ORIGINAL ARTICLE

Robust fault diagnosis of a high-voltage circuit breaker via an ensemble echo state network with evidence fusion

Xiaofeng Li[1] · Xiaoying Zheng[2] · Tao Zhang[1] · Wenyong Guo[1] · Zhou Wu[3]

Abstract

Reliable mechanical fault diagnosis of high-voltage circuit breakers is important to ensure the safety of electric power systems. Recent fault diagnosis approaches are mostly based on a single classifier whose performance relies heavily on expert prior knowledge. In this study, we propose an improved Dempster–Shafer evidence theory fused echo state neural network, an ensemble classifier for fault diagnosis. Evidence credibility is calculated through the evidence deviation matrix and the segmented circle function and employed as credibility weights to rectify the raw evidence. Then, an improved Dempster–Shafer evidence fusion algorithm is proposed to fuse evidence from different echo state network modules and sensors. Unlike conventional classifiers, the proposed methodology consists of multiple echo state neural network modules. It has better flexibility and stronger robustness, and its model performance is not sensitive to network parameters. Comparative analysis indicates that it can handle the paradox evidence fusion analysis and thus can achieve better diagnostic performance. The superiority of the reported fault diagnosis approaches is verified with the experimental data of a ZN12 high-voltage circuit breaker.

Introduction

A high-voltage circuit breaker (HVCB) is crucial for ensuring the safety and stability of electric power systems. Due to mechanical, electrical, and environmental impacts, HVCBs suffer from functional deterioration or even failure. According to statistical data, more than half of major failures are caused by mechanical issues (e.g., joint clearance faults caused by wear and corrosion) [1, 2]. Traditional maintenance is generally based on manual schemes, which are quite costly and time-consuming. Furthermore, as more disassembly is involved in the process of manual maintenance, HVCBs might suffer from secondary damage. There

is indeed a requirement for mechanical intelligence fault diagnosis [3, 4].

With the rapid development of industrial informatization and machine learning technology, machine learning-based methods for fault diagnosis of HVCBs have received attention from the industry. Feature extraction and data classification are the main tasks in fault diagnosis. More specifically, in these methods, characteristic data of different mechanical states are collected and then used to train and test the classification model. However, most of these methods are typically based on a single classifier and sensor, whose performance is sensitive to network parameters. In other words, the performance quality needs to be ensured by the expert's prior knowledge of optimal model configurations.

Although various machine learning techniques have been applied in the mechanical fault diagnosis of HVCBs, there are still two open issues. First, as prior knowledge of model structures and parameter selection are indispensable in recent approaches, insufficient expert knowledge may reduce the diagnostic accuracy of these methods. Second, the input data in the most recent fault diagnosis models are collected from a single sensor, and the diagnostic decisions are based on an individual classifier. Thus, the diagnosis models have

✉ Zhou Wu
wuzhsky@gmail.com

1 College of Power Engineering, Naval University of Engineering, Wuhan 430000, China

2 Department of Operational Research and Programming, Naval University of Engineering, Wuhan 430000, China

3 College of Automation, Chongqing University, Chongqing 400044, China

poor robustness, and prediction errors are more likely to occur. Synthesis fault diagnosis models with fusion analysis are rarely reported. Therefore, it is meaningful to explore effective information fusion techniques and further build an ensemble diagnosis model for enhancing the fault diagnosis performance of HVCBs.

To address the above problems, we report an improved Dempster–Shafer evidence theory-fused echo state neural network (IDS–ESN), a novel ensemble model that combines different intelligent classifiers through the improved DS evidence theory. More specifically, energy distributions in vibration intrinsic modal function (IMF) components by variational mode decomposition (VMD) are extracted as the feature vector for diagnostic model training and testing. Then, multiple ESN modules trained by features of different sensors are adopted as sub-classifiers. Moreover, considering the paradox issue, an improved DS evidence fusion algorithm is proposed by evaluating the deviation degree among multiple pieces of evidence. Finally, through evidence fusion, the ensemble IDS–ESN is obtained for the mechanical fault diagnosis of HVCBs.

In summary, the main contributions of this study are threefold.

1. To the best of our knowledge, the capability of ESN is investigated for the first time in the mechanical fault diagnosis of HVCBs. The comparative results demonstrate that this approach achieves promising performance improvement.
2. Typical DS is improved to fuse multisource information from multiple sensors and ESN sub-classifiers. The proposed IDS–ESN model can effectively enhance the accuracy and make the diagnostic model more robust than individual ESN models.
3. For IDS–ESN, a new training mechanism is investigated. In multiple scenarios, IDS–ESN is more robust and can achieve higher diagnostic accuracy than the traditional DS method.

The rest of our paper is structured as follows. "Related works" outlines the main related works. "Platform description and setup" describes the studied HVCB and the experimental setup. "Ensemble echo state network with evidence fusion" describes the theoretical background and our improved DS evidence theory. "Results and discussion" analyses the feature extraction of vibrations by the VMD method and the effectiveness of the improved DS evidence theory. Conclusions are given in "Conclusions".

Related works

In the past decade, artificial techniques have been widely applied in fault diagnosis, online monitoring, and intelligent decision-making in energy field [5, 6]. Feature extraction and machine learning algorithms are investigated in the fault diagnosis literature. For feature extraction, sound [7], contact travel curves [8, 9], electromagnet coil currents [10, 11], and vibrations [12–15] are typical signals used for fault diagnosis. Since abundant structural state-related information is contained in vibrations, most fault diagnosis studies are based on vibration analysis. A series of time–frequency methods have been proposed for signal processing and the subsequent feature extraction of vibration data, including empirical mode decomposition (EMD), local mean decomposition (LMD), empirical wavelet packet decomposition (WPD), the empirical wavelet transform (EWT), and VMD [16–20]. Furthermore, different types of amplitude and frequency vibration features, such as time–frequency entropy, permutation entropy, singular entropy, and energy entropy, were extracted as the input for fault diagnosis models. For multiclass issues, the extracted feature vector is generally high-dimensional and contains redundant components. Thus, the feature selection technique can be used to reduce the dimensionality of a raw feature vector by discarding redundant and irrelevant vectors, which can further enhance the model performance [21, 22]. These feature extraction methods have achieved promising performance with respect to both nonadaptive and adaptive analysis, which contributes to developing intelligent fault diagnosis methods for HVCBs.

For fault diagnosis, a robust and high-accuracy classifier is required to distinguish the type of fault. For example, Tao et al. [23] proposed a feature metric-based fault diagnosis approach under limited data conditions. In the study, a parametric optimization-based meta-learning network and a metric learning network were combined to extract optimization information to adapt between different domains and metric information for similarity discrimination, respectively. For the diagnosis of spring fatigue and oil damper leakage, Ma et al. [24] reported a fault diagnosis model based on a random forest (RF) model to reduce nonessential features. Zhang et al. [25] designed an asynchronous interval type-2 fuzzy approach to address the fault detection problem of a quarter-car suspension system. In particular, a fault diagnosis model based on a support vector machine (SVM) that relied less on the sample size was reported [26]. By optimizing kernel and penalty parameters, advanced SVM methods have been reported and have achieved performance improvements. Biclass and multiclass SVM methods

have been investigated for mechanical fault diagnosis of HVCBs, in which the looseness of a base screw, electromagnet immobility, and overtravel have been diagnosed successfully [27, 28]. Compared with other neural networks, echo state networks (ESN) have strong self-adaptability in storing nonlinear input–output mapping relationships. Only the output weights of the ESN need to be trained by a linear regression algorithm, which avoids gradient disappearance and high computational complexity. Therefore, ESN models have been widely applied in control [29, 30], pattern recognition [31, 32], and nonlinear time series prediction [33, 34]. As a novel research field, machine learning and swarm intelligence approaches have been successfully combined, and outstanding results have been obtained in different areas.

Dempster–Shafer (DS) evidence theory could be utilized as a fusion approach for fault diagnosis, target recognition, and condition monitoring [35]. However, while there are probable conflicts of different evidence, the traditional DS evidence fusion method could produce a contradictory result. Considering that there might be deviations in sensor data, Murphy [36] collected the evidence multiple times and calculated the average to weaken the conflict caused by these deviations. On this basis, Deng et al. [37] further introduced the Euclidean distance function to calculate the support degree of specific evidence by other evidence, which improved the robustness of the original method. A new evidence fusion rule was reported for bearing and gear fault diagnosis by Li et al. and Zhang et al. [38, 39], in which evidence credibility was obtained by substituting the evidence distance matrix with a modified Gini index function. To prevent conflicting evidence fusion in the conventional DS approach, this fusion rule is achieved directly by evaluating the correlation between different evidence data. The DS evidence theory has been successfully applied in the mechanical fault diagnosis of rotating parts, such as bearings, gearboxes, and rotors, providing valuable guidance for fault diagnosis of HVCBs.

