Progress In Electromagnetics Research M, Vol. 91, 197–211, 2020

Transformer Fault Diagnosis Model and Method Based on DBNI in Photoelectric Sensors Diagnosis System

Xuewei Zhang[1, 2, *], Hanshan Li[2], Liping Lu[2], and Xiaojuan Sun[2]

Abstract—In order to improve the efficiency of transformer fault diagnosis and monitoring in power systems, and to realize fault diagnosis of unmanned remote adaptive transformer equipment, we present a method of multi-sensor and multi-direction optical image integrated monitoring in this paper. By monitoring and collecting transformer fault information combined with the changing characteristics of transformer temperature and electrical signals, we establish a transformer calculation model based on multi-level fault and multi-characteristic parameters. According to the characteristics of transformer faults, we use a deep belief network identification (DBNI) algorithm for the transformer and construct the training samples of the transformer diagnosis model using an optimum weight fusion algorithm. The experimental results show that the DBNI model can fully explore the characteristics of large samples, analyze multiple faults information, and extract the hidden features of fault samples. The DBNI model has higher fault diagnosis accuracy than a BP neural network and a single DBN without data fusion and SVM. The DBNI's fault diagnosis accuracy reaching 99.45%. The experimental results show that this model has good robustness of interference ability and can be used intuitively to carry out remote on-line unattended transformer fault diagnosis and information feedback.

1. INTRODUCTION

In recent years, the rapid development of intelligent transformers has played a very important role in smart grids' structure. Because of transformers' complex and changeable working conditions, complex faults, and high incidence rates, the maintenance cost of transformers is high. Therefore, timely and effective fault prediction for transformer windings and other key components has become an important way to improve the reliability of transformer operation and reduce maintenance costs [1, 2]. In view of the huge amount of transformer data, mining the rules and characteristics and transforming them into fault detection and state analysis of key transformer components has become a key research area [3, 4].

At present, in transformer diagnosis, dissolved gas-in-oil analysis (DGA) and three-ratio algorithm are mostly used. In [5], a DGA and a deep belief network (DBN) are used to diagnose transformer faults. In [6], based on DGA technology, the k-nearest neighbor (KNN) algorithm and classification and regression trees (CART) algorithm are used to build a transformer fault diagnosis model based on a single intelligent algorithm. However, for existing transformer faults, traditional methods still have some shortcomings. The main ones are that the three-ratio algorithm lack of coding and the coding boundaries are too absolute; DGA algorithm cannot monitor in real time; and the degree of intelligence is not high. Reference [7] adopts a multi-sensor fusion algorithm to measure the transformer's state temperature information, and the measurement error is small, at the same time, which has a fast detection speed.

Received 7 January 2020, Accepted 6 April 2020, Scheduled 5 May 2020

* Corresponding author: Xuewei Zhang (xueweizhang163@163.com).

[1] School of Mechanical and Electrical Engineering, Xi'an Technological University, Xi'an 710021, China. [2] School of Electronic and Information Engineering, Xi'an Technological University, Xi'an 710021, China.

According to the conclusion of [7], in this paper we use temperature information and discharge signal information fusion to diagnose transformer faults. At the same time, we combine these with a deep belief network to build a transformer fault diagnosis system based on deep belief network identification (DBNI). Based on the original DBN [8, 9], in this paper we improve the training sample data and model, analyze the characteristic rules of the transformer data, extract the transformer state detection value, and realize fault detection. Then, we test and analyze the performance of the DBNI-based transformer fault diagnosis method using sample data obtained from a project site and compare our method with fault diagnosis methods based on artificial neural networks [10] and SVMs [11, 12].

2. DEEP BELIEF NETWORK IDENTIFIER AND DIAGNOSIS MODEL ON TRANSFORMER

2.1. Transformer Diagnosis Method in Photoelectric Sensors Diagnosis System

Photoelectric sensor diagnostic systems are composed of multiple infrared temperature sensors, discharge circuit sensors, and computer processing modules. Figure 1 shows the diagnosis testing diagram of a photoelectric sensor diagnostic system.

In Figure 1, the infrared temperature sensor and discharge circuit detection module are used to monitor the temperature and discharge signal of the transformer. The infrared temperature sensor is used to obtain the temperature of different parts of the transformer, and the discharge circuit detection modules are installed on the wall of the transformer box to obtain partial discharge signal information. Based on this platform, we use a deep belief network identifier to diagnose transformer faults.

In Figure 1 and the corresponding Figure 2, the infrared temperature sensors monitor the transformer fault from different angles of the transformer. The discharge circuit detection modules are installed on 6 different positions of the transformer, and each sensor can monitor the position

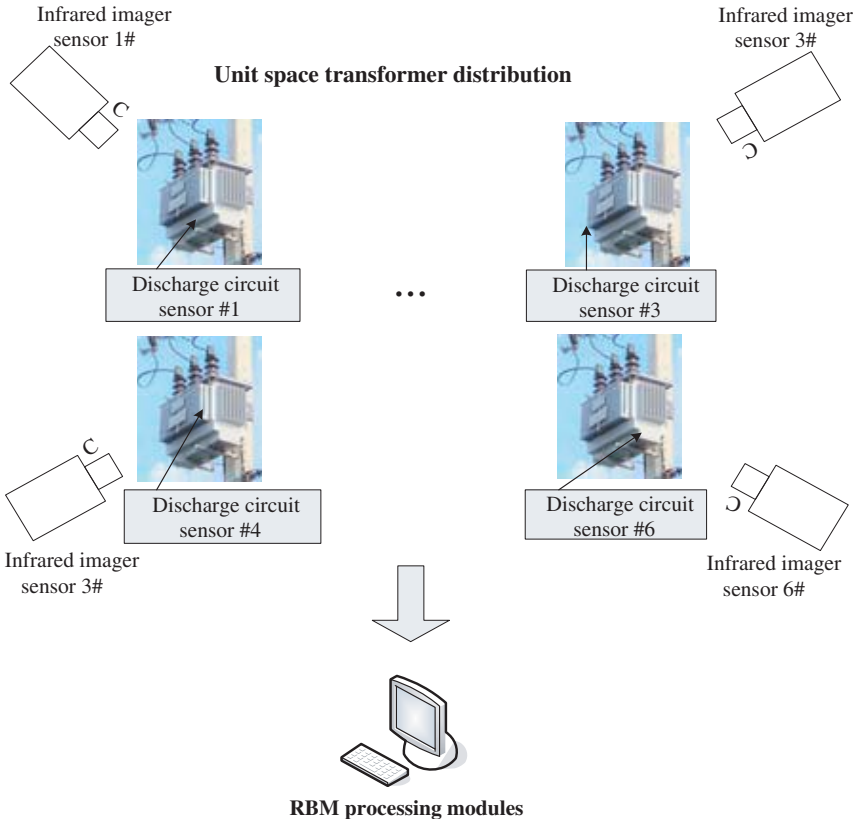

Figure 1. The testing diagram of a photoelectric sensor diagnostic system.

