Gearbox fault diagnosis through quantum particle swarm optimization algorithm and kernel extreme learning machine

Shuo Meng[1], Jianshe Kang[2], Kuo Chi[3], Xupeng Die[4]
Shijiazhuang Branch, Army Engineering University, Shijiazhuang, China
[1]Corresponding author
E-mail: [1]*864748906@qq.com*, [2]*jskang201206@126.com*, [3]*dynamicck@emails.imau.edu.cn*,
[4]*diexupeng@163.com*

Check for updates

Abstract. Gearbox is the key component of mechanical transmission system. Accurate fault diagnosis of gearbox is of great significance to ensure the operation of rotating machinery. Based on the comprehensive simulation test-bed in the laboratory, a gearbox fault diagnosis method based on QPSO-KELM is proposed. Firstly, the fault pre planting experiments of gear fault, bearing fault and gear bearing mixed fault are carried out on the comprehensive simulation test-bed. Then, the vibration signals collected are preprocessed by TSA to eliminate noise. The time domain, frequency domain and NASA feature parameters of the preprocessed signals are taken as training samples and test samples of QPSO-KELM. The experimental results show that the proposed method can effectively solve the problem of gearbox fault pattern recognition, and the fault diagnosis accuracy is higher than traditional methods, so the research has certain reference significance and engineering application value.

Keywords: quantum particle swarm optimization, kernel extreme learning machine, gearbox, feature extraction, fault diagnosis.

1. Introduction

Gearbox is widely used in fans, machine tools, vehicles and other equipments. As a key component of mechanical transmission system, gearbox will cause huge economic losses and even casualties in case of failure during equipment operation. Therefore, the earlier the fault is found in the gearbox and the earlier the maintenance is carried out, the more loss can be reduced. The main components of gearbox are shaft, bearing and gear, among which gear failure and bearing failure are quite common.

At present, the mechanical transmission system fault diagnosis research is extensive, and the method is diverse. As the most mainstream classification method, intelligent classification algorithm has achieved good results in the field of equipment fault diagnosis. Common intelligent classification algorithms include neural network [1, 2], support vector machine (SVM) [3-5], kernel extreme learning machine (KELM) [6, 7], deep learning [8-10] and other methods. Wang [2] proposed a fault diagnosis method based on RDGWPR-MSE and PNN, which is used to realize the automatic fault identification of electric submersible pump. Wang and Yan [4] used the energy of IMF component after SVD decomposition as feature parameters to input and train SVM model, so as to realize bearing fault diagnosis. Saari et al. [5] proposed a fault diagnosis method for wind turbine bearing based on one class SVM. Iosifidis A et al. [7] studied the classification method of KELM, and achieved good results. Yang Yu et al. [9] proposed a structure adaptive DBN algorithm to solve the problem of difficult to determine the structure of DBN, which was successfully applied to the fault classification of rolling bearings. Shao et al. [10] successfully diagnosed the bearing fault by combining the dual tree complex wavelet packet transform with the adaptive DBN algorithm.

Compared with these methods, KELM has a stronger comprehensive advantage in sample

size, network generalization, calculation speed and accuracy, so KELM is more suitable for solving gearbox fault diagnosis problems that require higher calculation speed and accuracy. Since the classification capability of KELM network is greatly affected by the value of kernel parameters and penalty coefficient, it is necessary to optimize the above structure parameters. Currently, genetic algorithm [11], fish swarm algorithm [12], whale algorithm[13], ant colony algorithm [14], wolf algorithm [15], particle swarm algorithm [16] and other intelligent algorithms have been widely applied in the field of parameter optimization, such as Xu [13] used whale algorithm to optimize wavelet kernel extreme learning machine (WKELM) and Wang [17] used the modified WKELM wolves algorithm in diagnosing rolling bearing fault; Wu [18] used ant colony algorithm for fault diagnosis of rotating mechanical equipment. Pei [19] used particle swarm optimization algorithm to optimize KELM for transformer fault diagnosis. Liu [20] proposed the method of combining the variational mode decomposition (VMD) with the improved KELM for engine fault diagnosis. However, most of the above parameter optimization methods have the problem of long calculation time and easy to fall into the local optimum. Therefore, the QPSO-KELM (KELM-based on Quantum Particle Swarm Optimization) [21-23] was put forward. With strong global search ability, the QPSO can find the optimal parameters of KELM, thereby, the learning speed and classification accuracy of KELM can be improved and the accuracy of gearbox fault diagnosis will also be improved.

2. Fault diagnosis strategy of gearbox based on QPSO-KELM

The fault diagnosis process of gearbox based on QPSO-KELM is described as follows:

(1) Signal acquisition. In general, the closer the sensor is to the source of the vibration, the better the signal will be.

(2) Extract the fault characteristic parameters. The time-domain characteristic parameters, frequency-domain characteristic parameters and the characteristic parameters in the technical report of NASA were extracted from the collected signals. Then, the corresponding training set and test set were selected from all samples in a ratio of 3:1.

(3) The training of KELM. Aiming at maximizing the training accuracy, the training set is input to KELM, and the QPSO algorithm is used to search the optimal parameters. After that, the trained KELM can be obtained.

