A Sparse Auto-encoder-Based Deep Neural Network Approach for Induction Motor Faults Classification

Wenjun Sun, Siyu Shao, Rui Zhao, Ruqiang Yan, Xingwu Zhang, Xuefeng Chen

PII: S0263-2241(16)30064-1
DOI: http://dx.doi.org/10.1016/j.measurement.2016.04.007
Reference: MEASUR 3944

To appear in: *Measurement*

Received Date: 15 February 2016
Revised Date: 17 March 2016
Accepted Date: 5 April 2016

A Sparse Auto-encoder-Based Deep Neural Network Approach for Induction Motor Faults Classification

Wenjun Sun[1], Siyu Shao[1], Rui Zhao[2], Ruqiang Yan[1, 3,] *, Xingwu Zhang[3] and Xuefeng Chen[3]
[1]School of Instrument Science and Engineering
Southeast University, Nanjing 210096, China
Email: ruqiang@seu.edu.cn
[2]School of Electrical and Electronic Engineering
Nanyang Technological University, Nanyang Avenue, Singapore 639798.
[3] Collaborative Innovation Center of High-End Manufacturing Equipment
Xi'an Jiaotong University, Xi'an 710049, China
*Corresponding Author

Abstract

This paper presents a deep neural network (DNN) approach for induction motor fault diagnosis. The approach utilizes sparse auto-encoder (SAE) to learn features, which belongs to unsupervised feature learning that only requires unlabeled measurement data. With the help of the denoising coding, partial corruption is added into the input of the SAE to improve robustness of feature representation. Features learned from the SAE are then used to train a neural network classifier for identifying induction motor faults. In addition, to prevent overfitting during the training process, a recently developed regularization method called "dropout" which has been proved to be very effective in neural network was employed. An experiment performed on a machine fault simulator indicates that compared with traditional neural network, the SAE-based DNN can achieve superior performance for feature learning and classification in the field of induction motor fault diagnosis.

Keywords: *Sparse auto-encoder; Deep neural network; Fault diagnosis; Denoising; Dropout*

1. INTRODUCTION

Induction motor as one of the industrial power driving sources, occupies an important position in national economy and has been widely applied to driving many kinds of machinery and industrial equipment, such as lifting hoist equipment, mining equipment, machine tools, etc. In order to guarantee normal operation of induction motors with timely maintenance, and avoid unnecessary loss, fault diagnosis on them is necessary [1, 2]. However, due to environmental interference and inherent motor structure complexity, effective fault diagnosis for induction motors is challenging. Up to date, various sensing techniques [3-5] have been employed to measure certain physical quantities, including current, vibration, radial and axial flux, rotor speed, etc., for the purpose of identifying induction motor faults. This is based on the fact that properties of the induction motor may change when there are some damages occurred in it and such property change could be reflected in the measurement data. By extracting features from these measurement data, a classifier can then be built to distinguish different faults in induction motors. Therefore, the fault diagnosis problem can be converted into a classification problem which can be solved by machine learning-based algorithms such as neural network [6, 7] and support vector machine [8, 9].

In the field of fault diagnosis, the majority of machine learning algorithms are supervised learning which needs a large amount of labeled data for training. Since neural networks possess strong representation ability due to its stacked hidden layers, they have been widely used as classifiers for machine fault diagnosis [10-12]. But the training of neural networks also needs a large amount of high-quality labeled data, and if the training samples are limited or cannot cover the testing distribution, the neural network may be easily overfitted which leads to a poor generalization especially for a complex classification problem. The induction motor fault diagnosis belongs to such a challenging problem since the system itself is a complex electromechanical one. Therefore, some advanced signal processing technologies have been proposed for motor current or vibration signal analysis [13-15], to extract useful fault features for diagnosis. However, these state-of-the-art methods usually require the researchers to have a deep understanding on the induction motor system and the fault signals. In addition, the required expert knowledge is not easily obtained so that these methods are not general to address this task, which means these methods are not intelligent enough relative to the machine learning. In contrast, a new machine learning method called “deep learning” has been proposed to realize unsupervised feature learning [16, 17]. Since a deep “auto-encoder” network structure was trained to learn low-dimensional codes from high-dimensional input vectors [18], deep learning approaches have gained great interests and achieved remarkable results in many fields such as image recognition [19], speech recognition [20] and others [21, 22]. Furthermore, sparse auto-encoder (SAE) [23-25], which is one of the famous unsupervised feature learning methods, has been widely studied to realize deep learning, as it is highly effective for finding succinct and high-level representations of complex data.

Inspired by the prior research, the purpose of this paper is to present a SAE-based deep neural network (DNN) approach for induction motor fault diagnosis. In this presented approach, the SAE first learns succinct features automatically from high-dimensional data in an unsupervised manner. To improve the robustness of feature learning and prevent the identity transformation, the denoising coding is embedded into the SAE by masking noise randomly on its input to ensure the robustness of extracted features and improve the performance of the SAE. After that, the DNN utilizes the SAE to extract features which are used to train a classifier to diagnose various induction motor faults. In order to overcome the deficiency of overfitting during the training process of the DNN, the “dropout” technique is introduced into the whole DNN by randomly dropping neurons together with their connections for preventing co-adaptation of the neurons. Rest of this paper is organized as follows. Section II introduces the algorithm of the DNN for induction motor fault diagnosis, including SAE and denoising auto-encoder algorithm. Then section III discusses the details of experiments performed on a machine fault simulator and the corresponding results. Finally, section IV presents some conclusions drawn from this study.

2. DEEP NEURAL NETWORK

Unsupervised feature learning is able to learn discriminative and effective features from a large amount of unlabeled data [17]. In the field of fault diagnosis, labeled vibration signals are difficult to be acquired which need specific and elaborate experimental setting. Therefore, unsupervised feature learning may provide an effective solution to fault diagnosis. Here, one of the typical unsupervised feature learning algorithms-SAE is investigated. In the proposed framework, the SAE combined with denoising module is adopted to learning features from the vibration signals. Then, the learned features are fed into a neural network classifier with dropout. The detail of the framework is illustrated in the following sections.

