IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 17, 2018, accepted January 17, 2019, date of publication January 31, 2019, date of current version February 14, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2895909

An Efficient Fault Diagnostic Method for Three-Phase Induction Motors Based on Incremental Broad Learning and Non-Negative Matrix Factorization

SAI BIAO JIANG[1,2], PAK KIN WONG[2], RENCHU GUAN[3,4], (Member, IEEE), YANCHUN LIANG[3,4], AND JIA LI[1]

[1]School of Electronic Information Engineering, Zhuhai College of Jilin University, Zhuhai 519041, China
[2]Department of Electromechanical Engineering, University of Macau, Taipa 999078, Macau
[3]College of Computer Science and Technology, Jilin University, Changchun 130012, China
[4]School of Computer, Zhuhai College of Jilin University, Zhuhai 519041, China

Corresponding author: Pak Kin Wong (fstpkw@um.edu.mo)

This work was supported in part by the National Natural Science Foundation of China under Grant 61602207, Grant 61572228, and Grant 61472158, in part by the Zhuhai Premier Discipline Enhancement Scheme under Grant 2015YXXK02, in part by the Guangdong Premier Key-Discipline Enhancement Scheme under Grant 2016GDYSZDXK036, in part by the Youth Innovation Talents Program of the 2017 Guangdong University Provincial Key Platform and Major Research Projects under Grant 2017KQNCX252, and in part by the Student Science and Technology Innovation Training Special Fund of Guangdong University under Grant PDJHA0620.

ABSTRACT Three-phase induction motors (TPIMs) are prone to numerous faults due to their complicated stator and rotor conditions and require a fast response, accurate, and intelligent diagnostic system. Recently developed fault diagnostic systems for induction motors are based on machine learning approaches, but their complex structure typically results in long training time. Moreover, they need to be retrained from scratch if the system is not accurate. We apply incremental broad learning (IBL) method to the diagnosis of TPIM faults. The IBL can train and retrain the network efficiently due to its flexible structure. The new diagnostic framework also consists of feature extraction techniques (empirical mode decomposition and sample entropy) and a non-negative matrix factorization (NMF) IBL approach. The experimental results demonstrate that the IBL system is superior to some algorithms, such as deep belief networks, convolutional neural networks, and extreme learning machine. Moreover, the IBL simplified by NMF is more accurate than the IBL without NMF.

INDEX TERMS Fault diagnosis, feature extraction, incremental board learning, non-negative matrix factorization, three-phase induction motor.

I. INTRODUCTION

Three-phase induction motors (TPIMs) are ubiquitous in our daily lives due to their low cost, reasonably small size, robustness, and low maintenance [1]. Their increased popularity has resulted in numerous studies having appeared in the literature. Although TPIMs are reliable, they are prone to some undesirable stresses, which cause faults that result in failure. It is therefore necessary to constantly monitor their health to prevent any serious accidents. The previous work [2] has shown that the common problems with TPIMs include phase imbalances, short circuits, mechanical imbalances, bent rotors, broken rotor bars, defects in the bearing inner and outer raceway, and broken ball-bearings.

Currently, there are model-based, knowledge-based, and data-driven fault diagnostic methods for TPIMs [3]–[5]. Model-based diagnostic methods are typically based on a single analytical model, but no single model can diagnose all possible faults, given the wide variety of motor designs and configurations. The model-based method is therefore difficult to apply to motor diagnosis. The knowledge-based method reaches a diagnosis based on a decision tree consisting of a sequence of questions and answers. The diagnostic knowledge is vague and limited due to the complexity of TPIMs.

The associate editor coordinating the review of this manuscript and approving it for publication was Michael Lyu.

This method is therefore also not suitable for TPIM diagnoses. The data-driven method relies on signal-based analysis techniques that are highly suitable for analyzing a system whose domain information is insufficient or unknown, as is the case with TPIMs. As a result, we propose our approach based on these data-driven methods employing machine learning to build classifiers to deal with fault diagnostic problems.

Recently, several different kinds of machine learning methods have been applied to fault diagnosis: deep belief networks (DBN) [6], [7], deep Boltzmann machines (DBM) [8], support vector machine (SVM) [9], extreme learning machines (ELM) [10], [31], and convolutional neural networks (CNN) [11], [12]. They are particularly suited to AC motors where the relationship between motor current and speed is nonlinear [13]. Although deep learning is rather powerful, most of these networks suffer from time-consuming training processes because there typically are a great number of hyper parameters and complicated structures involved [14]. Recently, broad learning (BL) has been proposed to improve this training performance [15]. Unlike previous methods, broad learning contains two layers: an input layer that consists of mapped features and enhancement nodes, and an output layer. Despite this type of structure is simple, it can lead to a significantly improved performance by increasing the number of enhancement nodes. Therefore, BL has a potential to improve both induction motor diagnosis accuracy and training speed. As BL is not designed for fault diagnosis, the associated techniques, such as feature extraction and selection for signal data processing, should be carefully integrated with BL. In other words, this work is not a straightforward application of BL. The techniques are reviewed below.

Before training the BL to construct the fault diagnostic system, the data needs to be acquired and processed. To increase accuracy, useful fault features should be extracted using signal processing. While Fast Fourier Transform (FFT) is commonly used during signal processing, this method cannot account for non-stationary signals that are commonly found in TPIM diagnoses. The very nature of the signal requires a time-frequency analysis [16]. A commonly proposed time-frequency analysis method for processing non-stationary signals is the short-time Fourier transform (STFT) [17]. However, this method has the inherent defect of producing interrelated resolutions of time and frequency. Other methods include wavelet transform or empirical mode decomposition. Wavelet transform [18] suffers from energy leakage because any signal characteristic that is not well correlated with the shape of mother wavelet is either masked or completely ignored. Empirical mode decomposition (EMD), has been introduced as a useful tool for analyzing nonlinear and non-stationary signals [19]. It disassembles a signal into many intrinsic mode functions (IMFs) to extract the main features. The analysis of every IMF can lead to a more accurate and effective feature extraction from the original signal, but the disassembling produces a multi-dimensional input to the fault classifier. It is well known that the accuracy of the fault classifier can be decreased when its input becomes multi-dimensional. To resolve this problem, feature selection algorithms should be employed to reduce the input dimensionality to a manageable level. Sample Entropy is known as an effective feature selection algorithm [20] that is capable of finding regularities in some IMFs based on signal statistics. Apart from that, domain knowledge (DK) can also be employed to represent system features in the problem domain. Domain knowledge features are important information of a specific field. Taking motor current signal as an example, maximum current, minimum current and average current are DK features. The author's previous work in [10] showed that the addition of appropriate domain specific features could enhance the accuracy of the diagnostic system. Motivated by the above discussion, it is a good idea to integrate DK, EMD and Sample Entropy for feature extraction and selection, and produce representative and compact inputs (i.e. features) to BL.

In general, if the diagnostic model is built by processed training data and the existing machine learning algorithm, the diagnostic system model cannot be modified once the model has been trained. The same is true for the motor fault diagnosis system. If the diagnostic system is inaccurate, it requires a great effort to remodel the system, especially for deep learning machines [15]. The same applies to the original BL. To overcome this problem, an incremental broad learning (IBL) algorithm [15] is considered in this study after the system is trained by BL. This dynamic step-wise updating algorithm can update the output weight of newly added enhancement nodes. As the IBL is designed to quickly update the weights without having to re-run the entire training cycle, this research employs IBL to remodel the fault diagnostic system for TPIMs.

