Online detection method for inter-turn short-circuit fault of permanent magnet synchronous motor based on deep learning

Chengji Shi*, Xiaojuan Lu

School of Automation and Electrical Engineering, Lanzhou jiaotong University, lanzhou, gansu, 730070, China

*Corresponding author’s e-mail: 494307834@qq.com

Abstract. Aiming at the difficult problem of inter-turn short-circuit fault diagnosis of permanent magnet synchronous motors without stopping the machine, this paper takes timing anomaly detection as the starting point, and proposes an online fault detection method based on transfer learning deep self-coding network. First, a migration deep self-coding network model is constructed, and a combined sample of negative sequence current and electromagnetic torque is used to extract common features in different domains. Then, the online detection model of abnormal timing is established. Finally, the permutation entropy value of the normal state of the motor is used to construct an alarm threshold to improve the matching speed of abnormal sequences in the online data. Experimental results show that the method in this paper has better detection practicality.

1. Introduction

The inter-turn short-circuit fault of permanent magnet synchronous motor is one of the most common faults in the motor, and it is a very destructive motor fault. If the early slight turn-to-turn short circuit is not identified, it will cause a more serious phase-to-phase short circuit or single-phase ground fault. Therefore, it is of great significance to study the online detection method of inter-turn short circuit for the timely detection of early motor faults and reduce the accident rate [1].

At this stage, the identification of short-circuit between turns of permanent magnet synchronous motors (PMSM) is mostly based on vibration and other parameters [2] to determine the running state, determine the fault location and the degree of deterioration. This article focuses on the problem of early fault online detection under non-stop conditions, which helps to evaluate the working state of the motor in real time. At present, fault detection in online scenarios is to extract early fault features such as time-frequency [3] and wavelet [4] from the vibration signal, and then construct a learning model such as Fisher discriminator [5] and artificial neural network [6] to perform anomalies Tested and achieved certain results. However, the above method usually cannot adaptively extract features. The offline model is not fully applicable to online data due to noise interference and other reasons, which easily reduces the accuracy of the detection results. Secondly, the above method often uses anomaly detection algorithm, which does not fully consider the timing relationship of the samples, and it is easy to generate false alarms due to small fluctuations.

This paper presents an online fault detection method for deep transfer learning. Add the Laplace regularization term of the weighting coefficient to the objective function of the multi-layer autoencoder to obtain more early fault features that are more expressive, use the permutation entropy value of the offline data to construct the alarm threshold, and realize the timing faults in the online
data by matching the abnormal sequence Quick detection of patterns. Finally, the effectiveness of the method for online detection of inter-turn short circuit faults is verified through experiments.

2. Deep learning applications

2.1 Transfer learning
Transfer learning is a type of deep learning method that uses existing knowledge to solve problems in different but related fields. By using the feature information contained in the data of one domain to improve the performance of the prediction model in another related but different domain. Transfer learning has been successfully applied to the learning problem of insufficient training samples [7], using the domain adaptation method [8] to learn the common features of the source and target domains to achieve cross-domain information transfer of auxiliary data from different sources.

2.2 Deep self-coding network
This paper first selects the multi-layer autoencoder [9] as the basic model of deep network. The encoder is used to extract hidden features of input data, expressed as: \( H = f(WX + b) \). Where \( W \) is the weight matrix of the input layer and the hidden layer. The decoder is used to reconstruct the input data from the extracted features, expressed as: \( Z = g(W'\hat{H} + b') \). The activation function is: \( \text{sigmoid}(x) = 1/(1+e^{-x}) \). The goal of the autoencoder is to minimize the following reconstruction errors: \( L_{AE} = \frac{1}{2n}\|Z - X\|_F^2 \), Where \( \|\cdot\|_F \) represents Frobenius norm. Secondly, in order to eliminate the difference in data distribution between the source and target domains, this paper introduces the maximum mean discrepancy (MMD) regular term in the loss function. The MMD distance is defined as [10]: \( \text{MMD}(S,T) = \frac{1}{N(N-1)}\sum_{i=1}^{N} \sum_{j=1}^{N} \phi(x_i)\phi(x_j) - \frac{1}{n(n-1)}\sum_{i=1}^{n} \sum_{j=1}^{n} \phi(x_i')\phi(x_j') \), Where \( \phi \) represents the reproducing kernel Hilbert space, \( S \) and \( T \) represent the input sample sets in the source and target domains respectively, and \( N \) and \( n \) represent the number of samples in the source and target domains, respectively.

In order to strengthen the distinction between normal state data and early fault data, the Laplace regular term of the weight matrix is constructed as follows: \( L_w = \sum_{k=1}^{K} \exp(-\|\Delta W_k\|_F^2 / \sigma) \), Where \( K \) is the number of weight matrices in the multi-layer autoencoder, and \( \sigma \) is the penalty factor, \( \Delta = D_1 \otimes I_1 \otimes I_2 \otimes D_2 \), \( I_1 \) and \( I_2 \) are identity matrices, and \( D_1 \) and \( D_2 \) are Laplace operators. Integrating these three parts together, we finally get the objective function of the multi-domain migration deep autoencoder: \( L = L_{AE} + \lambda L_{MMD} + \frac{\mu}{2} L_w \), Where \( \lambda \)、\( \mu \) > 0 is the regularization parameter.

