Overtemperature fault diagnosis of front bearing for main spindle based on CNN + LSTM

Wang Wei1*, Kang Ruiqing1, and Zhang Yu1

1 School of Automation and Electrical Engineering, University of Science and Technology Beijing, 100083 Beijing, China

wangwei_link@qq.com

Abstract. The main spindle is an important transmission component of the wind turbine. The overtemperature fault of the front bearing of the main spindle is caused due to mechanical wear, grease failure and other reasons. A neural network based on convolutional neural networks (CNN) and long short memory network is built (LSTM) to judge the early fault. Method used in this paper can find the fault in advance. Compared with BP neural network, support vector machine, the accuracy of the model used in this paper is higher, which is up to 99.77%. The mechanism model of spindle operation will be established to analyse the manifestations of various faults and improve the accuracy in the future.

1. Introduction
The over-temperature accident of the front bearing will cause the main spindle to be scrapped. At present, the SCADA system used in the wind field can only shut down the blade’s rolling according to the set threshold temperature. However, when the machine shut down in this way, the spindle has been severely damaged and scrapped. If the faults can be warned in advance, it can effectively reduce loss. There has been some researches on the main spindle, but are most on the structure and physical properties of the material. The researches on main spindle components mainly focus on the vibration during rolling. In the actual, the signal acquisition frequency is low, and the offline data is mostly at the level of seconds or minutes. If the data used for vibration analysis, it is often difficult to achieve better results. The temperature of the wind turbine is slow, and the change does not exceed 0.1°C per minute in the normal. Therefore, if the temperature is used as the data source, the minute-level data can meet the demand.

The actual environment and industrial production process are complex to obtain the mechanism model directly. Traditional fault diagnosis methods often rely on process models, but data-driven fault diagnosis methods can overcome this shortcoming and gradually become a hot research topic [1]. The SCADA system stores a large amount of historical data. How to select suitable features has become a problem that needs to be solve. Common feature selection methods such as Pearson correlation analysis [2], grey relational analysis (GRA) and Relief algorithm [3]. Pearson correlation uses the covariance coefficient between features to determine the degree of correlation, but it can only analyze linear correlation. Gray correlation analysis requires manual determination of related parameters, which has serious subjective effects. Han improved GRA and applied it to the correlation analysis between two-dimensional vectors. The vector projection method was used to calculate the connection vector between the input point and the output point which measured the degree of correlation [4]. The ReliefF algorithm can handle the correlation of multi-feature data, but only two types of samples of the multi-type data...
can be extracted each time, the samples are not uniform [5]. DTW algorithm is widely used in the field of natural language processing. Firstly, the samples are regularized in the time domain, and then the calculation is performed. This method maintains the characteristics of features in the time domain and solves the problem of feature correspondence between features [6]. Cheng used the DTW algorithm to calculate the distance between the synthetic aperture radar image and the contours of different main images. The shorter the distance, the higher the similarity [7]. SOH used the DTW algorithm to adjust the air quality data, the processed data and the time features relationship were more accurate. And then used the deep neural network for training, the processed data converged faster and the error was reduced [8].

To establish a fault diagnosis model, traditional methods have many ways such as: frequency domain analysis method [9]. Zhang used wavelet transform to diagnose the faults of the circuit components of the wind turbine inverter. Wavelet transformed the collected voltage signal, decomposed and reconstructed the original signal for fault diagnosis and achieved good results [10]. Modern machine learning methods are widely used, including decision trees, mean clustering, artificial neural networks (Artificial Neural Network, ANN). Among them, ANN are widely used including Deep Belief Network (DBN), LSTM, CNN. Li Bin used the improved DBN to process the vibration signals of rolling spindles and obtained a good classification effect [11]. Zhu Lin improved CNN by genetic algorithm to predict the remaining life of turbofan engines [12]. Lei used the LSTM network to distinguish 11 types of faults of wind turbines [13]. Jiang used a multi-scale convolution kernel CNN to analyze the gearbox faults of wind turbines, and the results were better than the origin CNN [14]. Among the above fault diagnosis methods, the frequency domain analysis method can only be used for high-frequency data, while some data in the actual production process are low-frequency data, so the frequency domain analysis is not easy to produce results. At the same time, Most of the network models used are single type, so they can not make full use of the spatial and temporal characteristics in the data, resulting in the accuracy of fault diagnosis is lower than that of complex models.

