Analysis of Turbine Rotor Fault Diagnosis Method

Zhibiao Shi and Wenzhuo Dong*

School of Mechanical Engineering, Northeast Electric Power University, Jilin, China, 132012

*2274971760@qq.com

Abstract: The role of steam turbine in generating set is very important. During the operation, the rotor state is monitored by high-frequency vibration sensor. With the rapid development of computer, sensor and communication technology, modern machinery and equipment is developing towards the direction of electromechanical integration, as a result of the monitoring equipment group of large scale and the required point, single point more high sampling frequency and data collection time span is long, so the monitoring data of exponential growth, mechanical equipment health monitoring field enter the era of “big data”. The traditional method of turbine rotor fault diagnosis has encountered some difficult technical problems and bottlenecks. This paper first introduces the fault mechanism of turbine rotor. Then, the fault diagnosis methods of turbine rotor proposed by many scholars this year are analyzed, the fault characteristics of turbine rotor are summarized and the challenges are discussed, and the direction of further research is prospected.

Key words: Turbine rotor; Failure mechanism; Fault diagnosis

1. Introduction

The turbine plays an important role in the generator set. During the operation, the rotor state is monitored by high-frequency vibration sensor. Because the rotor running environment is complex and the fault features are easy to be submerged in the noise, the fault diagnosis results are affected and the fault occurrence is difficult to predict. Therefore, it is very important to study the fault diagnosis of turbine rotor accurately and quickly[1].

Turbine rotor fault diagnosis is mainly divided into two parts, one is the selection of the detection signal, the other is the signal processing process. At present, the main signal used in turbine rotor fault diagnosis is vibration signal[2]. Signal processing includes two core stages. The first stage is to extract features from the original signal, and the second stage is to classify the extracted features.

2. Study on fault vibration mechanism of turbine rotor

2.1 mechanism of rotor imbalance fault

Due to the quality of the material distribution, installation error, machining error and rotor wear, corrosion and deposition in the operation process factors on the influence of the rotor of the rotating machinery parts, such as the rotor rotates center do not coincide with the location of the normal, causes the damage of eccentric mass of rotor or rotor, causing the rotor fault.
2.2 Mechanism of rotor misalignment fault
Misalignment faults exist in many forms. Usually, the rotor has many shafts connected with each other, so misalignment of shafting is one of the main faults. The rotating shafts are connected by coupling to transfer torque and motion. In the process of high-speed operation and load bearing for a long time, it is easy to cause deformation and bending of rotor, uneven thermal expansion of bearing seat, uneven settlement of foundation, etc., as well as installation errors, all of which will lead to misalignment faults.

2.3 Mechanism of rotor rubbing fault
Rotating machinery is a periodic rotary movement, in the movement when the axis of rotation occurs bending deformation or because the coupling is not medium cause the axis deviated from the normal range, it will cause the motion and static parts of the gap, and produce friction phenomenon.

3. Research status of turbine rotor fault diagnosis
This paper mainly explores the core stage of signal processing: pattern recognition, i.e. classification according to extracted features. Fault diagnosis based on pattern recognition develops rapidly. Artificial neural network (ANN) and support vector machine (SVM), as intelligent recognizers, have received the most extensive attention and research in the field of mechanical fault diagnosis. But they also have some disadvantages. In order to solve or reduce the inherent defects of SVM, Tipping proposed a new sparse probability model algorithm similar to SVM in 2000: relevance vector machine. Although RVM has better sparsity and generalization ability, and has achieved relatively ideal results in prediction probability, it also has its own shortcomings in the training process. Due to its high computational complexity and large overhead, RVM is only suitable for handling small sample problems. In order to solve this problem, Tipping proposed a fast marginal likelihood algorithm in 2003, which significantly improved the speed of model training[3]. In 2005, Thayananthan A popularized this model, which solved the training problem of multiple output regression and classification[4]. In 2006, Rasmussen C E from Cambridge, UK and Max Planck, Germany published gaussian process machine learning, which made gaussian radial basis (RBF) kernel function become a widely used RVM kernel function[5]. In 2007, with the development of multi-core RVM research, multi-core RVM began to enter the application field. In 2008, Damoulas T and Girolami M applied multi-core RVM method to protein folding recognition and homogeneous detection, and obtained higher classification accuracy than single-core RVM[6]. In 2010, Psorakis I et al. summarized the sparsity and accuracy of solutions of multi-classification RVM method, which extended the original RVM to the application of multi-classification data by introducing auxiliary variable Y[7]. In 2010, Wu Lianghai used particle swarm optimization (PSO) to optimize the parameters of RVM and train RVM to get the optimal parameters by using the method of PSO[8]. In 2010, Li Gang used genetic algorithm to optimize the core parameters of RVM[9]. In 2011, Liu Junshun put forward the study on fast classification algorithm of relevant vector machine based on clustering. In addition, Bit Reduction and SMOTE algorithm were adopted to pretreat the original training samples with clustering, and then multiple local classifiers were constructed for various data clusters, and the feasibility of the algorithm was verified by experiments[10]. In 2012, Zhao Tong deeply studied the optimization method of correlation vector machine. According to the low proportion of “correlation vector”, he proposed a fast estimation and optimization algorithm of correlation vector machine, and proposed a solution to the problem of low detection accuracy caused by the use of a single kernel function[11]. In 2013, Liu Changyuan improved the “one-to-one” classifier, which is the most widely used and has the highest classification accuracy among multiple classification problems of related cameras, and greatly improved the classification time of the algorithm[12]. In 2015, Ouyang Ting proposed a fast multi-classification algorithm based on clustering related vector machines, which improved the computational speed of multi-classification problems[13].

