.com
Beijing, China

Jinglun Fu
Institute of Engineering
Thermophysics & University of
Chinese Academy of Sciences,
CAS
fujinglun@mail.etp.ac.cn
Beijing, China

ABSTRACT

There are many fault diagnosis methods by artificial intelligence. Most of them are similar: build multi-dimensional feature space by multiple signal in a single moment, and then diagnosing the faults by the classical machine learning method. These methods have achieved good results, but they do not consider the continuity of time. The time series of signal is an important basis which is reflecting equipment faults. For this problem, we transform the different time series of one fault signal to image set by GAF method, and then classify different classes from image set by CNN method. Random Forest, Support Vector Machine and Artificial Neural Network are applied as the contrast. Experimental results show that: "GAF + CNN" method have an advantage over contrast methods in the same bearing vibration signal dataset.

INTRODUCTION

As a high-speed rotating machine, the gas turbine needs to complete high-strength, high-precision and high-complexity work, so perfect condition monitoring, Fault diagnosis and protection isolation systems must be configured. Bearing safety is an important guarantee for gas turbine operation, and the original vibration signal generated when the rolling bearing fails contains abundant fault information, which is conducive to the application of fault diagnosis technology. More and more attention has been paid to the fault diagnosis technology of bearing as an important means to maintain the safe and stable operation of gas turbine (Gao L et al. 2011), which has been paid more and more attention. The conventional fault diagnosis methods include model-based method, qualitative experience-based method and data-driven method. The traditional fault diagnosis method has been unable to meet the requirements of people because of the limitations of the physical model, and a large number of operation data of the equipment has

been gradually used for fault diagnosis because it can well reflect the state and mechanism of the operation of the equipment. Data-driven methods have gradually been the development direction of fault diagnosis. At present, data-driven methods mainly include statistical method, signal process and artificial intelligence-based method (Wang H et al. 2009) Among them, statistical method is limited by the statistical theory, and its scope of application is small. Signal processing technology is an important method of bearing fault analysis, which is represented by Fourier transform (J. Rosero et al. 2007), wavelet packet analysis (Gao L et al. 2011) Hilbert-Huang transform (Huang N E et al. 1998) and other methods. Signal processing method based on signal analysis technology extracts time-frequency domain characteristic parameters and characterizes the state of the system. This process will still have the information loss of data (Jasper J et al. 2010). The fault diagnosis method based on artificial intelligence technology does not need a specific mathematical model. Based on sufficient historical data, the fault diagnosis can be realized by means of computer learning, reasoning and decision-making. Among many fault diagnosis methods based on artificial intelligence, the typical representative methods are artificial neural network (ANN), support vector machine (SVM) and fuzzy logic. Using artificial neural network for the fault diagnosis, its principle is to establish neural network to realize the mapping of fault data and fault categories, and then to train and correct parameters of neural network based on the dataset with fault mark. Finally, a neural network which can accurately determine the fault type of fault data is obtained (Park DC et al. 1991) SVM overcomes the difficulty of limited data, so it is suitable for fault diagnosis in the case of small sample. The basic idea is to map the original linear and inseparable data to high dimensional space by using the kernel function. The hyperplane is constructed in the high dimensional space to classify the data (Vapnik V N, 1999), so that the fault

classification can be realized. Fuzzy logic was first put forward by Zadeh. The principle of fuzzy logic applied to fault diagnosis is expounded in detail, which is to fuzzify the collected data (Isermann R et al. 1998). Based on the fuzzy rule base, the probability of failure caused by these inputs is deduced. In the fault diagnosis method based on artificial neural network and support vector machine, each fault data is regarded as a single-time multivariable representation of the state of the system when using the data. Without considering the continuity of a single variable on the time scale, the time series is lost, which is an important feature of equipment failure. Using single variable time series to determine the fault type can be solved by statistical method and signal processing method, but both methods have their own shortcomings. Therefore, it is necessary to explore a fault diagnosis method based on time series and overcome the shortcomings of the above two methods.

1. RESEARCH METHOD

Research methods of this article mainly use the GAF theory and convolution neural network, through the GAF theory bearing fault signal can be the time series of transformation for the characterization of the fault condition of pictures, again by convolution neural network powerful ability to recognize images, classifying the generated images. This is an indirect classifying bearing fault signal.