Platform description and setup

High voltage circuit breaker

A spring-driven HVCB uses stored elastic potential energy to realize the operations of opening and closing and is now widely used in energy power systems. Figure 1 describes the mechanical system of the studied ZN12 HVCB. Its fast operation is generally controlled within tens of milliseconds. Thus, the structural dynamic characteristics are quite sensitive to clearance joints. Severe collisions between moving parts would be strengthened by clearance joints, thus deteriorating the operation quality. Under the influence of wear, corrosion, etc., the joint clearance size will increasingly deviate
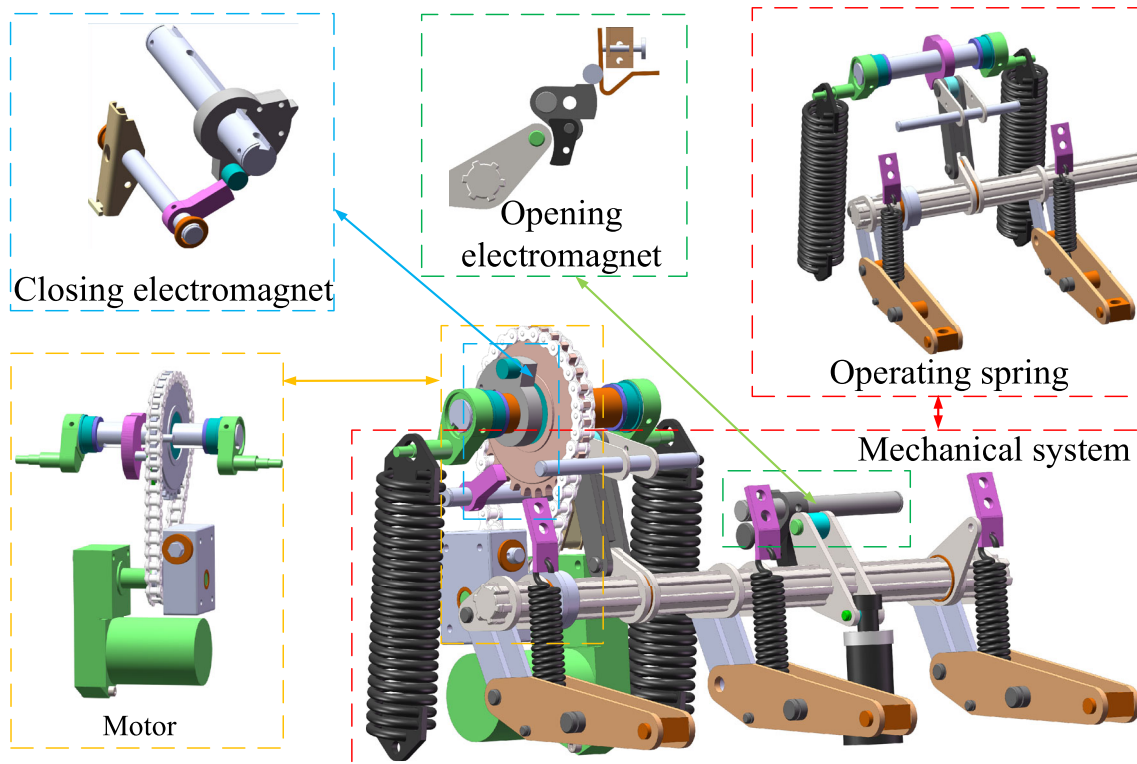
Fig. 1 Mechanical system of the ZN12 HVCB
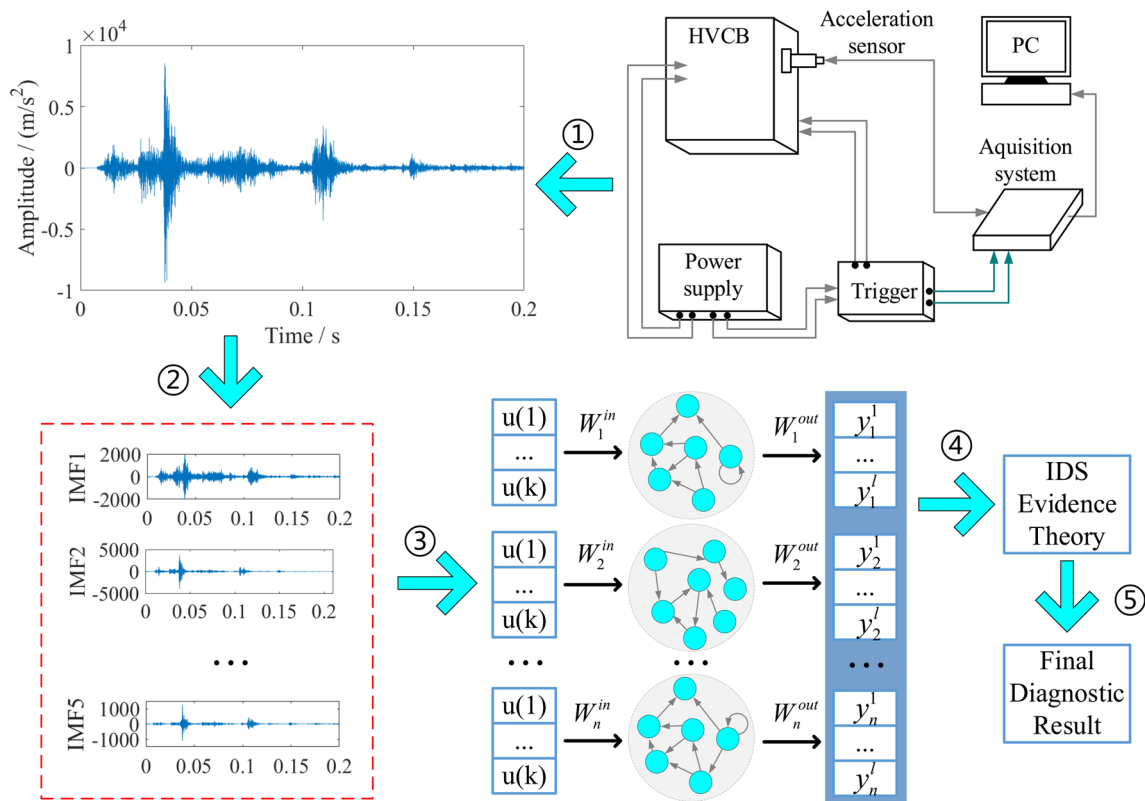
Fig. 2 Flowchart of the IDS–ESN-based fault diagnosis model

from the original design value, which leads to greater impact stress during mechanical operation and induces mechanical faults. Therefore, this paper focuses on diagnosing mechanical faults caused by clearance joints.

The vibration signal consists of a series of complex high-frequency vibration waves generated by various components (e.g., electromagnet and motor components, as shown in Fig. 1). Its amplitude, frequency, energy distribution, and other signal characteristics could contain large amounts of structural state-related information. In this paper, combining the energy distribution of vibration signals in different IMF components of VMD, an ensemble classifier is proposed to diagnose HVCB faults. Multiple ESN modules are adopted as sub-classifiers and fused via an improved DS evidence fusion theory. The overall technical flowchart is given in Fig. 2, and it illustrates the main steps of fault diagnosis in the paper.

Step 1 Repeat the operation experiments of a ZN12 HVCB in different mechanical states. Except for the normal condition, clearance joint faults at three locations are constructed. Furthermore, by changing the joint clearance size, we can obtain two fault conditions for each fault joint clearance. Each type of operation experiment under different fault conditions is carried out 100 times. Thus, a total of 400 groups of vibration signals of two measuring points under different operation conditions are collected.

Step 2 Decompose the obtained vibration signal by VMD method. Then, the energy distribution in different IMFs is extracted as the mechanical state-related feature of the HVCB. The most common way to evaluate a fault diagnosis model is to use training data to train the model and test data to test the model's performance. Therefore, the feature data obtained from the operation experiment are divided into four data sets, as listed in Table 1, for model training and testing.

Step 3 Perform model training of multiple ESN modules, in which random uniform sampling of the original accelerator data is applied. Taking data set A in Table 1 as an example, 25 training samples (50 data points for each fault condition) are randomly selected for each load condition to generate a training set. The remaining 25 samples are classified as test sets. It is important to note that to guarantee an objective result, the testing data should not appear in the training data set. Otherwise, over-optimistic problems could emerge in the diagnostic result.

Step 4 Employ the trained ESN modules with different network parameters (spectral radius and different numbers of reservoir neurons) as the sub-classifiers of the ensemble model. Their model training tasks are implemented on the above data sets in Table 1. In addition, raw diagnostic results

Table 1 Data set for model training and testing

Fault type	Normal		Fault I		Fault II		Fault III	
Clearance size (mm)	< 0.04	< 0.04	0.25	0.75	0.25	0.75	0.25	0.75
Data set A								
Training	25		25		25		25	
Testing	25		25		25		25	
Data set B								
Training		25		25		25		25
Testing		25		25		25		25
Data set C								
Training	50		50		50		50	
Testing		50		50		50		50
Data set D								
Training		50		50		50		50
Testing	50		50		50		50	

of the different mechanical states mentioned in step 1 then be obtained.