With the continuous in-depth research by scholars at home and abroad, we will find that for the same classification task, when we use machine learning algorithm to do it, we should first identify the
feature and label, then input this data into the algorithm for training, and finally save the model to predict the accuracy of classification. But this method exists the problem is that we need to determine beforehand good features, every feature is a dimension, characteristics of too little number, we may not be accurate classification, namely what we owe fitting, if the number of features is overmuch, may cause we are obsessed with a certain characteristics in the process of classification result in classification error, namely the fitting. As a result, we need to spend a lot of time and energy on feature engineering to achieve a good result of model training. However, the emergence of neural network makes it unnecessary for us to do a lot of feature engineering, such as designing the content of features or the number of features in advance. We can directly feed the data into it, let it train and “correct” itself, and then get a better effect.

In terms of fault diagnosis based on pattern recognition, deep learning, as a young intelligent pattern recognition algorithm, has been successfully applied in fault diagnosis and prediction, showing its advantages in many aspects. Compared with the traditional information processing technology, it saves money and reduces manpower. The processing speed of big data is also better than the traditional algorithm, which fully shows the huge computing power in the information age to automate the processing of data. With the help of big data from various aspects, deep learning has achieved success in various fields. With the structure and form of neural network, deep learning fully conducts feature mining of data. Meanwhile, with the advantage of massive data, the instability of neural network has been overcome to some extent. Applied to the fault diagnosis of mechanical and electrical equipment, it can not only effectively use the massive data collected by sensors, but also have a powerful feature mining ability.

Currently, the main deep learning algorithm is divided into convolution neural network and the neural network, and compared with the neural network cycles, convolution neural network has better adaptability and adjustability, so the convolution neural network under the researchers' study, there are already many improved methods, although ordinary one-dimensional convolutional neural network has a significant effect, but considering the factors under different working conditions of steam turbine rotor fault data accuracy is not high, is not sensitive to individual factors, in order to improve the practicability of diagnostic algorithm improve the convolutional neural network model, one of the most important principle is a complicate the convolutional neural network structure, improve its generalization ability.

4. Fault diagnosis based on convolutional neural network
Since 2006, HINTON et al[14]groundbreaking put forward the deep learning theory, it has attracted wide attention in academia and industry and opened up a new way for the utilization of big data. Convolution neural network (CNN) is the leader in the field of deep learning with its extraordinary feature learning and pattern recognition ability, and has achieved a series of breakthrough research results in the fields of speech recognition, image processing and target detection. In terms of traditional pattern recognition methods, fault diagnosis generally needs to rely on signal processing means to extract fault features and use pattern recognition algorithm for diagnosis. In practical application, these methods have their own limitations, and it is difficult to make full use of the increasing mass data of mechanical equipment. Therefore, CNN, as a “sharp tool” with strong feature learning and pattern recognition ability, has been widely concerned by field scholars and introduced into the field of mechanical equipment fault diagnosis.