1.1 GAF theory

To solve above problems, GAF (Gramian Angular Field) which can transform the time series to the image was got from reference (Wang Z et al. 2015). The coordinates of each pixel in the transformed image represent the relationship between the values of the corresponding time series variables. The transformation process is shown in figure 1.

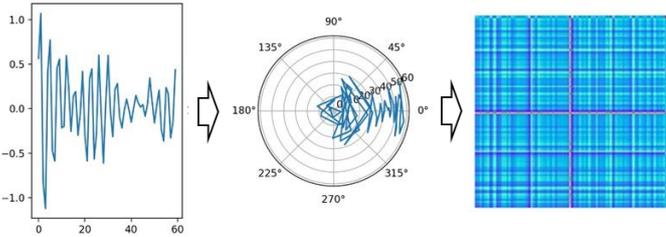


Figure 1 Series2image process by GAF theory

Given a time series $X = \{x_1, x_2, x_3, \dots, x_n\}$ of n real-valued observations, X should be rescaled so that all values fall in the interval $[-1, 1]$:

$$\tilde{x}_i = \frac{(x_i - \max(X)) + (x_i - \min(X))}{\max(X) - \min(X)} \quad (1)$$

Thus we can represent the rescaled time series \tilde{X} in polar coordinates by encoding the value as the angular cosine and time stamp as the radius with the equation below:

$$\begin{cases} \phi = \arccos(\tilde{x}_i), -1 \leq \tilde{x}_i \leq 1, \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, t_i \in \mathbb{N} \end{cases} \quad (2)$$

t_i is the time stamp and N is a constant for regularizing the span of the polar coordinate system. The transformation of time series from the Cartesian coordinate system to the polar coordinates system increases the way to understand the time series. As time increases, the corresponding values will take on a special shape in polar coordinate. The encoding map of equation 2 has two important properties. First, it is bijective as $\cos(\phi)$ is monotonic when $\phi \in [0, \pi]$. Given a time series, the proposed map produces one and only one result in the polar coordinate system with a unique inverse function. Second, as opposed to Cartesian coordinates, polar coordinates preserve absolute temporal relations.

After transforming the rescaled time series into the polar coordinate system and considering the trigonometric sum between each point, the temporal correlation within different time intervals can be identified in the angular perspective. The GAF is defined as follows:

$$G = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \cdots & \cos(\phi_1 + \phi_n) \\ \cos(\phi_2 + \phi_1) & \cdots & \cos(\phi_2 + \phi_n) \\ \vdots & \ddots & \vdots \\ \cos(\phi_n + \phi_1) & \cdots & \cos(\phi_n + \phi_n) \end{bmatrix} \quad (3)$$

$$= \tilde{X}' \cdot \tilde{X} - \sqrt{I - \tilde{X}'^2} \cdot \sqrt{I - \tilde{X}^2}$$

I is the unit row vector $[;1;1\dots,1]$. After transforming to the polar coordinate system, time series at each time step was taken as a 1-D metric space. By defining the inner product $\langle x, y \rangle = x \cdot y - \sqrt{1-x^2} \cdot \sqrt{1-y^2}$, G is a Gramian matrix:

$$\begin{bmatrix} \langle \bar{x}_1, \bar{x}_1 \rangle & \cdots & \langle \bar{x}_1, \bar{x}_n \rangle \\ \langle \bar{x}_2, \bar{x}_1 \rangle & \cdots & \langle \bar{x}_2, \bar{x}_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \bar{x}_n, \bar{x}_1 \rangle & \cdots & \langle \bar{x}_n, \bar{x}_n \rangle \end{bmatrix} \quad (4)$$

The GAF preserves the temporal dependency and contains temporal correlations, because $G_{(i,j)|i-j=k}$ represents the relative correlation by superposition of directions with respect to time interval k . Along the main diagonal, the time series can be reconstructed approximately from the high level features learned by the deep neural network. However, the GAF is large because the size of Gramian matrix is $n \times n$ when the length of the raw time series is n . To reduce the size of the GAF, the time series should be smoothed while keeping trends. Piecewise Aggregation Approximation was applied as a useful method. Both GASF and GADF are GAF method and have the similar reconstructed properties.

