Original Article

A multi-layer approach to DN 50 electric valve fault diagnosis using shallow-deep intelligent models

Yong-kuo Liu [a, b, *], Wen Zhou [b], Abiodun Ayodeji [b], Xin-qiu Zhou [b], Min-jun Peng [b], Nan Chao [b]

[a] State Key Laboratory of Nuclear Power Safety Monitoring Technology and Equipment, Shenzhen, 518172, China
[b] Fundamental Science on Nuclear Safety and Simulation Technology Laboratory, Harbin Engineering University, Harbin, 150001, China

A R T I C L E   I N F O

A B S T R A C T

Timely fault identification is important for safe and reliable operation of the electric valve system. Many research works have utilized different data-driven approach for fault diagnosis in complex systems. However, they do not consider specific characteristics of critical control components such as electric valves. This work presents an integrated shallow-deep fault diagnostic model, developed based on signals extracted from DN50 electric valve. First, the local optimal issue of particle swarm optimization algorithm is solved by optimizing the weight search capability, the particle speed, and position update strategy. Then, to develop a shallow diagnostic model, the modified particle swarm algorithm is combined with support vector machine to form a hybrid improved particle swarm-support vector machine (IPs-SVM). To decouple the influence of the background noise, the wavelet packet transform method is used to reconstruct the vibration signal. Thereafter, the IPs-SVM is used to classify phase imbalance and damaged valve faults, and the performance was evaluated against other models developed using the conventional SVM and particle swarm optimized SVM. Secondly, three different deep belief network (DBN) models are developed, using different acoustic signal structures: raw signal, wavelet transformed signal and time-series (sequential) signal. The models are developed to estimate internal leakage sizes in the electric valve. The predictive performance of the DBN and the evaluation results of the proposed IPs-SVM are also presented in this paper.

© 2020 Korean Nuclear Society, Published by Elsevier Korea LLC. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

One of the most important safety-critical components in a nuclear power system that has received increased safety verification is the electric valve. DN 50 electric valve is a key component that performs critical isolation and safety functions. In complex systems such as nuclear power plant, DN50 valves are extensively used for isolation, regulation, and discharge of the working fluid to achieve the balance of plant. Nuclear power plant valves operate in a high-temperature, high-pressure and radioactive environment, and the performance of the valves directly affects the reliable operation of other systems. Given the safety functions performed by the valve, and to ensure valve availability, it is important to develop a robust, high-fidelity monitoring and fault diagnosis system, dedicated to the valve.

The conventional approach to electric valve monitoring is via visual examination and a simple threshold. This approach is slow and prone to error. With the availability of a large amount of operational data, research effort is being directed to the development of data-driven fault diagnostic systems for industrial components. Such research efforts have yielded some data-driven algorithms developed into fault diagnostic systems [1,2]. Such algorithms span from shallow architectures, utilized for knowledge-based operator support systems [3] to deep architecture for heat exchanger leak rate monitoring [4]. Deep learning has many advantages that shallow architectures do not have, including the ability to process high-dimensional, nonlinear data, and approximate complex functions. In recent years, a deep feed-forward neural network such as convolution neural network, deep belief network (DBN), long short term memory (LSTM) and encoders have made rapid development in nuclear fuel assembly defect detection [5], pipe thinning model development [6], nuclear reactor pressurizer water level prediction [7] and steam generator tube flaw classification [8]. However, there is no expert consensus on the choice of one model or architecture over another.

* Corresponding author. Fundamental Science on Nuclear Safety and Simulation Technology Laboratory, Harbin Engineering University, Harbin, 150001, China.
E-mail address: lyk08@126.com (Y.-k. Liu).

In principle, support vector machines (SVM) have shallow architectures. Since Vapnik proposed the support vector machine (SVM), the method has gained a lot of traction and improvement within the machine learning research community. SVM has advantages in dealing with small data size, nonlinear and high-dimensional problems. SVM has been applied to predict nuclear plant process measurement drift [9] and for predictive maintenance of nuclear infrastructure [10]. However, SVM hyperparameter selection and optimization issues result in the below-optimal model [11]. To solve the SVM hyperparameter optimization issue, some research works integrate other soft computing algorithms such as genetic algorithm [12], sequential feature selection, and particle swarm optimization (PSO). The PSO optimization algorithm is used to search and optimize the SVM parameters to find the best combination of parameters and improve the accuracy of fault diagnosis. The PSO algorithm is a global search method. However, the algorithm suffers from the local optimal issue and may not be a suitable optimization algorithm for SVM if used directly. In the literature reviewed, the local optimal issue is not addressed before the application of PSO.

Moreover, a major weakness in the application of data-driven techniques to diagnose faults in the electric valve is that characteristic vibration and acoustic signals used for diagnosis are usually weak. This signal is corrupted by interference and background noise from other coupled components. To address this issue, Yang et al. [13] proposed a total variation denoising method to filter noise mixed in the nuclear plant sensors. Jinyang et al. [14]. also proposed a hierarchical discriminating sparse coding method to denoise acoustic signals and isolate interference. However, these methods have not been verified in real electric valve systems. Also, the complexity of the methods makes it unattractive for real-world applications.

Furthermore, in most application of the data-driven diagnostic system in complex systems such as nuclear power plants, the whole diagnosis is being performed by a single algorithm. Where a hybrid algorithm is used, the result relies on other qualitative reasoning approaches that are inadequate to handle the complexity involved in industrial systems [15]. Application of such model in complex systems such as nuclear power plants results in weak models with high false alarm [16]. Also, most of the research work on nuclear plant fault diagnosis focused on the whole plant diagnosis or fault isolations in other components, with no consideration for the specifics that are peculiar to DN50 electric valves. Further, it is also common knowledge that the credibility of the training dataset is critical to the performance of data-driven models. In most applications, the data used to train the models are synthetic and are not representative of the actual signal from real valves.

Consequently, to address the weaknesses identified above, this paper proposes an integrated shallow-deep architecture for the DN50 electric valve fault diagnosis system using vibration and acoustic signals. In this paper, based on the reported effectiveness in the real-world application of vibration and acoustic signal for fault diagnosis [17,18], we experimented with three different structural data acquisition/signal processing techniques in the development of the shallow and deep models. Towards the implementation of a robust shallow-deep intelligent fault diagnostic system, these experiments are conducted to verify which data structure best fits the model, and which data-driven architecture best solves the problem of fault detection and fault size estimation in the electric valve. The acoustic signals are acquired in raw form, sequential form (time series), and transformed (using wavelet packet technique.)

The shallow architecture is implemented with an improved particle swarm-support vector machine (IPs-SVM). The developed (IPs-SVM) model is trained with the wavelet transformed vibration signal. The deep architecture is implemented with deep belief network (DBN). The developed DBN model is evaluated on the three signal structures and the results are presented in this paper. The novelty in this paper is summarized below:

1. The local optimal issue of particle swarm optimization algorithm is solved by optimizing the weight search capability, the particle speed and position update strategy to achieve an improved particle swarm optimization algorithm.
2. The hyperparameter selection problem of SVM is solved by integrating the improved particle swarm optimization algorithm to obtain improved particle swarm-support vector machine (IPs-SVM).
3. It is observed that different data structure is used to develop data-driven models. For broader applicability, we experimented with three different kinds of data structures: raw data, time series (sequential) data and wavelet packet transformed data.
4. To solve the high false alarm problem in the implementation of data-driven models in complex systems, we split the fault diagnostic task into two: fault detection and classification is done with the IPs-SVM model, while fault size estimation is done with deep belief network (DBN).
5. The resulting IPs-SVM model's enhanced classification capability is used to classify phase imbalance and damaged valve fault. Phase imbalance anomaly and damaged valve packing experiments are conducted in a real DN50 electric valve, and the vibration signal from the valve is obtained and transformed using a wavelet packet reconstruction technique to reduce the effect of background noise and interference.
6. The DBN is utilized to estimate internal leakage size in the DN50 electric valve using the acoustic signal. The acoustic signal is obtained from six (6) different internal leakage size experiments conducted on the DN50 electric valve.

Considering the importance of the training data to any data-driven algorithm, we obtained real DN50 electric valve vibration and acoustic signals in an environment that closely represents the valve operation in the industry.

2. Background

2.1. Wavelet packet decomposition

Orthogonal wavelet packets are a set of functions by which standard orthogonal base in space can be constructed. A given signal can be divided into many frequency bands by a set of orthogonal filters combined with a low-pass and a high-pass filter. The energy of the signal $f(t)$ can be expressed as:

$$\|f(t)\|^2 = \int_{-\infty}^{\infty} |f(t)|^2 dt \tag{1}$$

The energy is orthogonally transformed onto the adjacent frequency band without overlapping by the wavelet packet transform. The sum of the squares of the coefficients obtained by the wavelet packet decomposition is equal to the energy of the original signal in the time domain, expressed as:

$$E_j = \sum_{i=1}^{N} \|n_i\|^2 \tag{2}$$

Where j denotes the corresponding frequency band, N is the number of data points, and n_i is the wavelet packet decomposition reconstruction coefficient. When there is a fault in the component,

the energy of each frequency band in the signal changes. Therefore, the ratio of the energy of each frequency band after decomposition to the total energy can be selected as the characteristic parameter for fault diagnosis. Then the wavelet packet decomposition energy ratio is defined as:

$$p_j = E_j \bigg/ \sum_{j=1}^{k} E_j \qquad (3)$$

Where k is the number of the frequency band.

2.2. Particle swarm optimization (PSO) and improved particle swarm optimization (IPSO)

(a) Particle swarm optimization (PSO)

The PSO algorithm treats each particle in the swarm population as a particle with no size and mass. They search in space at a certain speed and dynamically update the flight speed based on the current flight experience of the population and the individual.

In a j-dimensional search space, if there are m particles in the population, then x_{ij} is the position vector of the particle i, and v_{ij} is the velocity vector of the particle i. In the standard PSO algorithm, the speed and location of the particles in search are updated as follows:

$$v_{ij}(t+1) = \omega(t) \times v_{ij}(t) + c_1 \times rand() \times \left(p_{bj}(t) - x_{ij}(t) \right) + c_2 \times rand()$$
$$\times \left(g_{bj}(t) - x_{ij}(t) \right)$$
$$(4)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \qquad (5)$$

Where ω is called inertia weight; c_1 and c_2 are called acceleration factors; $rand()$ is a random number uniformly distributed in the interval [0,1]; p_{bj} is the best position which the current particle ever fly through; g_{bj} is the best position of the entire population in which the particle is located, which is the best position experienced by the current iteration of the entire population.