(4) Calculate the test accuracy. Input the test samples to the trained KELM, the correct classification times of test samples will be calculated and the test accuracy will be calculated.

The above process is shown in Fig. 1.
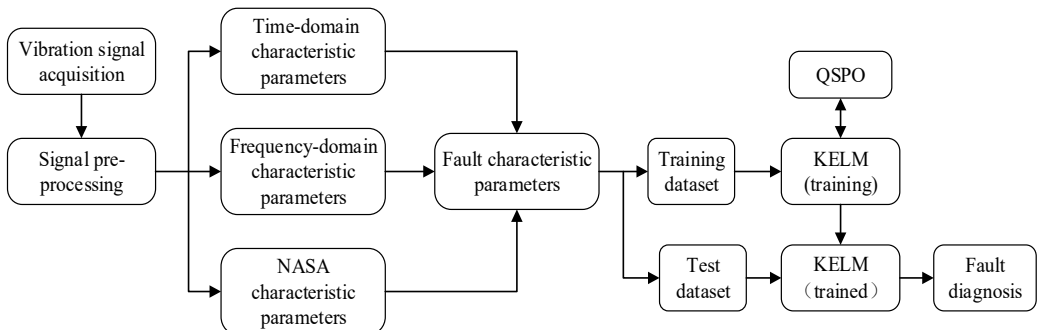
Fig. 1. The process of fault diagnosis of gearbox based on QPSO-KELM

In the Fig. 1, the gearbox fault diagnosis process also involves time-domain synchronous average algorithm [24], time-domain synchronous resampling algorithm [25, 26], pre-whitening algorithm [27], relevant fault characteristic parameters, KELM, QPSO and other relevant contents.

2.1. Signal preprocessing

During the signal acquisition process, the signal are easy interfered by other vibration signals, such as the vibration of the bearing of the equipment itself, the vibration of the motor, the meshing vibration of the gear independent of the synchronous shaft and the vibration from other machines and equipment. As a signal preprocessing technology, time synchronous averaging (TSA) technology can eliminate the signal components irrelevant to the synchronous shaft and enhance the fault signals of the synchronous shaft gear and its meshing gear, making the gear fault diagnosis easier. Where, the reference axis is the synchronous axis and the non-reference axis is the nonsynchronous axis. The theoretical equation of TSA is as follows:

$$y_n = \frac{1}{M} \sum_{m=0}^{M-1} x_{n+mN}, \quad n = 1, \dots N, \tag{1}$$

where, y_n is the TSA signal, x is the original signal, and M is the average times.

In the actual process, the implementation steps of TSA are as follows:

(1) Determine the zero crossing point of the synchronous axis according to the speed pulse signal, that is, the starting or ending position of each rotation of the synchronous axis.

(2) According to the position of zero crossing, divide the vibration acceleration signal into $M-1$ segments, and each segment represents the sampling signal of the integer rotation of the synchronous axis.

(3) Perform cubic curve interpolation for each signal segment, so that the sampling points of each signal segment are the same. The newly interpolated sampling points are $n = 2^{ceiling(log2(g))}$, where g is the mean of sampling points of each signal segment.

(4) Stack and average the resampled signals to obtain the time-domain synchronous average signal.

Based on the speed signal sampled at the same angle, TSA resamples the time-domain vibration acceleration signal to ensure that each revolution has the same number of sampling points, that is, equal angle sampling. The speed signal is usually obtained by the photoelectric sensor, and a rotation of the axis produces several pulses. The more pulses generated by a rotation, the more accurate the speed signal will be. The position of zero crossing can be obtained by rotating speed signal, and it is the key of the TSA algorithm. The gear fault data is mainly the fault characteristic parameters extracted from the time-domain synchronous average signal.

2.2. Extract feature parameters [28]

After preprocessing, the feature extraction of the preprocessed signal is carried out. The feature parameters are the basis for feature classification and fault recognition. The feature parameters adopted in this paper include time-domain feature parameters, frequency-domain feature parameters and some feature parameters mentioned in the report of NASA.

Time-domain characteristic parameters include: maximum (X_1), minimum (X_2), peak-to-peak (X_3), mean value (X_4), mean square value (X_5), root mean square (X_6), variance (X_7), standard deviation (X_8), energy (X_9), root square amplitude (X_{10}), mean square amplitude (X_{11}), mean square amplitude (X_{12}), kurtosis (X_{13}), skewness (X_{14}), waveform index (X_{15}), peak index (X_{16}), pulse index (X$_{17}$), margin index (X_{18}), clearance coefficient (X_{19}), etc:

$$X_1 = \max(x_i), \quad X_2 = \min(x_i), \quad X_3 = x_{\max} - x_{\min}, \quad X_4 = \frac{1}{N} \sum_{i=1}^{N} x_i,$$

$$X_5 = \frac{1}{N}\sum_{i=1}^{N} x_i^2, \quad X_6 = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}, \quad X_7 = \max|x_i|, \quad X_8 = \frac{1}{N}\sum_{i=1}^{N}(x_i - X_4)^2,$$

$$X_9 = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - X_4)^2}, \quad X_{10} = \sum_{i=1}^{N} x_i^2, \quad X_{11} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - X_4)^3}{X_6^3},$$

$$X_{12} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - X_4)^4}{X_6^4}, \quad X_{13} = \left(\frac{1}{N}\sum_{i=1}^{N}|x_i|^{1/2}\right)^2, \quad X_{14} = \frac{1}{N}\sum_{i=1}^{N}|x_i|,$$

$$X_{15} = \frac{X_6}{|X_{14}|}, \quad X_{16} = \frac{X_7}{X_6}, \quad X_{17} = \frac{X_7}{X_{14}}, \quad X_{18} = \frac{X_7}{X_{13}}, \quad X_{19} = \frac{X_1}{X_3}.$$