The migration self-coding network structure is shown in Figure 1:

![Migration deep self-coding network structure diagram](image-url)
3. Online anomaly detection model

The process of online detection requires that the fault detection model should have good accuracy and real-time. This section uses permutation entropy [11] to construct a detection model by means of abnormal sequence detection. First, the depth features extracted from the migration self-encoding network are divided into sequences with a length of 100 samples, and then the permutation entropy is calculated for each sequence. It is necessary to set a reasonable threshold in the process of line anomaly detection. Using deep transfer learning method to extract public features, and then use these offline data to calculate a more reasonable alarm threshold. The detection process is as follows.

First, use the common features extracted by \( H = f(WX + b) \) to calculate the arrangement entropy value of the motor data in the normal state. The specific calculation steps refer to [12].

1. Perform phase space reconstruction on sequence fragments to obtain matrix \( K_y \).
2. Sort each row of \( K_y \) in ascending order to construct its index matrix \( S_y \).
3. Calculate the permutation entropy of the sequence:
   \[
   H_{PE}(j) = - \sum P_i \ln P_i, \quad \text{where } P_i \text{ is the probability of occurrence of each arrangement in } S_y.
   \]
   In order to facilitate the calculation, \( H_{PE} \) is normalized. Secondly, determine the minimum value of permutation entropy as the threshold for online detection. Finally, the sliding window is used to generate online sequence fragments from the online detection data, and the depth features are extracted using the migration depth self-encoding network, and then the permutation entropy value is calculated to match it with the threshold criteria. The specific process is shown in Figure 2.

4. Analysis of results

The method mentioned in this article: the offline stage is mainly the construction of deep transfer learning models and the construction of anomaly detection models, and the online stage uses public feature representation to extract the features of the target and perform anomaly recognition. Using MATLAB / Simulink platform to build a three-phase permanent magnet synchronous motor model, the motor parameter settings are shown in Table 1.

| parameter                  | Numerical value | parameter                  | Numerical value |
|----------------------------|-----------------|----------------------------|-----------------|
| Rated power /kW            | 5.5             | Number of pole pairs       | 4               |
| Rated voltage /V           | 220             | Number of stator slots     | 24              |
| Rated frequency /Hz         | 100             | Parallel winding turns     | 148             |
By setting the number of working turns of the parallel winding of the motor, the inter-turn short-circuit fault is simulated to obtain 6000 sets of characteristic sample data sets. Part of the data of the feature sample is shown in Table 2.

### Table 2 Partial feature sample data set

| A     | B     | C     | T     | label | A     | B     | C     | T     | label |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.01  | 0.04  | 0.04  | 3.84  | 1     | 0.89  | 0.27  | 1.03  | 4.22  | 2     |
| 0.02  | 0.07  | 0.05  | 3.97  | 1     | 1.13  | 0.19  | 1.34  | 4.32  | 3     |
| 0.45  | 0.39  | 0.81  | 4.1   | 2     | 1.97  | 0.42  | 1.75  | 4.41  | 3     |

In Table 2, A, B, and C are the amplitudes of the 3rd harmonic current of Phase A, Phase B, and Phase C, respectively; T is the electromagnetic torque. The label value indicates the motor status. 1 means normal; 2 means slight inter-turn short circuit; 3 means severe inter-turn short circuit.

In this paper, a combination of negative sequence current and electromagnetic torque samples is used to form unrelated feature items into a sample set, so as to avoid diagnostic errors due to the error of a certain feature, while improving the learning ability of deep neural networks and avoiding overfitting.

### 4.1 Migrating deep self-coding networks

In order to ensure the reliability of the deep migration self-coding learning model, this paper optimizes from three indicators: hidden layer, learning rate and embedded optimization algorithm. In this experiment, the number of trainings was set to 12000, or the training was terminated when the training loss was less than 0.01. Determine the optimal situation by optimizing the objective function.

The training algorithm will also have a great impact on the training process. According to the comparison results in Figure 3, Adam is selected as the optimal algorithm of the network. According to Fig. 4, comparing the loss values corresponding to different hidden layer states, the effect of 3-layer hidden layer training is the best. Figure 5 shows that the loss value is the smallest when the learning rate is 0.01, the learning rate is too high for the overall network training efficiency, and too low makes the training process slow.
Among them, the hidden layer number in Fig. 3 is 3, and the learning rate is 0.01; the learning rate
in Fig. 4 is 0.01, and the Adam optimization algorithm is used.

In order to verify the convergence effect of the migrating self-encoding network, Figure 6 shows
the variation of training error during the above training process. It can be seen from Figure 6 that the
experiment has been carried out 1200 trainings. The training error has converged obviously when the
number of trainings is about 100. In the subsequent training process, although the training error
continued to decline, the rate of decline gradually slowed. The whole tends to converge.