Equation (4) consists of three parts [19]: the first part is the product of the current flight speed and inertia weight in the current iterative process of the particle, and this product reflects the particle's inheritance from its current motion state. The particles perform inertial motion according to their flight speed. The second part represents the learning cognitive mode of the particle itself. The last part is the influence of the social model of the population, indicating the exchange and cooperation of the particles in the population.

In Equation (5), the particles are adjusted in the searching process by mutual influence and mutual learning. The particle swarm algorithm flow chart is shown in Fig. 1.

(b) IPSO

While PSO is running, each particle is regarded as a feasible solution to the optimization problem in the search space and the flight behavior of the particles can be treated as the search process of all individual particles. However, since the convergence depends on a random search, the algorithm cannot be guaranteed to converge to a local optimum [20]. The standard PSO algorithm introduced above belongs to the global PSO algorithm because its speed updating is based on its historical optimal value and the global optimal value of the particle population. The updating speed

of the local particle is based on its historical optimal value and the optimal value of the particles in its neighborhood. The global PSO algorithm converges faster, but it is easier to get the local best, and the local PSO algorithm converges slower than the global PSO algorithm, but it is not easy to obtain the local best. Hence, for the standard PSO algorithm, it is easy to fall into the local optima [19]. This constraint is addressed by developing an improved PSO algorithm, which combines the global PSO search method with the local PSO algorithm. The PSO algorithm is improved in the following aspects:

(c) Improvement of particle's update strategy for velocity and position based on the fusion of the global PSO algorithm and the local PSO algorithm. The distance between the particle and other particles in the iteration is calculated. If the ratio of the distance between the particle and the maximum distance is less than a threshold, indicating that the particle belongs to the neighborhood, the improved algorithm is utilized to update the velocity and position of the particle; If the ratio of the distance between the particle and other particles equals, the standard PSO algorithm is used. Also, if the ratio of the distance between the current particle and other particles is less than the threshold, then the particle updates its speed according to the particle's historical optimal value p_{bj}, the optimal value l_b of particles in the adjacent area, and the global optimal value of the population g_{bj}.

For each iteration, the distance between the particles and the remaining particles are computed, and the distance between any two particles a and b are set to d_{ab}, the maximum distance is d_{max}, and the ratio d_{ab}/d_{max} is obtained. The threshold ξ changes according to the number of iterations, and it is expressed as :
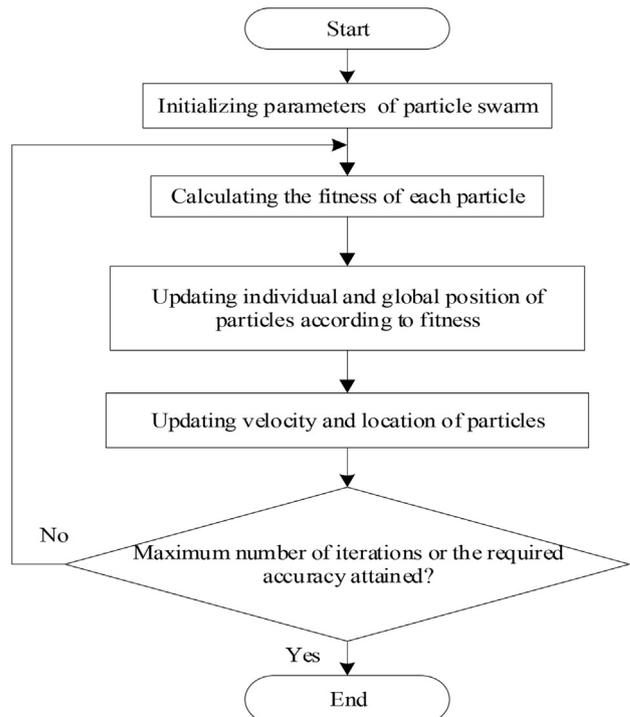

Fig. 1. The flow chart of the PSO algorithm.

$$\xi = \frac{0.3t + 0.6t_{max}}{t_{max}} \tag{6}$$

Where t is the number of iterations, and t_{max} is the maximum number of iterations. When $\xi < 0.9$, if the ratio $d_{ab}/d_{max} < \xi$, the particle b is considered to belong to the neighborhood of the particle a. At this point, the speed and position of the particles are updated as follow:

$$v_{ij}(t+1) = \omega(t) \times v_{ij}(t) + c_1 \times rand() \times$$
$$\left(p_{bj}(t) - x_{ij}(t) \right) + c_2 \times rand() \times \left(g_{bj}(t) - x_{ij}(t) \right) + c_3 \times rand()$$
$$\times \left(l_{bj}(t) - x_{ij}(t) \right) \tag{7}$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \tag{8}$$

Where t is the number of iterations; ω is the inertia weight; c_1, c_2, and c_3 are learning factors.

If $\xi > 0.9$ or the ratio $d_{ab}/d_{max} > \xi$, the particle velocities are updated as shown in equation (4).

(d) Improvement on the inertia weight search method

In the standard PSO algorithm, the weight ω is linearly reduced, which also gradually reduces the search step size, and converges the iteration to the extreme point. The disadvantage of this method is that the algorithm easily falls into local minima. To overcome this constraint, the weight ω is reduced by the sigmoid function to ensure that the population searches at a faster speed in the early stage of the search, and the search speed decreases rapidly in the middle, making the particles to easily converge to the global optimal value. In the last stage, the particles converge at a certain speed until the final convergence. The expression of the weight in the IPSO algorithm is:

$$\omega = \frac{\omega_{max} - \omega_{min}}{1 + \exp(2e \bullet t/t_{max} - e)} + \omega_{min} \tag{9}$$

Here, ω_{max} and ω_{min} are the maximum and minimum inertia weights respectively; e is the control coefficient used to adjust the speed.

2.3. IPs-SVM

The SVM algorithm diagnose faults by classifying sample data. For multi-class implementation, different kinds of samples are divided into different regions by the classification model, and the corresponding state of the data is determined by the region that each sample belongs [21].

For a classification problem, the SVM algorithm determines the maximum distance between each class and a parametric hyperplane so that the distance between the data point and the hyperplane is as large as possible. Consider a linearly separable problem having a sample set $S = \{(x_i, y_i)\}, i = 1, 2, ..., N, x \in R, y_i \in \{-1, 1\}$, the classification hyperplane is defined as:

$$\omega \bullet x + b = 0 \tag{10}$$

Normalizing equation (8) to satisfy the sample set, we obtain:

$$y_i[(\omega_i \bullet x) + b] - 1 \geq 0 \quad i = 1, 2, ..., N \tag{11}$$

In Equation (11), ω is the weight; b is the threshold; and N is the total number of samples.

Here, the classification interval is $2/\|\omega\|$. To maximize the interval, $\|\omega\|$ is minimized. Therefore, the classification surface that satisfies equation (11) and minimizes ω is the optimal classification surface. The above process can be transformed into a constrained minimum problem, including the objective function and the restriction.

$$\begin{cases} \min \dfrac{1}{2}\|\omega\|^2 \\ s.t. y_i[(\omega \bullet x_i) + b] - 1 \geq 0 \end{cases} \tag{12}$$

Where ω is an independent variable, the objective function is a quadratic function of ω, and the restriction is a linear function of ω. This is a convex quadratic programming problem, which can be solved by introducing a Lagrange multiplier. The new objective function is:

$$L(\omega, b, \alpha) = \frac{1}{2}(\omega, \omega) - \sum_{i=1}^{n} \alpha_i \{y_i[(\omega \bullet x_i) + b] - 1\} \tag{13}$$

Equating the derivatives of ω and b respectively to zero (0), we have:

$$\begin{cases} \dfrac{\partial}{\partial \omega} L(\omega, b, \alpha) = 0 \\ \dfrac{\partial}{\partial b} L(\omega, b, \alpha) = 0 \end{cases} \tag{14}$$

From (14):

$$\begin{cases} \omega = \displaystyle\sum_{i=1}^{N} \alpha_i y_i x_i \\ \displaystyle\sum_{i=1}^{N} y_i \alpha_i = 0 \end{cases} \tag{15}$$

Substituting (13) into (11) gives:

$$\begin{cases} \min \dfrac{1}{2} \displaystyle\sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j (x_i \bullet x_j) - \sum_{i=1}^{N} \alpha_i \\ s.t. \displaystyle\sum_{i=1}^{N} y_i \alpha_i = 0, \alpha_i \geq 0, i = 1, 2, ..., N \end{cases} \tag{16}$$

Where α_i is the Lagrangian multiplier corresponding to sample i. Then the optimal classification function expression is:

$$f(x) = \text{sgn}\{(\omega \bullet x) + b\} = \text{sgn}\left(\sum_{i=1}^{N} y_i \alpha_i (x_i, x) + b \right) \tag{17}$$

Where sgn($\bullet$) is the function sign, and the classification of sample x can be determined by the positive or negative value of the function.

For samples that are not linearly separable, a slack variable $\xi_i \geq 0 (i = 1, 2, ..., N)$ is introduced. That is, a small amount of error is allowed to exist, such that the restriction in equation (12) becomes:

$$y_i[(\omega \bullet x_i) + b] \geq 1 - \xi_i \tag{18}$$

Then the optimization problem becomes:

$$\begin{cases} \min \dfrac{1}{2}\|\omega\|^2 + C \displaystyle\sum_{i=1}^{N} \xi_i \\ s.t. y_i[(\omega \bullet x_i) + b] \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, ..., N \end{cases} \tag{19}$$

Where C is the penalty factor. Constructing the Lagrangian equation

results in the following dual Lagrangian operator:

$$
\begin{cases}
\min \dfrac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j (x_i \bullet x_j) - \sum_{i=1}^{N}\alpha_i \\
s.t. \sum_{i=1}^{n} y_i \alpha_i = 0_i, 0 \le \alpha_i \le C, i = 1,2,...,N
\end{cases}
\tag{20}
$$

Here, the solution process of the optimal classification hyperplane is the same as the linearly separable case.

The nonlinear case can be solved by utilizing non-linear mapping. The original samples are mapped to a linearly separable higher-dimensional feature space, and the best classification hyperplane is constructed in the high-dimensional feature space. In this case, the inner product operation (x_i, x_j) in the linear case needs to be changed to $(\varphi(x_i), \varphi(x_j))$.