1.2 CNN theory

Inspired by the biology of human brain to sense and process the stimulation of external information, Le Cun et al. proposed a deep neural network model with convolution structure in 1994 (Lawrence S et al. 1997; Ketkar N, 2017). Convolution neural network (CNN) is a mature network architecture for deep learning and application. It has achieved many unexpected effects in the field of computer vision. In principle, a standard convolutional neural network architecture is mainly composed of convolution layers. The main function of convolution layer is to extract the features of the input data, and the convolution core in the convolution layer is used to complete this function. We can think of the convolution kernel as a scanner with a specified window size. The scanner extracts the features from the data by scanning the input data over and over again. If we input image data, we can recognize the important features of the image after processing the convolution kernel. The pooling layer in CNN can be regarded as a way to extract the core features of input data in CNN, which not only achieves the compression of the original data, but also greatly reduces the parameters involved in the calculation of the model. In a sense, the computational efficiency is improved. The most commonly used pooling layer method is the average pooling layer and the maximum pooling layer. The input data processed by the pooling layer is usually the characteristic graph generated after convolution operation. The main function of the full connection layer is to compress the features extracted from the input image after convolution and pooling, and to complete the classification function of the model according to the compressed features. The core features of the input image extracted from the convolution layer and the pooling layer are multiplied by the weight parameters defined in the fully connected layer, and finally compressed into the only 10 output parameters, which are actually a result of a classification. After further processing of the activation function, our classification prediction results can be more obvious.

1.3 Contrast methods

Although the continuity of variables on the time scale is not considered, the traditional classification method for time series problems still achieves some effects. For the application of mechanical fault diagnosis, Han T (T Han et al 2018) explored the application of random forest, support vector machine and artificial neural network etc. In this paper, three typical classification methods were referred to: random forest, support vector machine and artificial neural network.

1.3.1 Random Forest (RF)

As the simple principle and low computational cost, RF has become a representative technique of Ensemble learning and has been applied in various regression and classification problems. The key of this algorithm is that improving the generalization performance of Bagging greatly by introducing "random feature selection". The algorithm flow is shown in figure 2:

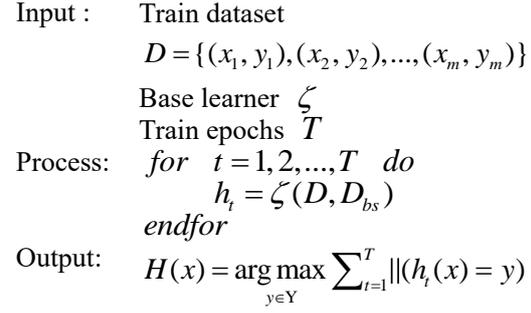


Figure 2 RF algorithm flow

1.3.2 Support Vector Machine (SVM)

Based on kernel function and convex optimization method, nonlinear support vector machine can solve multi - classification problems well. The algorithm flow of support vector machine is shown in the figure 3:

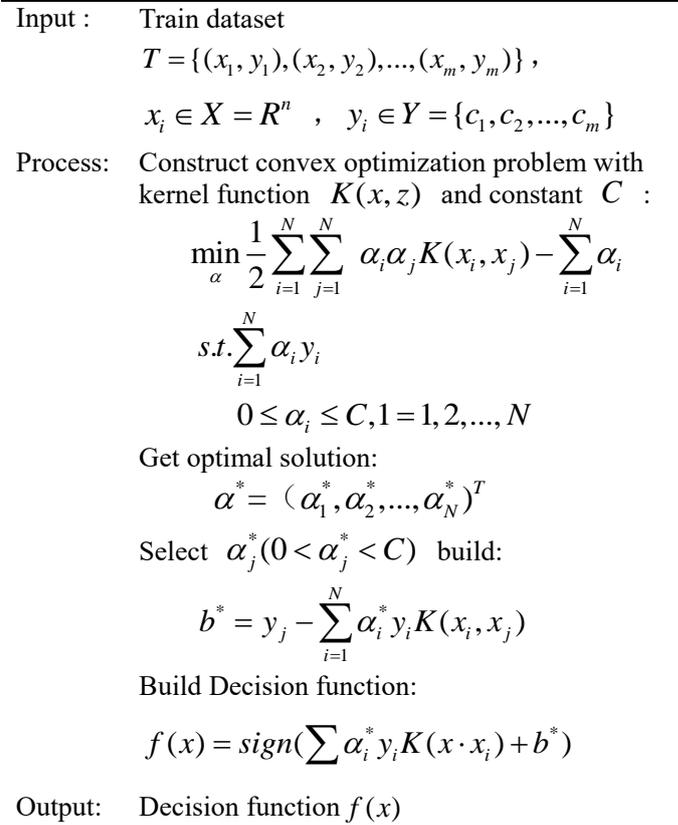


Figure 3 SVM algorithm flow

1.3.3 Artificial Neural Network (ANN)

Network models can be constructed by building multiple layers of neurons and transferring data between adjacent layers. Setting nonlinear function can realize the nonlinear transformation of input and output. Through constant data input and parameter modification, ANN can output sample categories according to sample features. The topological structure of ANN is shown in Figure 4.