Nevertheless, the IBL system may contain some redundancy nodes after the system completes training, which may result in poor diagnostic precision. Thus, this diagnostic system needs to be simplified through the use of low-rank approximations, for which several approaches exist, including singular value decomposition (SVD) [15] and non-negative matrix factorization (NMF) [21]. The disadvantage of the SVD algorithm is that it is difficult to interpret, and its processing capability for noisy data is unstable. Especially when dealing with high-dimensional data, decomposition speed and precision will be very poor. The NMF algorithm provides a new method to solve matrices based on a simple iteration. This method converges quickly. The left and right non-negative matrices typically require only a small storage space. Therefore, this research is the first study to apply NMF to simplify the network structure after IBL in order to reduce the system error.

This research is organized as follows. Section II states the proposed diagnostic framework while broad learning and the related theory are illustrated in Section III. The experimental setup and data per-processing are presented in Section IV. Section V presents the experimental results and comparisons

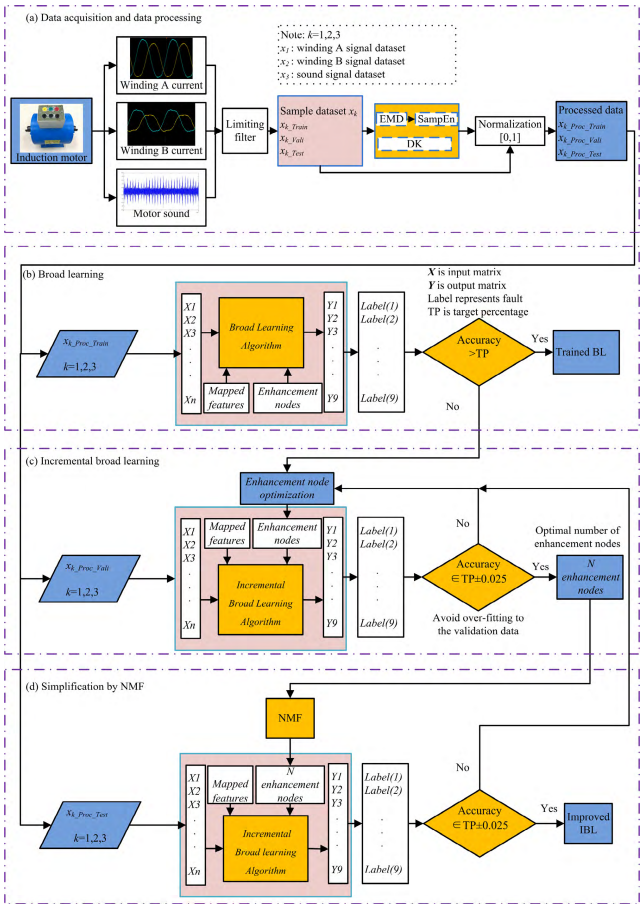


FIGURE 1. Proposed diagnostic framework for TPIMs based on incremental broad learning (from data processing, over training, to testing).

with previous approaches. Finally, conclusions are drawn in the last section.

II. PROPOSED DIAGNOSTIC FRAMEWORK

The proposed TPIM fault diagnostic framework consists of four sub-modules: (a) data acquisition and data processing, (b) broad learning, (c) incremental broad learning, (d) structure simplification by NMF (Fig. 1).

A. DATA ACQUISITION

The data acquisition sub-module digitally detects four different signals: sound waves, as well as windings A, B, and C currents which are denoted as $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$, and $\boldsymbol{x}_4$, respectively. As the TPIM is supplied with three-phase symmetrical currents, it only needs two stator currents [22]. Thus, three only signals are eventually used for demonstration purposes. They are denoted as $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and $\boldsymbol{x}_3$, respectively. A band-limiting filter is also used to reduce interference [23]. The sound-signal $\boldsymbol{x}_1$, for instance, is divided into three independent groups: (i) training dataset, (ii) validation dataset, and (iii) test dataset, all of which are collected from experiments.

B. DATA PROCESSING

We use empirical mode decomposition (EMD) to decompose raw signals. The training dataset is denoted as $\boldsymbol{x}_{\boldsymbol{k}-EMD-Train}$, the validation dataset as $\boldsymbol{x}_{\boldsymbol{k}-EMD-Vali}$, and the test dataset as $\boldsymbol{x}_{\boldsymbol{k}-EMD-Test}$. To remove irrelevant and redundant information from the extracted features, an effective statistical algorithm called Sample Entropy (SampEn) is applied to select the features of interest from $\boldsymbol{x}_{\boldsymbol{k}-EMD-Train}$, $\boldsymbol{x}_{\boldsymbol{k}-EMD-Vali}$, and $\boldsymbol{x}_{\boldsymbol{k}-EMD-Test}$. The results are then saved as $\boldsymbol{x}_{\boldsymbol{k}-SE-Train}$, $\boldsymbol{x}_{\boldsymbol{k}-SE-Vali}$ and $\boldsymbol{x}_{\boldsymbol{k}-SE-Test}$, respectively. Features are normalized to ensure that they contribute equally. Finally, the processed training dataset, validation dataset and test dataset are re-named to $\boldsymbol{x}_{\boldsymbol{k}-Proc-Train}$, $\boldsymbol{x}_{\boldsymbol{k}-Proc-Vali}$ and $\boldsymbol{x}_{\boldsymbol{k}-Proc-Test}$ respectively. In order to diagnose faults efficiently, proper domain knowledge (DK) features [24] are added to the signals for sound, winding A current, and winding B current, respectively.

C. BROAD LEARNING

The broad learning model is trained using the processed training dataset $\boldsymbol{x}_{\boldsymbol{k}-Proc-Train}$. The output of the trained BL model is evaluated by quantifying the training accuracy. The model will complete successfully if the training accuracy reaches a manually chosen target percentage. Otherwise, the model is entered the phase of increment broad learning by increasing its number of enhancement nodes.

D. INCREMENTAL BROAD LEARNING

The validation dataset $\boldsymbol{x}_{\boldsymbol{k}-Proc-Vali}$ serves as input to the incremental broad learning sub-module. The number of enhancement nodes is increased dynamically to increase accuracy. The IBL model leads to over-fitting if the number of enhancement nodes is very large. To avoid this, an optimal number of enhancement nodes, N, should be determined. This process is manually done by trial and error until the validation accuracy meets the target percentage.

E. STRUCTURE SIMPLIFICATION

It should be noted that once the learning system completes, the network may contain some redundancy nodes (due to its broad expansion) which can be removed by simplifying the system through low-rank approximations. The NMF is applied to compress the structure of IBL in order to further reduce the system error. To ensure the diagnostic accuracy after compression, the system is returned to the incremental broad learning sub-module if the diagnostic accuracy does not meet the range of $[\mathrm{TP} \pm 0.025]$.

III. BROAD LEARNING AND RELATED THEORIES

A. BROAD LEARNING THEORY

Broad learning theory [15] is based on the traditional random vector functional-link neural network (RVFLNN). Unlike the typical RVFLNN that takes its input directly and establishes enhancement nodes, BL uses the input to construct a set of mapped features. This BL model presents the input data $\boldsymbol{X}$ and projects the data, using $\phi_i(\boldsymbol{X}\boldsymbol{W}_{ei} + \boldsymbol{\beta}_{ei})$, to become the ith mapped feature $\boldsymbol{Z}_i$, where $\boldsymbol{W}_{ei}$ are the random weights. The concatenation of the first i group of mapping features is

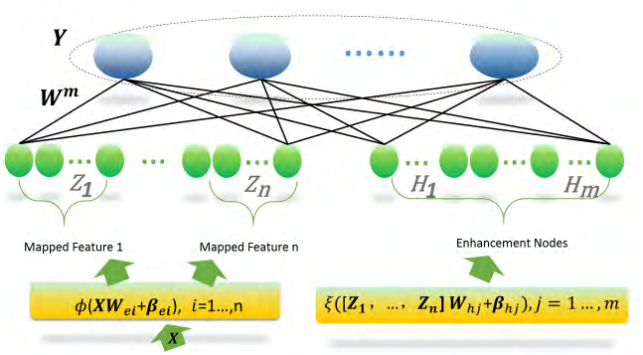


FIGURE 2. Broad learning network.