2.1Sparse Auto-encoder

An auto-encoder is a symmetrical neural network that can learn the features in an unsupervised learning manner by minimizing reconstruction errors [21]. The basic structure of an auto-encoder is shown in Figure 1, where it tries to learn an approximation in the hidden layer so that the input data can be perfectly reconstructed at the output layer. However, the intrinsic problems of the auto-encoder, such as simply copying input layer to hidden layer, make it not effective to extract meaningful features even though its output can be a perfect recovery of the input data. SAE, as an extension of the auto-encoder, can learn relatively sparse features by introducing a sparse penalty term inspired by the sparse coding [26] into the auto-encoder. It can improve the performance of traditional auto-encoder and exhibits more practical application values.

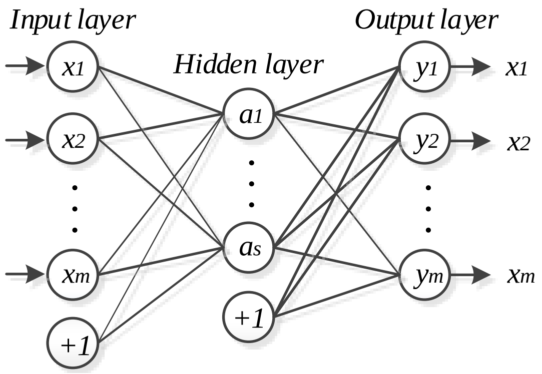


Figure 1. The structure of an auto-encoder neural network

From the measured vibration signals, an $N\times M$ data set can be constructed as $X=\{x(1), x(2), , . . ., x(i), . . ., x(N)\}$, $x(i) \in R^M$, where N is the number of data samples for each working condition, and M is the length of each data sample. This data set is used as the input matrix of the SAE. Then a three layer neural network similar to that in Figure 1 can be constructed, where the sigmoid function is chosen as activation function to the network. For the unlabeled input data matrix X, the goal is to learn and obtain a feature expression $h(x(i), W, b) = \sigma(Wx(i) + b), i = 1, \cdots, N$, at the hidden layer so that the output $\sigma(W^T h(x(i), W, b) + c)$ is close to or refactoring the input. In the meantime, the sparse penalty term is added to the objective function of the auto-encoder so that the learned features are of the constraint rather than simply repeating input. The sparse penalty term actually works on the hidden layer to control the number of "active" neurons. In practice, if the output of a neuron is close to 1, the neuron is considered to be "active", otherwise it is "inactive". It is better to keep the neurons of the hidden layer "inactive" in most of the time. Suppose that $a_j(x)$ denotes the activation of hidden unit j. In the forward propagation process, for a given input X, the activation of the hidden layer can be denoted as $a = sigmoid(WX + b)$, where W denotes the weights between the input layer and the hidden layer and b is the biases. Then the average activation of the hidden unit j can be given as:

$$\rho_j = \frac{1}{n}\sum_{i=1}^{n}\left[a_j(x(i))\right] \tag{1}$$

In this process, the average activation of each hidden neuron ρ_j is expected to be close to zero, namely the neurons of the hidden layer is mostly "inactive". To achieve this, the sparse term is added to the objective function that penalizes ρ_j if it deviates significantly from ρ. The penalty term is expressed as:

$$P_{penalty} = \sum_{j=1}^{S_2} KL(\rho \| \rho_j) \tag{2}$$

where S_2 is the number of neurons in the hidden layer. $KL(\cdot)$ is the Kullback-Leibler divergence (KL divergence) [27], which can be written as:

$$KL(\rho \| \rho_j) = \rho log\frac{\rho}{\rho_j} + (1-\rho)log\frac{1-\rho}{1-\rho_j} \tag{3}$$

This penalty function possesses the property that $KL(\rho \| \rho_j) = 0$ if $\rho_j = \rho$. Otherwise, it increases monotonically as ρ_j diverges from ρ , which acts as the sparsity constraint.

The cost function of the neural network is defined as:

$$C(W,b) = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{2}\left\|h_{w,b}(x(i)) - y(i)\right\|^2\right)\right] + \frac{\gamma}{2}\sum_{l=1}^{m_{l-1}}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(W_{ij}(l)) \tag{4}$$

Adding the sparse penalty term to the cost function, it can be modified as:

$$C_{sparse}(W,b) = C(W,b) + \beta\sum_{j=1}^{S2} KL(\rho \| \rho_j) \tag{5}$$

where β is the weight of the sparsity penalty.

During the coding process, the optimal parameters of W and b need to be identified. Since the sparse cost function shown in equation (5) is directly related to the parameters W and b, it can be solved by minimizing $C_{sparse}(W,b)$ to obtain these two parameters. This can be realized using the back-propagation algorithm [28], where the stochastic gradient descent approach is used for training and the parameters W and b in each iteration can be updated as:

$$W_{ij}(l) = W_{ij}(l) - \varepsilon \frac{\partial}{\partial W_{ij}(l)} C_{sparse}(W,b) \tag{6}$$

$$b_i(l) = b_i(l) - \varepsilon \frac{\partial}{\partial b_i(l)} C_{sparse}(W,b) \tag{7}$$

where ε is the learning rate. A forward pass on all training examples is used to compute the average activation ρ_j to get the sparse error, then the back-propagation algorithm works to update the parameters. After that, the effective sparse feature representations can be learned by the SAE.

2.2Denoising Auto-encoder

Vincent et al. [29] proposed a denoising auto-encoder learning algorithm based on the idea of making the learned feature representation robust by adding partial corruption to the input pattern, which can be used to train stacked auto-encoders to initialize the deep architectures. It has been known that this type of learning can help to obtain better feature representation with denoise coding, which motivates the introduction of denoising module into our study.