Compared with other methods in this paper, it is necessary to manually set some features to represent the differences between different series. In this paper, 16 time-domain

statistical features are set for the experimental link of artificial neural network method. (T Han et al 2018)

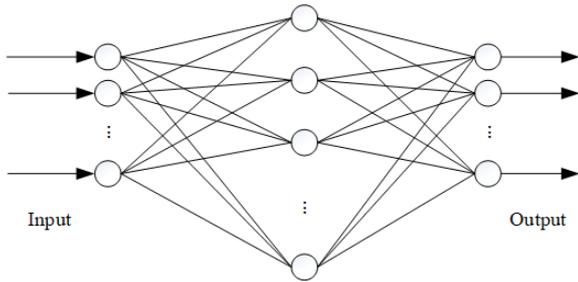


Figure 4 The topological structure of ANN

2. DESIGN OF BEARING FAULT DIAGNOSIS METHOD BASED ON GAF AND CNN

2.1 Algorithm design of GAF transformation

GAF transform method has detail mathematical deduction process, which is beneficial to programming. Extracting time series from bearing fault datasets and setting the length, Standard data set were got. Then, calling the GramianAngularField method from Pyts package can get 2 typical methods: GASF (Gramian Angular Summation Field) and GADF (Gramian Angular Difference Field), GASF was used in the transformation and image generation. Algorithm design of GAF transformation was shown in Figure 5.

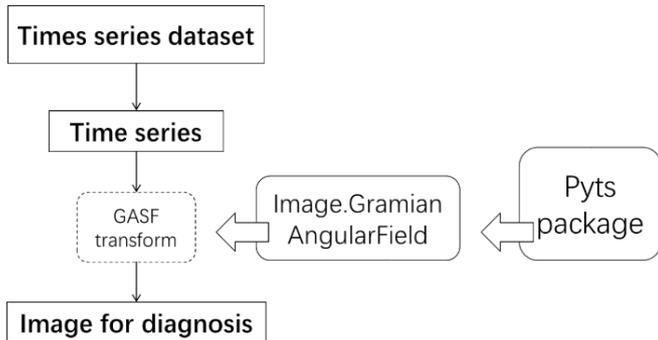


Figure 5 The algorithm design of GAF transformation

2.2 The design of CNN for image classification

The fault diagnosis algorithm based on CNN is different from the traditional intelligent algorithm. It integrates feature extraction and classification steps to achieve automatic feature extraction and classification, and abandons the complex process of artificial feature extraction before fault recognition. The structure of the convolutional neural network model designed in this paper is shown in figure 6. The neural network structure used is composed of three convolutional pooling units and one output unit.

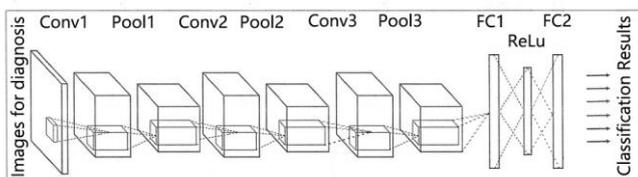


Figure.6 The structure of the convolutional neural network model

According to the convolutional neural network structure in figure 3, the implementation process of the convolutional neural network in this paper is shown in figure 7.