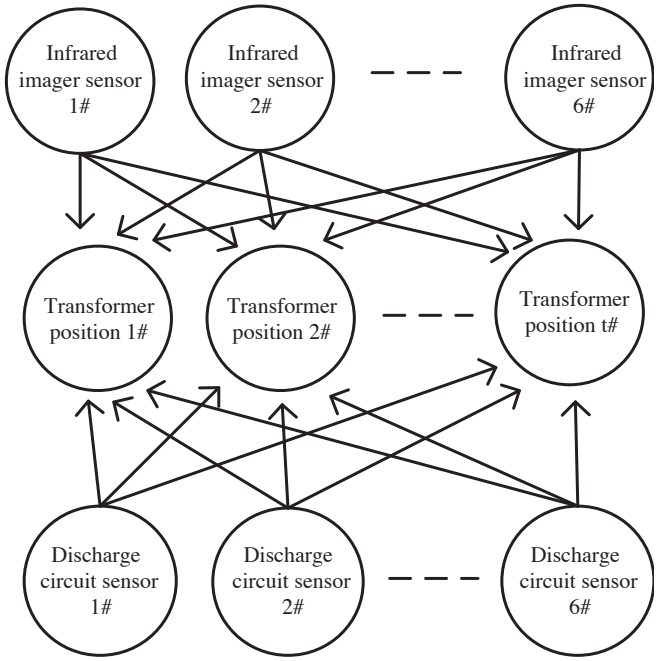

Figure 2. Transformer fault detection based on multi-sensor fusion.

information at t, as shown in Figure 2. In the photoelectric multi-sensor information acquisition system, the uncertainty of the measured data is caused by the influence of sensor precision, transmission errors, environmental noise, and interference. In Figure 2, we use photoelectric multi-sensor data fusion to diagnose transformer faults.

2.2. Deep Belief Network Identifier Model

A deep belief network identifier (DBNI) is an improved deep belief network (DBN). A deep belief network (DBN) is a generative structural graph model with multiple hidden layers. It is composed of several Restricted Boltzmann Machines (RBM) and a BP neural network stack. An RBM is a two-layer recurrent neural network. Random binary inputs and outputs are connected using symmetrical weights, as shown in Figure 3.

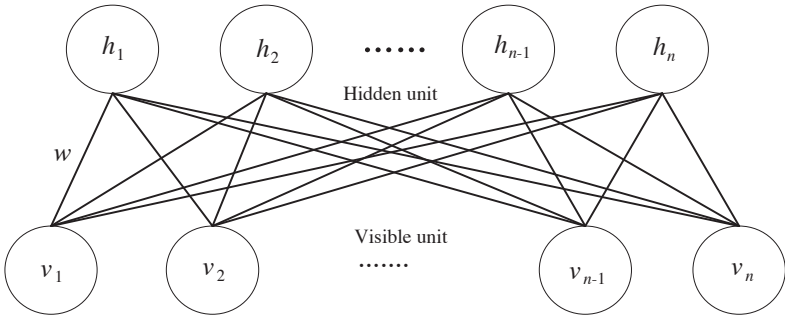

Figure 3. RBM structure.

For a given set of states $(v,\ h)$, the joint allocation energy function of the visible and hidden elements is:

$$E(v,h|\theta) = -\sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{m} b_j h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} v_i w_{ij} h_j \tag{1}$$

In Eq. (1), $\theta = (w, a, b)$ is the parameter of the RBM model [13]; v_i and a_i are the state and offset of the ith visible unit, respectively; h_j and b_j are the state and offset of the jth hidden unit; and w_{ij} is the connection weight between the ith visible unit and the jth hidden unit. The state probability of the RBM obeys the regular distribution.

Based on Eq. (1), the joint probability distribution of any group (v, h) can be obtained using Equation (2).

$$P(v, h|\theta) = \frac{1}{Z(\theta)} \exp(-E(v, h|\theta)) \tag{2}$$

In Eq. (2), $Z(\theta) = \sum_{v,h} \exp(-E(v, h|\theta))$ is a normalization factor. The activation probability of the jth neuron in the hidden layer is expressed by Equation (3).

$$P(h_{j=1}|v, \theta) = \delta\left(b_j + \sum_i w_{ij} v_i\right) \tag{3}$$

In Eq. (3), δ is the sigmoid activation function.

Similarly, when the state h of a hidden unit is given, the activation states of the visible units are independent of each other, so the activation probability of the ith visible unit is:

$$P(v_i = 1|h, \theta) = \delta\left(a_i + \sum_j w_{ij} h_j\right) \tag{4}$$

For a training sample set with capacity S, $S = \{v^1, v^2, \ldots, v^s\}$, by maximizing the logarithmic likelihood function $L(\theta)$ of RBM on the input training set to obtain the model parameter θ, we can fit the training data set, that is, the hidden layer can be used as the feature of the visual layer data, as expressed through Equation (5).

$$\theta^* = \arg\max_\theta L(\theta) = \arg\max_\theta \sum_{k=1}^{s} \log P(v^k|h, \theta) \tag{5}$$

To train an RBM, the logarithmic likelihood function is derived from parameter θ of the model. The derivation method is based on the contrastive divergence (CD) algorithm.

$$\frac{\partial \ln(L(\theta, v))}{\partial \theta} = \frac{\partial \sum \ln P_\theta(v)}{\partial \theta} = \sum \left\{ E_{P_\theta(h|v)}\left[-\frac{\partial E_\theta(v, h)}{\partial \theta}\right] - E_{P_\theta(v,h)}\left[-\frac{\partial E_\theta(v, h)}{\partial \theta}\right]\right\} \tag{6}$$

In Eq. (6), $E_{P_\theta(h|v)}[-\partial E_\theta(v, h)/\partial\theta]$ represents the expectation of $-\partial E_\theta(v, h)/\partial\theta$ in probability $P(h|v)$.