For the input matrix X, the process of denoising auto-encoder is shown in Figure 2, where the initial input X is corrupted to get a partially destroyed version $\mathcal{X}$ by applying a stochastic mapping $\mathcal{X} \sim qD(\mathcal{X}|X)$, then $\mathcal{X}$ is coded to y. Here distribution $qD(\bullet)$ can be described as follows: for each input $x(i)$, a fixed number of components are chosen randomly and their values are set to 0, while the others are kept unchanged. Then a joint distribution is defined as:

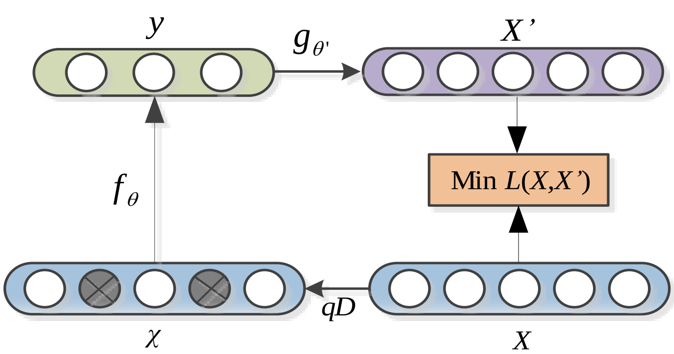


Figure 2. Structure of denoising auto-encoder

$$q^0(X, \mathcal{X}, y) = q^0(X) qD(\mathcal{X}|X) \delta_{f_\theta(\mathcal{X})}(y) \tag{8}$$

where $\delta_{f_\theta(\mathcal{X})}(y)$ is set to 0 if $f_\theta(x) \neq y$. $q^0(X)$ denotes the empirical distribution associated with N

training inputs. Thus y is a deterministic function of $\mathcal{X}$. Let θ be one parameter of the joint distribution $q^0(X, \mathcal{X}, y)$, it can be considered as a link parameter between $\mathcal{X}$ and y. To obtain a reconstructed version (denoted as X') of the input matrix X from y, an objective function can be set as

$$arg\,min_{\theta,\theta'}\, E_{q^0(X,\mathcal{X})}\,[L(X, X')] \tag{9}$$

where $L(X, X')$ is the reconstruction error. The reconstructed matrix X' can be obtained by minimizing (9) using a stochastic gradient descent approach.

Specifically, to implement the denoising coding, some elements of the input X are randomly set to zero for generating the corrupted input $\mathcal{X}$. Then it is mapped with the basic auto-encoder to a hidden representation y, where $y = f_\theta(\mathcal{X}) = sigmoid(W\mathcal{X} + b)$ with $\theta = \{W, b\}$, from which X' can be reconstructed as $X' = g_{\theta'}(y) = sigmoid(W_1 y + b_1)$. The reconstruction error $L(X, X')$ is computed from the difference between X' and X, and can be minimized by solving the cost function (9).

It should be noted that the motivation behind the corruption of the values of some components in X shares the same idea of "salt noise" in image processing. Because unsupervised initialization of layers with an explicit denoising criterion can help to capture meaningful structure in the input distribution, the learned intermediate representations can contribute a lot to the subsequent learning tasks such as classification and clustering. Furthermore, it was indicated that denoising autoencoder can improve the robustness of features and the sparsity of weights, which both make the discovery of interesting and prominent pattern easier [30].

2.3Dropout

Dropout is a technique that can help to reduce "overfitting" when training a neural network with limited training data set [31]. Generally, when the known training data set is small, the overfitting problem will occur which lead to a poor performance on the test data. In this study, the dropout technique is applied to training the SAE-based DNN to prevent complex co-adaptations on the training data and avoid the extraction of the same features repeatedly. Technically, the "dropout" can be realized by setting the output $a = sigmoid(WX + b)$of some hidden neurons to zero so that these neurons will not be involved in the forward propagation training process. It should be noted that there are some differences between the training process and testing process with dropout. The dropout is turned off during testing, which means the outputs of all hidden neurons will not be masked during testing. This will help to improve the feature extraction and classification capability of the SAE-based DNN.

2.4Proposed Framework

By taking advantage of the unsupervised learning, the denoising SAE is used to learn features from unlabeled data and initialize the DNN structure. Then the learned sparse feature representation of the auto-encoder is utilized to train a neural network classifier with a dropout module for induction motor fault diagnosis. Instead of direct utilization of learning representation by denoising SAE, the parameters of the NN classifier with the corresponding parameters are initialized in the well-trained SAE and then updated further. This adaptation enables a further fine-tuning of learned parameters so that the learned representation can capture more discriminative components in raw vibration signals. In addition, the previous study [32] indicated that the number of hidden neurons in the deep model is as important as the choice of learning algorithm and the depth of the model for achieving high classification performance, which means that single-layer network in unsupervised feature learning can also perform well. Therefore, in this study only one-layer SAE is utilized to realize the DNN for the induction motor fault diagnosis. The training procedure of the DNN is shown in Figure 3 and it can be described as follows:

The unlabeled induction motor vibration data *X1* are first used to train the SAE though following steps:

1) Set up the learning rate, sparse rate and denoising parameters, dropout rate, etc., and initialize the weight *W* and *b* randomly.
2) Use the stochastic batches training method in the forward propagation algorithm to compute the average activation ρ_j for sparsity.
3) Compute the sparse cost function as:

$$C_{sparse}(W,b) = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{2}\left\|h_{w,b}\big(x(i)\big) - y(i)\right\|^2\right)\right] + \beta\sum_{j=1}^{S2} KL(\rho\|\rho_j) \tag{10}$$

4) Update the parameters *W* and *b* based on Eq.(6) and Eq. (7).

Then the labeled induction motor vibration data (*X2*,*Y2*) are used to train the DNN for classification.

1) Use the parameters of the SAE to initialize the first layer of the DNN.
2) Set up the training parameters and dropout rate, and conduct the forward propagation algorithm to extract the labeled features for classification.
3) Compute the mean square error for the cost function of the DNN using Eq. (4).
4) Conduct the back-propagation algorithm the same as before except for the sparse term (set sparse penalty term to 0) to update the weights and fine-tune the entire network.

Finally, the test data set (*X3, Y3*) are used to verify the effectiveness of the presented SAE-based DNN.