Some common statistical features in the frequency domain are shown below. Feature p_1 namely mean frequency. Features p_2-p_4, p_6, and p_{11}-p_{13} describe the convergence of the spectrum power, reflecting the energy of frequency spectrum. Features p_5 and p_7-p_9 show the change of main frequencies which are dominant in the frequency spectrum. The features in frequency-domain are calculated from the spectrum of the original signal, which contains more effective information than time-domain features:

$$p_1 = X_{mf} = \frac{\sum_{k=1}^{K} s(k)}{K}, \quad p_2 = \frac{\sum_{k=1}^{K}(s(k) - p_1)^2}{K - 1}, \quad p_3 = \frac{\sum_{k=1}^{K}(s(k) - p_1)^3}{K(\sqrt{p_2})^3},$$

$$p_4 = \frac{\sum_{k=1}^{K}((s(k) - p_1))^4}{Kp_2^2}, \quad p_5 = X_{fc} = \frac{\sum_{k=1}^{K} f_k s(k)}{\sum_{k=1}^{K} s(k)}, \quad p_6 = \sqrt{\frac{\sum_{k=1}^{K}(f_k - p_5)^2 s(k)}{K}},$$

$$p_7 = X_{rmsf} = \sqrt{\frac{\sum_{k=1}^{K} f_k^2 s(k)}{\sum_{k=1}^{K} s(k)}}, \quad p_8 = \sqrt{\frac{\sum_{k=1}^{K} f_k^4 s(k)}{\sum_{k=1}^{K} f_k^2 s(k)}}, \quad p_9 = \frac{\sum_{k=1}^{K} f_k^2 s(k)}{\sqrt{\sum_{k=1}^{K} s(k)\sum_{k=1}^{K} f_k^4 s(k)}},$$

$$p_{10} = \frac{p_6}{p_5}, \quad p_{11} = \frac{\sum_{k=1}^{K}(f_k - p_5)^3 s(k)}{Kp_6^3}, \quad p_{12} = \frac{\sum_{k=1}^{K}(f_k - p_5)^4 s(k)}{Kp_6^4},$$

$$p_{13} = \frac{\sum_{k=1}^{K}(f_k - p_5)^{1/2} s(k)}{K\sqrt{p_6}}, \quad p_{14} = \sqrt{\frac{\sum_{k=1}^{K}(f_k - p_5)^2 s(k)}{\sum_{k=1}^{K} s(k)}}.$$

In addition to the common time-domain and frequency-domain features, NASA also proposed some features, such as FM0, ER, FM4, FM4*, M6A, M6A*, M8A, M8A*, NA4, NA4*, NB4, NB4*, ER, EOP, etc. These features are mainly used to assess the gear failure:

$$FM0 = \frac{PP_x}{\sum_{n=0}^{H} P_n}, \quad FM4 = \frac{N\sum_{i=1}^{N}(d_i - \bar{d})^4}{\left[\sum_{i=1}^{N}(d_i - \bar{d})^2\right]^2}, \quad FM4^* = \frac{1/N\sum_{i=1}^{N}(d_i - \bar{d})^4}{\left[1/M'\sum_{j=1}^{M'}1/N\sum_{k=1}^{N}(d_{jk} - \bar{d}_j)^2\right]^2},$$

$$M6A = \frac{N^2\sum_{i=1}^{N}(d_i - \bar{d})^6}{\left[\sum_{i=1}^{N}(d_i - \bar{d})^2\right]^3}, \quad M6A^* = \frac{1/N\sum_{i=1}^{N}(d_i - \bar{d})^6}{\left[1/M'\sum_{j=1}^{M'}1/N\sum_{k=1}^{N}(d'_{jk} - \bar{d}'_j)^2\right]^3},$$

$$M8A = \frac{N^3\sum_{i=1}^{N}(d_i - \bar{d})^8}{\left[\sum_{i=1}^{N}(d_i - \bar{d})^2\right]^4}, \quad M8A^* = \frac{1/N\sum_{i=1}^{N}(d_i - \bar{d})^8}{\left[1/M'\sum_{j=1}^{M'}1/N\sum_{k=1}^{N}(d'_{jk} - \bar{d}'_j)^2\right]^4},$$

$$NA4 = \frac{N\sum_{i=1}^{N}(r_i - \bar{r})^4}{\left[1/M\sum_{j=1}^{M}\sum_{k=1}^{N}\left(r'_{jk} - \bar{r}'_j\right)^2\right]^2}, \quad NA4^* = \frac{1/N\sum_{i=1}^{N}(r_i - \bar{r})^4}{\left[1/M_1\sum_{j=1}^{M_1}\sum_{k=1}^{N}\left(r'_{jk} - \bar{r}'_j\right)^2\right]^2},$$