Step 5 Produce the fused fault diagnosis result. In the fusion process of the ensemble model, the output of each ESN module is considered as the raw evidence and fused by DS evidence fusion algorithm for the final fault diagnosis result by our improved DS evidence fusion algorithm. It is worth mentioning that for objective fault diagnosis evaluation, the established models often need to be tested many times for average accuracy, and multiple sampling steps are necessary. Taking data set A in Table 1 as an example, for each training circle, the 25 data points in the training set need to be randomly sampled from data set A, and the original testing set is replaced with the remaining 25 data points.
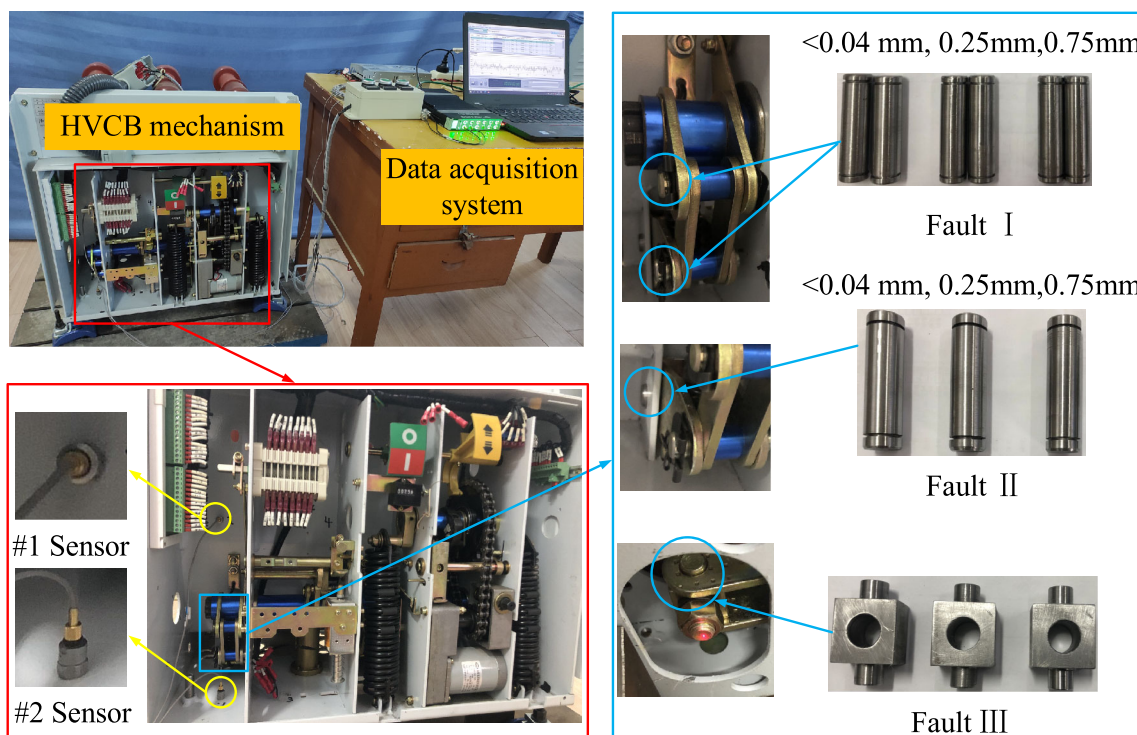
Fig. 3 Experimental setup

Table 2 Fault position, raw vibration signal, and its FFT spectra

Type	Position	Raw signal	Spectra
Normal	/		
Fault I			
Fault II			
Fault III			

Testing platform

The HVCB is generally in a long-term static state in its service period, resulting in a shortage of sample data in recent fault diagnosis research. Limited by insufficient fault data, clearance joint fault diagnosis has rarely been reported in previous studies. In this paper, an experimental setup based on a real ZN12 HVCB is built. As shown in Fig. 3, in addition to the normal operating condition, joint fault conditions at three positions are simulated by changing the joint clearance size (fault sizes of 0.25 mm and 0.75 mm). For the normal condition, its joint clearance size is controlled within 0.04 mm, which is negligible.

Two sets of monoaxial CCLD/IEPE shock accelerometers (Brüel and Kjær, type 8339) are screwed into the HVCB to record vibrations during the closing operation. The measurement range, sensitivity, and upper cutoff frequency of our accelerometer are $\pm$ 10,000 m/s^2, 0.25 mv/g, and 20 kHz,

respectively. In addition, a Brüel & Kjær type 3053-B-120 signal acquisition card is adopted for vibration collection. It is worth mentioning that to prevent mutual influence between two adjacent mechanical operations of the HVCB, an interval of more than 3 min for each opening and closing operation test is needed. Furthermore, our multiple comparison tests have shown that the installation location of the accelerometer is significant for effective vibration measurement. To ensure that the extracted vibration signal contains sufficient state-related information, the accelerometers should be installed near the failure point, but the distance between the two measuring points should be as far as possible. It is ideal for the accelerometer axis lines to be perpendicular to each other.

During the test, the sampling frequency of the acquisition card is set to 65,536 Hz with a sampling time of 1 s. For each normal case and for the three types of joint clearance faults, there are 100 groups of vibration signals from the two accelerometers, and each signal contains 65,536 data points.
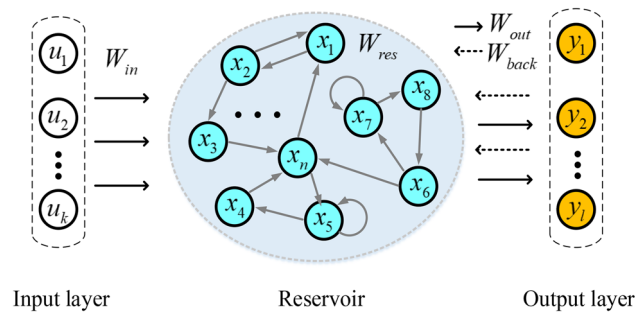
Fig. 4 Regular network architecture of ESN

Table 2 lists the three joint fault locations and the relative FFT spectrum of vibration signals from sensor #1 under the four operating conditions.

Table 2 shows that the FFT spectrum of the obtained vibration signals covers the whole frequency range (0–32,768 Hz), which ensures that the measured data contain all state-related information of the HVCB. On the whole, the vibration signal has the characteristics of a narrow time domain, wide frequency band, and high amplitude. In addition, its vibration FFT spectrum characteristics under different working conditions, including the evolution trend and the amplitude peak point versus frequency, have significant similarities, which makes fault differentiation difficult. Next, the ESN-based technique is applied to the aforementioned fault diagnosis.

Ensemble echo state network with evidence fusion

Echo state network

ESN is a simple and efficient recurrent neural network that is gaining popularity in time series forecasting. As shown in Fig. 4, the hidden layer in the ESN is replaced by a dynamic reservoir, which consists of a large number of sparsely connected neurons.

For illustration, taking an ESN module composed of K input neurons, N reservoir neurons, and L output neurons as an example, the basic equations can be described as follows:

$$x(n + 1) = \tan h(W^{\text{in}}u(n + 1) + W^{\text{res}}x(n) + W^{\text{back}}y(n)), \tag{1}$$

$$y(n + 1) = W^{\text{out}}u(n + 1), \tag{2}$$

where $x(n) = (x_1(n), ...x_N(n))^T$, $y(n) = (y_1(n), ...y_L(n))^T$ and $u(n) = (u_1(n), ...u_k(n))^T$ represent the input vector, the reservoir state, and the output vector, respectively. W^{in},

W^{res}, and W^{out} denote the input-to-hidden, hidden-to-hidden, and hidden-to-output connection weight matrices, respectively. In the ESN, W^{out} is the only parameter that needs to be trained, and other weights are randomly determined and do not change with the training process. In the traditional method, the parameter W^{out} is determined by minimizing the training error as

$$W^{\text{out}} = \arg\min_{W} \|WX - D\|_2^2, \tag{3}$$

where X and D are the $L_{tr} \times N$ reservoir state matrix and $L_{tr} \times L$ signal label matrix during the training stage, respectively, and L_{tr} is the training length. X and D can be calculated as follows:

$$X = \begin{bmatrix} x_1(1) & x_2(1) & ... & x_N(1) \\ x_1(2) & x_2(2) & ... & x_N(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(L_{tr}) & x_2(L_{tr}) & ... & x_N(L_{tr}) \end{bmatrix}, \tag{4}$$

$$D = \begin{bmatrix} d_1(1) & d_2(1) & ... & d_L(1) \\ d_1(2) & d_2(2) & ... & d_L(2) \\ \vdots & \vdots & \vdots & \vdots \\ d_1(L_{tr}) & d_2(L_{tr}) & ... & d_L(L_{tr}) \end{bmatrix}, \tag{5}$$

where $d(t) = [d_1(t), ... d_L(t)]$ denotes the tth teacher signal. In general, Eq. (3) can be directly solved by the pseudoinverse method. However, an ill-posed solution and overfitting problems can occur due to the pseudoinverse technique for high-dimensional internal states. In that case, a two-norm penalty term of regularization parameter ρ is introduced to the cost function:

$$W^{\text{out}} = \arg\min_{W} \|WX - D\|_2^2 + \rho\|W\|_2^2. \tag{6}$$

Then, the solution of Eq. (6) is given by

$$(W^{\text{out}})^T = (X^T X + \rho I)^{-1} X^T D, \tag{7}$$

where I is the identity matrix. In the current study, the output weights are trained by the ridge regression mechanism. The training process of the traditional ESN module is outlined in Algorithm 1 [34].