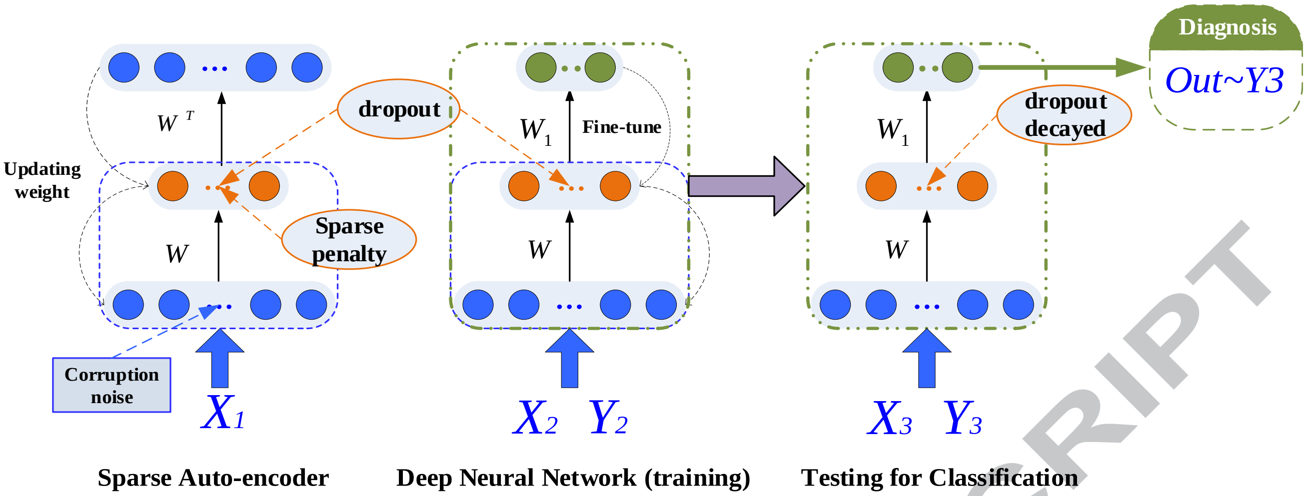


Figure 3. Structure of the sparse auto-encoder-based deep neural network

3. Experimental Study

3.1Experimental Setup

To verify the effectiveness of the SAE-based DNN for induction motor fault diagnosis, the experiments have been conducted on a machine fault simulator as illustrated in Figure 4. The vibration signals can be acquired by an acceleration sensor when the motor operates under six different conditions. The vibration data of the simulator under the power supply frequency of 50 Hz is collected with the sampling frequency of 20 kHz. The descriptions of different motor working conditions are listed in Table 1.

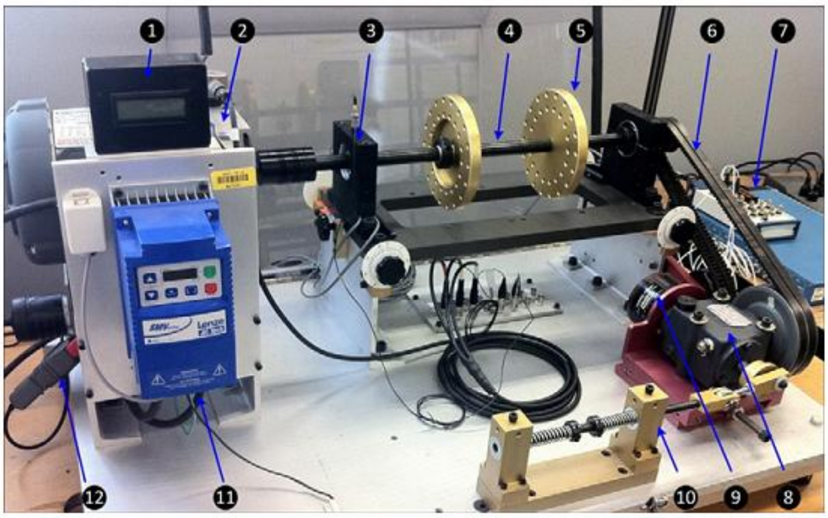


Figure 4. Experimental setup: (1) Tachometer, (2) Induction Motor, (3) Bearing, (4) Shaft, (5) Load Disc, (6) Belt, (7) Data Acquisition Board, (8) Bevel gearbox, (9) Magnetic Load, (10) Reciprocating Mechanism, (11) Variable Speed Controller, (12) Current Probe

Table 1. Motor condition descriptions

Condition		Description
HEA	Normal motor	Healthy, no defect
BRB	Broken bar	Three broken rotor bars
BRM	Bowed rotor	Rotor bent in center 0.01”
RMAM	Defective bearing	Inner race defect bearing in the shaft end
SSTM	Stator winding defect	3 turns shorted in stator winding
UBM	Unbalanced rotor	Unbalance caused by 3 added washers on the rotor

The task defined here is to classify different motor working conditions based on vibration data. In this study 600 data samples with 2000 features from each induction motor working condition were obtained, in which each feature corresponds to one sampled data point of the vibration. For each working condition, 400 of the samples were chosen randomly for training and the rest for testing. When 2400 training samples and 1200 testing samples are obtained with feature dimensionality of 2000, they are all normalized to set the value between -1 and 1.

3.2Compared Approaches

Several neural network-based methods are compared here with the proposed approach. Firstly, to verify the effectiveness of the SAE, normal neural networks without SAE are considered and three different model structures are considered. Two of them contain one hidden layer, which is the same as the proposed approach. Their hidden nodes are set as 100 and 600, respectively. They are denoted as NN(100) and NN(600). In addition, the neural network with two hidden layers, whose hidden nodes are 600 and 100, respectively, and denoted as NN(600-100), is also considered here. Then, two classifiers built on top of the SAE are also considered, i.e., support vector machine (SVM) and logistic regression (LR) [33, 34]. The features learned by the SAE, i.e., the hidden output of the SAE, is fed into both the SVM and the LR classifiers. For a fair comparison, the SAE is the same as the one used in the proposed approach, but the features learned by the SAE cannot be fine-tuned during training for both SVM and the LR classifiers.

In the proposed SAE-based DNN framework, the input size, hidden node and output size are set to 2000, 600 and 2000, respectively. Considering the vibration signals represent six different working conditions, the layer sizes for DNN is set as [2000 600 6]. The hidden nodes of the DNN and SAE are same so that the features learned by the SAE can be utilized by the DNN and fine-tuned further during supervised training. The three hyper-parameters of the proposed SAE-based DNN including sparse parameter, denoted as sparse rate β, denoising rate and dropout rate are set to be 0.4, 0.1 and 0.3, respectively, via cross validation for comparison with other baseline methods. All parameters with their values of the proposed SAE-based DNN are listed in Table 2 for better understanding.