2. CNN+LSTM model

LSTM network and convolutional network have their own advantages. LSTM network is widely used in time series data, including natural language processing, which can extract the information data in time series. CNN network is used to solve image classification problems. Combine the two networks, extract the original information with the two networks respectively, and then merge the extracted information, and classify and output the final result by the BP network finally. The structure is shown in Figure 1:
3. Data preparation

3.1. SCADA data introduction

The experimental data comes from the real data of SCADA from a wind farm in northwest China, which is recorded every minute. First, select 48 features which may be related to the temperature of the front bearing for the spindle from more than 2,000 features in the SCADA database.

In order to determine the fault samples, it is necessary to establish the evaluation rules for the over-temperature fault of the front bearing of the spindle.

3.2. Analysis of DTW algorithm sequence matching degree

DTW algorithm is to find the shortest path to minimize the cumulative distance $D$ from $w_i$ to $w_m$, so as to realize the realignment of two time series data, as Figure 2 shown.

$$ W = w_1, w_2, \ldots, w_m $$

![Figure 2. DTW algorithm](image)

Calculate the distance between two time series. The shorter the distance between the two series, the higher the matching degree in time. The calculating method is shown in the following:

$$ D_{i+1} = D_i + \sum_{j=1}^{m} \sum_{j-1}^{n} \min \left\{ d_{i+1,j}, d_{i,j+1}, d_{i+1,j+1} \right\} $$

In Table 1, the results show that the average distance between the engine room control cabinet temperature and the front bearing temperature of the main spindle is the shortest, so the matching degree with the front bearing temperature of the spindle is the highest.

| Order | Features                          | Pearson | DTW   |
|-------|-----------------------------------|---------|-------|
| 1     | Engine room control cabinet temp. | 0.912   | 0.034 |
| 2     | Engine room temp.                 | 0.891   | 0.053 |
| 3     | Hub temp.                         | 0.864   | 0.056 |
| 4     | Ambient temp.                     | 0.908   | 0.065 |
| 5     | Spindle rear bearing temp.        | 0.798   | 0.035 |
| 6     | 2# blade driver temp.             | 0.839   | 0.064 |
| 7     | 3# blade driver temp.             | 0.842   | 0.065 |
| 8     | 1# blade driver temp.             | 0.830   | 0.064 |
| 9     | 1# blade control cabinet temp.    | 0.757   | 0.086 |
| 10    | 3# blade control cabinet temp.    | 0.760   | 0.089 |
| 11    | 2# blade control cabinet temp.    | 0.775   | 0.090 |
3.3. Determination of fault temperature threshold

The simpler way to establish evaluation rules is to determine a threshold temperature of the spindle in general state. And when the temperature of the front bearing exceeds the threshold, it is viewed as a fault condition. Using the data of the normal condition, record the maximum point of the front spindle temperature at the interval of 0.1°C. After traversing all the intervals, a set of points is obtained and are linearly regressed as the threshold temperature line. By using the least square method, the following corresponding relationship is obtained:

$$T_{\text{threshold}} = 0.86T_{\text{control cabinet}} + 21.05°C$$  \hspace{1cm} (3)

In order to retain an appropriate margin for determination and make sure all points in normal conditions are placed below the fitting line, the fitting curve is shifted upward. The corresponding formula is shown in formula 4:

$$T_{\text{threshold}} = \min\left(0.86T_{\text{control cabinet}} + 27.05°C, 50°C\right)$$  \hspace{1cm} (4)