According to the Hilbert-Schmidt principle, the kernel function satisfies the Mercer condition: for any function $g(x)$, when $\int_a^b g^2(x)dx$ is limited, $K(x_i, x_j)$ is the inner product of space $(\varphi(x_i), \varphi(x_j))$, and the optimization problem becomes:

$$
\min \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^{N}\alpha_i
\tag{21}
$$

The classification decision function is:

$$
\mathrm{sgn}\left(\sum_{i=1}^{N} y_i \alpha_i K(x_i, x) + b\right)
\tag{22}
$$

In the utilization of IPSO and development of IPs-SVM, the first step is that the parameters that need to be optimized, such as the SVM penalty factor C and parameters γ of the RBF kernel function, are taken as the values of the elements of the particle position vector. Secondly, the fitness of the particles and the accuracy of the cross-validation results are computed. The IPs-SVM algorithm flowchart is shown in Fig. 2.

2.4. Deep belief network (DBN)

The DBN is a generative neural network consisting of several Restricted Boltzmann Machine (RBM). An RBM consists of a hidden layer and a visible layer. The hidden layer consists of hidden cells, and the visible layer consists of visible cells. Both the visible unit and the hidden unit are binary variables whose state is 0 or 1. The units between adjacent layers have connections, and the units in the same layer are not connected [22]. The RBM unit stack and structure is shown in Fig. 3. The state of the bottom unit is the visible input data vector, and the output of the hidden unit is the input of the visible layer unit of the next RBM unit.

The training of DBN is generally divided into two processes: unsupervised pre-training and supervised fine-tuning. If the DBN is used for classification, the training of the DBN is first performed in an unsupervised manner to help the DBN effectively mine the fault features in the input data. Then the network can be supervised using back-propagation approach. In the unsupervised pre-training phase, also known as the feature learning phase, the basic purpose is to calculate the hidden layer output corresponding to the input signal v, that is, the hidden layer feature signal h, so as to maximize the joint probability distribution $P(v, h)$. In a more intuitive expression, it is a process of reconstructing the visible layer signal through the hidden layer feature signal h, making the error between v and $\widehat{v}$ minimal [23]. When the input signal and parameters of the network are given, the parameter initialization network model between layers is [24]:

$$
\begin{cases}
\mathbf{h} \sim P(\mathbf{h}|\mathbf{v},\mathbf{W},\mathbf{a},\mathbf{b}) = \dfrac{1}{P(\mathbf{v})\bullet Z}\bullet e^{(\mathbf{b}^T\bullet\mathbf{h}+v^T\bullet\mathbf{W}\bullet\mathbf{h})} \\
\widehat{\mathbf{v}} \sim P(\mathbf{v}|\mathbf{h},\mathbf{W},\mathbf{a},\mathbf{b}) = \dfrac{1}{P(\mathbf{h})\bullet Z}\bullet e^{(\mathbf{a}^T\bullet\mathbf{v}+v^T\bullet\mathbf{W}\bullet\mathbf{h})}
\end{cases}
\tag{23}
$$

Where v is the input of the visible layer, h is the output of the hidden layer, W is the connection weight matrix of the visible layer to the hidden layer, a is the offset of the visible layer, and b is the offset of the hidden layer. In practical application, parameters W, a, b are initialized, and the equation relating the hidden layer h to the visible layer v is:

$$
\begin{cases}
P(h(i) = 1|v) = \sigma\left(v^T \bullet \mathbf{W}_i + b_i\right) \\
P(h(i) = 0|v) = 1 - \sigma\left(v^T \bullet \mathbf{W}_i + b_i\right)
\end{cases}
\tag{24}
$$

Here, $\sigma(\bullet)$ is the activation function, which is usually a sigmoid function, $h(i)$ is the output value of the node i of the hidden layer, W_i is the connection weight, and b_i is the offset of the hidden layer node. The equation for calculating the visible layer v according to the hidden layer h is:

$$
\begin{cases}
P(\widehat{v}(j) = 1|\mathbf{h}) = \sigma(\mathbf{W}_j \bullet h + a_j) \\
P(\widehat{v}(j) = 0|\mathbf{h}) = 1 - \sigma(\mathbf{W}_j \bullet h + a_j)
\end{cases}
\tag{25}
$$

The parameters are updated according to the error between v and $\widehat{v}$, and the updated relation for the parameter θ (W, a, b) is:

$$
\begin{cases}
\theta^{(t+1)} = \theta^{(t)} + \eta\dfrac{\partial \log P(v)}{\partial \theta}\big|_{\theta^{(t)}} = \theta^{(t)} + \eta \times \varDelta\theta^{(t)} \\
\dfrac{\partial \log P(v)}{\partial \theta} = \langle h(v - \widehat{v})\rangle
\end{cases}
\tag{26}
$$

Where η is the learning rate or step size; t is the number of iterations; $\langle h(v - \widehat{v})\rangle$ is the mean value of the product between the maximized probability of visible layer v' and its corresponding hidden feature vector.

3. The proposed IPs-SVM/DBN fault diagnosis system

To solve the problems identified in section 1, the following steps are implemented to develop a robust IPs-SVM model for fault detection and DBN model for fault size estimation.

Step1. (a): In the signal acquisition and feature extraction phase, the vibration signal and the acoustic emission signal from the electric valve is collected. The vibration signal obtained is based on the phase imbalance in the valve motor drive and damage valve packing faults simulated using the experimental setup described in Section 4. To reduce the interference and background noise from other components, the signal is processed using the wavelet packet decomposition approach. The energy ratio of the transformed signal is then utilized to train the proposed IPs-SVM used for fault detection.

(b) The acoustic signal obtained is based on six different internal leakage faults simulated on the experimental setup. The acoustic emission signal is used for detecting the occurrence of the leakage fault in the valve and predicting the leakage rate. The collected signal is extracted using three different processes: raw signal extraction, time series (sequential) signal extraction and wavelet packet transformed signal decomposition. We experimented with these three signal structures by training three different DBN models using each of the signal structure.
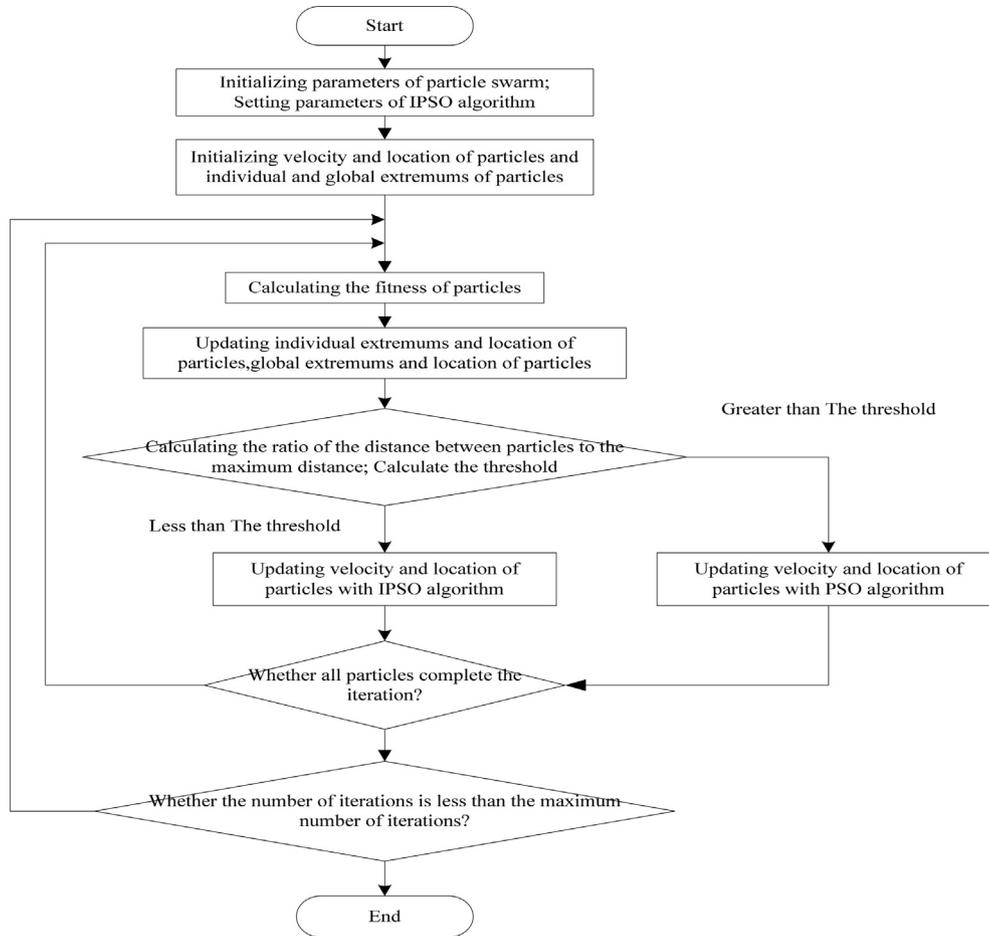
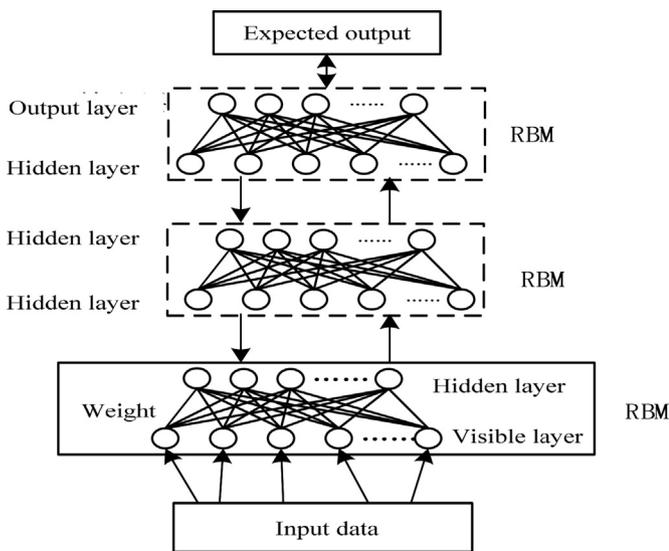**Fig. 2.** The flowchart of IPs-SVM algorithm.

Fig. 3. The structure of DBN

Step 2. In the model training phase, the data samples comprising the characteristic parameters are partitioned and used as the model training samples and test samples. The samples obtained in Step 1(a) is used to train the IPs-SVM model utilized for fault detection and classification. The samples obtained in Step 1(b) is used to train

the DBN fault size estimation models for leak degree evaluation.

Step3. The performance of the proposed IPs-SVM for fault detection is evaluated against the conventional support vector machine (SVM) and a hybrid of (un-improved) particle swarm optimization algorithm with the conventional SVM (PSO-SVM). Moreover, the result of the fault size estimation (DBN models) developed using different signal structure is analyzed. The proposed method for the electric valve fault diagnosis is shown in Fig. 4.