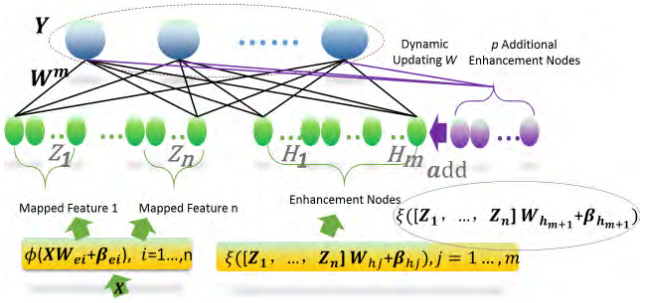


FIGURE 3. Increment of p additional enhancement nodes.

denoted as $\boldsymbol{Z}^i \equiv [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_i]$. In a similar fashion, the jth group of enhancement nodes, $\xi_j(\boldsymbol{Z}^i\boldsymbol{W}_{hj} + \boldsymbol{\beta}_{hj})$, is denoted as $\boldsymbol{H}_j$, and the concatenation of the first j group of enhancement nodes is denoted as $\boldsymbol{H}^j \equiv [\boldsymbol{H}_1, \ldots, \boldsymbol{H}_j]$. In practice, i and j can be selected differently depending upon the complexity of the modeling task. Furthermore, ϕ_i and ϕ_k can be different functions for $i \neq k$. Similarly, ξ_j and ξ_r can be different functions for $j \neq r$. Without loss of generality, the subscripts of the ith random mappings ϕ_i and the jth random mappings ξ_j are omitted.

For the BL model to take advantage of the sparse auto encoder characteristics, it applies the linear inverse problem and adjusts the initial $\boldsymbol{W}_{ei}$ to obtain better features. The details of the algorithm are given below.

Let $\boldsymbol{X}$ be the input dataset that consisting of Q samples, each with M dimensions, and $\boldsymbol{Y}$ is the output matrix with $\boldsymbol{Y} \in \mathbb{R}^{Q\times C}$. For n feature mappings, it can be represented as:

$$\boldsymbol{Z}_i = \phi(\boldsymbol{X}\boldsymbol{W}_{ei} + \boldsymbol{\beta}_{ei}), \quad i = 1 \ldots, n \tag{1}$$

where $\boldsymbol{W}_{ei}$ and $\boldsymbol{\beta}_{ei}$ are randomly generated. If we denote all feature nodes as $\boldsymbol{Z}^n \equiv [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n]$ and denote the mth group of enhancement nodes as

$$\boldsymbol{H}_m \equiv \xi(\boldsymbol{Z}^n\boldsymbol{W}_{hm} + \boldsymbol{\beta}_{hm}) \tag{2}$$

then the broad learning model can be represented as

$$\boldsymbol{Y} = [\boldsymbol{Z}^n|\boldsymbol{H}^m]\boldsymbol{W}^m \tag{3}$$

$$\boldsymbol{W}^m = [\boldsymbol{Z}^n|\boldsymbol{H}^m]^+\boldsymbol{Y} \tag{4}$$

$\boldsymbol{W}^m$ are the connecting weights for the broad structure and can be computed easily through the ridge regression approximation of $[\boldsymbol{Z}^n|\boldsymbol{H}^m]^+$ using Eq.(4) [15]. Fig. 2 summarizes the broad learning network.

B. THEORY OF INCREMENT ENHANCEMENT NODES

Under certain conditions, additional enhancement nodes can be embedded in a learning network to improve its outcome in case of failing and to accomplish an ideal performance. Let $\boldsymbol{A}^m = [\boldsymbol{Z}^n|\boldsymbol{H}^m]$, and $\boldsymbol{A}^{m+1} \equiv [\boldsymbol{A}^m|\xi\left(\boldsymbol{Z}^n\boldsymbol{W}_{h_{m+1}}\right) + \boldsymbol{\beta}_{h_{m+1}})]$, where $\boldsymbol{W}_{h_{m+1}} \in \mathbb{R}^{nk\times p}$, and $\beta_{h_{m+1}} \in \mathbb{R}^p$, where k denotes the number of nodes of each mapping feature and p is the number of additional enhancement nodes. Linking weights and biases between mapped features and the additional enhancement nodes are created at random. The pseudoinverse of the new matrix is thus:

$$(\boldsymbol{A}^{m+1})^+ = \begin{bmatrix} (\boldsymbol{A}^m)^+ - \boldsymbol{D}\boldsymbol{B}^T \\ \boldsymbol{B}^T \end{bmatrix} \tag{5}$$

$$\boldsymbol{D} = (\boldsymbol{A}^m)^+\xi\left(\boldsymbol{Z}^n\boldsymbol{W}_{h_{m+1}} + \boldsymbol{\beta}_{h_{m+1}}\right) \tag{6}$$

$$\boldsymbol{B}^T = \begin{cases} (\boldsymbol{C})^+ & if\ \boldsymbol{C} \neq 0 \\ \left(1 + \boldsymbol{D}^T\boldsymbol{D}\right)^{-1}\boldsymbol{B}^T\boldsymbol{B}(\boldsymbol{A}^m)^+ & if\ \boldsymbol{C} = 0 \end{cases} \tag{7}$$

$$\boldsymbol{C} = \xi\left(\boldsymbol{Z}^n\boldsymbol{W}_{h_{m+1}}\right) + \boldsymbol{\beta}_{h_{m+1}}) - \boldsymbol{A}^m\boldsymbol{D} \tag{8}$$

where the superscript $+$ represents a pseudoinverse. Again, the new weights are

$$\boldsymbol{W}^{m+1} = \begin{bmatrix} \boldsymbol{W}^m - \boldsymbol{D}\boldsymbol{B}^T\boldsymbol{Y} \\ \boldsymbol{B}^T\boldsymbol{Y} \end{bmatrix} \tag{9}$$

The incremental broad learning structure is shown in Fig. 3. It is noteworthy that the pseudoinverse of the related matrix is a result of the regularization approach. More specifically, fast incremental learning can be achieved when the algorithm calculates the pseudoinverse of the additional enhancement nodes rather than the entire $\boldsymbol{A}^{m+1}$.

C. STRUCTURE SIMPLIFICATION THEORY

Incremental broad learning with increased number of enhancement nodes may run the risk of producing redundancy. Generally speaking, the structure can be simplified by a series of low rank approximation methods. As mentioned previously, NMF is employed to provide the structural simplification.