Algorithm 1 Training procedures of ESN

1: **Input**: u;

2: **Output**: W^{out}

3: Data preprocessing: normalize the dataset, and randomly divide

it into a training set and a test set;

4: Network initialization: $\{W^{in}, W^{res}, W^{back}, K, N, L, \rho\}$;

5: **for** i←1 to L_{tr}

6: Compute $x(i)$ as in Eq. (1)

7: **end**

8: X←$(\mathbf{x}(1),...,\mathbf{x}(L_{tr}))$

9: D←$(\mathbf{d}(1),...\mathbf{d}(L_{tr}))$

10: $W_{out}{}^T$←$(X^T X + \rho I)^{-1} X^T D$

It is worth mentioning that whether the reservoir has the echo state property is very important for the performance of the ESN. The meaning of the echo state property is that the influence of previous input on the future state should gradually tend to disappear. During training, the reservoir gradually removes the historical information to reach an asymptotically stable state. Overall, the main features of the ESN module can be summarized as follows: (1) the core structure is a randomly generated and unchanged sparsely connected reservoir. (2) The output weight is the only part that needs to be trained and adjusted. (3) Its training task can be completed by simple linear regression.

Dempster–Shafer evidence theory

DS evidence theory can be employed to combine information from multiple sensors or evidence sources to produce a synthesis decision. For example, suppose a set comprises N mutually independent and exclusive elements that describe a certain system and the so-called discernment frame.

$$\Theta = \{A_1, A_2, \ldots A_i, \ldots, A_N\}. \tag{8}$$

In fault diagnosis of HVCBs, elements $A_1 \sim A_N$ in the discernment frame generally represent the possible working conditions, including the normal and fault states. Then, $m(A)$ is defined as the basic probability assignment function, which reflects the confidence degree of a specific subset of Θ. This function obeys the following limitations.

$$\begin{cases} m(\phi) = 0 \\ 0 \le m(A) \le 1, \forall A \subset \Theta \ . \\ \sum_{A \subset \Theta} m(A) = 1 \end{cases} \tag{9}$$

Assume that there is a finite BFA function $m_1, m_2, \ldots, m_n$; then, the evidence fusion law is defined as follows:

$$(m_1 \oplus m_2 \cdots \oplus m_n)(A) = \frac{1}{1-k} \sum_{A_1 \cap A_2 \ldots \cap A_n = A} m_1(A_1) \\ * m_2(A_2) \ldots * m_n(A_n), \tag{10}$$

where k is called the conflict coefficient, which is applied to evaluate the degree of evidence conflict and can be expressed as follows.

$$k = \sum_{A_1 \cap A_2 \ldots \cap A_n = \emptyset} m_1(A_1) * \ldots m_n(A_n) \\ = 1 - \sum_{A_1 \cap A_2 \ldots \cap A_n \ne \emptyset} m_1(A_1) * \ldots m_n(A_n). \tag{11}$$

DS evidence fusion theory provides a good method to synthesize information from different sources, but there are also some disadvantages. As k [refer to Eq. (11)] increases, the conflict degree between the different pieces of evidence becomes more serious, and the result obtained by the fusion method will contradict the intuitive judgment. As shown in Table 3, the common evidence fusion paradox issues, including the complete conflict paradox, 0 trust paradox, 1 trust paradox, and high conflict paradox, limit the applicability of DS theory [38]. For fault diagnosis issues, $A \sim E$ in Table 3 can be treated as concerned fault types. $m_1 \sim m_4$ can be the produced fault prediction results of each classifier module. For instance. Module m_1 in the complete conflict paradox line completely diagnoses the input data as fault type A. However, module m_2 diagnoses it as fault type B with 100% certainty. Therefore, the diagnostic results obtained by m_1 and m_2 are self-contradictory, which leads to the complete conflict paradox issue. Modules m_3 and m_4 produce probability distributions of the three fault types.

Improved Dempster–Shafer evidence theory

In the conventional DS evidence fusion algorithm, inaccurate fusion results could be caused by conflicts from different evidence sources. In this section, the evidence deviation degree is first calculated by the Euclidean distance and then substituted into a segmented circle function to evaluate the evidence's credibility. Evidence credibility is further employed as the weight to amend the original evidence and decrease evidence conflict.

First, the BPA matrix $B_{M \times N}$, where M and N represent the kinds of evidence and proposition in the frame of discernment Θ, is defined. To obtain the deviation degree of a particular piece of evidence, $p_i = (m_i(A_1), m_i(A_2), \ldots m_i(A_3))$ is defined

Table 3 Paradox issues in convolutional DS fusion

Paradox	Evidence	Proposition				
		A	B	C	D	E
Complete conflict paradox	m_1	1	0	0	–	–
	m_2	0	1	0	–	–
	m_3	0.8	0.1	0.1	–	–
	m_4	0.5	0.2	0.3	–	–
0 trust paradox	m_1	0.5	0.2	0.3	–	–
	m_2	0.5	0.2	0.3	–	–
	m_3	0	0.9	0.1	–	–
	m_4	0.5	0.2	0.3	–	–
1 trust paradox	m_1	0.9	0.1	0	–	–
	m_2	0	0.1	0.9	–	–
	m_3	0.1	0.15	0.75	–	–
	m_4	0.1	0.15	0.75	–	–
High conflict paradox	m_1	0.7	0.1	0.1	0	0.1
	m_2	0	0.5	0.3	0.1	0.2
	m_3	0.6	0.1	0.15	0	0.15
	m_4	0.55	0.1	0.1	0.15	0.1
	m_5	0.6	0.1	0.2	0	0.1

as the ith row in the BPA matrix. Then, the Euclidean distances between p_i and other evidence points are added to obtain the deviation degree of the ith evidence point:

$$d_i = \sum_{j=1}^{M} \|p_i - p_j\| = \sum_{j=1}^{M} \sum_{k=1}^{N} |m_i(A_k) - m_j(A_k)|, \quad (12)$$

where d_i is the so-called absolute deviation degree coefficient of the ith evidence point. In Eq. (12), the range of $m_i(A_j)$ is [0,1], and the sum of $m_i(A_j)$ meets the criteria of $\sum_{j=1}^{N} m_i(A_j) = 1$. Furthermore, the relative deviation degree coefficient is obtained as follows:

$$\varepsilon_i = \frac{d_i}{2M} \quad \varepsilon_i \in [0, 1], M \geq 3, \quad (13)$$

where M represents the data category, that is, the categorical series of mechanical faults in the paper. Parameter ε_i reflects the similarity between the ith evidence point and the other evidence points. In other words, a smaller ε_i value (closer to 0) means that evidence m_i has greater consistency with the other evidence and is more objective; thus, better evidence credibility should be given and vice versa. Therefore, the evidence credibility should be a decreasing function of ε_i. For that purpose, we propose a segmented circle function as follows:

$$\varepsilon_i^* = \begin{cases} 0.5 + \sqrt{0.25 - \varepsilon_i^2} & (0.5 > \varepsilon_i \geq 0) \\ 0.5 - \sqrt{0.25 - (\varepsilon_i - 1)^2} & (1 \geq \varepsilon_i \geq 0.5) \end{cases}, \quad (14)$$
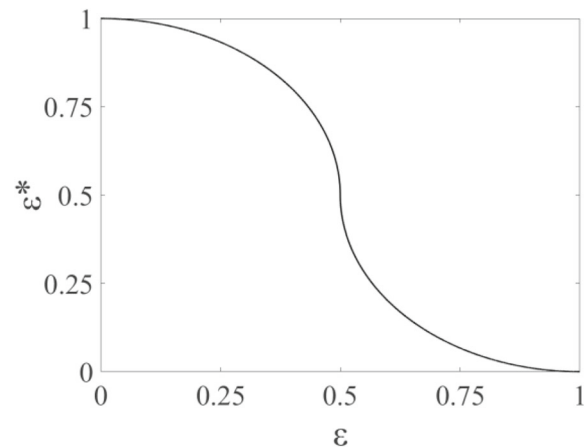
Fig. 5 Segmented circle function curve

where ε_i^* represents the final deviation degree coefficient. It can be detected that within the argument range [0, 1], Eq. (14) is a typical decreasing function that varies from 1 to 0. In addition, at the start and end of the curve, the slope is smaller (refer to Fig. 5), which retains higher credibility of evidence with a low deviation degree and reduces the influence of abnormal evidence on the final fusion result.