3.1. Experimental testbed for data acquisition

3.1.1. Experimental setup

To implement a functional representation of the DN50 electric gate valve obtainable in the real valve operation, an experimental testbed is set up. The experimental equipment is a set of simple water circulation circuit with the electric valve. The purpose of the experiment is to collect the operation data of the electric valve under normal and fault conditions, used to verify the feasibility of the intelligent method. The mechanical vibration and acoustic signals contain the information that can reflect the working state and fault condition of the equipment. Two kinds of sensors area arranged to collect the comprehensive valve operation information which lays the foundation for the subsequent analysis.

The testbed consists of a vertical circulating power centrifugal pump, pressure sensor, electromagnetic flowmeter, differential pressure transmitter, and the Z941H−25P electric gate valve. The

schematic diagram of the testbed and the actual laboratory layout are shown in Fig. 5 and Fig. 6 respectively. As shown in Fig. 5, the main circulation pipe has a diameter of 50 mm, three electric gate valves are arranged in parallel on the pipeline, and a static pressure transmitter is installed in each of the upstream and downstream circuits of the valve. On the total circuit downstream of the valve, a differential pressure transmitter and an electromagnetic flowmeter are also installed. The above sensors can measure the pressure before and after the valve, the pressure difference and the circulating flow of the circuit.

The range of the pressure sensor is 0—2.5 MPa, the range of the differential pressure sensor is 0—500 KPa, and the range of the electromagnetic flowmeter is 0.2~5 m^3/h. The centrifugal pump speed used in the experiment can be controlled by the frequency converter. The centrifugal pump pumps water from the water tank for circulation. The water tank is made up of two stainless steel tanks in series with a volume of about $3m^3$. Only one parallel circuit is enabled at a time during the experiment so that only one gate valve is operated in the main channel. In the experimental setup, the valve switch can be controlled at any time by the computer system.

Also, two acceleration sensors are arranged orthogonally on the motor drive of the valve to measure the vibration signal of the valve during operation. One of the accelerometers is mounted on the motor coil housing and the other is mounted on the gear reducer housing. The two accelerometers used in the experiment are AD50T self-amplifying accelerometers. The amplifiers have a frequency range of 0.5—15000Hz, and a working temperature of −40—120 °C, and there are special magnetic fixtures and acquisition cards of MPS-060602 with two acquisition channels. The acceleration sensor arrangement is as shown in Fig. 7.

The testbed uses the SR40 M acoustic emission sensor, with dimension $\Phi22 \times 36.8$ mm, a frequency range of 15—70 KHz, and an operating temperature of −20—120 °C. The acoustic emission pre-amplifier uses a PA I broadband preamplifier. The bandwidth is 10 KHz~2.0 MHz, the working temperature is −20—65 °C; the acoustic emission acquisition card is SAEU2S type acoustic emission acquisition card, USB interface, with two channels, the sampling rate is 769 kHz—10 MHz.There are many options for the installation position of the acoustic emission sensor. The basis for the selection is that the signal with a large amplitude and energy can be obtained at this position. After consulting the literature, the optimal measurement point for the leakage inside the valve is the downstream flange. However, due to the large weight of the electric valve selected in the experiment, the bench beam shields the flange surface during the installation process. Consequently, the acoustic emission sensor is mounted on the valve base shell, as shown in Fig. 8. The sensor mounting surface is smoothened with a sand-paper toremove impurities. It is adhered to the specified position with a special coupling agent and is reinforced with a tape.

3.1.2. Experimental procedure

The internal leakage of the valve is observed externally as follows: the valve position indicator shows that the valve is in the closed position, the flowmeter has no obvious indication of abnormality, and there is still a small amount of working fluid flowing through the closed valve, which is the internal leakage state of the valve. The internal leakage produces a jet flow downstream of the valve block. The impact of the jet flow on the valve causes the molecular stress wave, which is the acoustic emission.

When the valve position is fully open, the frequency converter is adjusted to control the rotation speed of the main pump at 15hz-870rev/min. when the instrument is normal and the water circulation has been established, the driving pressure of the working medium is 0.26Mpa, and the local resistance of the gate valve is very small. The two static pressure indications are initially
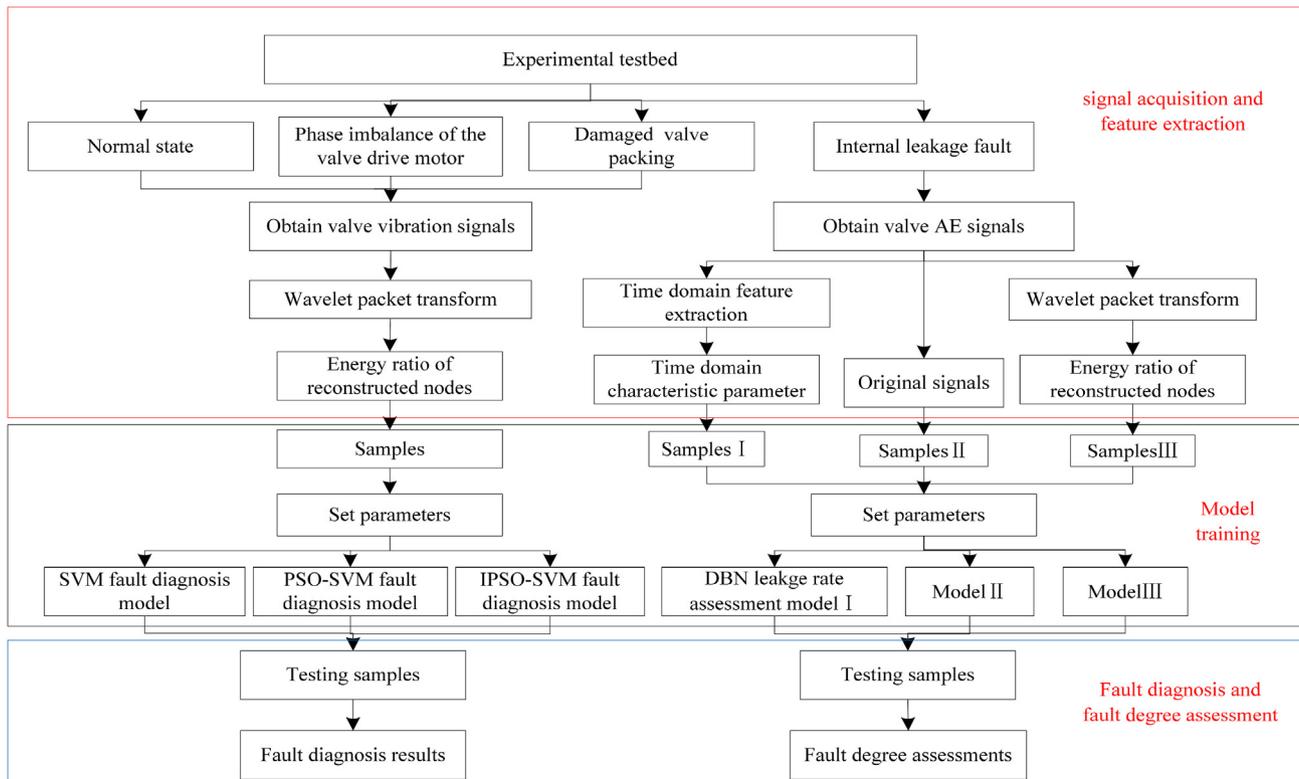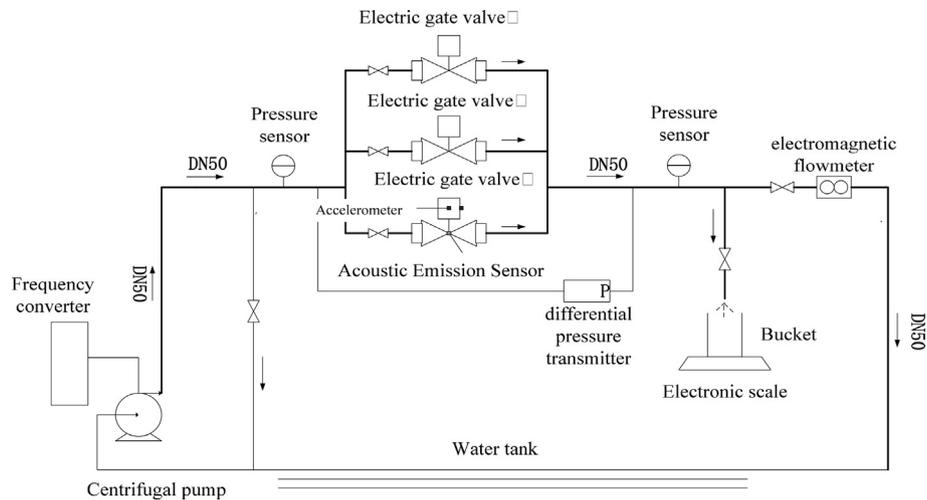

Fig. 4. Flow chart of the proposed method.

Fig. 5. Schematic diagram of the experimental setup.

(a) valve   (b) Water tank and computer

Fig. 6. A picture of the real experimental testbed.

0.00Mpa, the differential pressure manometer indicates 5.00kpa−8.00kpa, and the total flow of the pipeline is 3.03m^3/h. At this time, the acoustic emission sensor does not receive any triggering signal.

Due to the limitation of the experiment, the internal leakage mode of this experiment is simulated with the sealing surface not tightly closed. In this case, to work around the limitations of the experiment, the total flow of the pipeline needs to be reduced to less than 0.2m^3/h. Then, the flowmeter no longer reads, and flow is established by rotating the valve motor handwheel. The acoustic emission signals and the internal leakage flow rate for different internal leakage degree under stable pressure are obtained using the following procedure:

① Push on the gear lever of the motor to make the handwheel mesh with the reducer.
② Rotate the handwheel and observe the flowmeter at any time until the flowmeter is slightly higher than 0.2m^3/h.
③ Adjust the remaining valve position, for example, by turning the handwheel 30° at a time, and measure the flow and acoustic emission signal at each time.
④ For flow measurement, the flow measurement valve is opened first, and then the downstream manual ball valve is closed. The outgoing working medium can be contained in a water box. The quantity of the collected water is measured for 1min, i.e. the internal leakage flow.
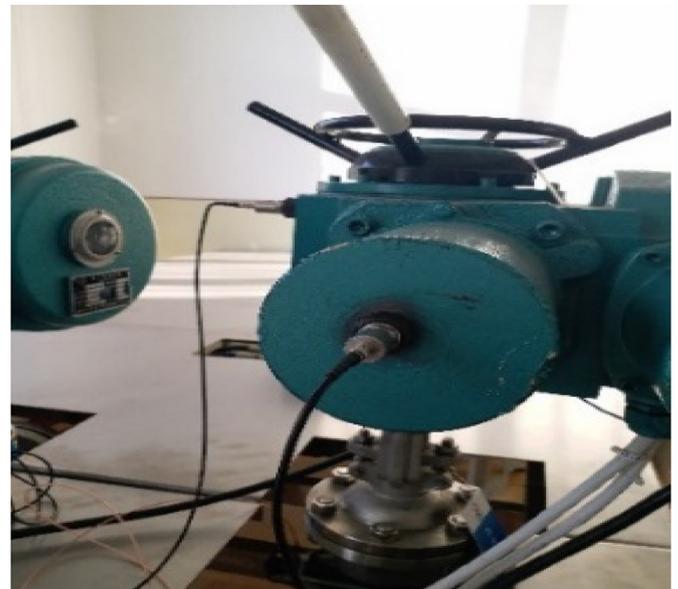

Fig. 7. Acceleration sensor installation.