Table 2. Parameters setting up of SAE-based DNN

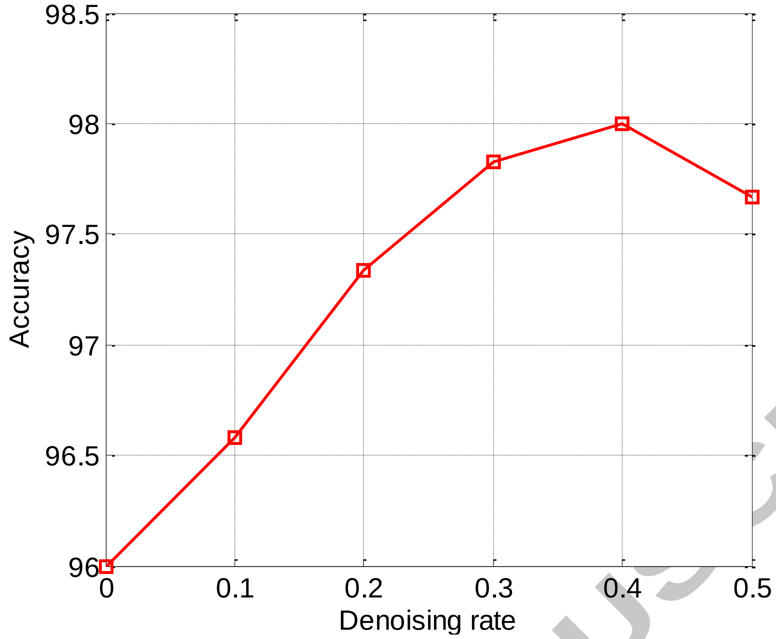
SAE			DNN		
M	2000	input nodes	M	2000	input nodes
S_2	600	hidden nodes	S	600	hidden nodes
out	2000	output nodes	*out*	6	output nodes
ρ	0.08	sparse target	*dropout*	0.3	dropout rate
β	0.4	sparse rate	ε	1	learning rate
denoise	0.1	denoising rate			
dropout	0.3	dropout rate			
ε	1	learning rate			

Here for all NN-based methods, the cross-entropy error is calculated based on training samples and back-propagated through layers to update model parameters, in which the stochastic gradient descent method was used. The simulation environment for all algorithms performed in the experiments is as follows: MATLAB R2011b software environment.

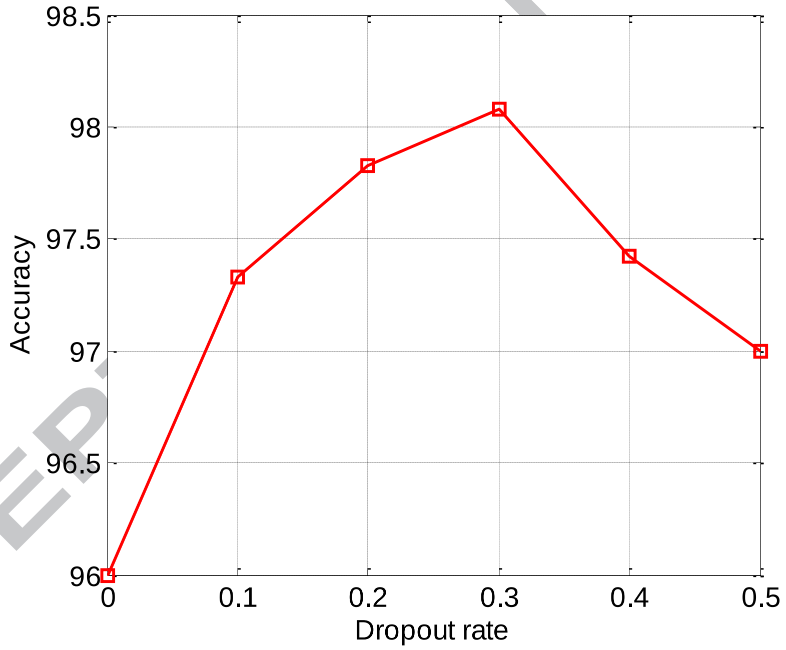
3.3Analysis of Denoising Coding and Dropout

Since the level of corruption noise is a hyper-parameter in the SAE, it is meaningful to study the effect of corruption noise strength on the final classification performance. Here, the noise levels are changed from 0 to 0.5 with a step size of 0.1. Figure 5(a) shows the classification performance with different corruption noise. It can be seen that proper denoising coding in the SAE could improve its performance, but too heavy corruption will degrade the input data quality, leading to a decrease of the classification performance. This verified that the SAE trained with appropriate noisy data can extract more robust features than traditional one and it is of great significance in practical complex environment.

In addition, the effect of dropout on performance of the DNN is also studied. Here, dropout rate is changed from 0 to 0.5 with a step size of 0.1. Figure 5(b) shows the classification performances with different dropout rates. The results showed that the best classification performance was obtained at a dropout rate close to 0.3 and the performance would decrease when the dropout rate was more than 0.3. This indicated that dropout can improve the performance of the DNN, but too much dropout may loss some important neurons for feature representation. Based on results shown in Figure 5, it can conclude that the SAE-based DNN approach presented in this study can realize the induction motor fault diagnosis and have a good classification accuracy at about 96% at least.



(a) Effect of denoising coding



(b) Effect of dropout

Figure 5. Studies on denoising coding and dropout

Then to verify the performance of presented approach that combines the sparse coding with denoising coding in the auto-encoder to learn features for the DNN and utilizes the "dropout" technique to overcome overfitting during the training process, a comparison study was carried out, where traditional SAE is used as the baseline. The classification performance under different labeled training data set sizes were investigated, where the size of training data were changed from 80 to 400 with a step size of 80. The results are shown in Figure 6.