If v describes the visual state of the original data of the likelihood function, $E_{P_\theta(h|v)}[-\partial E_\theta(v, h)/\partial\theta]$ expresses the expected value of the partial derivative of the energy function of the original data under its own distribution. For this reason, partial derivatives of w, a, and b are obtained using Eq. (6). The update criterion of $\theta = \{w, b, a\}$ is determined by fast gradient ascent method as:

$$\Delta w_{ij} = \varepsilon\{\langle v_i h_j\rangle_{P_\theta(h|v)} - \langle v_i h_j\rangle_{recon}\} \tag{7}$$

$$\Delta b_i = \varepsilon\{\langle v_i\rangle_{P_\theta(h|v)} - \langle v_i\rangle_{recon}\} \tag{8}$$

$$\Delta a_j = \varepsilon\{\langle h_j\rangle_{P_\theta(h|v)} - \langle h_j\rangle_{recon}\} \tag{9}$$

Here, ε is the learning rate; $\langle\bullet\rangle_{P_\theta(h|v)}$ is the expectation of the partial derivative under the $P_\theta(h|v)$ distribution; and $\langle\bullet\rangle_{recon}$ is the expectation of the partial derivative under the reconstruction model distribution [14].

After the pre-training is completed, each RBM layer obtains the initial parameters, which constitute the preliminary framework of the DBN. Then, the DBN needs to be tuned to further optimize the parameters of the network layer, which makes the discriminant performance of the network better [15]. Compared with the BP neural network, the training speed is faster, and the convergence time is shorter [16].

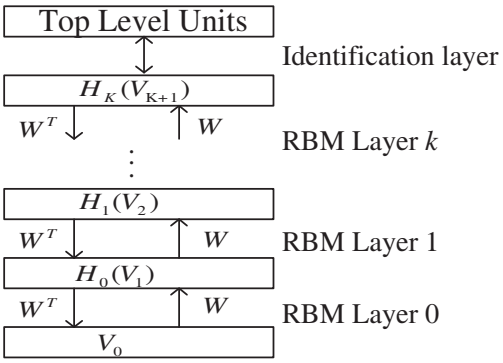

Figure 4. DBNI model.

Based on this, we construct the deep belief network identification (DBNI) model, which is composed of several layers of RBM stacked at the bottom, while the top layer is added to represent the last layer of the expected output variable, namely the identification layer, as shown in Figure 4.

This method uses a large number of unlabeled samples in the project site for pre-training and model parameter optimization and a small number of labeled samples for further optimization, which solves the problem of transformer fault classification effectively and improves the accuracy of fault diagnosis. Finally, the method is validated using an engineering example and comparing the fault diagnosis method with BPNN and SVM.

The BPNN constructed is a three-layer network model, which has an input layer, a hidden layer, and an output layer [17]. The input of the model is 6 sensor values, so it has 6 neuron nodes. As shown in Figure 1, the infrared temperature sensors are the input layer. The output of the model is the probability value of the samples belonging to the above six states, so the output layer is 6 neuron nodes, while the number of neuron nodes in the hidden layer is set to the empirical value of 100. The DBNI presented in this paper is used to identify Iris, which is a typical classification data set. The network parameters W, a, and b are initialized as small random values obeying the Gauss distribution, and the initial learning rate ρ is set to 0.1.

For the SVM method, because one SVM can only give two classification results, in this paper we use four SVM classifiers. The classification process is represented by a binary tree as shown in Figure 5.

Here, the fault types are divided into five states: normal, medium low temperature overheating, high temperature overheating, low energy discharge, and high energy discharge. For SVM 1 classifiers, when the output is 1, it is diagnosed as normal state, while when the output is -1, it is diagnosed as a fault state. The classifiers are independent of each other and can work in parallel.

BPNN, SVM, and DBNI were used to identify Iris from typical data sets. The maximum number of iterations of BPNN training was 2500, and the learning rate was 0.01. The regularization coefficient of the SVM was 500, and the kernel function parameter was set to 0.5. The initial learning rate of DBNI was 0.1, in the network, the number of RBM layer is 6. The performance of the three methods on different training sets is shown in Table 1.

As can be seen from Table 1, when the number of training samples is small, the average accuracy of the identification results of DBNI is basically the same as that of SPNN and SVM. With the increase of the number of training samples, the average recognition rate of DBNI is higher than that of BPNN and SVM.

In order to make the experiment more convincing, we added another group of experiments, as shown in Figure 6.

In Figure 6, in the test, we found that with the increase of training set and model network layer number, the average diagnostic accuracy of the DBNI fault diagnosis method increases, and the growth trend will gradually slow down. However, the average diagnostic accuracies of BPNN and SVM remain unchanged after the number of samples in training set reaches about 200, so it is difficult to improve the value of their parameters.

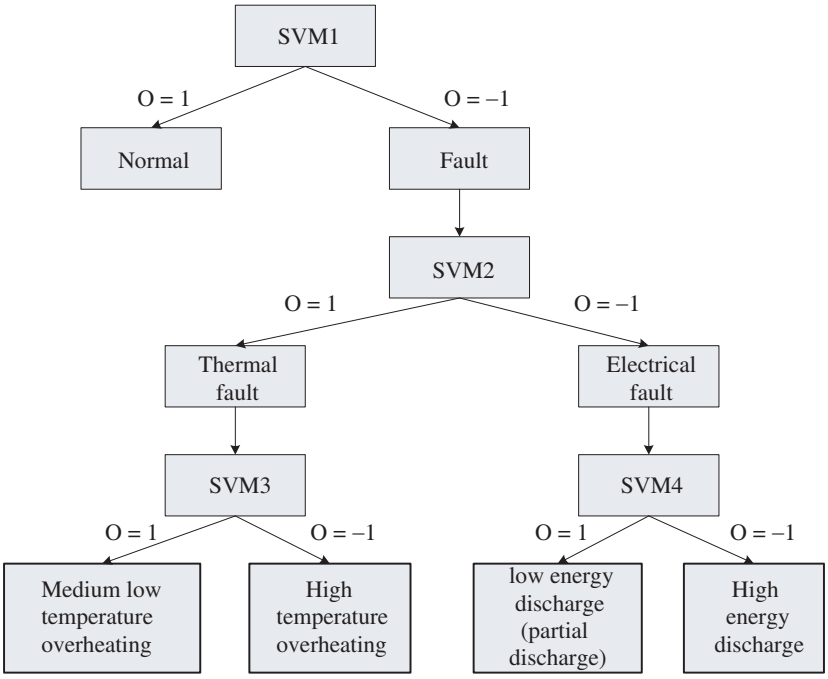

Figure 5. Fault diagnosis model based on SVM.

Table 1. BPNNSVM and DBNI identification case on the different.

Algorithm	Training set	Test set	Average accuracy
BPNN	40	20	72.37%
SVM	40	20	77.69%
DBNI	40	20	75.31%
BPNN	80	20	82.23%
SVM	80	20	86.45%
DBNI	80	20	88.71%
BPNN	120	20	89.11%
SVM	120	20	90.37%
DBNI	120	20	93.19%

3. TRANSFORMER FUSION FAULT DIAGNOSIS ALGORITHM BASED ON DBNI

3.1. Transformer Fault Diagnosis Model Based on DBNI

Based on the second part, we apply the DBNI algorithm to diagnose transformer faults.