Since the input dataset is normalized, the connecting weight matrix $\boldsymbol{W}^m$ of IBL is non-negative. With a connecting weight matrix of $\boldsymbol{W}^m \in \mathbb{R}^{n\times m}$, a non-negative matrix $\boldsymbol{I} \in \mathbb{R}^{n\times r}$, and another non-negative matrix $\boldsymbol{W}^r \in \mathbb{R}^{r\times m}$, the following equation is obtained:

$$\boldsymbol{W}^m \approx \boldsymbol{I}\boldsymbol{W}^r \tag{10}$$

where $\boldsymbol{W}^m$ can be decomposed into two small matrices [25]. m is the dimension of enhancement feature values, n is the number of samples, and r is the reduced rank. $\boldsymbol{W}^m$ is an original matrix. The right matrix $\boldsymbol{W}^r$ is called the coefficient

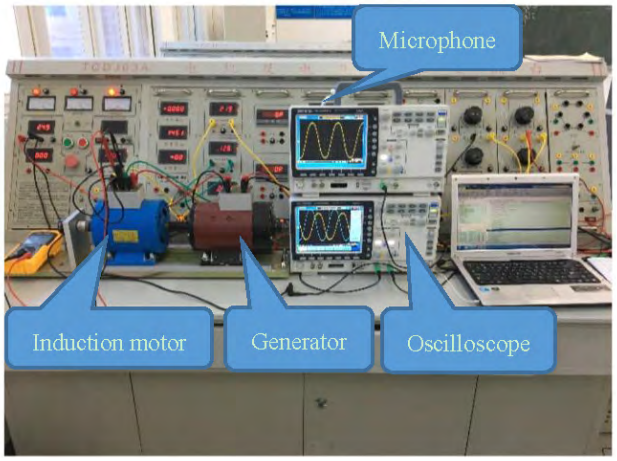


FIGURE 4. Test bench of TPIM.

matrix. The left matrix $\boldsymbol{I}$ is called the basic matrix. The column vectors of the original matrix are the weighted sums of all column vectors in the left matrix, while the weight coefficient is the element of the corresponding column vectors of the right matrix. In general, r should be chosen to be smaller than m, thus, the dimension reduction of the original matrix can be realized and the dimension reduction matrix of data features can be obtained by replacing the original matrix with the coefficient matrix $\boldsymbol{W}^r$:

$$\boldsymbol{W}^r \approx \boldsymbol{I}^+ \boldsymbol{W}^m \tag{11}$$

The resulting connecting weight matrix $\boldsymbol{W}^m$ is thus significantly reduced through NMF.

IV. EXPERIMENTAL SETUP AND DATA PREPROCESSING

A. TEST RIG

We use the motor test bench, TCDJ-03A, as our experimental platform. The experimental setup (Fig. 4) consists of a platform that supplies the three-phase power to a TPIM which in turn drives a generator. This test bench can measure motor voltage, motor current, and motor speed. The TPIM (220V and 1.1A) is wired using a star connection style. The generator is connected to the TPIM through a coupler and the generator behaves as a load that can be adjusted through a variable resistance. A GDS-2202A oscilloscope is used to measure TPIM currents and convert the current signal to digital data. A microphone (Lenovo P121), connected to a computer (Samsung R429), is used to acquire the acoustic data. According to the National Standards of the People's Republic of China (GB3806-81), the distance between the measuring point and the reflecting surface should be no less than one meter and the angle between measuring points in the microphone and airflow direction should be 45°.

B. TEST SCHEME

The literature mentions mainly nine different fault cases for TPIMs [26], [27], with only one normal case. The fault cases are phase imbalance, short circuit, mechanical imbalance, bent rotor, broken rotor bar, bearing outer raceway defect, bearing inner raceway defect, and broken bearing ball. In order to offer training data to the classification system, each of the nine possible faults should be simulated through artificial destruction (see Fig 5).

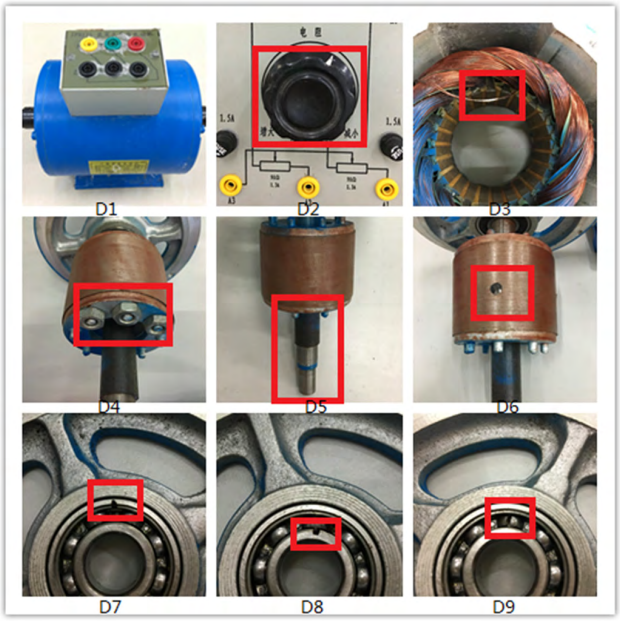


FIGURE 5. Nine fault cases of a TPIM. D1: Normal. D2: Phase imbalance. D3: Short circuit in stator winding A. D4: Mechanical imbalance. D5: Bent rotor. D6: Broken rotor bar. D7: Bearing outer raceway defect. D8: Bearing inner raceway defect. D9: Broken bearing ball.

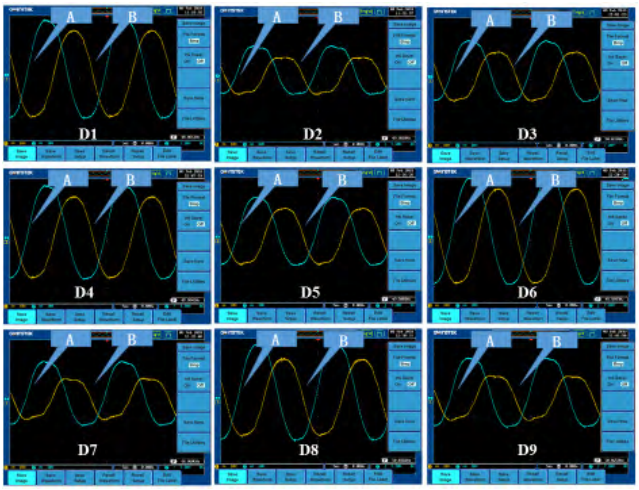


FIGURE 6. Waveform of nine different TPIM winding currents A & B under rated-load conditions (2.5 waveform periods).

Each case is tested for three electric load conditions: (i) an underload condition with a 0.8× rated current; (ii) a rated load, and (iii) an overload condition with a 1.2× rated.

C. SAMPLE DATA ACQUISITION

To construct and test the proposed diagnostic method, the sample data must be acquired first. The raw signals, x_k

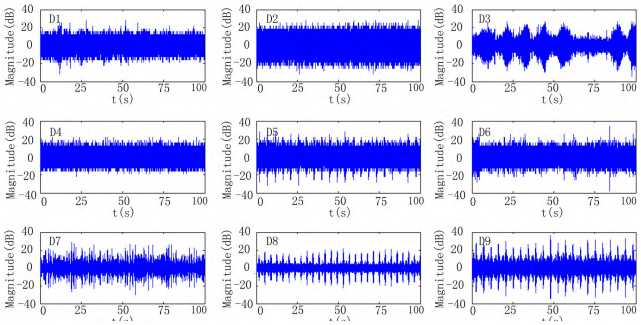


FIGURE 7. Raw acoustic data of nine fault cases under rated load condition (data size: 8 × 10^5 data in 100 seconds).

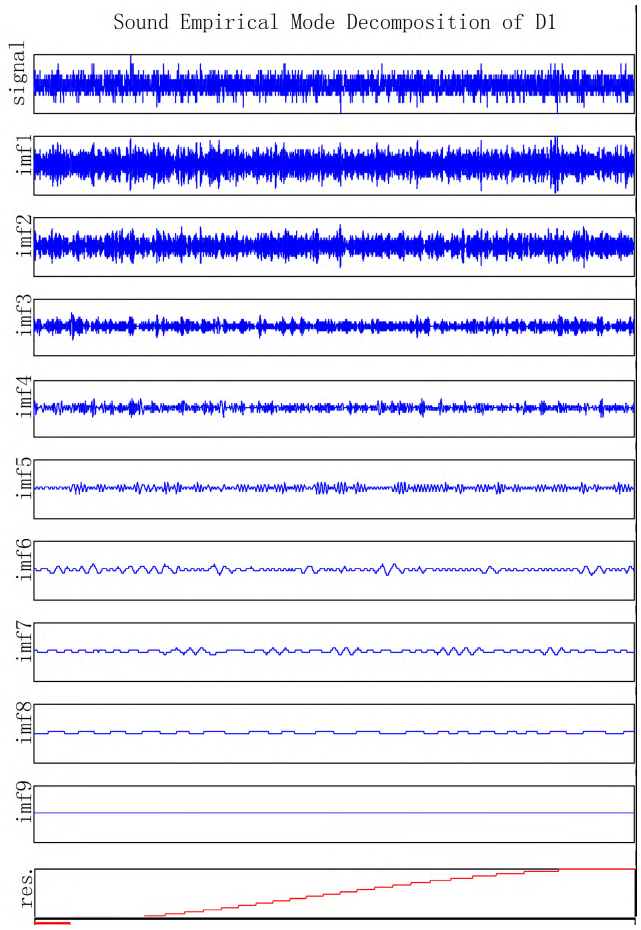


FIGURE 8. Empirical mode decomposition of acoustic signal in normal condition D1 (showing one sample with 1000 data points).