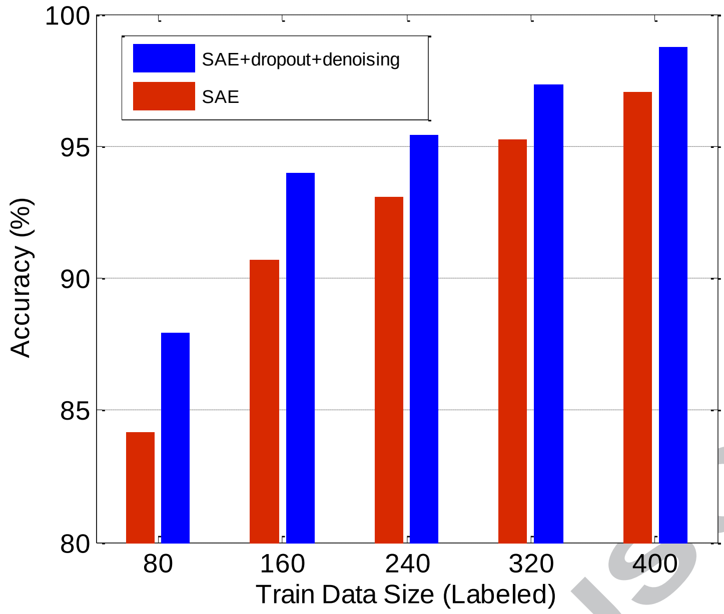


Figure 6. Comparison of two methods

It can be seen that the presented approach has shown better performance than the SAE alone. In addition, the performance improvement achieved by the proposed approach for a small training data size is more significant than the one for a large training size, which verifies the robustness of the proposed approach. It may be explained that corruptions of input samples can be regarded as expanding data samples artificially. Furthermore, dropout can improve the performance of the deep neural network by preventing complex co-adaptations between the hidden neurons, especially for small training data size.

3.4Result and Discussion

Comprehensive comparison results of the classification performance are shown in Table 3. It is shown that the proposed SAE-based DNN framework achieves the best performance in all scenarios.

For the three NN models without SAE, the best one is NN with two hidden layers. This observation can be explained by the fact that the increased number of layers contributes to the learning capability of the neural network. However, NN(600-100) still performs worse than the proposed approach with one hidden layer. This result verifies the effectiveness of the SAE. The robust feature learned by the SAE can contribute to performance improvement of the following DNN.

Then, the proposed approach is also compared with two classifiers built on the top of the SAE. Different from the proposed approach, SVM and LR cannot fine-tune the features learned by the SAE. This difference may explain their incomparable performance. It can be seen in Table 3 that both SVM and LR classifiers can classify the features learned by the SAE well into the corresponding fault categories and classification accuracy for either of them is much high than the NN with one hidden layer. This result is consistent with that the SAE can learn effective features for the fault diagnosis.

Table 3. Comprehensive comparison of classification accuracy

Classifier	Health	BRB	BRM	RMAM	SSTM	UBM	Average
NN(100)	78.00%	92.50%	70.40%	81.70%	87.40%	70.25%	80.04%
NN(600)	74.00%	97.50%	73.50%	71.00%	95.00%	35.00%	74.33%
NN(600-100）	88.78%	99.56%	100.00%	99.94%	92.72%	97.33%	96.39%
SVM	92.50%	96.50%	100.00%	100.00%	91.00%	98.50%	96.42%
LR	83.50%	94.00%	100.00%	100.00%	81.50%	97.50%	92.75%
DNN	**92.68%**	**99.91%**	**100.00%**	**100.00%**	**93.50%**	**99.55%**	**97.61%**

In practice, the quality of the vibration data measured from the induction motor will be affected by environmental factors. Thus it is necessary to perform stability analysis of the SAE-based DNN for dealing with data under disturbance. Partial corruption at different levels to the data was added into the input of both the neural networks and the SAE-based DNN, and their influence on the classification accuracy was shown in Figure 7. When the degree of the data corruption increased from 0 to 40%, even though the performance of the SAE-based DNN decreases, the overall classification accuracy is still larger than 90%. In comparison, the performance of the neural networks decreases significantly, i.e. from 96.58% to 76.58% for the neural network with two hidden layers, and from 73.83% to 52.25% for the neural network with one hidden layer, reflecting that the SAE-based DNN possesses good stability against disturbance for induction motor fault diagnosis.

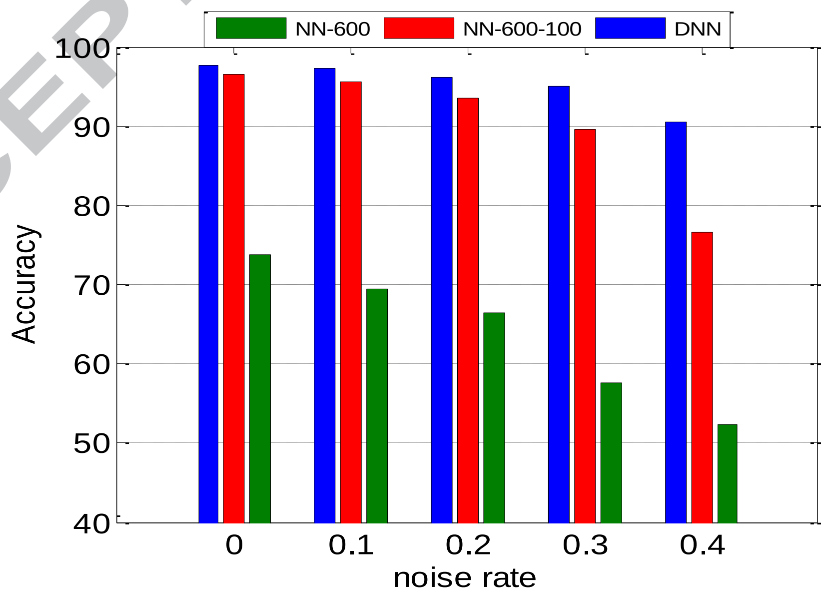
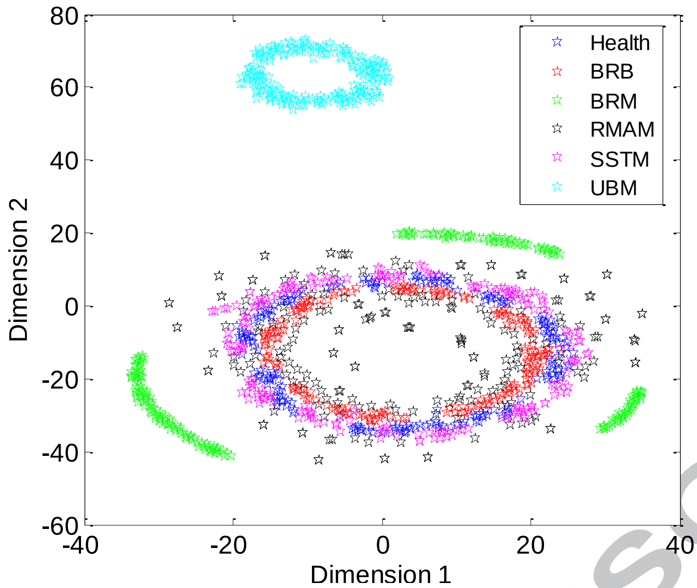