Table 2 shows their respective codings.

Transformer fault diagnosis is a multi-layer identification task. The diagnosis results are divided into six types: normal, medium low temperature overheating, high temperature overheating, partial discharge, low-energy discharge, and high-energy discharge.

The processing flow of transformer fault data based on DBNI is shown in Figure 7.

The input of the DBNI model is the discharge signal and temperature value monitored by photoelectric sensors on-line (after fusion and standardization), as shown in Figure 1. After the top-level of the model, the Softmax classifier outputs the probability of fault state. For example, medium low temperature overheating fault probability is 80%; high temperature overheating fault probability is

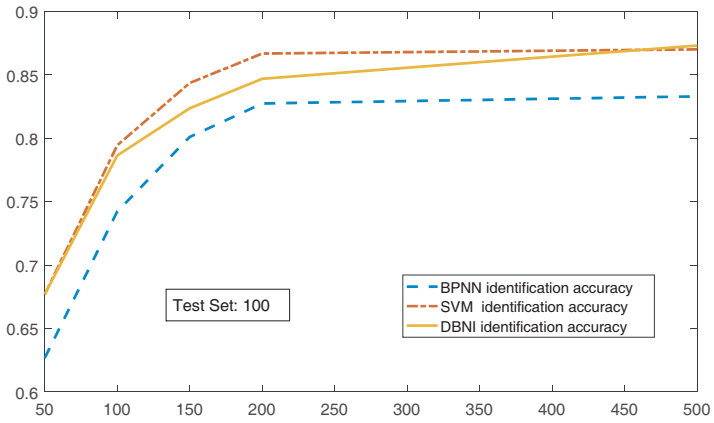

Figure 6. BPNN, SVM, DBNI algorithm identification accuracy.

Table 2. Transformer status codes.

Transformer condition	Code
Normal	(000001)
Mid low temperature overheating	(000010)
High temperature overheating	(000100)
Partial discharge	(001000)
Low energy discharge	(010000)
High energy discharge	(100000)

12%; low energy discharge failure probability is 8%.

Figure 8 shows the transformer fault diagnosis model based on DBNI, which combines the DBN model with the data fusion model based on the optimum weighted algorithm. In the data fusion model, we fuse the data collected by each sensor. Then, we extract the performance parameters of the fault samples from the DBNI model, and the transformer fault analysis results are obtained.

3.2. Fusion of Transformer Fault Information Based on Multi-Sensor

According to Figure 1, collect real-time data of transformer and recorded as x_i. According to Figure 2, perform data fusion calculation on x_i, then we get $\hat{x}$. Take $\hat{x}$ as the input data of Figure 8. The process of optimized weighted fusion algorithm is as follows.

We refer to the variance of the n sensors as $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, respectively. The true value to be estimated is x. The measured values of each sensor are $x_1, x_2, \ldots, x_n$. Suppose that $x_1, x_2, \ldots, x_n$ are independent of each other and that x is an unbiased estimate namely,

$$\begin{aligned} E\hat{x} &= E\left[\sum_{i=1}^{n} g_i x_i\right] = E\left[g_1 x_1 + g_2 x_2 + \ldots + g_n x_n\right] \\ &= g_1 E x_1 + g_2 E x_2 + \ldots g_n E x_n = (g_1 + g_2 + \ldots g_n)x = x \end{aligned} \tag{10}$$

The observation equation of the n sensors for a certain system state is

$$Z = Lx + e \tag{11}$$

In Eq. (11), x is a one-dimensional state variable, and Z is an n-dimensional measurement vector, where $Z = (z_1, z_2, \ldots, z_n)^T$, L is a known n-dimensional constant vector. Suppose $L = (11, \ldots, 1)^T$, and

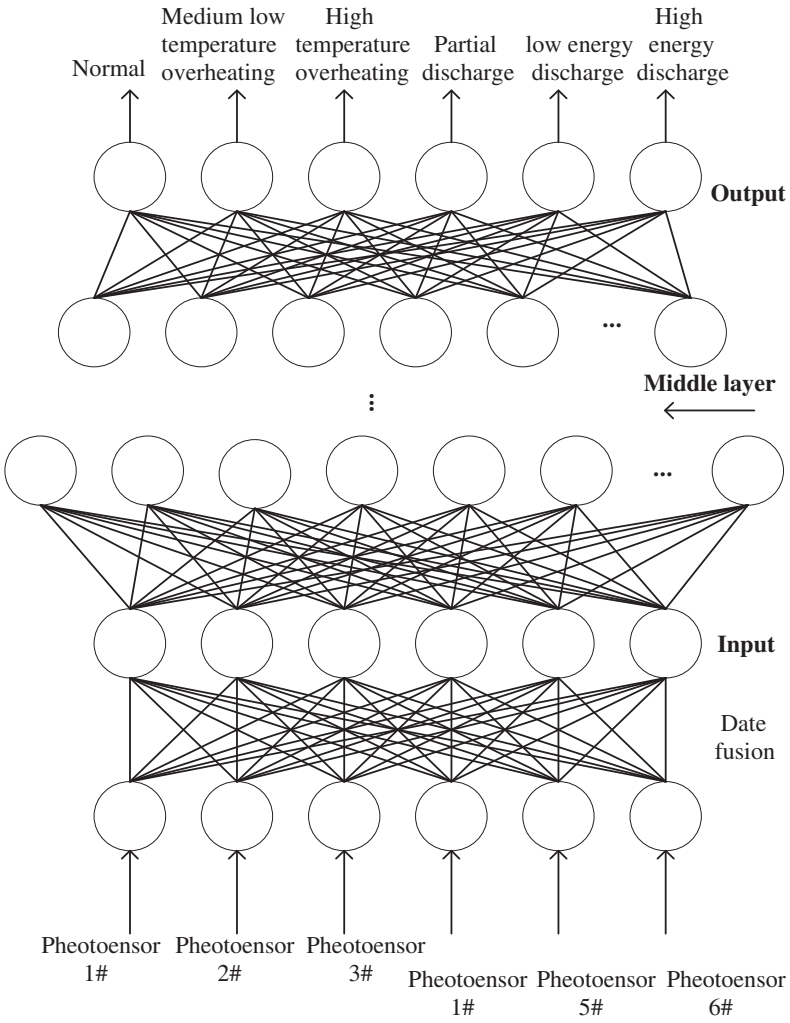

Figure 7. Data processing flow based on DBNI.