$$NB4 = \frac{1/N\sum_{i=1}^{N}(s_i - \bar{s})^4}{\left[1/M\sum_{j=1}^{M}1/N\sum_{k=1}^{N}\left(s_{jk} - \bar{s}_j\right)^2\right]^4}\bar{s}, \quad NB4^* = \frac{1/N\sum_{i=1}^{N}(s_i - \bar{s})^4}{\left[1/M'\sum_{j=1}^{M'}1/N\sum_{k=1}^{N}\left(s_{jk} - \bar{s}_j\right)^2\right]^4},$$

$$ER = \frac{RMS_{d_i}}{RMS_{d'_i}} = \sqrt{\frac{1/N\sum_{i-1}^{N}(d_i)^2}{1/N\sum_{i-1}^{N}(d_i')^2}}, \quad EOP = \frac{N\sum_{i=1}^{N}(re_i - \bar{re})^4}{[\sum_{i=1}^{N}(re_i - \bar{re})^2]^2}.$$

3. Research on classification method based on QPSO-KELM

3.1. KELM

In KELM algorithm, the kernel function is introduced into ELM, and the input weights and offsets in ELM are replaced by kernel function mapping, which makes the output of the ELM more stable and solves the over learning problems. Therefore, the KELM is widely used in classification identify areas.

ELM is a single hidden layer feedforward neural network, and its hidden layer weight does not need to be adjusted by feedback regulation. The structure of the ELM is shown in the Fig. 2, including an input layer, a hidden layer and an output layer.

Suppose $\mathbf{Z} = \{\mathbf{x}_i, \mathbf{t}_i\}_{i=1}^{N}$ are N training sample sets, where $x_i = [x_{i1}, x_{i1}, \cdots, x_{im}]^T \in \mathbf{R}^m$ is the input characteristic parameter and $t_i = [t_{i1}, t_{i1}, \cdots, t_{in}]^T \in \mathbf{R}^n$ is the sample label, then the output function expression of ELM training model is:

$$\sum_{i=1}^{M} \beta_i\, g(\mathbf{w}_i \cdot \mathbf{x}_i + b_i) = o_i, \quad i = 1,2,\cdots,N, \tag{2}$$

where, $w_i = [w_{1i}, w_{2i}, \cdots, w_{mi}]$ and $\beta_i = [\beta_{i1}, \beta_{i2}, \cdots, \beta_{in}]^T$ represent the input and output weight matrix of the ith hidden layer node respectively; $o_j = [o_{i1}, o_{i2}, \cdots, o_{in}]^T$ represents the output layer vector of the ELM.
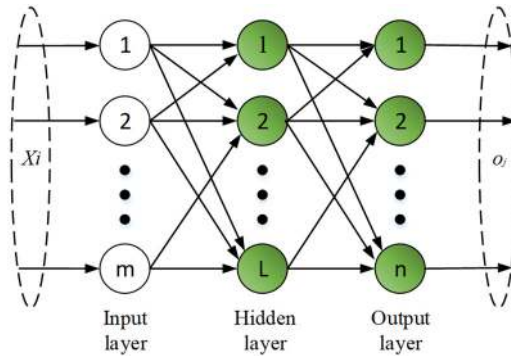

Fig. 2. ELM model

The goal of ELM training is to make the error between the actual output of the training sample and the sample label close to 0, the equation is:

$$\sum_{i=1}^{N} \|o_i - t_i\| = 0, \quad \beta_i w_i b_i \sum_{i=1}^{M} \beta_i\, g(\mathbf{w}_i \cdot \mathbf{x}_i + b_i) = t_i, \quad i = 1,2,\cdots,N, \quad H\beta = T. \tag{3}$$

So the existence of β_i, w_i and b_i can be used to obtain the Eq. (4), as below:

$$\sum_{i=1}^{M} \beta_i \, g(\mathbf{w}_i \cdot \mathbf{x}_i + b_i) = t_i, \quad i = 1,2, \cdots, N. \tag{4}$$

The Eq. (4) can be expressed as a matrix:

$$H\beta = T, \tag{5}$$

where, H is the output matrix of the hidden layer, which can be expressed in the following equation:

$$H(w_1, \cdots, w_M, b_1, \cdots, b_M, x_1, \cdots, x_N) = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_M \cdot x_1 + b_M) \\ \vdots & \cdots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_M \cdot x_N + b_M) \end{bmatrix}_{N \times M}, \tag{6}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_M^T \end{bmatrix}_{M \times n}, \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times n}. \tag{7}$$

For β, the equation is as follows:

$$\beta' = H^+ T, \tag{8}$$

where, H^+ represents the inverse matrix of H, β' is the optimal quadratic solution to β.