⑤ After measuring the leakage flow, the downstream manual ball valve is then opened, while the flow measurement valve is closed and the acoustic emission measurement starts.

Due to the gate valve design, there is stroke error. Also, the flexible connection between stem and gate, plus the interference of viscosity and surface tension of working medium constraintthe valve handwheel from rotating in the opposite direction, so it must move from the opening position to the closing position at a point. Moreover, because the valve itself is only set with two scales, the valve position cannot be determined through the scale, so the measurementis taken to indirectly correspond to the leakage degree of the valve through the flow under the leakage condition at constant pressure.

3.1.3. Vibration signal processing

The experiment includes the vibration signal acquisition from the valve motor drive during the normal operation, phase imbalance in the valve motor, and damaged valve packaging fault. Phase imbalance fault is simulated by connecting a 50Ω and a 25Ω aluminum alloy resistor in parallel to the electric valve motor terminal. The damaged valve filler is simulated by repeatedly adjusting the hand-wheel switch valve. The experimental sampling frequency is 100 KHz, the trigger voltage is 0.1V, and the sampling time is 44s. The original vibration signal waveform of the motor drive in the three states of the valve are shown in Fig. 9. The figure shows the data points of the signal between 2s–2.1s.

After many experiments, the db12 wavelet is chosen to obtain a 6-layer transformation of the valve vibration signal, and the signal is decomposed into $2^6 = 64$ nodes. The frequency width of each frequency band is $100/(64*2) = 0.78125$ KHz,and the signal waveform of the reconstructed node in the first 8 frequency bands of the last layer of signal decomposition is drawn. The valve vibration signal wavelet packet decomposition node reconstruction signal of three valve states including the normal state, the valve motor drive phase imbalance fault and the valve packing fault is shown in Figs. 10–12.

Based on signal wavelet packet decomposition, the wavelet packet transform coefficients of the reconstructed signal are used to calculate the wavelet packet energy ratio of each band in the six layer. The energy is mainly concentrated in the first 8 segments, so
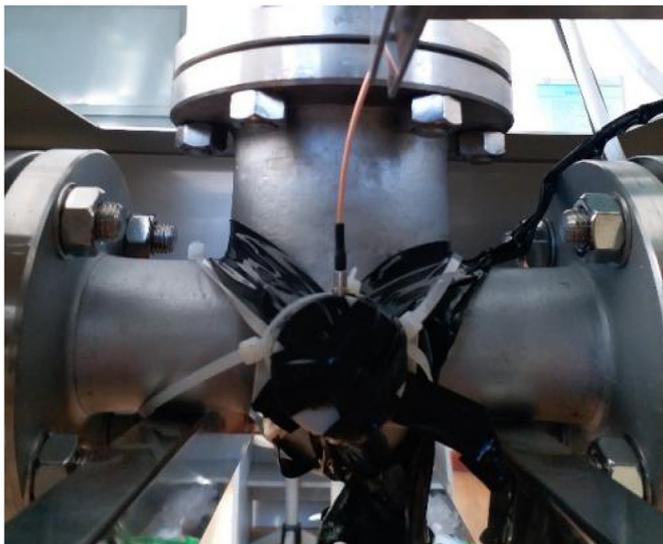

Fig. 8. Acoustic sensor installation.

only the first 8 nodes are taken. The results are shown in Fig. 13 below:

3.1.4. Acoustic emission (AE) signal processing

The fast Fourier transform is performed on the acquired AE signal, and it is found that the signal energy is mainly concentrated at 0–100 KHz, and the peak appears at 20–30 KHz. In the AE detection of the internal leakage of valves, optimum wavelet signal decomposition characteristics are obtained using the db8 wavelet [25]. The signal sampling frequency is 5 MHz.The simulated leakage rate of the valve are: 0.025 m^3/h, 0.055 m^3/h, 0.085 m^3/h, 0.115 m^3/h, 0.145 m^3/h and 0.175 m^3/h. Taking an instance of the valve internal leakage AE signal with a leakage rate of 0.025 m^3/h, we selected 5000 data points, and make a 6-layer transformation with a db8 wavelet. The original signal and wavelet packet decomposition node reconstruction signals are shown in Figs. 14 and 15.

Based on the wavelet packet decomposition in Fig. 15, the energy ratio of the wavelet packet nodes in each frequency band is calculated, and the first 10 frequency bands are taken as the signal characteristic vector. The characteristic parameters of the AE signal extraction under six kinds of leaking rate are as shown in Fig. 16.

4. Results and analysis

4.1. Fault classification

In the experiment, the signals indicating the normal state, valve motore drive phase imbalance fault, and damage of valve packing are obtained. For training samples, 30 signals are taken for each state, 15 of which are acquired when the valve is turned on, and 15 are acquired when the valve status is turned off. For the test samples, 12 samples are randomly selected for each signal. The composition of the final sample is shown in Table 1 and the IPSO algorithm parameter settings are shown in Table 2.

The range of the RBF function parameter γ is [0.01, 200], and the penalty factor C ranges from 0.1 to 500. In the process of training the model, three-fold cross-validation is performed. The training samples are randomly divided into three parts, one is randomly selected as the evaluation set, and the remaining two are used as training and testing sets. Consequently, the maximum fitness value for all particles is obtained, and the training result of the IPs-SVM model is shown in Fig. 17.

From Fig. 17, the training result is 96.25%. That is, the combination of the parameter values obtained by the optimization results in a model with 96.25% predictive accuracy. The best penalty factor C is 3.4783, and the best RBF kernel function parameter γ is 52.9270. Using the test sample on the IPs-SVM model obtained after the training, the model correctly identified 336 test samples with 93.3333% accuracy.

To quantify the effect of the modifications implemented in IPs-SVM, we conducted additional experiments to test whether there is a significant change from the result obtained from using direct PSO-SVM and just SVM algorithm without the PSO influence. For PSO-SVM, the PSO algorithm is first used to find the optimal parameter combination, and the obtained PSO-SVM model training result is shown in Fig. 18.

The fitness result of PSO-SVM model training is 95.97%, which indicates that the obtained model has 95.97% classification accuracy. The penalty factor C is 3.6788, and the RBF kernel function parameter γ is 50.0000. Using the test sample on the PSO-SVM model, the model correctly identified 335 test samples with an accuracy of 93.0556%.

In the implementation of SVM only, we applied the general grid search SVM parameter selection (enumeration) method. The main idea is to set the range of the penalty factor C and the kernel
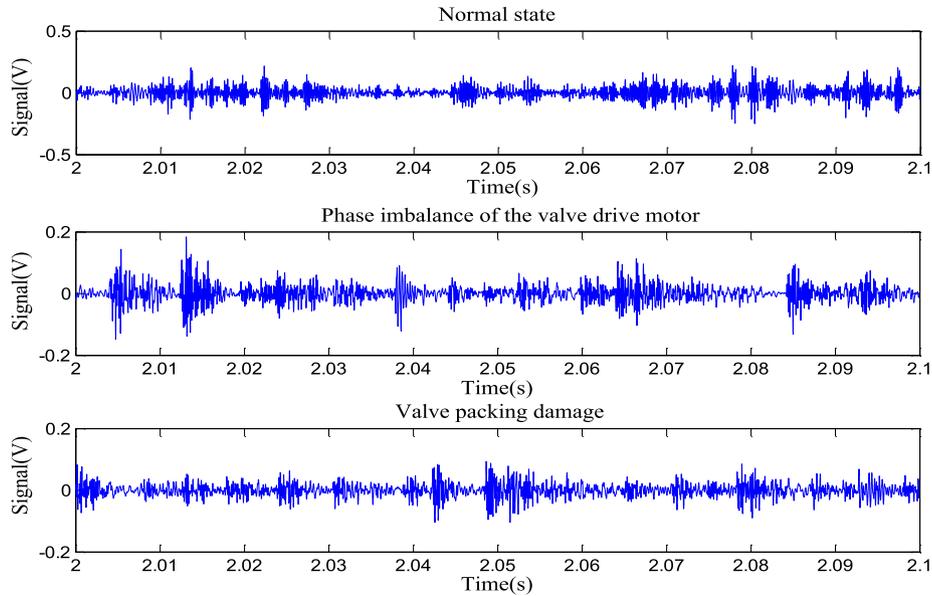
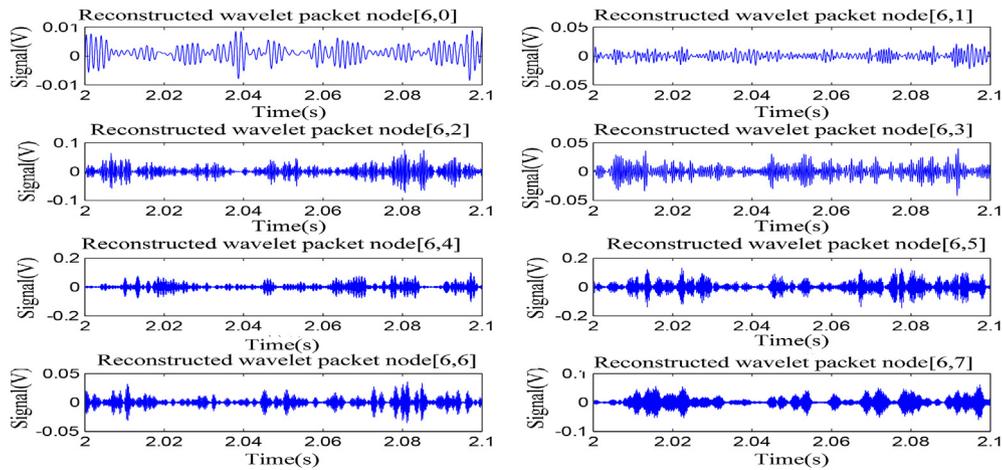**Fig. 9.** The raw waveform of the motor drive vibration signal.