Furthermore, the credibility of the obtained evidence is employed as a weight coefficient to amend the original evidence:

$$\begin{cases} m_i^*(A_j) = m_i(A_j)\varepsilon_i^* \quad (j = 1, 2 \ldots N) \\ m_i^*(\Phi) = 1 - \sum\limits_{j=1}^{N} m_i^*(A_j) \end{cases}, \qquad (15)$$

where Φ represents the uncertain proposition and $m_i^*(\Phi)$ describes the uncertainty degree of the ith proposition of the discernment frame. It should be noted that in the improved DS evidence theory, we calculate the summation of the Euclidean distances between the selected evidence and others and then obtain the evidence credibility by the segmented circle function. This unique combination can give higher confidence to the evidence that has better similarity with others. Furthermore, it weakens the negative effects of conflicting evidence, which will improve the accuracy of the fusion results.

Results and discussion

While using VMD, the final IMF number (I) needs to be predetermined and is vital for the effectiveness of this signal analysis method, since it directly affects the performance of VMD [40]. In this section, the centre frequency observation result is first analyzed to determine the optimal parameter I. Then, we perform two sets of experimental result analyses. In the first set, we evaluate how well our IDS evidence fusion algorithm solves the paradoxes by comparing it with several existing fusion approaches. In the second set, we analyze the performance of IDS–ESN for mechanical fault diagnosis of HVCBs and compare its performance with other widely applied classifiers and DS–ESN.

State-related feature extraction

In general, the centre frequency interval of adjacent IMFs gradually shrinks with increasing I and eventually leads to the over-decomposition problem. On the other hand, too small a predetermined I might bring about the modal aliasing issue, which causes the failure of state-related information separation. At present, there is no standard method for the selection of this parameter, and it is selected mainly based on expert experience. Although there are some reported techniques, such as introducing the cuckoo search algorithm into VMD to determine the mode number [5], they suffer from high computational complexity and are time-consuming. The centre frequency observation method is directly applied to determine the optimal mode number in this paper. Therefore, multiple VMD attempts are conducted to determine I. For the convenience of centre frequency observation, the different FFT spectra of IMFs (I varies in a range of 4–7) are given in Fig. 6.

The original vibration is well-decomposed into 5 IMFs with relatively uniform centre frequency intervals. Hence,

no obvious centre frequency increase occurs in the last IMF. For instance, when I is set to 7, the highest centre frequency is 26697 Hz, which is only an increase of 125 Hz in the case of $I = 5$. Moreover, as I increases ($I > 6$), the centre frequency difference between adjacent IMFs decreases, and the modal aliasing phenomenon occurs. In summary, the optimal mode number I of VMD in this paper is 5. Moreover, as the feature value of the HVCB's health state, the energy distribution of the vibration signal on the 5 IMFs is calculated as follows:

$$P_i = \frac{E_i}{E}, \qquad (16)$$

$$\begin{cases} E_i = \int\limits_{t_0}^{t_i} |A(t)|^2 \mathrm{d}t \\ E = \sum\limits_{1}^{5} E_i \end{cases}, \qquad (17)$$

where i is the serial number of the IMFs. t_0 and t_i represent the start and end times of the collected vibration signal, respectively. $A(t)$ and E_i denote the amplitude at different timepoints and the signal energy of each IMF, respectively.

Discussion of the improved Dempster–Shafer model

To demonstrate the superiority of the improved DS (IDS) algorithm, we compare the paradox evidence fusion ability of our IDS with that of the traditional and modified DS algorithms from Murphy [36], Deng [37], and Li [38]. Taking the data in Table 3 as an example, the comparative fusion results are listed in Table 4.

The complete conflict paradox ($k = 1$) makes the denominator of Eq. (10) zero and eventually leads to fusion failure of the convolution DS fusion method. In the 0 trust paradox, since evidence m_3 is equal to zero in column A, the final BPA of proposition A is always zero regardless of how other evidence supports A. For the 1 trust paradox, the final BPA results are $m(A) = m(C) = 0$ and $m(B) = 1$. This is contrary to the actual situation, as there is very little supporting evidence for proposition B. Regarding the high conflict paradox, the final fusion results are $m(B) = 0.3571$, $m(C) = 0.4286$, and $m(E) = 0.2143$. Due to the high conflicts among evidence, proposition C has a larger BPA value than proposition B, which is also contrary to the correct situation. Table 4 shows that all four modified methods can deal with paradoxical evidence fusion problems but with different BPA values. In the complete conflict paradox, 0 trust paradox, and 1 trust paradox, the maximum BPA values in IDS and the algorithm from Li are similar and obviously higher than those of the other methods. In the high conflict paradox, IDS has a maximum BPA value of 0.9870. For the maximum BPA value, the closer to 1 it is, the more confidence it presents

Fig. 6 FFT spectra with varying I values

(a) I=4

(b) I=5

(c) I=6

(d) I=7
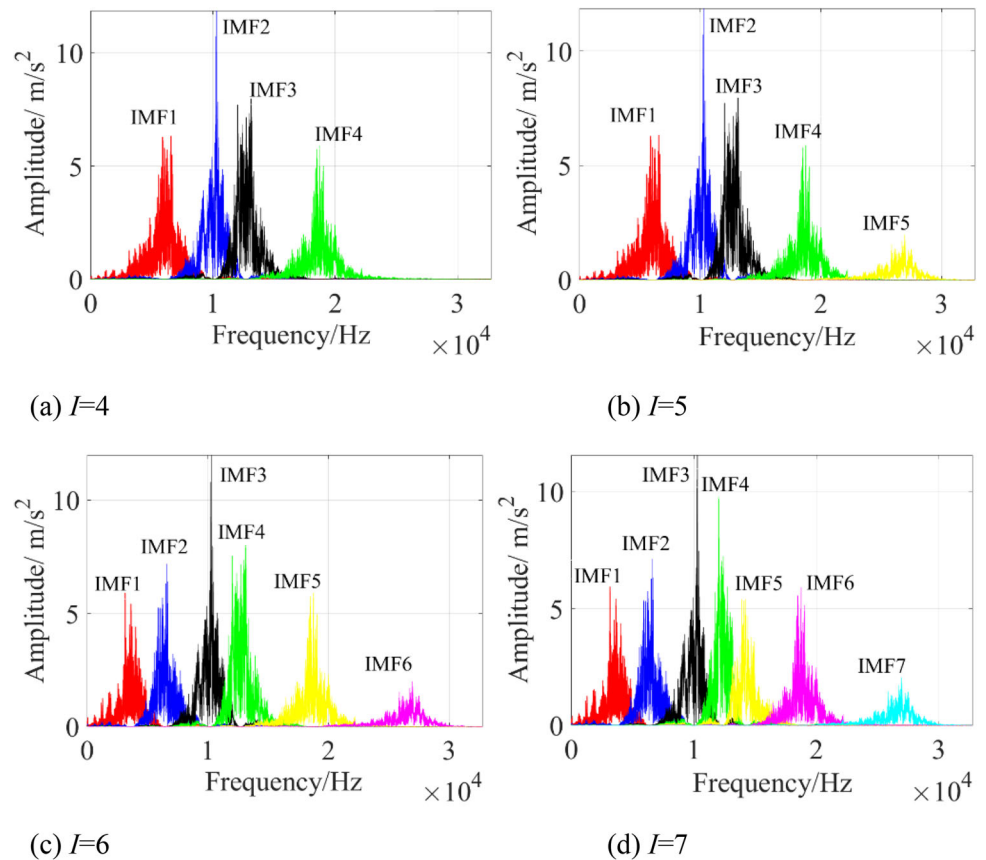
intuitively. Therefore, IDS is superior to other algorithms in high-conflict paradox evidence fusion cases. In addition, it is important to note that, compared with other faults, such as spring fatigue and overtravel faults, the state characteristic of clearance joint faults in this paper is weaker. Thus, a high conflict paradox among different sensors and classifiers is more likely to appear. In summary, our proposed IDS model can achieve the best overall performance while addressing paradoxical evidence for fault diagnosis of HVCBs.

Fault diagnosis result

In this paper, the average accuracy and stability at multiple times are selected as the performance evaluation indicators. First, multiple comparative experiments with several other state-of-the-art classifiers, namely, the back propagation (BP) neural network, radial basis (RBF) neural network, SVM model, extreme learning machine (ELM), and RF network are conducted to evaluate the performances of individual ESN modules. To verify the adaptability of the fault diagnosis model, these classifiers are trained and tested with the same input, corresponding to the different fault joint clearance sizes. Since the data in data sets $A \sim D$ (refer to Table 1) can come from single or multiple sensors, there are three combinations of training and test sets in this paper. To prevent

random sampling errors, the same model training and testing process is repeated 100 times, and the average accuracy is taken as the final result, as shown in Fig. 7.

In Fig. 7, 1 and 2, represent the experimental data of 0.25 mm and 0.75 mm fault joint clearance sizes, respectively, in the subsequent sections. $1 \rightarrow 2$ means training the model on the input of the 0.25 mm fault clearance size and then testing it on that of the 0.75 mm fault clearance size. Through Fig. 7, the following observations can be made. First, in Fig. 7a–c, the ESN has the highest diagnostic accuracy. In addition, according to our literature review, ESN is used in the field of mechanical fault diagnosis of HVCBs for the first time, which points out a new direction for this field. Second, the mechanical fault diagnosis performance results of different classifiers with multiple sensors as input are better than those with a single sensor. This indicates that the state-related information of the vibrations at different positions has complementary features. Combining different sensor information should provide more useful information for fault diagnosis. In addition, the diagnosis accuracy on single sensor data can be negatively affected by the sensor quality, signal processing method, environmental effects, and so on, which limits the diagnostic accuracy. Third, in ESN, the diagnostic accuracy values for $1 \rightarrow 1$ and $2 \rightarrow 2$ can reach 95.3% and 95.69% from double-sensor input, respectively.