According to the structural optimization and ERM criteria, the output weight β of ELM can be determined, that is:

$$Minimize: \quad L_{P_{ELM}} = \frac{1}{2} \|\beta\|^2 + C \frac{1}{2} \sum_{i=1}^{N} \|\xi_i\|^2, \tag{9}$$

$$Subject\ to: \quad h(x_i), \quad \beta = t_i^T - \xi_i^T, \quad i = 1, \cdots, N,$$

where, $\xi_i = [\xi_{i,1}, \cdots, \xi_{i,m}]^T$ represents network calculation error, $\frac{1}{2}\|\beta\|^2$ is the structural error, $\frac{1}{2}\sum_{p=1}^{N}\|\xi_p\|^2$ is the empirical error, and C is the penalty factor.

Through the optimization of the solution, the optimal solution of β can be obtained as follows:

$$\beta = H^T \left(\frac{I}{C} + HH^T \right)^{-1} T. \tag{10}$$

Therefore, the output model function of ELM is as follows:

$$o_i = h_i H^T \left(\frac{I}{C} + HH^T \right)^{-1} T. \tag{11}$$

By using kernel operation to replace matrix operation H in ELM, the KELM classification model can be established. The kernel matrix is defined as:

$$\Omega = HH^T : \Omega_{i,j} = h(x_i) \cdot h(x_j) = K(x_i, x_j), \tag{12}$$

where Ω is the symmetric matrix of $N \times N$ and $K(x, y)$ is the kernel function.

Eq. (11) can be expressed by kernel function as follows:

$$o_i = [K(x_i, x_j), \cdots, K(x_i, x_j)]\lambda, \tag{13}$$

where, λ represents the output weight matrix of KELM.

Different support vector machines can be constructed by selecting different kernel functions. Among the commonly used kernel functions, compared with the polynomial kernel function and Sigmoid kernel function, the radial basis kernel function has the advantages of simple parameters and strong adaptability to randomly distributed samples. Therefore, a SVM based on radial basis kernel function is established in this paper, and its expression is shown as follows:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{\sigma^2}\right). \tag{14}$$

According to Eq. (9) and (14), it can be seen that KELM contains kernel parameter σ and penalty coefficient C, and it is necessary to find the optimal parameter to make the classification effect of KELM the best.

3.2. QPSO-KELM

By quantizing the iterative updating process of particles of PSO algorithm, the QPSO can reduce the algorithm complexity, improve algorithm convergence speed and global search ability. The basic principles of QPSO are as follows:

Assume that Ω is the d-dimensional search space, and the population number of particles in the space is b, then the position of the ith particle can be expressed as:

$$V_i = (v_{i1}, v_{i2} \cdots v_{id}). \tag{15}$$

Suppose the individual optimal position of the particle is p_{ibest}, and the global optimal position of the particle is p_{gbest}, the p_{ibest} and p_{gbest} are as follows:

$$p_{ibest} = (p_{i1}, p_{i2} \cdots p_{id}), \tag{16}$$
$$p_{gbest} = (p_{g1}, p_{g2} \cdots, p_{gd}). \tag{17}$$

The particle can find and update its individual optimal position p_i and population optimal position p_g through iterative operation, and the average optimal position $mbest$ can be introduced as the population optimal center. Then the particle optimization process can be expressed as:

$$mbest = \frac{1}{M}\sum_{i=1}^{M} p_{ibest} = \left[\frac{1}{M}\sum_{i=1}^{M} p_{i1}, \frac{1}{M}\sum_{i=1}^{M} p_{i2} \cdots \frac{1}{M}\sum_{i=1}^{M} p_{id}\right], \tag{18}$$
$$p_{id} = \varphi p_{id} + (1 - \varphi)p_{gd}, \quad \varphi \in (0,1), \tag{19}$$
$$v_{id} = p_{id} \pm \alpha|mbest - v_{id}|\ln(1/u), \quad u \in (0,1), \tag{20}$$

where, α is the contraction expansion factor, which is dynamically adjusted in the iterative operation according to the Eq. (21):

$$\alpha = (\alpha_1 - \alpha_2) \times \frac{N - t}{N} + \alpha_2, \tag{21}$$

where, N represents the maximum number of iterations. α_1 and α_2 are the initial and final values of α respectively, usually the two parameters are set as: $\alpha_1 = 0.5$, $\alpha_2 = 1$.

According to the basic principle of KELM and QPSO, the network training process which is used QPSO to optimize the network structure parameters of KELM is as follows:

(1) Initialize the position of population particles, and set parameters such as particle swarm size, iteration step size, termination conditions and so on.

(2) Initialize the current position of each particle as p_{ibest}, define the classification accuracy of KELM as fitness function, and calculate the fitness value of each particle. The position of the particle with the maximum fitness value was initialized as p_{gbest}.

(3) Update the particle position according to Eq. (18-20).

(4) The fitness value of each particle is calculated, and the individual optimal position p_{best}, the group optimal position p_{gbest} and the group optimal center $mbest$ are updated based on the optimal fitness value.

(5) Judge whether the termination conditions are met. If so, stop the calculation and output the result, if not, return to step (3).