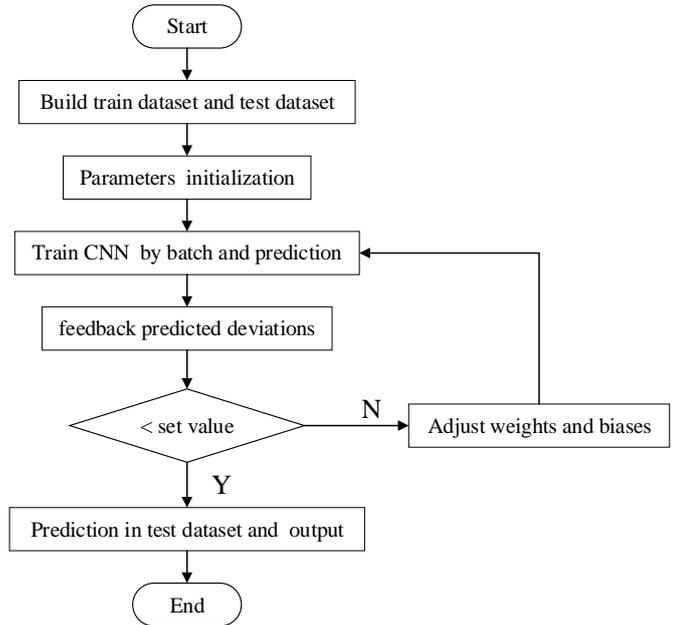


Figure 7 The implementation process of the convolutional neural network

The implementation of the algorithm is divided into two parts: the training phase and the test phase. The specific steps are as follows:

Step 1: the input signal is pre-processed to construct training set and test set.

Step 2: the training stage

- (1) initializes the parameters of the convolution neural network model, sets up the training termination condition,
- (2) carries on the batch training to the training set, then passes through the convolution layer, the pool layer, the whole connected layer processing, obtains the output;
- (3) calculate the error value of the actual output and the predicted output, if the error value meets the training termination condition, then stop the training, on the contrary, go to the step (4);
- (4) the error value of backpropagation is updated according to the error value of the network, and (5) back to step (3) to continue training.

Step-3: test phase

- (1) randomly sampling test samples from the test set, input to the model for testing, and get the output;
- (2) compare the output result with the test set true category, and use it to statistics the classification result;
- (3) return to step (1), re-extract the test samples for testing, until the completion of the test of all samples, the end.

3. EXPERIMENTAL VERIFICATION

3.1 Introduction of experimental dataset

Three data sets are included in the data packet (IMS-RExnord Bearing Data.zip). Each data set describes a test-to-failure experiment. Each data set consists of individual files that are 1-second vibration signal snapshots recorded at specific intervals. Each file consists of 20,480 points with the sampling rate set at 20 kHz. The file name indicates when the data was collected. Each record (row) in the data file is a data point.

3.2 Data pre-processing

This paper reconstructed the data set based on this data set, which contains 792 samples. Each sample is composed of a vibration signal numerical value with a length of 6000 that is continuously sampled according to time series, and a label reflecting bearing failure. There are 10 failure types in total. The information of experimental samples and their corresponding fault labels are shown in figure 8.

		Size of time series = 6000				
Size of dataset = 792	1	0.56365	1.069229	-0.83776	7
	2	0.061333	0.05883	0.056952	0
	3	0.035736	0.010964	-0.16487	9
	4	-0.0467	0.060913	0.00934	9
	7
	792	0.091613	0.037847	0.063025	3

Figure. 8 The information of experimental samples and corresponding fault labels

3.3 Experimental result

The convolutional neural network model proposed in this paper is established in the python-based Pytorch deep learning library, and the computer hardware configuration information is Intel(R) Xeon(R) w-2133 processor 8GB memory Windows 10 system.

Under the same training set and data set, the accuracy rate of "GAF + CNN" method is 52.52%, and that of random forest, support vector machine and artificial neural network is 39.92%, 31.52% and 22.17% respectively. The single time series sample of the data set has a time length of 6000, which greatly increases the difficulty of classification and makes the classification effect very bad. However, such data set also makes a good distinction between different methods in the long series classification. The experimental results of 4 methods was shown in figure 9.

The Torch deep learning framework has a various image transformation function, which can ensure that the trained model has good generalization performance. For the image classification problem, Adam optimizer is adopted according as a matter of experience. Some parameters of the train process refer to the experience value. The influence of these parameters on the result is not specifically investigated, which is also the focus of the follow-up work.