e is the n-dimensional measurement noise vector, which contains the internal noise and environmental interference noise of the sensor, where $e = (e_1, e_2, \ldots, e_n)^T$

$$E[e_i] = 0 \quad i = 1, 2, \ldots, n \tag{12}$$

$$E[e_i^2] = E[(x - y_i)^2] = \sigma_i^2 \quad i = 1, 2, \ldots, n \tag{13}$$

In Eq. (13), σ_i^2 is the measurement variance of the ith sensor. The n sensors are divided into m groups, denoted by $Z_1 = (z_{11}, z_{12}, \ldots, z_{1n_1})$, $Z_2 = (z_{21}, z_{22}, \ldots, z_{2n_2}), \ldots, Z_m = (z_{m1}, z_{m2}, \ldots, z_{mn_m})$. The average value measured by each group of sensors is

$$z_i = \frac{1}{n_i} \sum_{j=1}^{n_i} z_{ij} \quad i = 1, 2, \ldots, m \tag{14}$$

The random weighted estimates of each group of measurements in the system observation Equation (11) can be described as follows:

$$Z = Lx + e \tag{15}$$

In Eq. (15), 'Z' (which has m dimensions) is the random weighted vector of each group of sensor measurements.

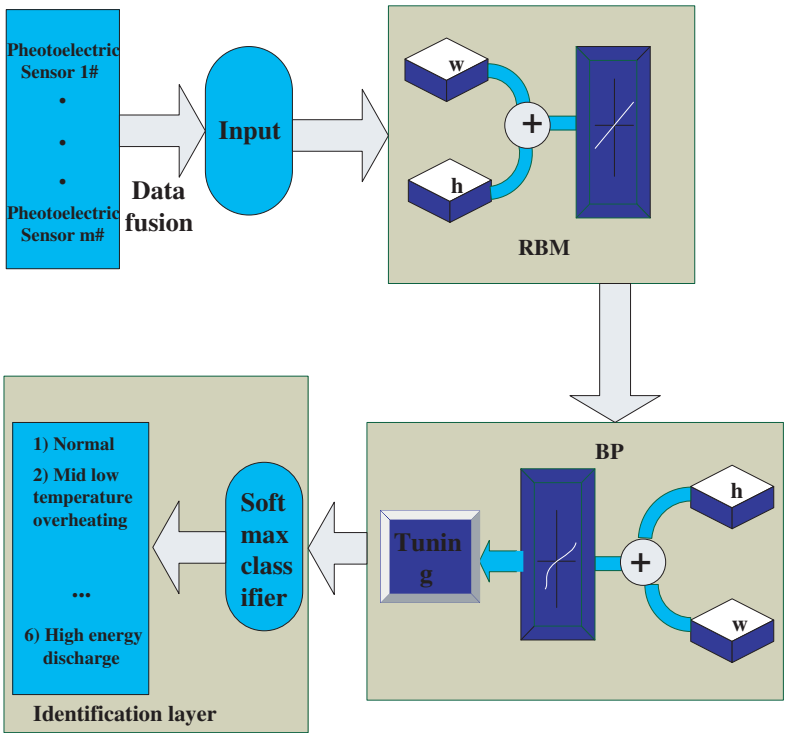

Figure 8. Transformer fault fusion diagnosis and identification model based on DBNI.

We denote $Z = (z_i, z_i, \ldots, z_i)$, and L is a known m-dimensional constant vector. $L = (11, \ldots, 1)^T$, e is a random weighted vector of m-dimensional measurement noise. $e = (e_1, e_2, \ldots, e_m)^T$, in which

$$e_i = \sum_{j=1}^{n_i} g_j e_{ij}, \quad i = 1, 2, \ldots, m \tag{16}$$

g_j is a random weighting factor. From Equation (12) we can deduce

$$E[e_i] = 0 \quad i = 1, 2, \ldots, m \tag{17}$$

If we assume that the total variance is σ^2, then

$$\sigma^2 = E[(x - \hat{x})^2] = E\left[\sum_{i=1}^{n} g_i x - \sum_{i=1}^{n} g_i x_i\right]^2 = E\left[\sum_{i=1}^{n} g_i^2 (x - x_i)^2 + 2 \sum_{i=1, j=1, i \neq j}^{n} (x - x_i)(x - x_j)\right] \tag{18}$$

In Eq. (18), g_i is a random weighting factor. The equation satisfies $\sum_{i=1}^{n} v_i = 1$. Because $x_1, x_2, \ldots, x_n$ are independent of each other, and x_i is an unbiased estimate of x, that is $E\hat{x}_i = x$,

$$E(x - x_i)(x - x_j) = 0 \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, n, \quad i \neq j \tag{19}$$

So σ^2 can be expressed as

$$\sigma^2 = E\left[\sum_{i=1}^{n} v_i^2 (x - x_i)^2\right] = \sum_{i=1}^{n} v_i^2 \sigma_i^2 \tag{20}$$

In Eq. (20), σ^2 is the variance of each sensor.

Equation (20) shows that the total variance σ^2 is a quadratic function of multiple variables with random weighting factors g_i. The minimum value of σ^2 is obtained by the extremum theorem of

multivariate functions and denoted as $\sigma^2_{\min}$.

$$\begin{cases} \sigma^2_{\min} = \min\left[\sum_{i=1}^{n} g_i^2 \sigma_i^2\right] \\ \sum_{i=1}^{n} g_i = 1 \end{cases} \tag{21}$$

When the total variance σ^2 is minimized using Eq. (20), the optimal random weighting factor is

$$\overset{*}{g}_i = \frac{1}{\left[\sigma_i^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right]} \quad i = 1, 2, \ldots, n \tag{22}$$

We collect transformer fault sample data to build a training matrix, a sample label matrix, and a test sample matrix. The sample matrix is a $p * q$ matrix, where q is the number of state parameters. According to the data fusion algorithm, the sample data are fused to obtain more accurate data. To build a DBNI network, we set the number of network layers and the number of nodes, where the maximum number of nodes is $2\sqrt{pq} + q$, and the initial state parameter is set to its minimum value. After that, RBM training and the BP neural network are used to optimize the model parameters. Finally, DBNI training is completed, and the diagnostic results are output. The accuracy rate is obtained by comparing the diagnosis result with the actual diagnosis result.

The specific implementation steps of the proposed DBNI-based transformer fault diagnosis method are shown in Figure 9.