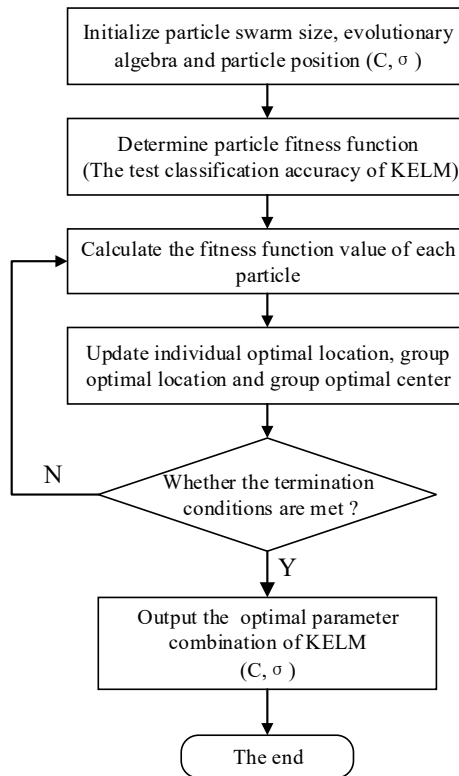According to the above steps, the modeling flow chart of QPSO-KELM can be drawn as shown in Fig. 3.

Fig. 3. Modeling flow chart of QPSO-KELM

4. Experimental verification and discussion

To verify whether the QPSO-KELM is available, the pre planting fault tests are carried out for gear fault, bearing fault and gear-bearing mixed fault. The fault simulation test rig is shown in Fig. 4. The test rig consists of four parts: the power and control part, the bearing fault simulation part (not used), the gear fault simulation part and the data acquisition part (not shown). This section mainly uses the gearbox fault simulation part, which is mainly composed of a two-stage reduction spur gearbox, a magnetic powder brake (supply the load) and a magnetic powder brake controller

(control the load change).

The perspective and internal structure of the two-stage reduction spur gearbox are shown in Fig. 5. The teeth number in the gearbox from high speed shaft to low speed shaft is 41, 79, 36 and 90. The fault gear is the intermediate shaft pinion gear (gear 3 with 36 teeth). The fault types include wear, break and miss teeth. The fault bearing is located in the deep groove ball bearing at the end cover of the gear 3 side of the medium-speed shaft (Fig. 5(b)). The deep groove ball bearing type is ER-16k, and its size parameters are shown in Table 1. The fault types include inner race fault, outer race fault and single roller fault.
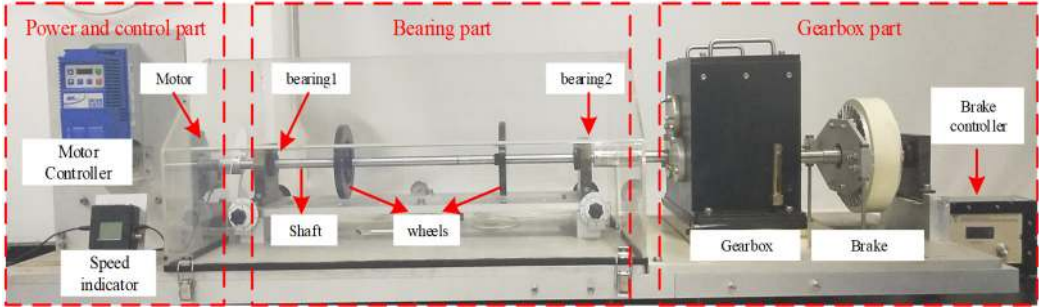

Fig. 4. The mechanical fault simulation test bench

In the test, we set two vibration acceleration sensors in vertical and horizontal directions respectively. Both sensors are installed outside the end cover of the fault bearing (Fig. 5(c)), which are used to collect the vibration acceleration data in vertical and horizontal directions. The data acquisition parameters are set as: sampling frequency $fs = 20.48$ kHz, sampling time $t = 48$ s, motor speed 30 r/s (the actual motor speed is about 29.602 r/s). According to the speed of the motor, the number of gear teeth of the gearbox and the main dimension parameters of the bearing ER-16K (Table 1), the main relevant frequency of the gearbox can be calculated by the calculation formula of bearing fault characteristic frequency (Table 2). The calculation results are shown in Table 3.
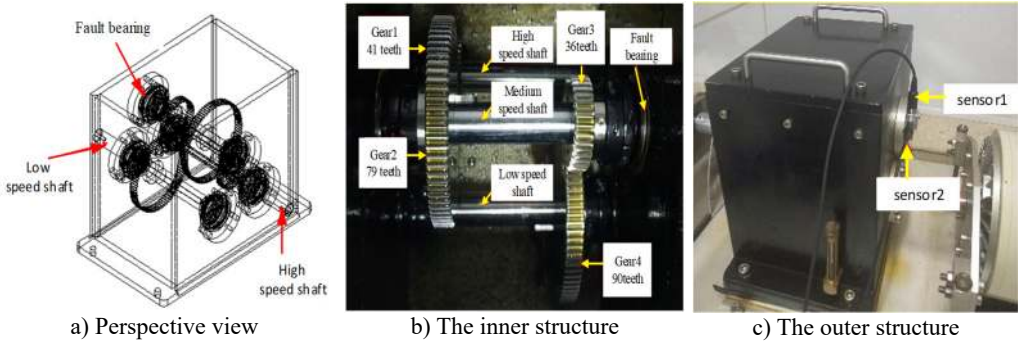

| a) Perspective view | b) The inner structure | c) The outer structure |

Fig. 5. The structure of gearbox

4.1. Bearing fault diagnosis

Firstly, the single fault of bearing is analyzed. In the experiment, the motor speed is set to 30 r/s, and the signal sampling frequency is 20.48 kHz. The analysis data is mainly collected by sensor1 (vertical direction). The test data types include bearing normal data, rolling element fault data, inner ring fault data and outer ring fault data, each type contains 240 groups of data, and 960 groups of data is in total. The sampling time of each group of data is 2 seconds.