($k = 1, 2, 3$), are simultaneously recorded by lab equipment. FIGURE 6 shows the current waveform of windings A & B. FIGURE 7 shows the raw acoustic data of the nine fault cases under the rated load condition.

For each fault case, every electric load condition is tested for 100 seconds generating 800,000 data points. In other words, one case of each signal has 800,000 data points under one load condition, which are then divided into 800 distinct samples (of 1000 data points per sample).

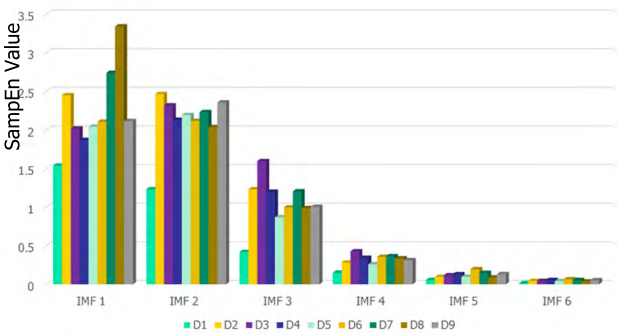


FIGURE 9. SampEn value for different fault conditions and six IMFs.

D. FEATURE EXTRACTION BY EMD AND SAMPEN

EMD decomposes the acquired signal into q intrinsic mode functions (IMFs) and a residual signal leaving the dimension of each IMF unchanged (Fig. 8). Each classifier therefore has a large number of input dimensions, which results in a poor fault classification accuracy [28]. To overcome this problem, an effective feature selection method is considered which reduces the number of input dimensions. We use Sample Entropy (SampEn) as an effective statistical algorithm to calculate the representative features from each IMF, such that the input dimension of each classifier is reduced from 1000 to 9. After several trials, it is found that the main features of the motor faults are closely related to the first six IMFs (Fig. 8). The SampEn values of IMF1 to IMF6 are much greater than those of IMF7 to IMF9. Theory suggests that a large value of SampEn signifies a high information content [29]. We therefore only retain the SampEn of IMF1 to IMF6 to represent the signal features, while the others are discarded to reduce the input dimension to each fault classifier.

The SampEn values for the nine fault cases exhibit different SampEn values for each IMF (Fig. 9). The feature is expected to be observable. It is very useful for classification.

E. DOMAIN KNOWLEDGE FEATURES

After completing the feature extraction by EMD and SampEn, additional domain features are extracted for the diagnosis as they contain important information for a specific field. The motor stator current signal, minimum amplitude, maximum amplitude, and average current voltage are used as domain knowledge (DK) features. In addition, ten statistical time-domain features are employed as DK features to analyze the motor acoustic signal [30] (Tab. 1).

F. DISTRIBUTION OF PROCESSED DATA

In general, the raw data are divided into three subsets: training dataset, validation dataset, and test dataset. In order to improve the diagnostic performance, the raw data are processed by EMD, SampEn, DK, and normalization. Then, each sample data set, $\boldsymbol{x_k}$, is divided into different subsets (see Tab. 2).

TABLE 1. Definition of common statistics in time-domain for acoustic signal [31].

Time-domain feature	Equation	Time-domain feature	Equation
1.Mean	$\frac{1}{N}\sum_{i=1}^{N} x_i$	6.Kurtosis	$\frac{\sum_{i=1}^{N}(x_i - x_m)^4}{(N-1)x_{std}^4}$
2.Standard deviation	$\sqrt{\frac{\sum_{i=1}^{N}(x_i - x_m)^2}{N-1}}$	7.Crest factor	$\frac{x_{pk}}{x_{rms}}$
3.Root mean square	$\sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$	8.Clearance factor	$\frac{x_{pk}}{\left(\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert\right)}$
4.Peak	$max\lvert x_i\rvert$	9.Shape factor	$\frac{x_{rms}}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$
5.Skewness	$\frac{\sum_{i=1}^{N}(x_i - x_m)^3}{(N-1)x_{std}^3}$	10.Impulse factor	$\frac{x_{pk}}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$

TABLE 2. Division of sample dataset into different subsets in one fault case where *Dk-Proc-Train* denotes a processed training dataset, *Dk-Proc-Vali* a processed validation dataset, and *Dk-Proc-Test* a processed test dataset.

	Type of dataset	Name of dataset (Sample size×Three load conditions)
Raw sample dataset $x_k(k=1,2,3)$	Training dataset	*Dk-Train(600×3)*
	Validation dataset	*Dk-Vali(100×3)*
	Test dataset	*Dk-Test(100×3)*
Processed sample dataset based on EMD+SampEn, DK and Normalization	Processed training dataset	*Dk-Proc-Train(600×3)*
	Processed validation dataset	*Dk-Proc-Vali(100×3)*
	Processed test dataset	*Dk-Proc-Test(100×3)*

V. EXPERIMENTAL RESULTS AND DISCUSSION

A. RESULT AND DISCUSSION OF DIFFERENT FEATURE EXTRACTION TECHNIQUES

Four kinds of feature extraction techniques are examined: FFT, STFT, EMD + SampEn, and EMD+SampEn+DK. Several parameters must be set *a priori*. In terms of DK, winding currents A and B and the sound data have 3, 3, and 10 DK features, respectively. For FFT, the feature size is set to 8000 as suggested by [32]. With STFT, the window function is a hamming window [17]. In the EMD+SampEn extraction, m and r need to be set in SampEn and we select their values as 2 and 0.2, respectively [19]. TABLE 3 indicates the number of extracted features by using EMD + SampEn and DK. As shown in TABLE III, the sound signal contains 16 features which consist of 6 EMD+SampEn features (i.e. IMF1, IMF2, IMF3, IMF4, IMF5, IMF6) and 10 DK features which is calculated by the equations in TABLE I. The winding A current signal contains 9 features which consist of 6 EMD+SampEn features (i.e. IMF1, IMF2, IMF3, IMF4, IMF5, IMF6) and 3 DK features. The three DK features are maximum, minimum and average values of winding A current respectively. The type of features of winding B current signal is the same as that of winding A current signal. To ensure that all features contribute equally, all reduced features are normalized [33].

TABLE 3. Number of extracted features from different signal types.