Table 4 Comparisons of the fusion results

Paradox	Methods	Proposition					
		A	B	C	D	E	Φ
Complete conflict paradox ($k = 1$)	DS	–	–	–	–	–	–
	Murphy	0.8204	0.1748	0.0048	–	–	–
	Deng	0.8166	0.1164	0.0670	–	–	–
	Li	**0.9284**	0.0716	0.0000	–	–	–
	IDS	**0.9260**	0.0290	0.0322	–	–	0.0128
0 trust paradox ($k = 0.99$)	DS	0.0000	0.7273	0.2727	–	–	–
	Murphy	0.4091	0.4091	0.1818	–	–	–
	Deng	0.4318	0.2955	0.2727	–	–	–
	Li	**0.7418**	0.2582	0.0000	–	–	–
	IDS	**0.7252**	0.0888	0.1858	–	–	0.0002
1 trust paradox ($k = 0.99$)	DS	0.0000	1.0000	0.0000	–	–	–
	Murphy	0.1676	0.0346	0.7978	–	–	–
	Deng	0.1388	0.1318	0.7294			–
	Li	0.0594	0.0000	**0.9406**	–	–	–
	IDS	0.0045	0.0138	**0.9811**	–	–	0.0006
High conflict paradox ($k = 0.9999$)	DS	0.0000	0.3571	0.4286	0.0000	0.2143	–
	Murphy	**0.7637**	0.1031	0.0716	0.0080	0.0538	–
	Deng	0.5324	0.1521	0.1462	0.0451	0.1241	–
	Li	0.6210	0.1456	0.1308	0.0000	0.1026	–
	IDS	**0.9870**	0.0030	0.0064	0.0000	0.0036	0.0000

The maximum BPA as bolded means the highest confidence

However, in the cases of $1 \rightarrow 2$ and $2 \rightarrow 1$, the accuracy values decrease to 86.6% and 79.1%, respectively.

Notably, the diagnostic accuracy of each classifier in this paper is lower than that in previous literature. This can be explained by the fact that previous works tend to investigate the diagnosis of totally different types of mechanical faults; thus, the feature is more differentiated. In this paper, the studied mechanical faults are in different positions, but all mechanical faults are essentially caused by abnormal joint clearance; these faults may have in-depth correlations, which increase the difficulty of diagnosis. Moreover, to reflect the diagnostic ability of ESN, the span of the fault joint clearance size is enlarged (from 0.25 to 0.75 mm), so the characteristics of fault data of different clearance sizes at the same position have greater dissimilarity.

In actual circumstances, the joint clearance size changes along with the service time of the HVCB. Therefore, it is more desirable that the obtained model can adapt to different fault clearance size conditions. From Fig. 7d, we can see that in the cases of $1 \rightarrow 2$ and $2 \rightarrow 1$ in the 100 testing circles, although some accuracy values are greater than 90%, the accuracy rate is not stable, and the amplitude fluctuation is more than 20%, which deteriorates the

model performance. This phenomenon can be explained by the characteristics of the ESN algorithm itself. In the ESN, the internal sparse connection structure of the neurons in the reservoir, the connection weight matrices of the reservoir input, and the reservoir-to-reservoir input are randomly initiated without adjustment in training. Therefore, the network structure and parameters of the ESN module obtained from different training processes are different, resulting in unstable diagnosis results. To address this problem, we initialize multiple ESN modules with different spectral radii and reservoir neuron numbers. Then, their output is fused by the aforementioned IDS algorithm. The average accuracy of 100 test circles for normal and three mechanical fault types is listed in Table 5. In addition, the total average accuracies of the individual ESN modules and ensemble ESN modules by IDS with input from multiple sensors vary and are described in Fig. 8.

Some interesting observations can be made from Table 5. In the case of $1 \rightarrow 2$, the total average accuracy of the individual ESN module is 86.56%. Through fusion analysis, the DS–ESN and IDS–ESN modules improve the accuracy by 2.16% and 4.27%, respectively. However, in the case of $2 \rightarrow 1$, the accuracy values of the DS–ESN and IDS–ESN modules

Fig. 7 Fault diagnosis results of different individual classifiers on data sets $A \sim D$

(a) First sensor

(b) Second sensor

(c) Multiple sensors

(d) Accuracy fluctuation
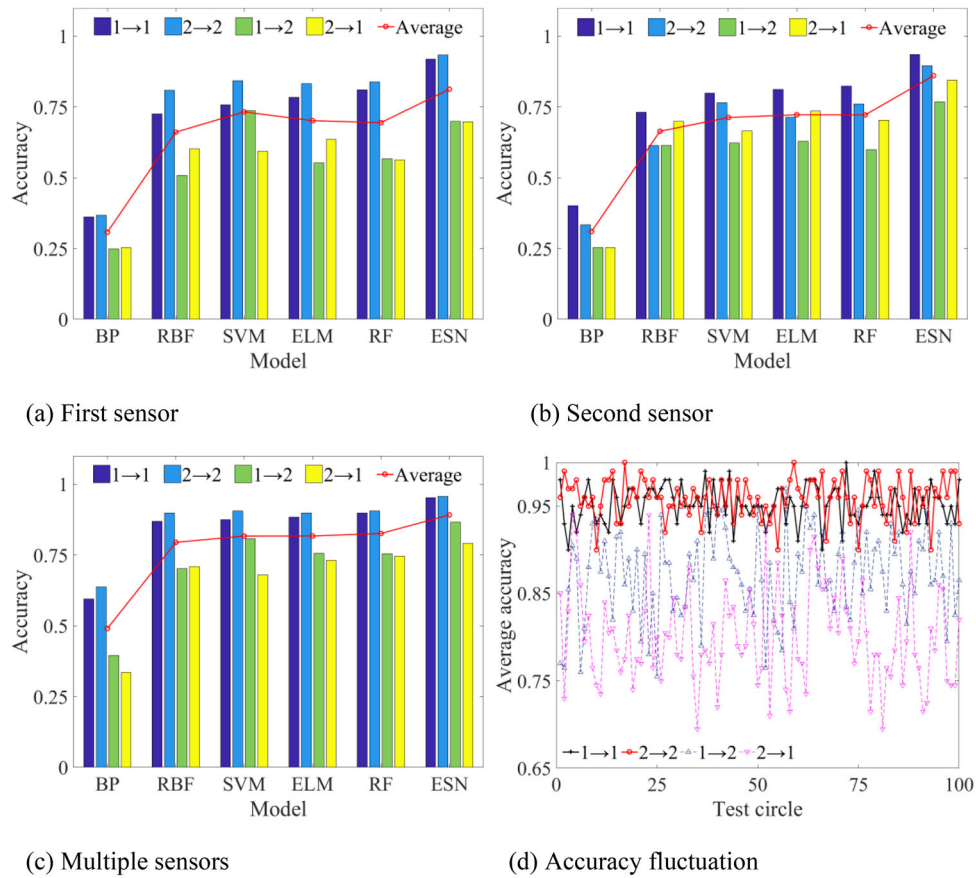
Table 5 Fault diagnosis results of individual and ensemble ESN modules

Case	Classifier	Normal (%)	Fault I (%)	Fault II (%)	Fault III (%)	Ave (%)
$1 \rightarrow 2$	ESN	99.04	88.50	85.18	72.96	86.42
	DS–ESN	98.96	95.68	87.88	72.36	88.72
	IDS–ESN	99.76	95.06	88.12	80.38	90.83
$2 \rightarrow 1$	ESN	98.14	88.10	95.70	32.54	78.62
	DS–ESN	98.24	89.56	96.16	31.54	78.87
	IDS–ESN	98.20	89.52	95.92	37.24	80.22

Fig. 8 Accuracy fluctuation of individual and ensemble ESN models

(a) from dataset C

(b) from dataset D
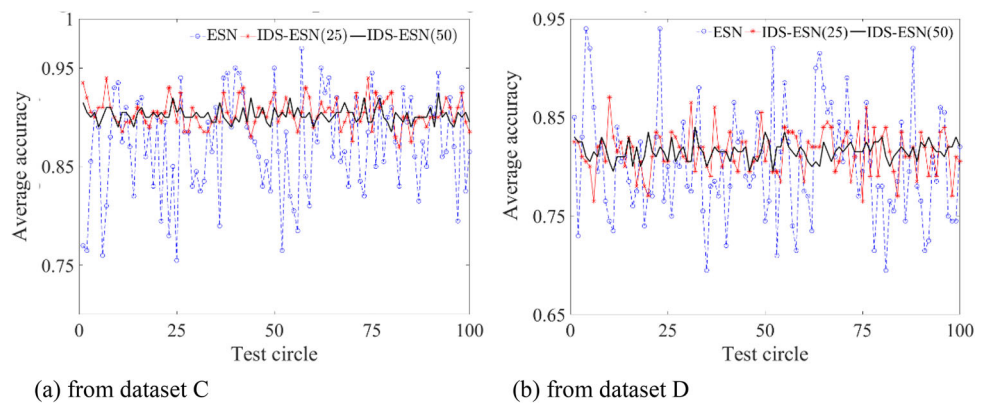
Fig. 9 Average accuracy with different percentages of mixed input