Fig. 10. Wavelet packet decomposition signal for valve normal state.

function parameter γ in advance, and iteratively adjust the values to reach the optimum. Finally, the combination of parameters that give the best classification accuracy is selected as the optimum combination of parameters. Here, the grid search method is applied to optimize the parameters. The range of the penalty factor C and the kernel function parameter γ is set to $[2^{-10}, 2^{10}]$, and the index step size is 0.8. The obtained parameter selection result is shown in Fig. 19.

The optimal penalty factor C obtained by the grid search method is 1.3195, and the kernel function parameter γ is 111.4305. Using these parameters in the SVM model, the best classification accuracy rate during training is 95.6944%, and the test result is 91.9444%. The results of the three classifiers evaluated are shown in Table 3.

It can be seen from the table above that the classification accuracy for the IPs-SVM model is the best for both training and test sample. It was observed that in the grid search SVM model, the penalty factor C and the kernel function γ take a number in the reported range at every discrete distance. This may account for the failure to attain an optimal result. Conversely, the

accuracy of the training results and test results obtained by IPs-SVM and PSO models is higher than that of the grid search SVM, and the IPs-SVM algorithm has higher accuracy than the PSO-SVM model, which shows that the IPs-SVM model is a more intelligent, more efficient classifier and has a higher generalization ability. As seen in Table 3, the parameter optimization effect of the IPSO algorithm shows a modest percentage improvement over that of the PSO algorithm. However, in the actual implementation, such modest improvement has a significant effect and much obvious advantage in the rate of false alarm generated by the model.

4.2. Fault degree assessment

Because the sampling frequency of the AE signal for valve leakage is high and the signal size is huge, we train the DBN network with the samples consisting of original signals to adaptively extract the signal characteristics and evaluate its prediction accuracy. When the valve has internal leakages with different leak rates, the valve AE signal parameters will also change, so the time
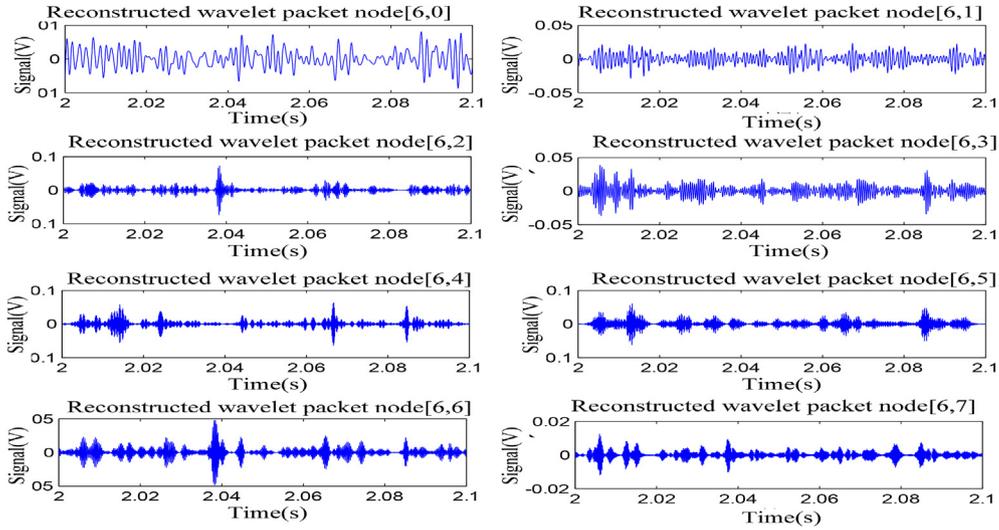
Fig. 11. Wavelet packet decomposition signal for valve motor drivephase imbalance.
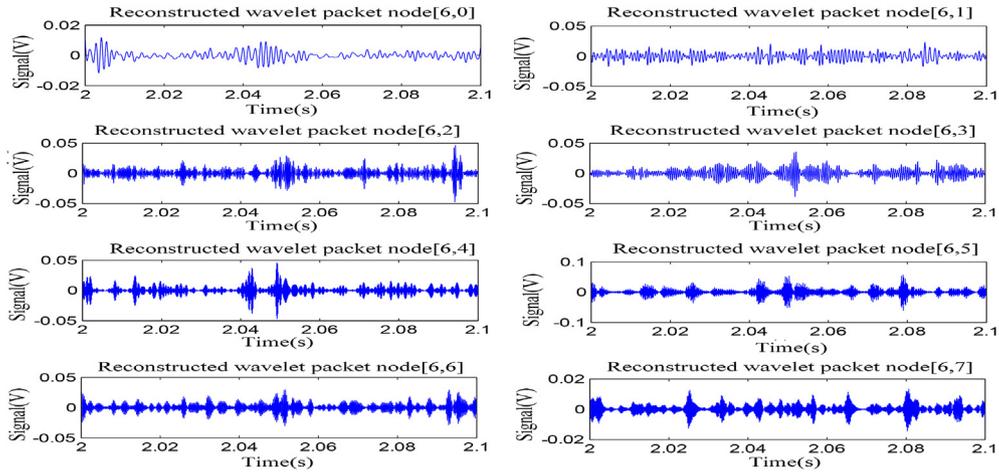

Fig. 12. Wavelet packet decomposition signal for damaged valve packing.
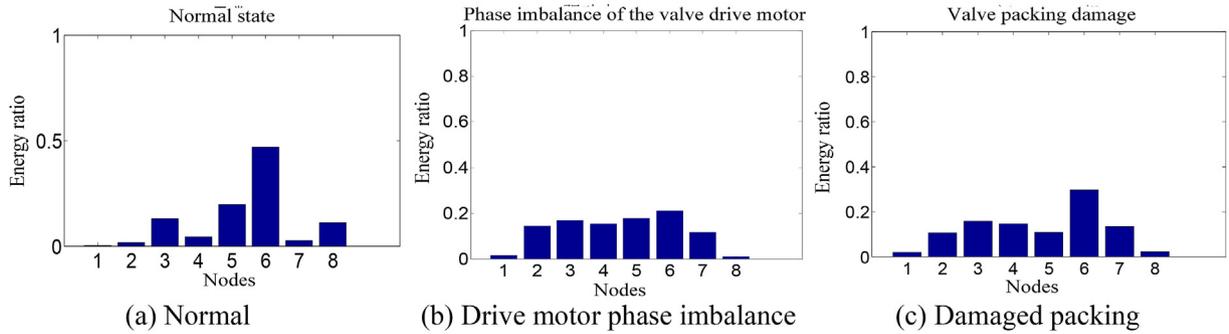

Fig. 13. Results of feature extraction (a) Normal (b) Motor drive phase imbalance (c) Damaged packing.

domain characteristic parameters of the AE signal are extracted as the characteristic parameters of the sample. The details of the signal acquisition, processing and model evaluation can be found in Refs. [26]. The evaluation process involved and the result obtained in the development of the DBN models using different data samples is summarized as follow:

a. DBN performance on the original (raw) data

We utilized 500 data points from the original signal, and select 1800 samples for each leak rate. In each selected sample (1800), 1500 samples are used as training data, while 300 are used as test data. The total number of training data from all leak samples is 9000 and the number of test data is 1800. The sample is normalized
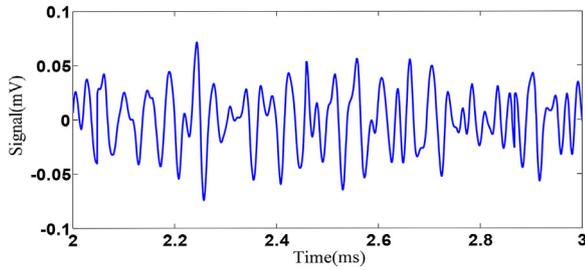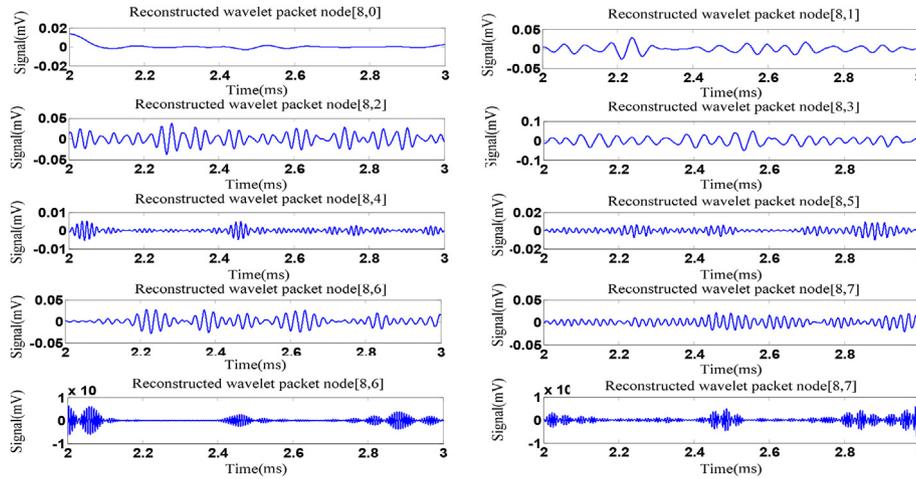
Fig. 14. AE original signal for internal leakage fault.

Fig. 15. Wavelet packet decomposition node reconstruction signal.

test samples, and the samples are normalized. The total number of the training sample is 9000, and the number of test samples is 1800. We utilized a similar pre-processing and training method used when the original sample was applied. The best performing network parameters obtained are as shown in Table 5.

After training the network, the model is tested, and the predictive output for the leak severity in the DBN model is shown in Fig. 22, and the absolute error is shown in Fig. 23.

As shown in Fig. 22, most of the data points are clustered around the actual leak rate, indicating that the evaluation result is close to the actual leak rate. For the sample with a leak rate of 0.025 m^3/h,

and used to develop the DBN network. First, we pre-train the network using the unsupervised method. Then we fully train the network with supervised learning. The network parameters utilized during the two training processes are as shown in Table 4.

After training the network, the test samples are used to evaluate the model. The leak rate in the test samples is 0.025 m^3/h, 0.055 m^3/h, 0.085 m^3/h, 0.115 m^3/h, 0.145 m^3/h, 0.175 m^3/h. The predicted leak size by the DBN model is shown in Fig. 20, and the absolute error is shown in Fig. 21.