4.1 Fault diagnosis principle of convolutional neural network
As one of the most important models of deep learning, CNN provides an end-to-end learning architecture. After years of development, CNN has derived multiple versions, and its capability of feature learning and modal recognition is also increasing day by day[15]. Usually, CNN first calculates its prediction error by means of supervised learning through feedforward operation (convolution operation, pooling operation, etc.), and then updates parameters through the back propagation algorithm based on gradient. The gradient feedbacks layer by layer from back to front until updating
the first layer parameters of the network. Compared with the traditional neural network architecture, CNN introduces the concepts of weight sharing and receptive field, which greatly reduces the number of parameters to be learned and has stronger learning ability.

The thought method of diagnosis includes convolution layer, nonlinear activation layer, pooling layer, full connection layer, etc. Through the organic combination of these basic “components”, the original feature space can be mapped to the feature domain with more representational significance, so as to realize the prediction of samples. Convolution layer and pooling layer play an important role in CNN diagnosis model, which is directly related to the advantages and disadvantages of feature extraction.

Convolution layer: carry out convolution operation layer by layer through the action of several convolution kernels, and obtain different feature graphs of each layer of network through the nonlinear activation layer. Each layer of network “combines” the convolution kernel output of the upper layer to extract the topological structure features hidden in the data layer by layer.

Pooling layer: it uses the overall statistical characteristics of the adjacent outputs of a certain position to replace the output of the network at this position. When the input makes a small amount of translation, the expression of the input is approximately unchanged, that is, translation invariance. The maximum pooling and average pooling are shown in equations 1 and 2.

\[
p^{(l,i,j)} = \frac{1}{W} \sum_{i=-(j-1)W+1}^{W} \alpha^{(l,i,j)}
\]

\[
p^{(l,i,j)} = \max_{(j-1)W+(j-1)W} \{ \alpha^{(l,i,j)} \}
\]

Where: \( \alpha^{(l,i,j)} \) is the activation value of \( i \) th neuron in \( l \) th layer and \( i \) th frame; \( W \) is the width of the pool area; \( p^{(l,i,j)} \) is the width of the pool area.

Pooling can be regarded as adding an infinitely strong priori: the function learned at this layer has invariance to a small amount of translation, and pooling can greatly improve the statistical efficiency of the network. On the other hand, because the pooling operation counts the surrounding data, the output can represent the input with less data, which means that the neural network has the ability to compress data and improve the computation rate of the neural network.

Through layer by layer convolution, pooling and other operations, CNN takes the minimum loss as the learning goal, and by extracting the translation, rotation and constant scaling feature representation of input data, it can effectively realize fault diagnosis and recognition with the help of top-level classifier.

4.2 Research status of fault diagnosis of convolutional neural network

According to the published literature, the application of CNN in fault diagnosis mainly has two ideas: 1) use CNN as a classifier; 2) use CNN as a method of feature extraction and pattern recognition.

CNN-based diagnostic methods are good at processing massive data, picking up features from them and distinguishing different patterns of information contained in different data. In addition, with the aid of CNN’s unique advantages, the diagnosis model based on CNN characteristics can be adaptive extraction and pattern recognition for whole organic fusion, overcome traditional feature extraction method to extract the characteristics of shallow characterization of the problem of insufficient ability of fault signals, and to avoid the dependence on professional knowledge and experience, effectively reduce the complexity of the diagnostic process. It also has unique advantages in fault diagnosis. Therefore, CNN-based fault diagnosis technology is one of the development directions of intelligent diagnosis.

4.3 Challenges and prospects of fault diagnosis based on convolutional neural network

Due to the complexity and intelligence of modern mechanical equipment, it is more and more difficult
to obtain accurate, complete and effective fault information, so it is urgent to use intelligent diagnosis method that integrates signal analysis, modeling and knowledge processing[16-17]. In a sense, how to obtain the characteristic distribution that can represent the fault state has become the focus of current mechanical fault diagnosis, which directly affects the effectiveness and reliability of fault monitoring and diagnosis. The CNN-based fault diagnosis method, by building a deep model and utilizing its unique advantages, can directly learn fault features from a large amount of data adaptively, so as to achieve intelligent diagnosis in the context of “big data”, which is unified with the adaptive extraction of fault features and pattern recognition. After the development of recent years, the CNN-based fault diagnosis strategy has been published, but there are still some challenging problems in the application process that need to be further explored.