After adding and comparing the data of the failed units, if the units which above the threshold temperature are recognized as fault data. Since the temperature of the front bearing of the main spindle cannot increase linearly with the relevant characteristics, and the temperature of the normal unit in the current data does not exceed 50°C, it is specified that the temperature of the front bearing of the main spindle exceeding 50°C is a fault state, and the relationship shown in Figure 3 is obtained:

![Figure 3. Threshold temperature](image)

3.4. Experimental sample preparation

In order to provide more practical guidance to maintenance personnel of wind farms, the samples are divided into three types, among which the samples under normal working conditions are type 1, the data within 3 days before the shutdown occurs in the fault samples are type 3, and the rest of the fault data are type 2. Type 3 failures require emergency treatment, while Type 2 failures can be observed first, and maintenance measures can be taken when warnings are continuously issued.

A total of 12 features with Pearson correlation higher than 0.75 were selected as input samples, including engine room control cabinet temperature, engine room temperature and so on. By using the sliding window method, the window size and time interval have been tested many times in the research, and 6 pieces of data with a time interval of 5 minutes were selected as a sample. There were more than 11.91 million cases had been sorted.

Since the data came from the real measure of the wind farm, there was inevitably a problem of data imbalance. There were more data in the normal state, and the amount of abnormal data was small, causing serious data imbalance problem Type 1 data accounted for over 99%. In order to solve the problem, it was necessary to deal with when making the sample set. Type 1 had a large amount of data,
but a large amount of them are redundant. Using down-sampled method, data at different times are selected to be added to the training set. Type 3 data is relatively small, only more than 13,000 pieces. The SMOTE resampling method is used when solving sample imbalance problem to increase the number of minority samples. The sample size of Type 2 is more than 350,000 cases which meets the demand.

4. Model building and training
In order to verify the effect of the CNN+LSTM network, compared with the CNN, LSTM, BP, and SVM. Train CNN+LSTM, CNN, and LSTM network models separately. The ratio of training set to test set data is 7:3. The training data selection type is 400,000. For example, 300,000 cases of type 2 and 200,000 cases of type 3 data, a total of 900,000 cases. The training iterations is 100 and the process is shown in Figure 4. Among all of them, the CNN+LSTM network model has the fastest convergence speed.

![Figure 4. Training results](image)

5. Experimental results

5.1. Result analysis
The data of all 11.91 million cases were tested, and the results were shown in Figure 5. The results of the CNN+LSTM network model were the best, while the single LSTM and BP network performed poorly, and the results were lower than the traditional SVM classification method. The results of the CNN model were second only to CNN+LSTM. Using the excellent classification task capabilities of the CNN network, combined with the time series information extracted by the LSTM network, could make full use of the current data to analyse the operating status of the wind turbine. The accuracy of the model reached 99.77%, which was higher than other models.
Detailed analysis of the model for different types of samples, the accuracy rate reflected the correct ratio of the retrieved samples, as shown in Figure 6. The accuracy of the model was maintained at a high level in the three types of samples, and the accuracy and precision of the type 1 and type 3 samples were the best. The CNN network performs slightly inferior to the CNN+LSTM network in the type 2 and type 3 data, and its accuracy rate had declined. A single BP and LSTM network performed poorly in the type 2 and type 3 samples, but the LSTM network performed best in the type 2 samples, indicating that the correct proportion of the type 2 samples retrieved by the LSTM network was higher. The traditional machine learning method SVM performed better than BP and LSTM networks.

5.2. Actual sample analysis
Taking an overtemperature fault for the front bearing of the main spindle on September 22 as an example, the performance of the model in practice was analysed. Figure 7 showed the change of the temperature with time. The emergency shutdown was triggered when the temperature exceeded 70 °C.
Figure 7. Actual temperature

When the overtemperature fault of the front bearing of the main spindle reached 70 °C, it is about 8:00 on September 22. Before reaching 70 °C, according to the established normal temperature threshold, the temperature exceeded the threshold after September 15. Therefore, according to the rules formulated by the data division standard, the trained model was used to predict and obtain the output results of each model. The prediction of each model was shown in Figure 8.