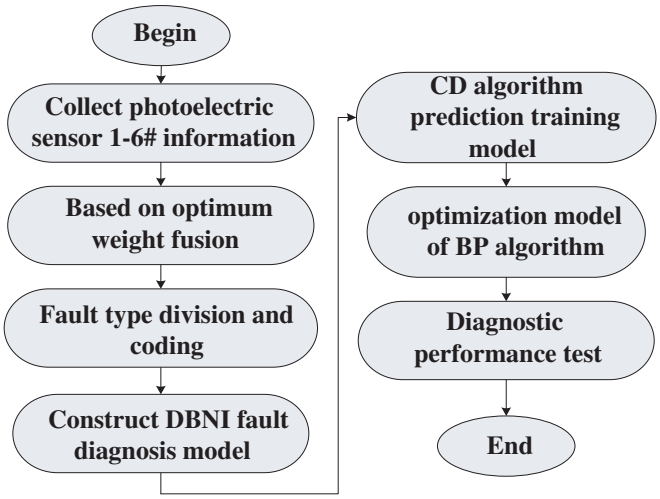

Figure 9. Flow chart of transformer fault diagnosis based on DBNI.

In this paper, we test the diagnostic performance of the DBNI fault diagnosis method using a transformer fault example. See Section 4 for detailed test results.

4. COMPUTATION AND EXPERIMENTAL ANALYSIS

3000 sets of oil-immersed power transformer temperature values in several engineering sites during a certain period of time before and after transformer faults were used as the pre-training set, and 600 sets of transformer electrical experimental data were used as the optimization set. The training/test set were sized using a 4 : 1 ratio. The fault factor was set to 0.1–0.5, and 20 samples were simulated every 0.1. The transformer fault analysis method based on DBNI was tested as follows.

1) Table 3 shows the temperature information at 6 locations of the oil-immersed power transformer in Figure 1, where S1–S6 represent the 6 temperature sensors, according to Figure 1. The temperature

Table 3. Temperature values at different positions of transformers.

Order number	S1	S2	S3	S4	S5	S6
1	102.2	102.4	103.5	103.7	102.8	102.9
2	97.8	97.3	97.8	96.9	975	97.2
3	36.8	36.9	37.6	37	37.3	37.2
4	51.1	51	51.7	52.1	51.4	51.9
5	49	48.7	49.3	49.9	49.4	48.9
6	63.2	63.3	64.1	63.3	63.9	63.4

values of the 6 sensors at the 6 locations were obtained through measurements of sensors. Suppose that the measurement noise of the six sensors is uncorrelated, and white noise variance is 0.04, 0.04, 0.06, 0.06, 0.02, and 0.02, respectively.

We take a temperature sensor as an example to discuss the data fusion of sensors.

Based on Equations (10)–(22),

$$
\begin{aligned}
e_{11} &= -1.3833, \quad e_{12} = -0.1833, \quad e_{13} = 0.9167, \\
e_{14} &= 1.1167, \quad e_{15} = 0.2167, \quad e_{16} = -0.6833 \\
e_{ij} &= \begin{bmatrix} -0.7166 & -0.5167 & 0.5833 & 0.7833 & 0.1167 & -0.0167 \\ 0.3833 & -0.1167 & 0.3833 & -0.5167 & 0.0833 & -0.2167 \\ -0.3333 & -0.2333 & 0.4667 & -0.1333 & 0.1667 & 0.0667 \\ -0.4333 & -0.5333 & 0.1667 & 0.5667 & -0.1333 & 0.3667 \\ -0.2000 & -0.5000 & 0.1000 & 0.7000 & 0.2000 & -0.3000 \\ -0.3167 & -0.2167 & 0.5833 & -0.2167 & 0.3833 & -0.1167 \end{bmatrix} \\
\overset{*}{g}_i &= \begin{bmatrix} 0.0864 & 0.1037 & 0.0877 & 0.0542 & 0.3721 & 0.3266 \end{bmatrix}^T \\
e_1 &= \sum_{j=1}^{n_i} \overset{*}{g}_j e_{1j} = 0.0161 \quad e_2 = -0.0132 \quad e_3 = 0.0645 \quad e_4 = 0.0228 \quad e_5 = -0.0460 \quad e_6 = 0.0941 \\
\sigma_1^2 &= 0.1833 \quad \sigma_2^2 = 0.1527 \quad \sigma_3^2 = 0.1805 \quad \sigma_4^2 = 0.2927 \quad \sigma_5^2 = 0.0445 \quad \sigma_6^2 = 0.0507
\end{aligned}
$$

According to the above calculation, the input values in Figure 7 are

$$\hat{x}_1 = \bar{x}_1 + e_1 = 102.9167 + 0.0161 = 102.9328$$

Similarly, the other 5 data fusion result sets are 97.4035, 37.1978, 51.5561, 49.154, and 63.6117, respectively.

We selected fifteen groups of sample data and simulated with MATLAB. Figure 10 shows the distribution of the measurement error using the optimum weighted fusion algorithm and a set of sensor information data.

From the calculation results, it can be seen that the sensor fusion error is less than that of single sensor measurements.

2) Relationship between fault identification, RBM layers, and pre-training sets of DBNI.

Figure 11 shows the calculation of DBNI-based transformer fault diagnosis for different numbers of RBM layers, while Figure 12 contains the results of the transformer fault identification cases on different RBM layers and pre-training sets.

When testing the number of different RBM training layers for the 3000 transformer temperature values, it can be seen from Figure 11 that when the number of RBM training layers is 3, the average accuracy of fault diagnosis is already high. Figure 12 shows that with the increase of the number of RBM layers, the accuracy rate increases and grows slowly when the pre-training layer is fixed.

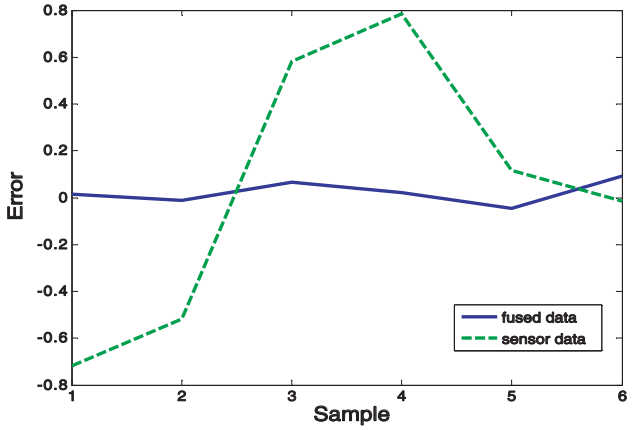

Figure 10. Distribution of measurement error.