(a) from dataset C

(b) from dataset D
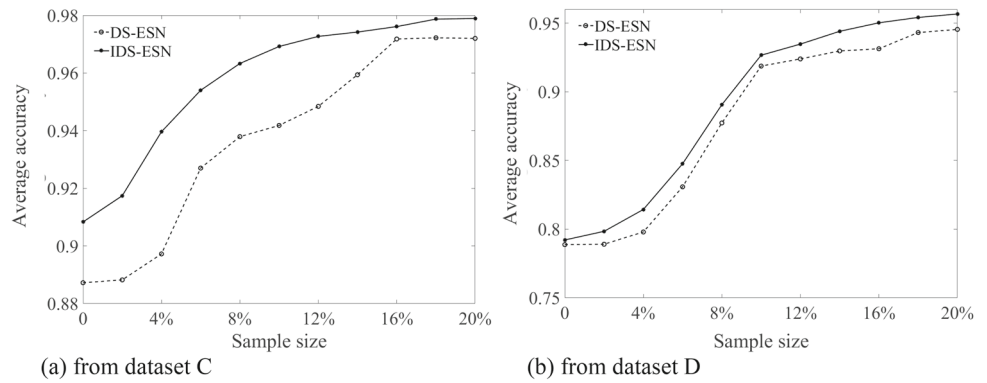
Fig. 10 Diagnosis statistics of the fusion model with 10% mixed input

(a) from dataset C

(b) from dataset D
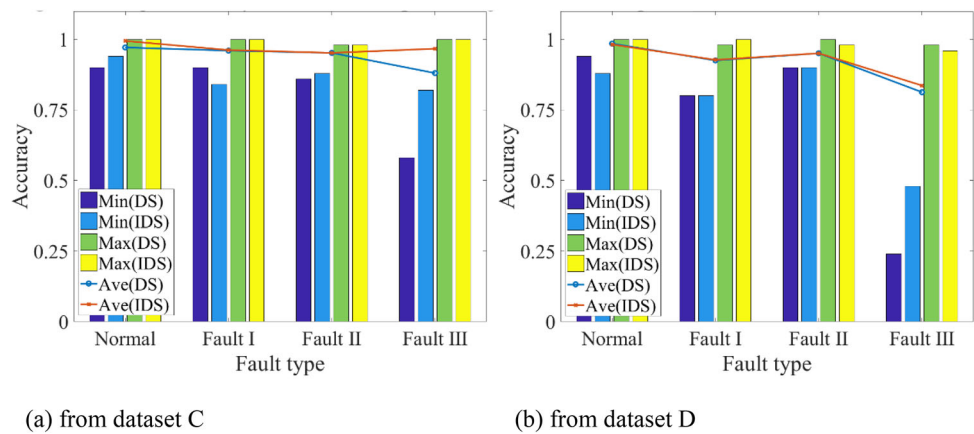
are almost constant (only a 1.6% increase in the IDS–ESN). In different circumstances, the diagnostic accuracy of Fault III is always significantly the lowest, especially in the $2 \rightarrow 1$ case, which eventually results in unsatisfactory final diagnostic results. This might be because the position of Fault III is far from the sensors in our experimental setup. Fault-related features are attenuated before the measuring points, making it more difficult to extract valid feature values from the measured signal. This phenomenon indicates that measuring points are crucial to the validity of the fault diagnosis results, and even the ensemble model may still have poor performance on a certain type of mechanical fault. In our future research, we will adopt other diagnostic methods to improve the diagnostic accuracy of this fault.

In Fig. 8, the red and black lines represent the average accuracy of the fusion model consisting of 25 and 50 ESN modules by IDS. We find that in IDS–ESN, the fluctuation amplitude of the average accuracy decreases significantly, and the degree of decline is positively correlated with the number of fused ESN modules. Therefore, main optimal configuration of our model is the number of ESN modules. More specifically, it is found when the number of the adopted echo state neural network modules increases to 50, the final output of IDS–ESN is sufficiently stable (the largest accuracy difference is within 4% of 100 statistical tests). In IDS–ESN, only the output weights of each ESN module need to be trained

by a linear regression algorithm, which avoids gradient disappearance and has high efficiency. Overall, the instability of the diagnostic accuracy is resolved. If all modules can be trained properly, the ensemble model is expected to better reflect the nonlinear mapping between the input and marked fault types. In this paper, for performance evaluation of unbalanced input, we train the IDS–ESN model with the mixing input, in which the training data contain a certain percentage of different clearance size data. The obtained results are described in Fig. 9. More specifically, the diagnostic results with a 10% mixed training sample are described in Fig. 10.

Figure 9 shows that the average accuracy of IDS–ESN is higher than that of DS–ESN. This can be attributed to the paradoxical evidence that the traditional D–S method fails at fusion, while our improved IDS fusion algorithm does not. In Fig. 10, three statistical parameters are selected (max, min, and average) according to the 100 statistical tests. After fusing multiple ESN module output results via our IDS approaches, the final average accuracy values of fault diagnosis based on multiple sensors and mixed training data can reach 96.92% and 92.66% in the $1 \rightarrow 2$ and $2 \rightarrow 1$ cases. Compared with the individual ESN results in Fig. 6, increases of 9.64% and 13.55% are observed.

In summary, the proposed IDS–ESN improves the accuracy of mechanical fault diagnosis of HVCBs and has many other advantages. On one hand, the reported IDS–ESN model

is more flexible and can deal with more complicated fault diagnosis issues, as extra ESN modules can be directly added and will not affect other ESN modules. On the other hand, the IDS–ESN model has stronger robustness. Even if several ESN modules suffer from strong interruption or disturbance, leading to the wrong output, through the fusion of IDS–ESN, the correct result can still be guaranteed with a certain probability in the end. Accurate mechanical fault diagnosis of HVCBs is crucial for the safety of electric power systems, and it requires a diagnosis model with high reliability and robustness under all environmental conditions. The proposed IDS–ESN model makes up for the weakness of single classifiers. Therefore, it can improve mechanical diagnosis accuracy, reduce accidents, and eventually bring significant economic benefits to electric power systems.

Last but not least, the main limitations of our model come from two aspects. One is state-related feature extraction. A signal processing approach named VMD with centre frequency observation is adopted to extract the energy distribution of vibration as the state-related feature. It belongs to the frequency domain feature extraction method. Hybrid features from the time domain and frequency domain combing feature selection algorithm are considered for fault diagnosis in some recent works, which improves diagnostic accuracy. Another limitation is the issue of unbalanced input, it is not analyzed currently. Therefore, in future work, hybrid state-related features and model performance in case of unbalanced input would be evaluated by the proposed methodology.

Conclusions

IDS–ESN is newly proposed in this study, and it consists of an ensemble classifier with multiple ESN modules and an improved DS evidence theory approach. Compared with other existing machine learning methods, it is first indicated that individual ESNs can achieve promising mechanical fault diagnosis of HVCBs. Then, an effective information fusion approach is explored. Multiple ESN modules are employed as sub-classifiers, whose fault classification output is fused to stabilize the classifier's diagnostic accuracy. The raw evidence in IDS–ESN is rectified by evidence credibility, which helps to address the paradoxical evidence fusion issue of the conventional DS technique. Comparative results show that our IDS–ESN model can fuse complementary evidence from different ESN modules and sensors for fault diagnosis. In addition, when the number of adopted ESN modules increases to 50, the final output of IDS–ESN is stable, and the largest accuracy difference of 100 statistical tests is within 4%. Therefore, the model output of IDS–ESN is not sensitive to the network parameters of a single ESN module. In other words, the proposed fault diagnosis approach has good robustness and does not rely heavily on expert prior knowledge for parameter setting. In future work, reservoir optimization of ESN modules in our IDS–ESN model will be explored to further enhance the feature mapping ability for more complex fault diagnosis tasks.