Induction motor signal type	EMD + SampEn	DK	Total features
Sound signal	6	10	16
Winding A current signal	6	3	9
Winding B current signal	6	3	9

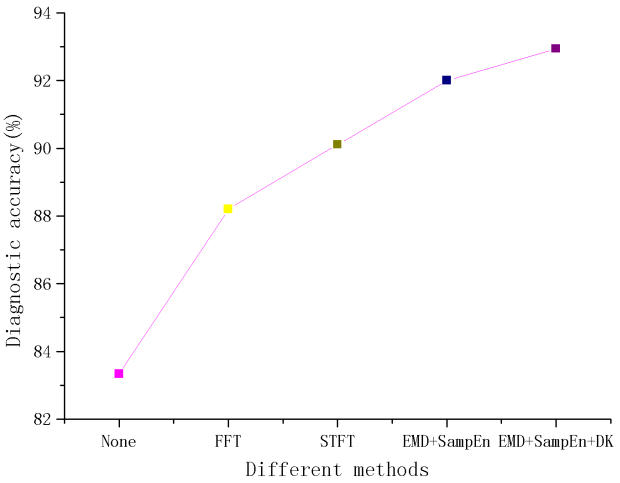


FIGURE 10. Diagnostic accuracies of different feature extractions.

In order to prove the effectiveness of the EMD + SampEn and DK techniques, we compare four different methods: FFT, STFT, EMD + SampEn, and EMD + SampEn + DK. The comparison is performed in IBL with 100 feature nodes and 240 enhancement nodes. Fig. 10 shows the best diagnostic accuracy is the EMD + SampEn + DK approach. The worst accuracy is obtained without the use of any feature extraction technique. The main reason for this result is that EMD is a self-adaptive time-frequency technique that can decompose the signal into several IMFs and adapt to the signal itself.

B. RESULT AND DISCUSSION OF DIFFERENT CLASSIFICATION METHODS

To prove the effectiveness of IBL, we compare the performance of the following different classification methods: deep belief network (DBN), convolutional neural network (CNN), extreme learning machine (ELM), broad learning (BL), and

incremental broad learning (IBL). For DBN, we employ a 5400-100-240-9 structure with 100 iteration and a learning rate of 0.1 [34]. For CNN, we use an original deep CNN structure (LeNet-5) which contains 7 layers. The kernel size of convolution layer 1 is 5×5. The pooling size of sampling layer 2 is 2×2. The kernel size of convolution layer 3 is 5×5. The pooling size of sampling layer 4 is 2×2. The kernel size of convolution layer 5 is 5×5. The number of nodes of the fully connected layer is 84. The number of nodes of the output layer is 9. The learning rate is set as 0.1, while the activation function is sigmoid [35]. For the extreme learning machine, the number of hidden node is set as 340; the activation function is configured as a sigmoid function [36]. As for broad learning, it follows the suggestion of [15]. In other words, the regularization parameter, λ, for the ridge regression is set to 10^{-8}. For the enhancement nodes, a sigmoid function is used to establish BL. Meanwhile, the associated parameters $\boldsymbol{W}_{ei}$ and $\boldsymbol{\beta}_{ei}$, for $i = 1, \ldots, n$ are drawn from a standard uniform distribution from within the interval [0, 1]. For the incremental broad learning, 30 additional enhancement nodes are added.

The results show that the highest accuracy is obtained for IBL reaching 92.94% (Tab. 4). The second highest is obtained for DBN with 92.71%, although the training cost is high with 378.7479 seconds. ELM with 340 hidden nodes produces the fastest training time but only low accuracy. The accuracy of ELM is improved to 92.52% by using 3000 instead of 340 hidden nodes, but the increase in hidden nodes comes at the expense of an increased training time. In contrast, BL reaches an accuracy of 91.61%, which is similar to DBN, while requiring a very short training time of just 1.0018 seconds, making BL the most efficient method in the comparison. This is because DBN or CNN can only improve their accuracy through a deep structure which is more time intensive for training. BL, on the other hand, only contains two layers which explains its much shorter training time. ELM has three layers, which also results in short training time. Even though the accuracy of BL is not the best (Tab. 4), its accuracy can be improved to 92.94 by using incremental broad learning (IBL) to add 30 enhancement nodes, at a low cost of an additional training time of 0.0497 seconds.

C. RESULT AND DISCUSSION OF DIAGNOSTIC ACCURACY

While IBL can improve accuracy efficiently by increasing the number of enhancement nodes, it suffers from over-fitting. Therefore, the number of enhancement nodes should be optimized.

In the beginning, we fix the number of feature nodes at 100 and start the optimization with an initial number of 30 enhancement nodes, to which the model is expanded with an additional 30 enhancement nodes at every iteration until reaching 300 nodes (after 10 iteration). While the initial accuracy of the BL model starts out as being the lowest (Fig. 11), it increases as the number of enhancement nodes is increased reaching a maximum value of 92.94% for 240 enhancement nodes. A further increase in nodes leads to a drop in accuracy due to over-fitting. The optimal number of enhancement nodes is thus around 240, which requires 1 original training and 7 incremental trainings to obtain.

In order to ensure that the best enhancement node number is close to 240, additional more number of enhancement nodes between 210 and 270 is also tested. Staring with 210, the model is further expanded with an additional 10 enhancement nodes at every iteration until reaching 270. The interval between 210 and 270 in Fig.11 shows the extra test result and the optimal number of enhancement nodes is indeed 240.

TABLE 4. Comparison of diagnostic accuracies and training time for different methods.

Method	Accuracy based on test dataset (%)	Total training time (s)
DBN	92.71	378.7479
CNN (1 epoch)	87.25	63.6417
CNN (10 epochs)	92.27	642.3905
ELM (340 hidden nodes)	72.30	0.3687
ELM (3000 hidden nodes)	92.52	65.5183
BL (310 nodes including 100 feature nodes and 210 enhancement nodes)	91.61	1.0018
IBL(100 feature nodes and 240 enhancement nodes)	**92.94**	**1.0018+0.0497 =1.0515**

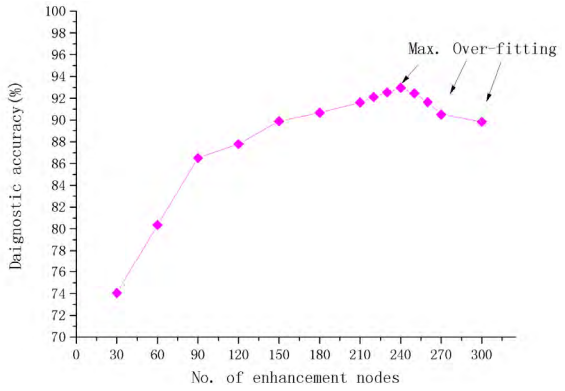


FIGURE 11. Accuracies of different numbers of enhancement nodes.

D. RESULT AND DISCUSSION OF RETRAINING TIME

To prove the effectiveness of retraining time of increment broad learning, the training time and retraining time are compared for the same number of enhancement nodes (100). The training time is the longest (0.71006 seconds) for the initial 30 enhancement nodes, while being much shorter for all subsequent iteration (Fig. 12). This is because the IBL algorithm is not necessary to change its former weight. It merely needs to calculate the weight of additional nodes which results in faster retraining times. The total training time for determining optimal accuracy is thus 1.0515 seconds requiring 8 iteration.

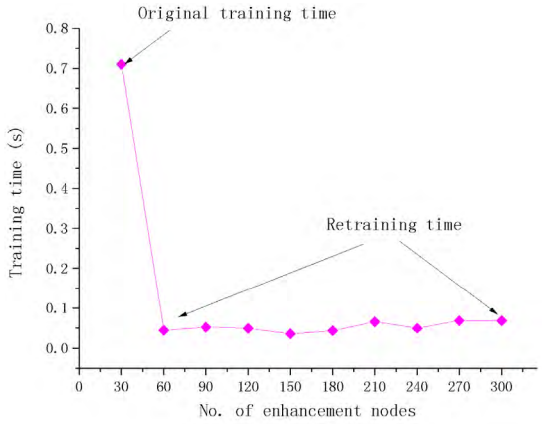


FIGURE 12. Training time of increment broad learning.