In Fig. 20, the data points are the assessment result (leakage rate) predicted by the DBN model, and the straight line is the actual (simulated) leak rate. For best performance, the data points are expected to cluster around the actual leak rate simulated. As observed in Fig. 20, this is not the case, as the data points are scattered. Hence, the evaluation result deviates from the actual leak rate, especially when the leakage rate is between 0.055 m^3/h~0.175 m^3/h, which shows that DBN model prediction is poor when trained with the raw signal from the valve. In Fig. 21, the maximum absolute error between the evaluation result and the actual result is 0.149 m^3/h, the mean value of absolute error is 0.0351 m^3/h, and the mean squared error of the evaluation result is 0.018. To achieve better results, we obtained a fresh time-series signal to train the DBN.

b. Time-series (sequential) signal characteristic parameter samples

The six characteristic parameters of the AE signal including amplitude, ringing count, rise time, energy, RMS and ASL with different leakage levels are taken as the characteristic parameters of the sample. In the preprocessing stage, 1800 samples were selected for each leak rate, including 1500 training samples, 300

the data point and the straight-line completely overlap, which shows that the predicted leak rate is the same as the actual leak rate. In Fig. 23, the maximum value of the absolute error between the evaluation result and the actual result is 0.02891 m^3/h, the mean value of the absolute error is 0.0022 m^3/h, and the mean squared error of the model is 1.5087e-05. The model result shows a significant improvement over the result derived from the original raw data.

c. DBN developed with wavelet extracted characteristic parameter

The energy ratio of the wavelet packet reconstructed node is extracted as demonstrated in section 3 and the characteristic parameters of the samples are used in the training of the DBN. Similar preprocessing and data sample sizes are also applied. The number of training samples is 9000, and the number of test samples is 1800. The network parameters obtained are as shown in Table 6. The predicted leak rate of the DBN model is also shown in Fig. 24, and the absolute error is shown in Fig. 25.

It is observed that the data points are scattered around the actual leak rate line, and there is a significant deviation between the predicted results and the actual leak rate. In Fig. 25, the maximum absolute error between the evaluation result and the actual result is 0.1523 m^3/h, the mean value of the absolute error is 0.0290 m^3/h, and the mean squared error of the evaluation result is 0.015. The evaluation result shows that the model is weak.

In summary, for fault detection tasks on the DN50 electric valve, the best performing model is found to be the IPs-SVM model trained with wavelet packet transformed vibration signal, using the parameter setting shown in Table 2. To estimate fault size in the valve, the best performing model is the DBN model trained with time-series signal, as indicated by the
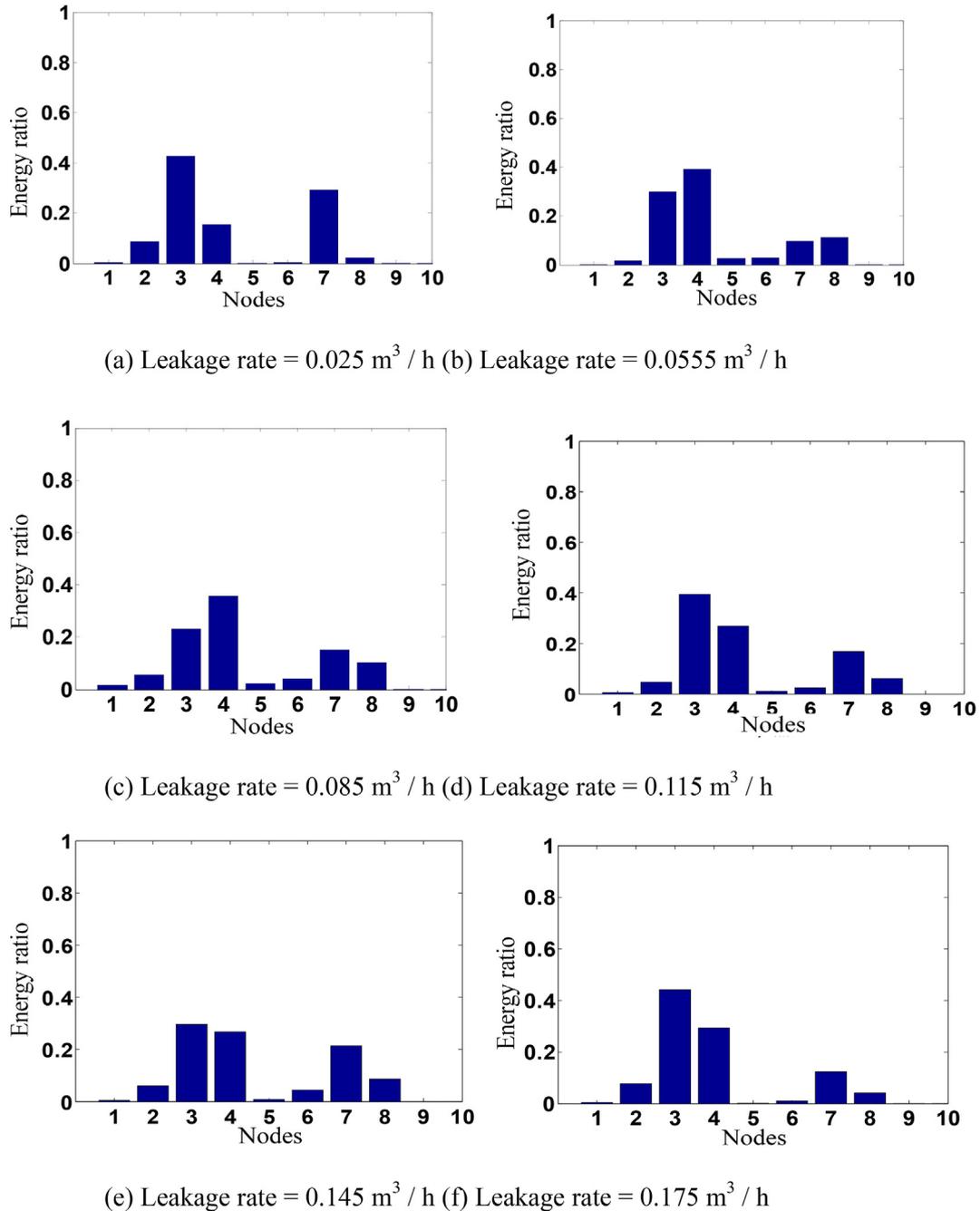
(a) Leakage rate = 0.025 m^3 / h (b) Leakage rate = 0.0555 m^3 / h

(c) Leakage rate = 0.085 m^3 / h (d) Leakage rate = 0.115 m^3 / h

(e) Leakage rate = 0.145 m^3 / h (f) Leakage rate = 0.175 m^3 / h

Fig. 16. Energy ratio of AE signal wavelet packet decomposition node.

Table 1
Composition of training samples and test samples.

Valve status	Number of training samples	Number of test samples	label
normal operation	240	120	1
Valve motor drive phase imbalance fault	240	120	2
Damaged valve packing	240	120	3

prediction plot, absolute error, and mean squared error. From this result, contrary to the perceived effectiveness of wavelet transformed data for developing deep learning models, we observed that the DBN fault size estimation model developed with wavelet transformed data performed worse than the model developed with the time-series signal.

5. Conclusion

This paper presents the development of a hybrid shallow-deep data-driven architecture for robust fault diagnosis in the DN50 electric valve. The shallow architecture is implemented with improved particle swarm support vector machine (IPs-SVM), while

Table 2
IPSO algorithm parameter settings.

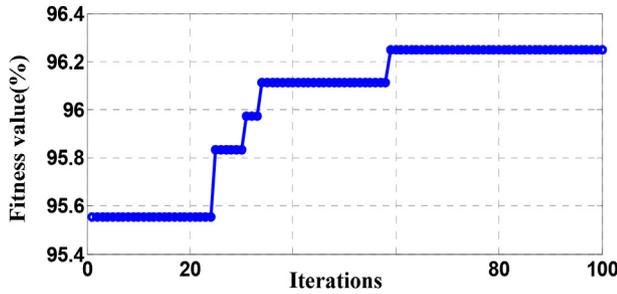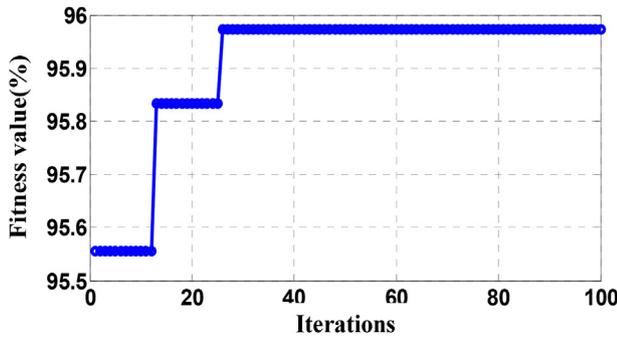| Parameter	Set value	Parameter	Set value
Learning factor c_1	1.5	Maximum inertia weight	1.2
Learning factor c_2	1.7	Minimum inertia weight	0.4
Learning factor c_3	1.5	Number of iterations	100
total group number	20	Control coefficient (e)	8

Fig. 17. IPs-SVM model training result.
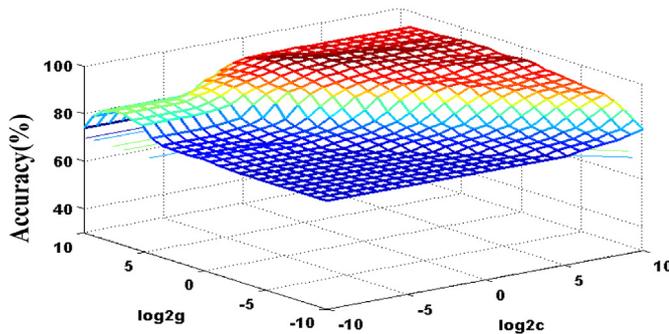

Fig. 18. PSO-SVM model training results curve.

Fig. 19. Grid search graph for hyper-parameter selection.

Table 3
Comparison of the results of three optimization methods.

Classifiers	Training result	Test Results	Penalty factor C	Kernel function parameter γ
Grid search SVM	95.6944%	91.9444%	1.3195	111.4305
PSO-SVM model	95.9722%	93.0556%	3.6788	50.0000
IPs-SVM model	96.2500%	93.3333%	3.4783	52.9270

the deep architecture uses a deep belief network (DBN). For efficient performance and to reduce false alarm rates, the model utilizes both vibration and acoustic signals. To achieve a robust

Table 4
DBN network parameter settings for the raw data sample.