The overall performance of the CNN+LSTM model was the best. There was only a partial data judgment error on September 16th. In this segment of data, the temperature of the spindle was reduced to 35°C close to the normal working temperature, which caused the misjudge. There was only a few minutes of delay during the transition period between Type 2 and Type 3. The overall performance of the model was excellent. The CNN model and the SVM model were generally more accurate, but there were many judgment errors, and they were concentrated in the temperature drop stage. The judgment accuracy of the models was lower than the CNN+LSTM model. The performance of LSTM and BP network were poor in actual samples, and prediction errors occurred in multiple time periods.

The results had large deviations from expectations, and were greatly affected by temperature changes. The results were consistent with the comprehensive test results of the model. In actual production, both CNN and SVM models could meet certain requirements, but there were more false positives than the CNN+LSTM model, which increased the operation and maintenance cost.

6. Summary
The research used CNN+LSTM network to establish fault diagnosis model and classify faults. The model could use the time and space information of the data at the same time to improve the ability of obtain information. At the same time, the data information was extracted in different dimensions. During the test, the accuracy of the model was as high as 99.77%, and the accuracy rate and recall rate were evaluated. In the future, depending on the operation of the wind farm, more data will be generated, increasing the diversity of data, and training models with richer data can further improve the reliability of the model.
Figure 8. Actual results

References
[1] X. Fu, J. Y. Jun, K. Cao. Overview of big data driven intelligent fault diagnosis technology in iron and steel, industry. CMC, 28(11), 1-6+26, (2020)
[2] Z. A. Ding, C. W. Le, L. L. Wu. PM based on cemd Pearson and deep LSTM hybrid model (2.5) concentration prediction method. CS, 47(S1), 444-449, (2020)
[3] Z. Y. Wang, Y. F. Zhang, X. P. Duan, P. Song, W. X. Yang, Fault feature parameter extraction method of wind turbine based on relief algorithm, NCEPT, 10, 57-62, (2017)
[4] M. Han, R. Zhang, T. Qiu, Multivariate chaotic time series prediction based on improved grey relational analysis, IEEE TS, 49(10), 2144-2154, (2017)
[5] Y. Zhao, J. W. Wu, S. L. Ma, Y. Chang, Y. Feng, N. Yang, D. G. Li. Integrated SVM method based on am ReliefF feature selection for mechanical fault diagnosis of high voltage circuit breaker, CJEE, 41(08), 2890-2901, (2021)
[6] Z. Q. Dong, Ultra short term wind speed prediction based on DTW-FCBF-LSTM model, EMI, 57(04), 93-98, (2020)
[7] L. Cheng, Y. Wang, L. Zhong. Technical framework of feature extraction based on pixel-level SAR image time series, AEORS, 8(4), 1665-1681, (2015)
[8] P W Soh, J W Chang, J W Huang. Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations. IEEE Access, 6, 38186-38199, (2018)

[9] H Li, G Feng, D Zhen, A normalized frequency-domain energy operator for broken rotor bar fault diagnosis, IEEE TIM, 70, 1-10, (2020)

[10] J Zhang, H Sun, Z Sun, Fault diagnosis of wind turbine power converter considering wavelet transform, feature analysis, judgment and BP neural network, IEEE Access, 7, 179799-179809, (2019)

[11] Z. Y. Yin, X. F. Wu, W. K. Zhong. Transformer fault identification and case analysis based on random forest, EMT, 43(23), 63-67, (2020)

[12] L. L. Chen, J. B. Si, Z. J. Dong. Application of improved deep confidence network in fault diagnosis of motorized spindle, MST, 3, 1-8, (2021)

[13] J Lei, C Liu, D Jiang. Fault diagnosis of wind turbine based on Long Short-term memory networks RE, 133(APR.), 422-432, (2019)

[14] Jiang G, He H, Yan J, Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox, IEEE TIE, 66(4), 3196-3207, (2018)

Acknowledgments
Thanks for the help of lab classmates and the wind farm who provide the whole database.