E. RESULT AND DISCUSSION OF STRUCTURE SIMPLIFICATION

After incremental broad learning, the expanded nodes are redundant. This complicated structure may increase the model diagnostic error. The NMF algorithm is applied to compress its structure and reduce its diagnostic error. This is illustrated by comparing the performance of IBL, SVD-IBL, and NMF-IBL (Tab. 5). In SVD-IBL and NMF-IBL, the network structure is compressed to 90% enhancement nodes. These three diagnostic systems are tested ten times. Then, the root mean square error (RMSE) and the standard deviation (SD) are calculated as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(x_{obs,i} - x_{model,i}\right)^2}{N}} \tag{12}$$

$$SD = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N-1}} \tag{13}$$

RMSE is very sensitive to the reflection of large or small errors in a group of measurements, and can therefore be a good indicator for the precision of the model. $x_{obs,i}$ denotes the accuracy of the i^{th} test. $x_{model,i}$ denotes the true model accuracy. The smaller the RMSE, the higher the precision of the model.

The RMSEs for these three systems are reduced as the total number of nodes is increased (Tab. 5). This is because the higher the number of nodes the higher the accuracy (except when over-fitting). After many trials, RMSE can be reduced to 0.2467 for a total number of nodes of 340 (100, 240). This means that it has enough nodes to ensure its accuracy. The SD for NMF-IBL is less than for IBL or SVD-IBL because NMF can effectively reduce the number of redundant nodes, leading to a more simplified structure. SVD-IBL can reduce the number of enhancement nodes but is unstable. The RSME of the NMF-IBL system is improved from 0.3878 to 0.2467 and the SD from 0.4628 to 0.2412 when the structure of nodes is simplified from (100, 240) to (100, 216).

TABLE 5. Performance comparison of IBL and NMF-IBL.

	Ω	RMSE	SD
IBL (Incremental broad learning without compression)	(100,120)-without compression	0.6398	0.6070
	(100,150)-without compression	0.5481	0.5621
	(100,180)-without compression	0.4871	0.5491
	(100,210)-without compression	0.4734	0.5251
	(100,240)-without compression	0.3878	0.4628
SVD-IBL [15] (Singular value decomposition-incremental broad learning)	(100,120)-(100,108)	0.6372	0.6047
	(100,150)-(100,135)	0.5475	0.5601
	(100,180)-(100,162)	0.4824	0.5489
	(100,210)-(100,189)	0.4731	0.5196
	(100,240)-(100,216)	0.3868	0.4559
NMF-IBL (non-negative matrix factorization incremental broad learning)	(100,120)-(100,108)	0.5886	0.4920
	(100,150)-(100,135)	0.5271	0.3967
	(100,180)-(100,162)	0.4460	0.3626
	(100,210)-(100,189)	0.3842	0.3528
	(100,240)-(100,216)	**0.2467**	**0.2412**

Remarks: (Number of initial nodes, Number of enhancement nodes) (·,·)-(·,·): Reduced from (·,·) to (·,·)

VI. CONCLUSIONS

In this research, a novel diagnostic framework to efficiently detect TPIM faults is proposed. A framework combining feature extraction, broad learning, incremental broad learning and NMF-IBL, is developed to successfully improve the accuracy, training and retraining time. Raw training data of the TPIM acoustic signal and winding A & B currents are acquired experimentally before processing the data using filters, EMD, SampEn, DK, and normalization. In a second step, the processed data are used for training a diagnostic system by using a broad learning algorithm. The model can be retrained through incremental broad learning if the desired accuracy is not achieved. Lastly, the system structure is simplified by applying NMF. The effectiveness of the proposed scheme is verified by using a test rig. The experimental results clearly demonstrate that incremental broad learning is very efficient

in terms of diagnostic accuracy and training speed. The main original results of this research can be summarized as follows:

1. A novel diagnostic framework based on IBL is proposed to diagnose both stator and rotor faults in TPIMs.
2. A feature extraction approach for processing raw data, in combination with EMD, SampEn, DK, is applied to improve diagnostic accuracy.
3. An incremental broad learning method is applied to retrain the diagnostic system and improve both accuracy and training speed. Although DBN and CNN are capable of improving system accuracy, their training time is very long. ELM can train quickly, but its accuracy is poor unless increasing the number of hidden nodes; even so, it still needs a long training time.
4. To our best knowledge, this is the first study to apply NMF to simplify the IBL structure in order to reduce RMSE and SD. The resulting NMF-IBL model is more accurate than the IBL without NMF.

The proposed method can also be applied to other similar motor diagnostic problems like direct current motors or permanent magnet synchronous motors. However, a lot of work still needs to be done, including the improvement of the EMD+SampEn and DK feature extraction method. We only compare four kinds of feature extraction methods and further work is needed to extend this comparison to include different methods. The number of incremental nodes is determined manually which can be improved by devising an automatic selection through particle swarm optimization or a similar algorithm. To improve accuracy, one could also try to increase the number of feature nodes or inputs instead of increasing the number of enhancement nodes. Lastly, NMF is a low-rank approximation to simplify the structure of IBL and additional study is needed to find a better alternative.

REFERENCES

[1] A. Siddique, G. S. Yadava, and B. Singh, "A review of stator fault monitoring techniques of induction motors," *IEEE Trans. Energy Convers.*, vol. 20, no. 1, pp. 106–114, Mar. 2005.

[2] R. N. Bell, D. W. Mcwilliams, P. O'Donnell, C. Singh, and S. J. Wells, "Report of large motor reliability survey of industrial and commercial installations. II," *IEEE Trans. Ind. Appl.*, vol. 21, no. 4, pp. 865–872, Jul. 1985.

[3] B. Cai, Y. Zhao, H. Liu, and M. Xie, "A data-driven fault diagnosis methodology in three-phase inverters for PMSM drive systems," *IEEE Trans. Power Electron.*, vol. 32, no. 7, pp. 5590–5600, Jul. 2017.

[4] X. Dai and Z. Gao, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2226–2238, Nov. 2013.

[5] S. Moon, H. Jeong, H. Lee, and S. W. Kim, "Interturn short fault diagnosis in a PMSM by voltage and current residual analysis with the faulty winding model," *IEEE Trans. Energy Convers.*, vol. 33, no. 1, pp. 190–198, Mar. 2018.

[6] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, pp. 1527–1554, Jul. 2006.

[7] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[8] R. Salakhutdinov and G. Hinton, "An efficient learning procedure for deep Boltzmann machines," *Neuro Comput.*, vol. 24, no. 8, pp. 1967–2006, Apr. 2012.

[9] P. Gangsar and R. Tiwari, "Multifault diagnosis of induction motor at intermediate operating conditions using wavelet packet transform and support vector machine," *J. Dyn. Syst. Meas. Control*, vol. 140, no. 8, 2018, Art. no. 081014.

[10] P. K. Wong, Z. Yang, C. M. Vong, and J. Zhong, "Real-time fault diagnosis for gas turbine generator systems using extreme learning machine," *Neurocomputing*, vol. 128, pp. 249–257, Mar. 2014.

[11] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.

[12] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016.

[13] I. Bravo-Imaz, H. D. Ardakani, Z. Liu, A. Garcia-Arribas, A. Arnaiz, and J. Lee, "Motor current signature analysis for gearbox condition monitoring under transient speeds using wavelet analysis and dual-level time synchronous averaging," *Mech. Syst. Signal Process.*, vol. 94, pp. 73–84, Sep. 2017.

[14] C. L. P. Chen, C.-Y. Zhang, L. Chen, and M. Gan, "Fuzzy restricted boltzmann machine for the enhancement of deep learning," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 6, pp. 2163–2173, Dec. 2015.

[15] C. L. P. Chen and Z. L. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 10–24, Jan. 2018.