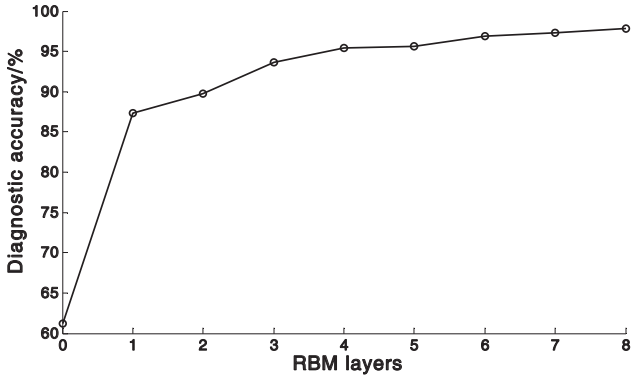

Figure 11. Transformer fault diagnosis based on DBNI with different RBM layers.

3) Relationship between structural network and reconstruction error of DBNI.

We consider the transformer's temperature detection as an example. The temperature signals collected by the photoelectric temperature sensors of 6 different bands were selected to construct a total of 6 DBNI networks with 3, 3, 4, 4, 5, and 5 layers, respectively. The total reconstruction errors of the samples were obtained by substituting all the training samples, as shown in Table 4.

Table 4. DBNI network structure and reconstruction error.

DBNI network structure	Reconstruction error
[6, 6, 8, 6]	0.072
[6, 8, 10, 6]	0.048
[6, 8, 9, 10, 6]	0.063
[6, 8, 10, 11, 6]	0.023
[6, 8, 9, 10, 12, 6]	0.038
[6, 6, 8, 10, 12, 6]	0.049

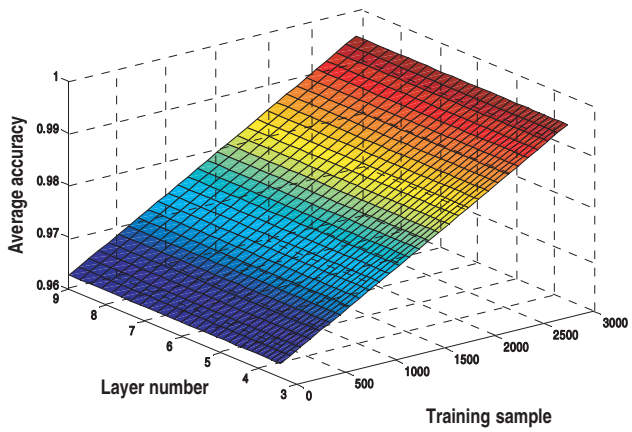

Figure 12. Transformer fault identification case on different RBM layers and pre-training sets.

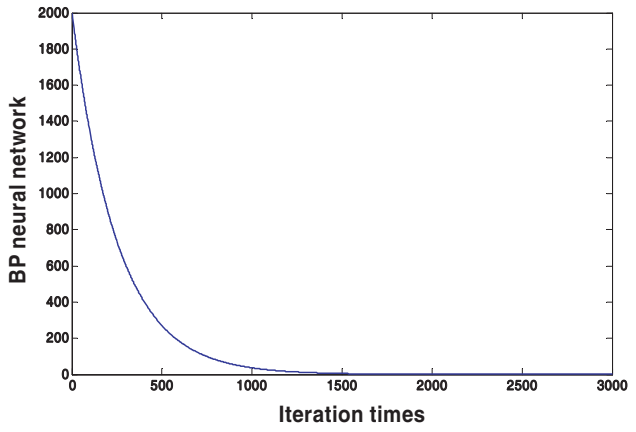

Figure 13. The distribution on BPNN error.

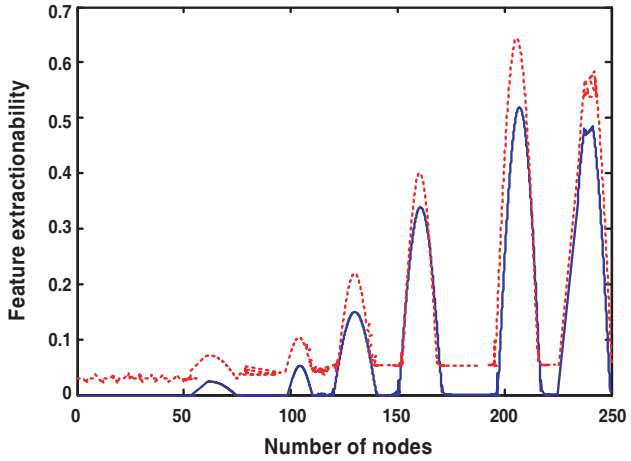

Figure 14. Reconstruction curve with 10 hidden layers.

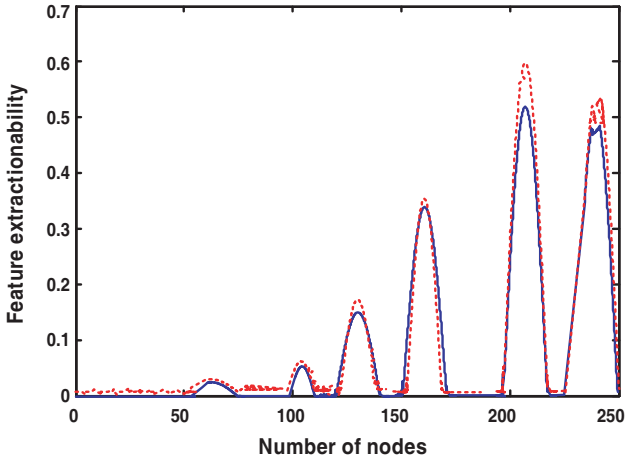

Figure 15. Reconstruction curve with 400 hidden layers.

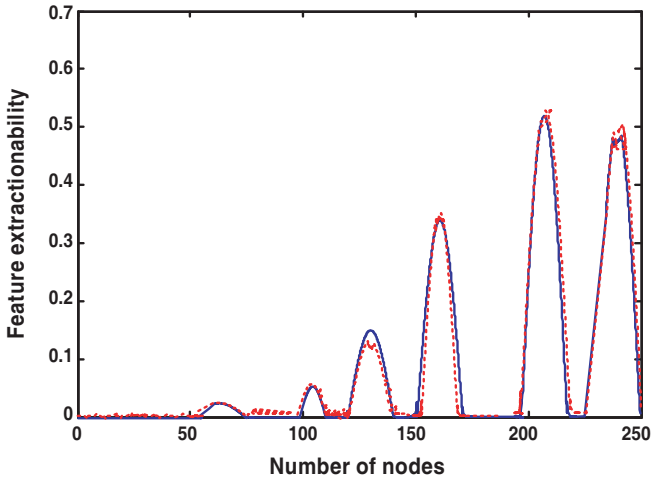

Figure 16. Reconstruction curve with 700 hidden layers.