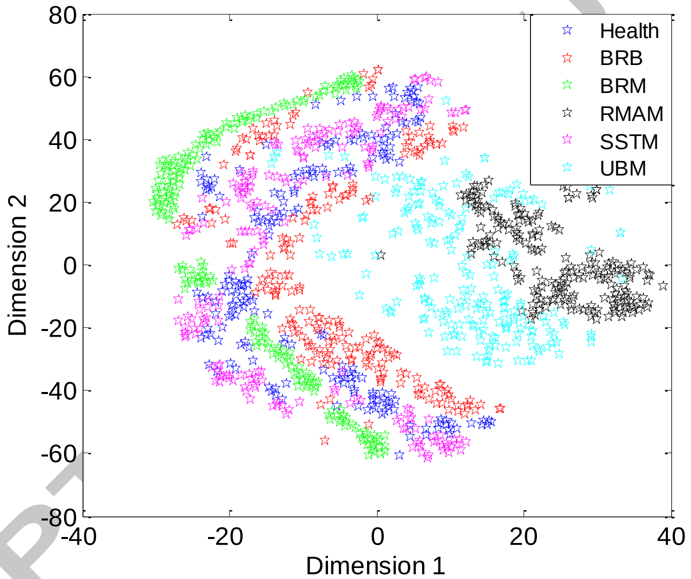
Figure 7. Histogram of classification accuracy with “noise” in input data

3.5Visualization of Learned Representation

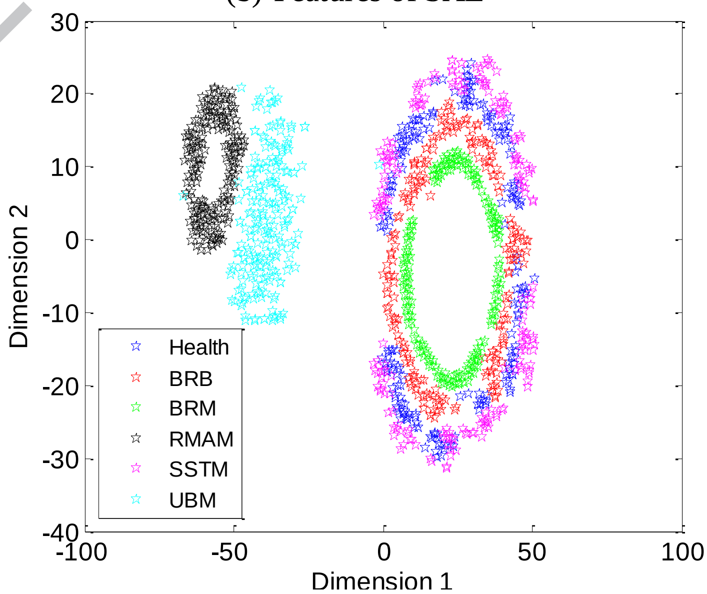
To demonstrate that the proposed approach is able to learn effective and discriminative representations for vibration signals automatically, the raw features, the features learned by the SAE and the features fine-tuned by the DNN are all visualized via a technique “t-SNE” [35, 36] which is an effective data visualization technique for high-dimensional data. Here, the principal component analysis (PCA) is used to reduce the dimensionality of the feature data to 50. This can speed up the computation of pairwise distances between the data points and suppresses some noise without severely distorting the interpoint distances. Then the dimensionality reduction technique “t-SNE” is used to convert the 50-dimensional representation to a two-dimensional map and the resulting maps as a scatterplot have been shown in Figures 8(a), 8(b) and 8(c), respectively. It can be seen that the SAE features fine-tuned by the DNN cluster the best where data points of different conditions are separated well. It presents the good separability of the features extracted by the SAE-based DNN in this study. The results also indicates that the features of the SSTM is the most similar with the features of the Health and they are hard to be separated well as others. This is also reflected as the classification accuracy of Health and SSTM in Table 3.



(a) Raw features



(b) Features of SAE



(c) Feature fine-tuned

Figure 8. Feature Visualization Maps

From the trend of the three maps (from Fig. 8a to Fig. 8c), it can be known that the SAE itself is an effective feature learning method for induction motor vibration signals and the SAE-based DNN approach that fine-tunes the features learned by the SAE can improve the learned representations of the SAE further. It also verifies the superior classification performance of the SAE-based DNN approach compared to other base-line methods. Therefore, the SAE-based DNN is an effective approach for induction motor feature learning and fault diagnosis.

4. Conclusions

A SAE-based DNN approach has been developed for induction motor fault diagnosis. This approach can learn features directly from raw data, which are discriminative over different working conditions of the induction motors. When the denoising coding is integrated with the SAE, it improves the robustness of the learned feature and the stability of the DNN against disturbance. The dropout technique integrated in the neural network classifier design reduces the overfitting in the training process and improves the performance of the DNN for induction motor fault classification. Experimental study and a detailed analysis has verified the effectiveness of the SAE-based DNN for induction motor fault diagnosis.

Acknowledgement

This work has been supported in part by the National Natural Science Foundation of China (51575102)

References

[1] M. E. H. Benbouzid, A review of induction motors signature analysis as a medium for faults detection. IEEE Trans. Ind. Electron. 47 (2000) 984-993.