Table 1. Main dimensions of bearing ER-16K

Number of rolling elements N_b	Ball diameter d / inch	Pitch diameter D / inch	Contact angle θ (°)
9	0.3125	1.516	0

Table 2. The formulas of fault characteristic frequencies to bearings

| Name	Formulas
Fault characteristic frequencies of outer race	$f_{BPFO} = \dfrac{N_b}{2} f_r \left[1 - \dfrac{d\cos(\theta)}{D} \right]$
Fault characteristic frequencies of inner race	$f_{BPFI} = \dfrac{N_b}{2} f_r \left[1 + \dfrac{d\cos(\theta)}{D} \right]$
Fault characteristic frequencies of ball spin	$f_{BSF} = \dfrac{D}{d} f_r \left[1 - \dfrac{d^2\cos^2(\theta)}{D^2} \right]$
N_b is the number of rolling elements, d means the Ball diameter, D means the Pitch diameter, θ means the Contact angle, f_r is the bearing rotation frequency	

Table 3. Relevant frequencies of the gearbox

Rotating frequency / r·s⁻¹			Meshing frequency / Hz		Fault feature frequency of bearing/ Hz		
f_{r1}	f_{r2}	f_{r3}	f_{m1}	f_{m2}	f_{BPFO}	f_{BPFI}	f_{BSF}
29.602	15.363	6.145	1213.662	553.061	54.882	82.448	35.681
f_{r1} is the rotating frequencies of high speed axis, f_{r2} is the rotating frequencies of medium speed axis, f_{r3} is the rotating frequencies of low speed axis; f_{m1} is the meshing frequencies of gear 1 and gear 2, f_{m2} is the meshing frequencies of gear 3 and gear 4; f_{BPFO} is the fault characteristic frequencies of the bearing outer race, f_{BPFI} is the fault characteristic frequencies of the bearing inner race, f_{BSF} is the fault characteristic frequencies of the bearing ball spin							

When diagnosing the bearing fault, follow the procedure shown in Fig. 1. Firstly, the time-domain, frequency-domain and NASA characteristic parameters of the bearing fault were extracted from the bearing data of each group, and 960 samples will be obtained. And each group of the samples is marked with the following rules: normal-1, roller-2, inner-3, outer-4. Then, the training set and test set are selected from all samples in a ratio of 3:1, and the training set is input to KELM for training. In order to maximize the test accuracy, QPSO is used to optimize the parameters of KELM. Finally, in order to verify the superiority and effectiveness of the method, the ELM and PSO-KELM are also applied to bearing fault diagnosis, and the experiment results of different methods are used to compare with each other.

Since the standard elm algorithm uses a single layer feedforward neural network structure, it is unnecessary to consider the parameter optimization of ELM. So the parameter optimization process of PSO-KELM and QPSO-KELM are analyzed, the results is shown in Fig. 6. Therefore, it is only necessary to analyze the parameter optimization process of PSO-KELM and QPSO-KELM.
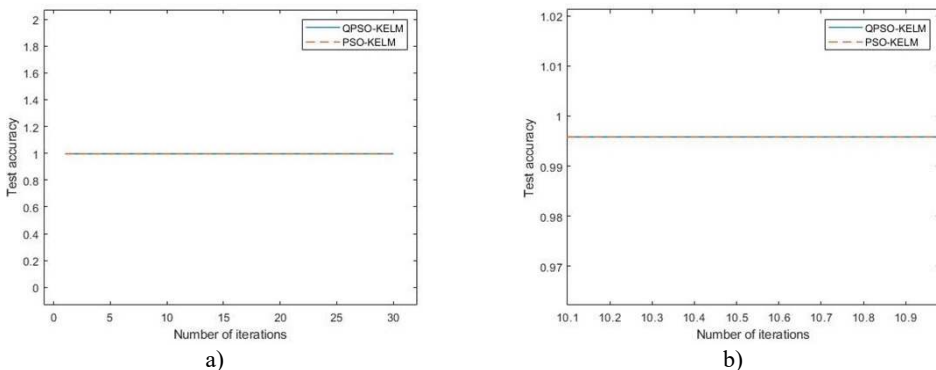

Fig. 6. a) Iterative process of the two methods, b) the local amplify of a)

Fig. 6(a) is iterative process of the two methods, Fig. 6(b) is the local amplify of Fig. 6(a). It can be seen from the figure that PSO-KELM and QPSO-KELM algorithm have the highest test accuracy from the beginning, which shows that the fault characteristics of bearing single fault are obvious, and it can achieve high classification accuracy without parameter iteration optimization. The initial parameters setting of QPSO-KELM is $C \in [0.01,1000]$, $\sigma \in [0.01,100]$, and the optomized parameters are $C = 1000$, $\sigma = 18$.