1) Feature extraction of diagnostic model
Feature learning of the CNN diagnostic model is not known, and it is difficult to analyze the learned features in an abstract way. In order to better apply CNN for fault diagnosis, how to define the meaning of learned features of the CNN diagnostic model is still a problem to be solved.

2) CNN model adaptive problem
Modern mechanical equipment produces a large amount of monitoring data, and its data types are diverse and inconsistent, which can reflect the state information of mechanical equipment from various aspects, such as vibration data and noise data. At the same time, signals generated by modern mechanical equipment are easily disturbed by the external environment and present complex and varied characteristics. It is difficult to ensure that the training data set used by the model is exactly the same as the distribution of the data set to be tested. Therefore, how to solve the adaptive problem of CNN model is also a direction worthy of research.

5. Conclusion
In this paper, the mechanism of turbine rotor fault is analyzed, the diagnosis method of traditional machine learning is summarized, the research idea and research status of the most advantageous CNN to realize mechanical fault diagnosis in deep learning are discussed, and the direction of further research is needed to realize fault diagnosis by CNN.

Reference
[1] He Qing, Xie Fangfang, Li Hong, et al. Vibration feature extraction of steam turbine units based on manifold learning method [J]. Vibration, test and diagnosis, 2014, 34 (4) : 705-708.
[2] Verma N K, Gupta V K, Sharma M, et al. Intelligent condition -based monitoring of rotating those using sparse auto - encoders [C]//Prognostics and Health Management (PHM), 2013 IEEE Conference on. IEEE, 2013:1-7.
[3] Faul A, Avenuse J t. Fast Marginal Likelihood Maximisation for failure Bayesian Models[J]. Proceedings of the Ninth International Workshop on Artificial Intelligence&Statistics, 2003,42 (22) : 3-6.
[4] Thayananthan a. temet-based pose estimation and tracking of 3D hand motion[D]. Cambridge: Degree thesis of Department of Engineering of University of Cambridge, 2005:25-30.
[5] Rasmussen C E, Williams C k. Gaussian Processes for Machine Learning[M]. Commonwealth of Massachusetts, MIT Press, 2006:211-244.
[6] Damoulas T, Girolami m. Probabilistic multi-class multi-kernel Learning: On Protein Fold Recognition And Remote Homology Detection[J]. Bioinformatics, 2008,24 (10) : 10-2390.
[7] Psorakis I, Damoulas T, Girolami M a. Multiclass relevance vector machines: sparsity and accuracy[J]. IEEE Neural Netw, 2010,21 (10) : 1588 -- 1598.
[8] Wu Lianghai. Network intrusion detection based on particle swarm optimization correlation vector machine [J]. Microelectronics and computer, 2010,27 (05) : 181-184.
[9] Li Gang, Wang Guilong, Xue Huifeng. Genetic algorithm optimization method for RVM kernel parameters [J]. Control engineering, 2010,17 (03) : 335-337.
[10] Liu junshun. Research on fast classification algorithm of correlation vector machine based on
clustering [D]. Master of science thesis of probability theory and mathematical statistics, south China university of technology, guangzhou, 2011:34-40.

[11] Zhao palm. Research on optimization methods of related vector machines [D]. Guangxi: master's thesis of computer application technology, guangxi university, 2012:24-27.

[12] Liu Changyuan. Research and application of multiple classification algorithm of correlation vector machine [D]. Harbin: engineering doctoral thesis of signal and information processing, Harbin engineering university, 2013:32-37.

[13] Ouyang Ting. Research on fast multi-classification algorithm of correlation vector machine based on clustering [D]. Master of science thesis of probability theory and mathematical statistics, south China university of technology, guangzhou, 2015:21-25.

[14] Hinton G E, Salakhutdinov R r. Reducing the dimensionality of data with neural networks[J]. Science, 2006,313 (5786) : 504-507.

[15] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C] //IEEE Conference on Computer Vision and Pattern recognition[J]. IEEE Computer Society, 2016: 770-778.

[16] Li Wei. Current situation and development trend of intelligent fault diagnosis technology for complex systems [J]. Computer simulation, 2004,21 (10) : 4-7.

[17] Chen Jin. Application of signal processing in fault diagnosis of mechanical equipment [J]. Vibration and shock, 1999(4) : 83-86.