Declarations

References

1. Liu Y, Zhang G, Zhao C, Qin H, Yang J (2021) Influence of mechanical faults on electrical resistance in high voltage circuit breaker. Int J Electr Power 129:106827. https://doi.org/10.1016/j.ijepes.2021.106827

2. Asghar R, Niayesh K (2021) Condition monitoring of high voltage circuit breakers: past to future. IEEE Trans Power Deliv 36:740–750. https://doi.org/10.1109/tpwrd.2020.2991234

3. Sun Q, Yu X, Li H, Fan J (2021) Adaptive feature extraction and fault diagnosis for three-phase inverter based on hybrid-CNN models under variable operating conditions. Complex Intell Syst 8:29–42. https://doi.org/10.1007/s40747-021-00337-6

4. Xu X, Cao D, Zhou Y, Gao J (2020) Application of neural network algorithm in fault diagnosis of mechanical intelligence. Mech Syst Signal Process 141:106625. https://doi.org/10.1016/j.ymssp.2020.106625

5. Wozniak M, Sikora A, Zielonka A, Kaur K, Hossain M, Shorfuzzaman M (2021) Heuristic optimization of multipulse rectifier for reduced energy consumption. IEEE Trans Ind Inform 18(8):5515–5526. https://doi.org/10.1109/TII.2021.3117976

6. Rudsari F, Razi-Kazemi A, Shoorehdeli M (2019) Fault analysis of high-voltage circuit breakers based on coil current and contact travel waveforms through modified svm classifier. IEEE Trans Power Deliv 34:1608–1618. https://doi.org/10.1109/TPWRD.2019.2915110

7. Wang L, Liu Z, Miao Q, Zhang X (2018) Complete ensemble local mean decomposition with adaptive noise and its application to fault diagnosis for rolling bearings. Mech Syst Signal Process 106:24–39. https://doi.org/10.1016/j.ymssp.2017.12.031

8. Forootani A, Afzalian A, Nadian G (2017) Model-based fault analysis of a high-voltage circuit breaker operating mechanism. Turk J Electr Eng 25:2349–2362. https://doi.org/10.3906/elk-1508-73

9. Wan S, Chen L (2019) Fault diagnosis of high-voltage circuit breakers using mechanism action time and hybrid classifier. IEEE Access 7:85146–85157. https://doi.org/10.1109/access.2019.2926100

10. Mei F, Pan Y, Zhu K, Zheng J (2017) On-line hybrid fault diagnosis method for high voltage circuit breaker. J Intell Fuzzy Syst 33:2763–2774. https://doi.org/10.3233/jifs-169325

11. Mei F, Mei J, Zheng J, Wang Y (2013) Development and application of distributed multilayer on-line monitoring system for high voltage vacuum circuit breaker. J Electr Eng Technol 8:813–823. https://doi.org/10.5370/jeet.2013.8.4.813

12. Chen L, Wan S (2020) Mechanical fault diagnosis of high-voltage circuit breakers using multi-segment permutation entropy and a density-weighted one-class extreme learning machine. Meas Sci Technol 31:5107. https://doi.org/10.1088/1361-6501/ab7deb

13. Gao W, Wai R, Qiao S, Guo M (2019) Mechanical faults diagnosis of high-voltage circuit breaker via hybrid features and integrated extreme learning machine. IEEE Access 7:60091–60103. https://doi.org/10.1109/access.2019.2915252

14. Yao Y, Wang N (2020) Fault diagnosis model of adaptive miniature circuit breaker based on fractal theory and probabilistic neural network. Mech Syst Signal Process 142:106772. https://doi.org/10.1016/j.ymssp.2020.106772

15. Gao W, Qiao S, Wai R, Guo M (2021) A newly designed diagnostic method for mechanical faults of high-voltage circuit breakers via SSAE and IELM. IEEE Trans Instrum Meas 70:1–13. https://doi.org/10.1109/tim.2020.3011734

16. Krishnan P, Raj A, Rajangam V (2021) Emotion classification from speech signal based on empirical mode decomposition and nonlinear features. Complex Intell Syst 7:1919–1934. https://doi.org/10.1007/s40747-021-00295-z

17. Huang N, Fang L, Cai G, Xu D, Chen H, Nie Y (2016) Mechanical fault diagnosis of high voltage circuit breakers with unknown fault type using hybrid classifier based on lmd and time segmentation energy entropy. Entropy 18:322. https://doi.org/10.3390/e18090322

18. Krishna R, Mishra S, Naik J, Dash P (2022) Adaptive VMD based optimized deep learning mixed kernel ELM autoencoder for single and multistep wind power forecasting. Energy 244:122585. https://doi.org/10.1016/j.energy.2021.122585

19. Liu Z, Yang S, Liu Y, Lin J, Gu X (2021) Adaptive correlated Kurtogram and its applications in wheelset-bearing system fault diagnosis. Mech Syst Signal Process 154:107511. https://doi.org/10.1016/j.ymssp.2020.107511

20. He D, Liu C, Jin Z, Ma R, Chen Y, Shan S (2022) Fault diagnosis of flywheel bearing based on parameter optimization variational mode decomposition energy entropy and deep learning. Energy 239:122108. https://doi.org/10.1016/j.energy.2021.122108

21. Shen C, Zhang K (2022) Two-stage improved Grey Wolf optimization algorithm for feature selection on high-dimensional classification. Complex Intell Syst 8:2769–2789. https://doi.org/10.1007/s40747-021-00452-4

22. Cao Y, Sun Y, Xie G, Li P (2021) A Sound-based fault diagnosis method for railway point machines based on two-stage feature selection strategy and ensemble classifier. IEEE trans Intell Transp Syst 23(8):2074–2083. https://doi.org/10.1109/TITS.2021.3109632

23. Tao H, Cheng L, Qiu J, Stojanovic V (2022) Few shot cross equipment fault diagnosis method based on parameter optimization and feature metric. Meas Sci Technol 33:115005. https://doi.org/10.1088/1361-6501/ac8368

24. Ma S, Chen M, Wu J, Wang Y, Jia B, Jiang Y (2019) High-voltage circuit breaker fault diagnosis using a hybrid feature transformation approach based on random forest and stacked autoencoder. IEEE Trans Ind Electron 66:9777–9788. https://doi.org/10.1109/tie.2018.2879308

25. Zhang X, Wang H, Stojanovic V, Cheng P, He S, Luan X, Liu F (2021) Asynchronous fault detection for interval type-2 fuzzy nonhomogeneous higher-level Markov jump systems with uncertain transition probabilities. IEEE Trans Fuzzy Syst 30:2487–2499. https://doi.org/10.1109/TFUZZ.2021.3086224

26. Yao L, Fang Z, Xiao Y, Hou J, Fu Z (2021) An intelligent fault diagnosis method for lithium battery systems based on grid search support vector machine. Energy 214:118866. https://doi.org/10.1016/j.energy.2020.118866

27. Yin Z, Hou J (2016) Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes. Neurocomputing 174:643–650. https://doi.org/10.1016/j.neucom.2015.09.081

28. Lin L, Wang B, Qi J, Chen L, Huang N (2019) A novel mechanical fault feature selection and diagnosis approach for high-voltage circuit breakers using features extracted without signal processing. Sensors 19:288. https://doi.org/10.3390/s19020288

29. Chen Q, Jin Y, Song Y (2022) Fault-tolerant adaptive tracking control of Euler-Lagrange systems—an echo state network approach driven by reinforcement learning. Neurocomputing 484:109–116. https://doi.org/10.1016/j.neucom.2021.10.083

30. Yao X, Fan S, Zhao B, Cao S (2020) Controller design based on echo state network with delay output for nonlinear system. Complexity 2020:1–6. https://doi.org/10.1155/2020/8643029

31. Lacy S, Smith S, Lones M (2018) Using echo state networks for classification: a case study in Parkinson's disease diagnosis. Artif Intell Med 86:53–59. https://doi.org/10.1016/j.artmed.2018.02.002

32. Ullah W, Hussain T, Khan Z, Haroon U, Baik S (2022) Intelligent dual stream CNN and echo state network for anomaly detection. Knowl Based Syst 253:109456. https://doi.org/10.1016/j.knosys.2022.109456

33. Yang C, Nie K, Qiao J, Wang D (2022) Robust echo state network with sparse online learning. Inf Sci 594:95–117. https://doi.org/10.1016/j.ins.2022.02.009

34. Wu Z, Li Q, Zhang H (2022) Chain-structure echo state network with stochastic optimization: methodology and application. IEEE Trans Neural Netw Learn 33:1974–1985. https://doi.org/10.1109/TNNLS.2021.3098866

35. Liu Z, Bi Y, Wang X, Sha L, Liu P (2021) A generalized TODIM-ELECTRE II method based on linguistic Z-numbers and Dempster-Shafer evidence theory with unknown weight information. Complex Intell Syst 8:949–971. https://doi.org/10.1007/s40747-021-00523-6

36. Murphy C (2000) Combining belief functions when evidence conflicts. Decis Support Syst 29:1–9. https://doi.org/10.1016/s0167-9236(99)00084-6

37. Deng Y, Shi W, Zhu Z (2004) Efficient combination approach of conflict evidence. J Infrared Millim Waves 23:27–32. https://doi.org/10.3321/j.issn:1001-9014.2004.01.006

38. Li S, Liu G, Tang X, Lu J, Hu J (2017) An ensemble deep convolutional neural network model with improved d-s evidence fusion for bearing fault diagnosis. Sensors 17:1729. https://doi.org/10.3390/s17081729

39. Jian X, Li W, Guo X, Wang R (2019) Fault diagnosis of motor bearings based on a one-dimensional fusion neural network. Sensors 19(1):122. https://doi.org/10.3390/s19010122

40. Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. IEEE Trans Signal Process 62:531–544. https://doi.org/10.1109/tsp.2013.2288675

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.