4.2 Analysis of online anomaly detection results
This experiment selects 10 sets of data, each set of data has 600 samples. Among them, 8 groups are
used for offline training, and 2 groups are used for target detection. All data are linearly normalized
before processing, and the input signal used as the deep learning model adopts the spectrum data after
FFT conversion.

The public feature representation is obtained by constructing the extraction of the deep transfer
learning model; on this basis, the sample set in the source domain is taken as offline data, the
permutation entropy value is calculated, and the abnormality detection threshold is determined; finally,
the remaining two sets of data are targeted. The object performs online detection, uses the trained
public feature representation to directly extract features, and calculates the permutation values in
sequence according to the sliding window method, compares with the national value, and judges
online whether the permanent magnet synchronous motor has a short circuit.

In order to further verify the effect of deep transfer learning on online fault detection results, a
reasonable threshold is set as the standard for anomaly detection according to the arrangement entropy
value of normal state data fluctuating within a relatively small range. In order to verify the role of
Laplace regular terms in extracting early fault features, this article randomly selects some sample data
of two motors in normal state as objects, and compares the features and corresponding entropy values
before and after adding Laplace regular terms. According to FIG. 7, it can be seen that the common
features mentioned after adopting the Laplace regular term fluctuate more when the state changes, and
the corresponding arrangement entropy value has a more obvious step phenomenon. This shows that
the Laplace regular term can effectively enhance the distinction between the normal state and the early
fault state, and the extracted features are more sensitive. It also shows that the permutation entropy has
a good discrimination effect on the fault.

![Figure 7 Comparison of sample characteristics before and after adding Laplace](image)

Randomly select a set of data in the source domain and the target domain, calculate the permutation
entropy, and determine the anomaly detection threshold. Take another set of data in the target domain
as the detection object. The characteristic trend of the detected sample data is shown in Figure 8. It can
be obtained from Figure 9 that the alarm value is about 0.38, and then the test sample is tested.
According to Fig. 9, it can be detected that the 42nd sample has an early slight turn-to-turn short-
circuit fault. Therefore, the validity of the migration deep self-coding network model used in this
paper to detect the inter-turn short-circuit fault of permanent magnet synchronous motors is verified.
5. Conclusion

Aiming at the problem of online detection of early turn-to-turn short-circuit faults of permanent magnet synchronous motors, this paper proposes an efficient method for abnormal sequence detection based on deep transfer learning. The use of Laplace regular terms to generate more sensitive early fault features not only helps to perform sequence anomaly detection based on permutation entropy, but also strengthens the robustness of the detection model and reduces the false alarm rate. This method has a simple model and reliable results, and is more suitable for online detection of early faults.

Acknowledgments

Authors acknowledge assistance financial support from Projects supported by the National Natural Science Foundation of China under Grant (51667013).

References

[1] Ren Hao, Qu Jianfeng, etc. Research status and challenges of deep learning in the field of fault diagnosis [J]. Control and Decision, 2017, 32(08): 1345-1358.
[2] Ding Feng, Qin Fengwei. Application of wavelet noise reduction and Hilbert transform in motor bearing fault diagnosis [J]. Journal of Motor and Control, 2017, 21(06): 89-95.
[3] Liu R N, Yang B Y. Time-frequency atoms-driven support vector machine method for bearings incipient fault diagnosis. Mechanical Systems and Signal Processing, 2016, 75: 345−370.
[4] Yang Rui, Li Hongkunetc. Extraction of early fault features of rolling bearings using optimal wavelet scale cyclic spectrum [J]. Journal of Mechanical Engineering, 2018, 54(17): 208-217.
[5] Li F. Weak fault diagnosis of rotating machinery based on feature reduction with Supervised Orthogonal Local Fisher Discriminant Analysis. Neurocomputing, 2015, 168: 505−519.
[6] Amar M, Gondal I, etc. Vibration spectrum imaging: A novel bearing fault classification approach. IEEE Transactions on Industrial Electronics, 2014, 62(1): 494−502.
[7] Zhuang Fuzhen, Luo Ping, He Qing, Shi Zhongzhi. Research progress of transfer learning [J]. Journal of Software, 2015, 26(01): 26-39.
[8] He Min, Tang Jian, etc. Soft measurement of load parameters of wet ball mill based on manifold regularization domain adaptive stochastic weight neural network [J]. Journal of Automation, 2019, 45(02): 398-406.
[9] Cho K, Courville A, Bengio Y. Describing multimedia content using attention-based encoder-decoder networks. IEEE Transactions on Multimedia, 2015, 17(11): 1875−1886
[10] Shao H D, Jiang H K, Zhang X. Rolling bearing fault diagnosis using an optimization deep belief network. Measurement Science and Technology, 2015, 26(11): 115002.
[11] Lu W N, Liang B, Cheng Y, Meng D S, Yang J, Zhang T. Deep model based domain adaptation for fault diagnosis. IEEE Transactions on Industrial Electronics, 2017, 64(3):2296−2305.
[12] Feng Fuzhou, Rao Guoqiang, etc. Anomaly detection and diagnosis of rolling bearing based on
permuation entropy and neural network [J]. Noise and Vibration Control, 2013, 33 (03): 212-217.