From Table 4, it can be seen that the DBNI network with the network structure of [6, 8, 10, 11, 6] has the smallest reconstruction error, which shows that the network structure can more accurately mine features hidden in the performance parameters.

The error curve of the BP neural network can be obtained by reverse adjustment of labeled samples, as shown in Figure 13. The error of the BP neural network reflects the matching degree between the hidden features of the state parameters and the fault types. The error of the BP neural network is the sum of the standard deviation between the eigenvalues of all training samples and fault types.

Figure 13 shows the relationship between the error of the BP neural network and the number of iterations. As can be seen, the BP neural network error is very large at the beginning of training, and the sample error is greatly reduced at about 2000 iterations. After 30,000 operations, the final BP error is only 0.39, which meets the requirements.

4) Effect of the number of hidden layer nodes on feature extraction ability of DBNI.

Experiments were carried out on 10, 400, and 700 nodes, respectively, as shown in Figures 14–16. The dashed lines represent reconstructed data, and the solid lines represent raw data.

It can be seen from Figure 16 that with the increasing number of network nodes, the coincidence degree between the reconstructed curve and original curve increases, as is RBM's ability to extract features from original data.

5) DBNI fault fusion diagnosis results.

We selected the DBNI fusion diagnostic network model with network structure [6, 8, 10, 11, 6] and simulated 600 test samples with fault status labels, with 100 samples from each fault type. To test the model's robustness to interference, the fault factor was set from 0.1 to 1.0. We then substituted the test sample into the fusion fault diagnosis model, compared the fault fusion diagnosis result with the fault state label, and then obtained the accuracy rate of the fusion fault diagnosis.

Next, test samples were substituted into the fusion fault diagnosis model based on DBNI and compared with the BP neural network and single DBN, calculating the accuracy rate of these algorithms. It can be seen that the accuracy rate of the fusion fault diagnosis model based on DBNI is 99.45%, which is slightly higher than that of the DBN model without data fusion. This shows that the model can improve the robustness to interference effectively and has a higher accuracy rate of fault diagnosis through the DBNI fusion diagnosis model.

5. CONCLUSIONS

This paper establishes a fusion fault diagnosis model based on DBNI. This model has simple structure and clear inputs and outputs. First, the state parameters of fault samples are fused with confidence data, and then the DBNI algorithm is used to identify multiple fault classifications and obtain the final diagnosis results.

(1) Compared with BP neural network and single DBN and SVM, the DBNI model has higher fault diagnosis accuracy. For example, the accuracy rate of fault diagnosis for transformer winding temperature is 99.45%.

(2) The DBNI method in this model can fully explore the characteristics of large samples, analyze multiple faults, and extract the hidden features of fault samples.

(3) The model uses the DBNI algorithm based on optimum weighted fusion to eliminate the accidental errors in the classification results of DBN faults. The experimental results show that this model has good robustness to interference.

ACKNOWLEDGMENT

This work has been supported by Project of the Key Programs of Shaanxi science and Technology Department (No. 2019GY-034) and Scientific Research Program funded by Shaanxi Provincial Education Department (No. 18JK0388).

REFERENCES

1. Dai, H. S., G. Sheng, and X. Jiang, "Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network," *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 24, No. 5, 2828–2835, Oct. 2017, doi: 10.1109/TDEI.2017.006727.
2. Kaur, A., Y. S. Brar, and G. Leena, "Fault detection in power transformers using random neural networks," *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 9, No. 1, 78–84, February 2019.
3. Mehdipourpicha, H., R. Bo, H. Chen, Md M. Rana, J. Huang, and F. Hu, "Transformer fault diagnosis using deep neural network," *2019 IEEE Innovative Smart Grid Technologies — Asia (ISGT Asia)*, 4241–4245, 2019, 10.1109/ISGT-Asia.2019.8881052.
4. Cabanas, M. F., F. Pedrayes González, M. G. Melero, C. H. Rojas García, G. A. Orcajo, J. M. Cano Rodríguez, and J. G. Norniella, "Insulation fault diagnosis in high voltage power transformers by means of leakage flux analysis," *Progress In Electromagnetics Research*, Vol. 114, 211–234, 2011.
5. Sun, Z. and Q. Cui, "The application of multiclass support vector machine in power transformer fault diagnosis," *Journal of Electronic & Information Technology*, No. 10, 25–28, 2019.
6. Mou, S. and T. Xu, "Fault diagnosis method of transformer based on adaptive deep learning model," *Journal of Software*, Vol. 12, No. 10, 14–18, 2018.

7. Zhang, X. and H. Li, "Research on transformer fault diagnosis method and calculation model by using fuzzy data fusion in multi-sensor detection system," *Optik*, Vol. 176, 716–723, January 2019.
8. Sun, Z. and L. Xue, "Marginal fisher feature extraction algorithm based on deep learning," *Journal of Electronics & Information Technology*, Vol. 35, No. 4, No. 4, 805–811, 2013.
9. Jiang, Y. and L. Huang, "Transformer internal fault diagnosis based on DGA and deep belief network," *Engineering Journal of Wuhan University*, Vol. 50, No. 5, 749–752, 2017.
10. Wang, L. and Y. Zhu, "Parallel phase resolved partial discharge analysis for pattern recognition on massive PD data," *Proceeding of the CSEE*, Vol. 36, No. 5, 1236–1244, 2016.
11. Zheng, Y. and Y. Zhu, "VPMCD method based on support vector regression and its application in partial discharge pattern recognition," *Journal of North China Electric Power University*, 2018.
12. Wang, X. and Y. Zhu, "On identification method of power capacitor dielectric loss angle based on deep learning," *Transaction of China Electrotechnial Society*, Vol. 32, No. 15, 145–150, 2017.
13. Shi, X. and Y. Zhu, "Application of deep learning neural network in fault diagnosis of power transformer," *Electric Power Construction*, Vol. 36, No. 12, 116–121, 2015.
14. Wang, L. and Y. Xie, "A fault diagnosis method for asynchronous motor using deep learning," *Journal of Xi'an Jiaotong University*, Vol. 51, No. 10, 128–134, 2017.
15. Geng, L. and X. Liang, "Real-time driver fatigue detection based on morphology infrared features and deep learning," *Infrared and Laser Engineering*, Vol. 47, No. 2, 2–8, 2018.
16. Zhao, Q. and Z. Li, "Deep multi-task learning for hierarchical classification," *Journal of Computer-aided Design & Computer Graphics*, Vol. 30, No. 5, 886–891, 2018.
17. Che, C. and H. Wang, "Fault fusion diagnosis of aero-engine based on deep learning," *Journal of Beijing University of Aeronautics and Astronauties*, Vol. 44, No. 3, 621–627, 2018.