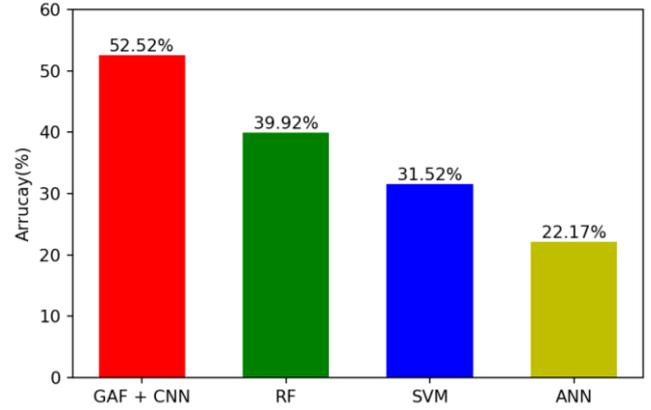


Figure. 9 The experimental results

4. CONCLUSION

This paper explores an indirect method to classify bearing fault signals, classifies bearing faults from the perspective of time series, and verifies the feasibility of this method. This method has the following advantages:

- (1) constructing feature images with time series as the entry point can obtain signal features that cannot be obtained at a single moment. It is beneficial to the diagnosis accuracy of bearing fault.
- (2) the image as the carrier of the characteristics of the signal, is conducive to the introduction of deep learning advanced feature extraction methods and classification methods to improve the accuracy of fault diagnosis. However, some parameters in the GAF transformation process and the training process of convolutional neural network refer to empirical values, and the influence of these parameters on the results has not been specifically investigated. The experimental results in this paper are not compared and analysed with the results of conventional bearing fault diagnosis methods. These will be the focus of follow-up work.
- (3) The "GAF + CNN" method fails to achieve the expected effect in the given problem, but it shows advantages compared with the classical classification methods RF, SVM and ANN.

REFERENCES

Gao, L., Yang, Z., Cai, L., Wang, H., & Chen, P. (2011). Roller bearing fault diagnosis based on nonlinear redundant lifting wavelet packet analysis. *Sensors*, 11(1), 260-277. Doi: 10.3390/s110100260

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., & Zheng, Q, et al. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings A*, 454(1971), 903-995. Doi: 10.1007/s10346-016-0714-6

Isermann, R. (1998). On fuzzy logic applications for automatic control, supervision, and fault diagnosis. *IEEE Transactions on Systems Man & Cybernetics Part A*, 28(2), 221-235. Doi: 10.1109/3468.661149

Jasper, J., & Othman, K. R. (2010). Feature extraction for human identification based on envelopegram signal analysis of cardiac sounds in time-frequency domain. *International*

Conference on Electronics & Information Engineering. Doi: 10.1109/ICEIE.2010.5559770

Ketkar, N. (2017). Introduction to PyTorch. Deep Learning with Python. Doi: 10.1007/978-1-4842-2766-4_12

Lawrence, S., Giles, C. L., & Tsoi, A. C. (2009). Convolutional neural networks for face recognition. *IEEE Transactions on Neural Networks*, 8(1), 98-113. Doi: 10.1109/72.554195

Cun, Y. L., & Bengio, Y. (1994). Word-level training of a handwritten word recognizer based on convolutional neural networks. *International Conference on Pattern Recognition, Vol 2-conference B: Computer Vision & Image Processing*. Doi: 10.1109/ICPR.1994.576881

LEI Yaguo, HE Zhengjia. (2011). Advances in applications of hybrid intelligent fault diagnosis and prognosis technique. *Journal of Vibration and Shock*, 30(9):129-135. Doi: 10.3969/j.issn.1000-3835.2011.09.028

Park, D. C., Elsharkawi, M. A., Marks, R. J. I., Atlas, L. E., & Damborg, M. J. (1991). Electric load forecasting using an artificial neural network. *IEEE Transactions on Power Systems*, 6(2), 442-449. Doi: 10.1109/59.76685

Rosero, J., Cusido, J., Ortega, J. A., Romeral, L., & Garcia, A. (2007). PMSM Bearing Fault Detection by means of Fourier and Wavelet transform. *Conference of the IEEE Industrial Electronics Society*. Doi: 10.1109/IECON.2007.4460324

T. Han, D. Jiang, Q. Zhao, L. Wang, and K. Yin,(2018). Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery," *Trans. Inst. Meas. Control*, vol. 40, no. 8, pp. 2681–2693

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Trans Neural Netw*, 10(5), 988-999. Doi: 10.1109/72.788640

Wang, H., Chai, T. Y., Ding, J. L., & Brown, M. (2009). Data driven fault diagnosis and fault tolerant control: some advances and possible new directions. *Acta Automatica Sinica*, 35(6), 739-747. Doi: 10.1016/s1874-1029(08)60093-2

Wang Z, Oates T. (2015) Encoding Time Series as Images for Visual Inspection and Classification Using Tiled Convolutional Neural Networks. *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*.