[16] A. Prudhom, J. Antonino-Daviu, H. Razik, and V. Climente-Alarcon, "Time-frequency vibration analysis for the detection of motor damages caused by bearing currents," *Mech. Syst. Signal Process.*, vol. 84, pp. 747–762, Feb. 2017.

[17] J. Burriel-Valencia, R. Puche-Panadero, J. Martinez-Roman, A. Sapena-Bano, and M. Pineda-Sanchez, "Short-frequency Fourier transform for fault diagnosis of induction machines working in transient regime," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 3, pp. 432–440, Mar. 2017.

[18] N. Bessous, S. E. Zouzou, W. Bentrah, S. Sbaa, and M. Sahraoui, "Diagnosis of bearing defects in induction motors using discrete wavelet transform," *Int. J. Syst. Assurance Eng. Manage.*, vol. 9, pp. 335–343, Apr. 2018.

[19] H. Han, S. Cho, S. Kwon, and S. B. Cho, "Fault diagnosis using improved complete ensemble empirical mode decomposition with adaptive noise and power-based Intrinsic mode function selection algorithm," *Electron.*, vol. 7, pp. 1–16, Jan. 2018.

[20] M. Ma, L. Guo, K. Su, and D. Liang, "Classification of motor imagery EEG signals based on wavelet transform and sample entropy," in *Proc. IEEE 2nd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Mar. 2017, pp. 905–910.

[21] Z. Yang, Y. Xiang, K. Xie, and Y. Lai, "Adaptive method for nonsmooth nonnegative matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 4, pp. 948–960, Apr. 2017.

[22] S. Chattopadhyay, S. Karmakar, M. Mitra, and S. Sengupta, "Symmetrical components and current Concordia based assessment of single phasing of an induction motor by feature pattern extraction method and radar analysis," *Int. J. Elect. Power Energy Syst.*, vol. 37, pp. 43–49, May 2012.

[23] J. G. Jones and M. J. Corbin, "Analytical redundancy using band-limiting filters," *IEE Proc. D Control Theory Appl.*, vol. 135, no. 4, pp. 257–267, Jul. 1988.

[24] A. Widodo *et al.*, "Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine," *Expert Syst. Appl.*, vol. 36, pp. 7252–7261, Apr. 2009.

[25] G. R. Naik, *Non-negative Matrix Factorization Techniques: Advances in Theory and Applications*. Berlin, Germany: Springer, 2015.

[26] M. Seera, C. P. Lim, D. Ishak, and H. Singh, "Fault detection and diagnosis of induction motors using motor current signature analysis and a hybrid FMM–CART model," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 1, pp. 97–108, Jan. 2012.

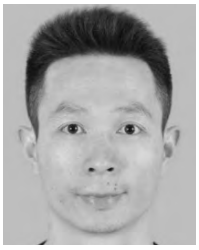
[27] B. S. Yang and K. J. Kim, "Application of Dempster–Shafer theory in fault diagnosis of induction motors using vibration and current signals," *Mech. Syst. Signal Process.*, vol. 20, no. 2, pp. 403–420, 2006.

[28] P. K. Wong, J. Zhong, Z. Yang, and C. M. Vong, "Sparse Bayesian extreme learning committee machine for engine simultaneous fault diagnosis," *Neurocomputing*, vol. 174, pp. 331–343, Jan. 2016.

[29] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol. Heart Circulatory Physiol.*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[30] C. Wu, L. Ning, R. Jiang, X. Wu, and J. Liu, "Intelligent identification of bearing faults using time domain features," in *Proc. 4th Int. Conf. Digit. Manuf. Automat.*, Qindao, China, Jun. 2013, pp. 713–716.

[31] J.-H. Zhong, P. K. Wong, and Z.-X. Yang, "Fault diagnosis of rotating machinery based on multiple probabilistic classifiers," *Mech. Syst. Signal Process.*, vol. 108, pp. 99–114, Apr. 2018.
[32] T. Vaimann and A. Kallaste, "Detection of broken rotor bars in three-phase squirrel-cage induction motor using fast Fourier transform," in *Proc. 10th Int. Symp. Topical Problems Field Electr. Power Eng.*, Pärnu, Estonia, 2011, pp. 52–56.
[33] P. K. Wong, J. H. Zhong, Z. X. Yang, and C. M. Vong, "A new framework for intelligent simultaneous-fault diagnosis of rotating machinery using pairwise-coupled sparse Bayesian extreme learning committee machine," *Proc. Inst. Mech. Eng., C, J. Mech. Eng. Sci.*, vol. 231, no. 6, pp. 1146–1161, 2016.
[34] H. Shao, H. Jiang, H. Zhang, W. Duan, T. Liang, and S. Wu, "Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing," *Mech. Syst. Signal Process.*, vol. 100, pp. 743–765, Feb. 2018.
[35] L. H. Wang, X. P. Zhao, J. X. Wu, Y. Y. Xie, and Y. H. Zhang, "Motor fault diagnosis based on short-time Fourier transform and convolutional neural network," *Chin. J. Mech. Eng.*, vol. 30, pp. 1357–1368, Nov. 2017.
[36] Z.-X. Yang, X.-B. Wang, and J.-H. Zhong, "Representational learning for fault diagnosis of wind turbine equipment: A multi-layered extreme learning machines approach," *Energies*, vol. 9, no. 6, pp. 379–396, 2016.



SAI BIAO JIANG received the M.Sc. degree from the Guangdong University of Technology, in 2011. He is currently pursuing the Ph.D. degree in electromechanical engineering with the University of Macau. He is also a Lecturer with the Zhuhai College of Jilin University. His research interests include data-driven fault diagnosis, feature extraction, and artificial intelligence.

PAK KIN WONG received the Ph.D. degree in mechanical engineering from The Hong Kong Polytechnic University, Hong Kong, in 1997. He is currently a Professor with the Department of Electromechanical Engineering and the Associate Dean (Academic Affairs) of the Faculty of Science and Technology, University of Macau. He has published over 218 scientific papers in refereed journals, book chapters, and conference proceedings. His research interests include automotive engineering, fluid transmission and control, artificial intelligence (AI), mechanical vibration, and AI for medical applications.

RENCHU GUAN received the Ph.D. degree in computer science and technology from Jilin University, Changchun, China, in 2010, where he is currently an Associate Professor. He was a Visiting Professor with the University of Trento, Italy, from 2011 to 2012, and with the University of Arkansas at Little Rock, from 2017 to 2018. He is also a Special Term Professor with the School of Computer, Zhuhai College, Jilin University. His research was featured in the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, *Knowledge-Based System*, and the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. His research interests include machine learning, text mining, and bioinformatics.

YANCHUN LIANG received the Ph.D. degree in applied mathematics from Jilin University, Changchun, China, in 1997, where he is currently a Professor with the College of Computer Science and Technology. He was a Visiting Scholar with The University of Manchester, U.K., from 1990 to 1991, a Visiting Professor with the National University of Singapore, from 2000 to 2001, a Guest Professor with the Institute of High Performance Computing of Singapore, from 2002 to 2004, a Guest Professor with Trento University, Italy, from 2006 to 2008, and a Visiting Professor with Missouri University, USA, from 2010 to 2018. He is also a Professor with the School of Computer, Zhuhai College of Jilin University, China. He has published over 400 papers. His research was featured in the IEEE TRANSACTIONS ON SMC, the IEEE TRANSACTIONS ON KDE, the *Journal of Micromechanics and Micro Engineering*, *Physical Review E*, and *Bioinformatics*. His research interests include computational intelligence, machine learning methods, and bioinformatics.

JIA LI received the M.Sc. degree from Kumamoto University, in 2011. She is currently a Lecturer with the Zhuhai College of Jilin University. Her research interests include digital image and biomedical signal processing, feature extracting and classification in artificial intelligence, and earth exploration technology and instrument.