[2] S. Huang, H. Yu, Intelligent fault monitoring and diagnosis in electrical machines, Measurement 46(9) (2013) 3640-3646.

[3] F.Filippetti, G. Franceschini, C. Tassoni, P. Vas, AI techniques in induction machines diagnosis including the speed ripple effect. IEEE Trans. Ind. Appl. 34 (1998) 98-108.

[4] P. J. Rodrigue, A. Belahcen, A. Arkkio, Signatures of electrical faults in the force distribution and vibration pattern of induction motors. IEE Proceedings-Electric Power Applications 153 (2006) 523-529.

[5] V. Fernão Pires, Manuel Kadivonga, J.F. Martins, A.J. Pires, Motor square current signature analysis for induction motor rotor diagnosis, Measurement 46(2) (2013) 942-948.

[6] A. Malhi, R. Yan, R. X. Gao, Prognosis of defect propagation based on recurrent neural networks. IEEE Trans. Instrum. Meas. 60 (2011) 703-711.

[7] F. Ahmadizar, K. Soltanian, F. Akhlaghiantab, I. Tsoulos, Artificial neural network development by means of a novel combination of grammatical evolution and genetic algorithm. Eng. Appl. Artif. Intell. 39 (2015) 1-13.

[8] X. Li, A. Zheng, X. Zhang, C. Li, L. Zhang, Rolling element bearing fault detection using support vector machine with improved ant colony optimization, Measurement 46(8) (2013) 2726-2734.

[9] D.J. Bordoloi, R. Tiwari, Support vector machine based optimization of multi-fault classification of

gears with evolutionary algorithms from time–frequency vibration data, Measurement 55 (2014) 1-14.

[10] R. J. Kuo, Intelligent diagnosis for turbine blade faults using artificial neural networks and fuzzy logic. Eng. Appl. Artif. Intell. 8 (1995) 25-34.

[11] A. Monica, Analysis of induction motor fault diagnosis with fuzzy neural network. Applied Artificial Intelligence 17 (2003) 105-133.

[12] H. Su, K. T. Chong, R. R. Kumar, Vibration signal analysis for electrical fault detection of induction machine using neural networks. Neural Comput. Appl. 20 (2007) 183-194.

[13] J. Zarei, J. Poshtan, Bearing fault detection using wavelet packet transform of induction motor stator current. Tribology International 40 (2007) 763-769.

[14] M. Blödt, D. Bonacci, J. Regnier, M. Chabert, J. Faucher, On-line monitoring of mechanical faults in variable-speed induction motor drives using the Wigner distribution. IEEE Trans. Ind. Electron. 55 (2008) 522-533.

[15] R. Puche-Panadero, M. Pineda-Sanchez, M. Riera-Guasp, J. Roger-Folch, E. Hurtado-Perez, J. Perez-Cruz, Improved resolution of the MCSA method via Hilbert transform, enabling the diagnosis of rotor asymmetries at very low slip. IEEE Trans. Energy Convers. 24 (2009) 52-59.

[16] J. Schmidhuber, Deep learning in neural networks: an overview. Neural Networks 61 (2014) 85–117.

[17] A. M. Cheriyadat, Unsupervised feature learning for aerial scene classification. IEEE Trans. Geosci. Remote Sens. 52 (2014) 439-451.

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

[19] A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25 (2012) 1106–1114.

[20] G. E. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process. Mag. 29 (2012) 82-97.

[21] H. C. Shin, M. R. Orton, D. J. Collins, S. J. Doran, M. O. Leach, Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4d patient data. IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 1930-1943.

[22] A. Droniou, S. Ivaldi, O. Sigaud, Deep unsupervised network for multimodal perception, representation and classification. Robotics & Autonomous Systems 71 (2015) 83-98.

[23] M. Ranzato, Y. L. Boureau, Y. Lecun, Sparse feature learning for deep belief networks. Adv. Neural Inf. Process. Syst. 2007, pp, 1185-1192.

[24] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, S. Yan, Sparse representation for computer vision and pattern recognition. Proc. IEEE. 98 (2010) 1031-1044.

[25] Y. L. Boureau, F. Bach, Y. Lecun, J. Ponce, Learning mid-level features for recognition. in Proc. IEEE Conf. CVPR, Jun 13-18, 2010, pp. 2559–2566.

[26] B. A. Olshausen, D. J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381 (1996) 607-609.

[27] S. Kullback, R. A. Leibler, On information and sufficiency. Annals of Mathematical Statistics 22 (1951) 79-86.

[28] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors. Nature 323 (1986) 533-536.

ACCEPTED MANUSCRIPT

[29] P. Vincent, H. Larochelle, Y. Bengio, P. A. Manzagol, Extracting and composing robust features with denoising autoencoders. In Proceedings of the 25th international conference on Machine learning, 2008, pp.1096-1103.

[30] L. L. Cao, W. B. Huang, F. C. Sun, Building feature space of extreme learning machine with sparse denoising stacked-autoencoder. Neurocomputing 174 (2016) 60-71.

[31] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R. R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580, 2012.

[32] A. Coates, A. Y. Ng, H. Lee, An analysis of single-layer networks in unsupervised feature learning. Jour. Mach. Learn. Res. 15 (2011) 215-223.

[33] G. Bouchard, Clustering and classification employing softmax function including efficient bounds. US, US 8065246 B2, 2011.

[34] M. Khashei, A. Z. Hamadani, M. Bijari, A novel hybrid classification model of artificial neural networks and multiple linear regression models. Expert Syst. Appl. 39 (2012) 2606-2620.

[35] L. V. D. Maaten, G. E. Hinton, Visualizing data using t-sne. Jour. Mach. Learn. Res. 9 (2008) 2579-2605.

[36] V. D. M. Laurens, Accelerating t-SNE using tree-based algorithms. Jour. Mach. Learn. Res. 15 (2015) 3221-3245.

HIGHLIGHTS

1) A sparse auto-encoder-based deep neural network is investigated for induction motor fault diagnosis
2) The deep neural network is of good stability against disturbance for fault diagnosis.
3) Denoising coding is added into the sparse auto-encoder for performance improvement.
4) Dropout technique is utilized to reduce data overfitting and generate good feature representations.