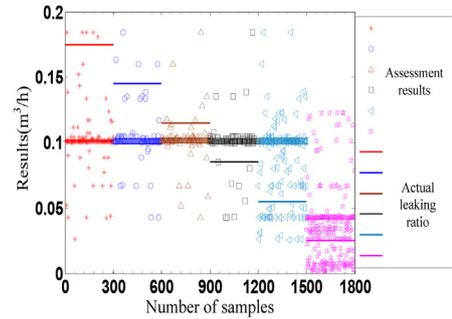|	Network structure	500-50-50
Unsupervised pre-training	Learning rate	0.7
	Momentum	0.01
	Batch size	60
	Number of training steps	150
Supervised training	Network structure	500-50-50-1
	Learning algorithm	Elastic BP algorithm
	Learning rate	0.6
	Number of training steps	20

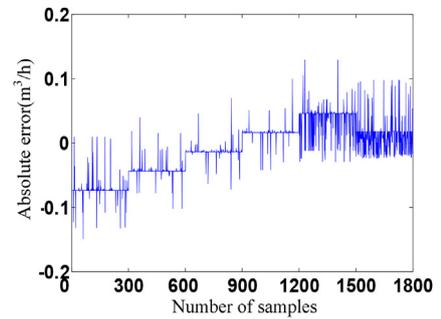Fig. 20. Leakage assessment result.

Fig. 21. The absolute error of the assessment result.

diagnostic system, the IPs-SVM uses the vibration signal for fault detection while the DBN utilizes an acoustic signal for fault size estimation. From the model evaluation results, we reached the following conclusions:

(1) The proposed IPs-SVM model solves the hyper-parameter selection problems of SVM and the local optimal issue of the conventional PSO. The model development using wavelet transformed signals also gives better performance and improved generalization and classification accuracy.

(2) For the DBN model, we observed that the model trained by time-series characteristic parameter samples achieve the minimum error and the best result. The model also shows

Table 5
DBN network parameter settings with time-series data.

	Network structure	6-6-6
Unsupervised pre-training	Learning rate	0.7
	Momentum	0.7
	Batch size	60
	Number of training steps	15
Supervised training	Network structure	6-6-6-1
	Learning algorithm	Elastic BP algorithm
	Learning rate	0.6
	Number of training steps	1500
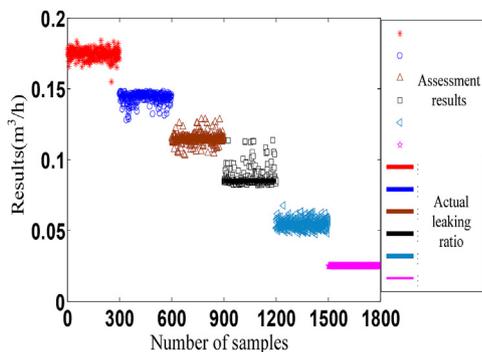
Fig. 22. Leak severity assessment result.
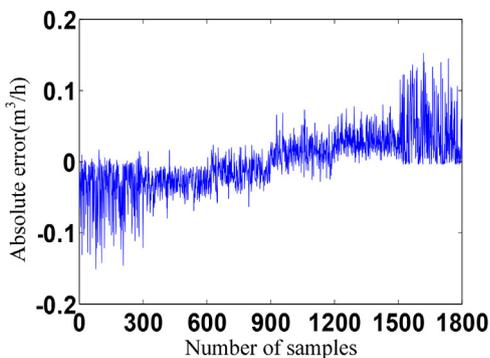
Fig. 23. The absolute error of the assessment result.

the capability to automatically extract the deep features of the signal, basically escaping the dependence on the traditional signal processing method and reducing the signal processing time.

(3) The experiment results show that this intelligent method applies to industrial-scale fault diagnosis of electric valves. Moreover, the accuracy of the classification results derived from these models is higher than that of the traditional fault

Table 6
DBN network parameter settings on wavelet packet transformed data.

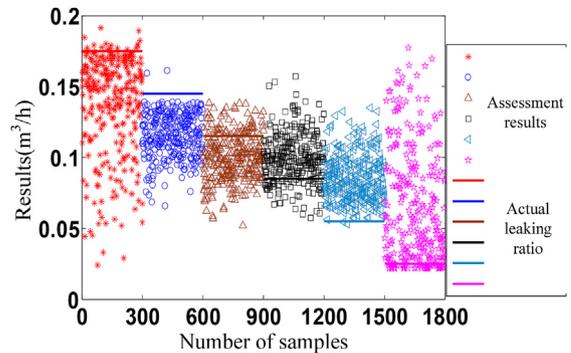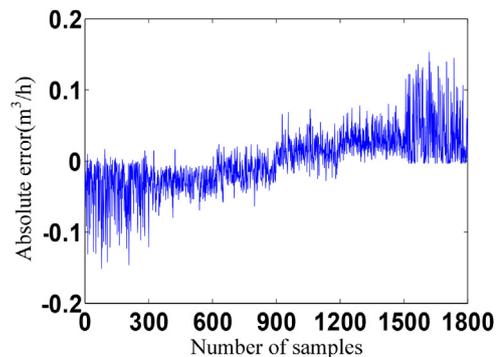|		
	Network structure	10-8-6
Unsupervised pre-training	Learning rate	0.7
	Momentum	0.7
	Batch size	100
	Number of training steps	50
Supervised training	Network structure	10-8-6-1
	Learning algorithm	Elastic BP algorithm
	Learning rate	0.6
	Number of training steps	1500

Fig. 24. Leakage assessment result.

Fig. 25. The absolute error of assessment result.

diagnosis method, which proves the feasibility and robustness of the method.

This research lays the foundation for subsequent research and engineering application of the technique in other industrial components. Our future research would focus on expanding the fault type and range, as well as integrating the shallow-deep model into a single system that is implementable in an operator support system.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.net.2020.07.001.

References

[1] A. Ayodeji, Y.-k. Liu, Support vector ensemble for incipient fault diagnosis in nuclear plant components, Nucl. Eng. Technol. 50 (2018) 1306–1313.
[2] M. Peng, et al., An intelligent hybrid methodology of on-line system-level fault diagnosis for nuclear power plant, Nucl. Eng. Technol. 50 (2018) 396–410.
[3] A. Ayodeji, Y.-k. Liu, H. Xia, Knowledge base operator support system for nuclear power plant fault diagnosis, Prog. Nucl. Energy 105 (2018) 42–50.
[4] A. Ayodeji, Y.-k. Liu, PWR heat exchanger tube defects: trends, signatures and diagnostic techniques, Prog. Nucl. Energy 112 (2019) 171–184.
[5] Z. Guo, et al., Defect detection of nuclear fuel assembly based on deep neural network, Ann. Nucl. Energy 137 (2019) 107078.
[6] K. Ryu, et al., Pipe thinning model development for direct current potential drop data with machine learning approach, Nucl. Eng. Technol. 52 (2020) 784–790.
[7] J. Zhang, et al., Application of cost-sensitive LSTM in water level prediction for nuclear reactor pressurizer, Nucl. Eng. Technol. 52 (2020) 1429–1435.
[8] J. Park, et al., MRPC eddy current flaw classification in tubes using deep neural networks, Nucl. Eng. Technol. 51 (2019) 1784–1790.
[9] J. Liu, et al., Nuclear power plant components condition monitoring by probabilistic support vector machine, Ann. Nucl. Energy 56 (2013) 23–33.
[10] H.A. Gohel, et al., Predictive maintenance architecture development for nuclear infrastructure using machine learning, Nucl. Eng. Technol. 52 (2020) 1436–1442.
[11] A. Ayodeji, Y.-k. Liu, SVR optimization with soft computing algorithms for incipient SGTR diagnosis, Ann. Nucl. Energy 121 (2018) 89–100.
[12] Z. Yangping, Z. Bingquan, W. DongXin, Application of genetic algorithms to fault diagnosis in nuclear power plants, Reliab. Eng. Syst. Saf. 67 (2000) 153–160.
[13] B. Yang, et al., Application of total variation denoising in nuclear power plant signal pre-processing, Ann. Nucl. Energy 135 (2020) 106981.
[14] J. Jiao, et al., Hierarchical discriminating sparse coding for weak fault feature extraction of rolling bearings, Reliab. Eng. Syst. Saf. 184 (2019) 41–54.
[15] J. Yang, J. Kim, An accident diagnosis algorithm using long short-term memory, Nucl. Eng. Technol. 50 (2018) 582–588.
[16] Y. Liu, et al., A cascade intelligent fault diagnostic technique for nuclear power plants, J. Nucl. Sci. Technol. 55 (2018) 254–266.
[17] A.R. Marklund, F. Michel, Application of a new passive acoustic leak detection approach to recordings from the Dounreay prototype fast reactor, Ann. Nucl. Energy 85 (2015) 175–182.
[18] W. Hwang, et al., Acoustic emission characteristics of stress corrosion cracks in a type 304 stainless steel tube, Nucl. Eng. Technol. 47 (2015) 454–460.
[19] Z.W. Jianguo, Support vector machine modeling and its intelligent optimization, in: Proceeding of International Conference on Information Computing and Automation Beijing, China, 2008. April 25-28.
[20] F. Van den Bergh, A.P. Engelbrecht, A study of particle swarm optimization particle trajectories, Inf. Sci. 176 (2006) 937–971.
[21] M. Guo, et al., Research on an integrated ICA-SVM based framework for fault diagnosis, in: SMC'03 Conference Proceedings. IEEE International Conference on Systems, Man and Cybernetics. Conference Theme-System Security and Assurance, Washington DC, USA, 2003, 8-8 October.
[22] L.Y. Jianwei, L. Xionglin, Research and development on deep learning, J. Comput. Appl. 31 (2014) 1922–1928.
[23] Y. Chen, D.-Q. Zheng, et al., Chinese relation extraction based on deep belief nets, Ruanjian Xuebao/J. Softw. 23 (2012) 2572–2585.
[24] Z.J. Licheng, et al., Deep Learning, Optimization and recognition(Chinese), Tsinghua University Press, Beijing, 2017, p. 58.
[25] C. Guoliang, Condition Recognition and Quantitative Analysis of Internal Leaks through Valves Based on Acoustic Emission Method (Chinese), China University of Petroleum Press, 2014, p. 17.
[26] A. Ayodeji, et al., Acoustic signal-based leak size estimation for electric valves using deep belief network, 5th International Conference on Computer and Communications (ICCC), Chengdu, China, Dec 6-9, 2019.