The final training accuracy and test accuracy of the methods are shown in Table 4. It can be seen from Table 4 that both of the training accuracy and the test accuracy of QPSO are the highest and reach 99.58 %. The QPSO-KELM classification results of the test dataset are shown in Fig. 7, in which there is only one misjudgment point, so the accuracy of the trained QPSO-KELM model is reliable.

Table 4. Bearing fault classification accuracy

	SVM	CNN	ELM	PSO-KELM	QPSO-KELM
Training accuracy	90.43 %	98.89 %	91.74 %	98.47 %	99.58 %
Testing accuracy	89.97 %	98.74 %	91.74 %	97.92 %	99.58 %
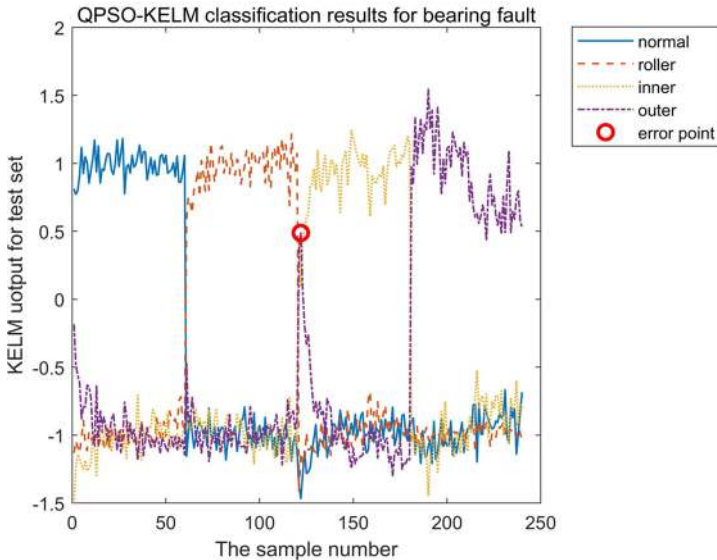
Fig. 7. Classification results of bearing test dataset based on QPSO-KELM

4.2. Gear fault diagnosis

Secondly, the single fault of gear is analyzed. The fault gear is the medium-speed shaft pinion in the gearbox, which is gear 3 in Fig. 5(b). In the experiment, the motor speed and the signal sampling frequency are the same as in section 4.1. The analysis data is mainly collected by sensor 1 (vertical direction). The test data types include gear normal data, worn teeth data, broken tooth data and missing tooth data, each type contains 240 groups of data, and 960 groups of data is in total. The sampling time of each group of data is 2 seconds.

When diagnosing the gear fault, follow the procedure shown in Fig. 1. Firstly, the characteristic parameters of gear fault in time-domain, frequency-domain and NASA are extracted from each group of gear data, and 960 samples will be gotten. And each group of the samples is marked with the following rules: normal-1, wear-2, miss-3, break-4. Then, the training set and test set are selected from all samples in a ratio of 3:1, and the training dataset is input to KELM for training. In order to maximize the test accuracy, QPSO is used to optimize the parameters of KELM. Finally, in order to verify the effectiveness of the method and highlight the advantages of the

proposed method. The ELM and PSO-KELM are also used for gear fault diagnosis to compare with QPSO-KELM.

The parameters optimization of PSO-KELM and QPSO-KELM are analyzed, the iteration process is shown in Fig. 8. It can be seen from the figure that the QPSO-KELM algorithm has the highest test accuracy from the beginning, which indicates that the QPSO-KELM has stronger learning ability, while PSO-KELM reaches the maximum value at the 11th iteration, which indicates that its learning ability is slightly inferior to the method proposed in this paper. The initial setting of parameters is the same as that in Section 4.1, and the settings after parameter optimization are as follows: $C = 1000$, $\sigma = 19$.

The final training accuracy and test accuracy of the methods are shown in Table 5. It can be seen from Table 5 that both of the training accuracy and the test accuracy of QPSO are the highest and the test accuracy reaches 98.75 %. The QPSO-KELM classification results of the test dataset are shown in Fig. 9, in which there are only three error points, so the accuracy of the trained QPSO-KELM model is available. From the iteration process and fault diagnosis accuracy, it can be seen that the fault features of bearing are more obvious than that of gear. Therefore, gear fault diagnosis needs less iteration times, but higher test accuracy can be achieved.
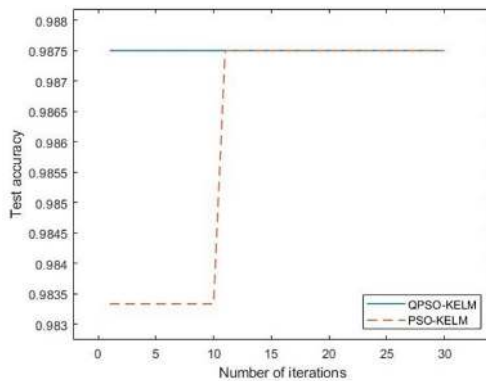

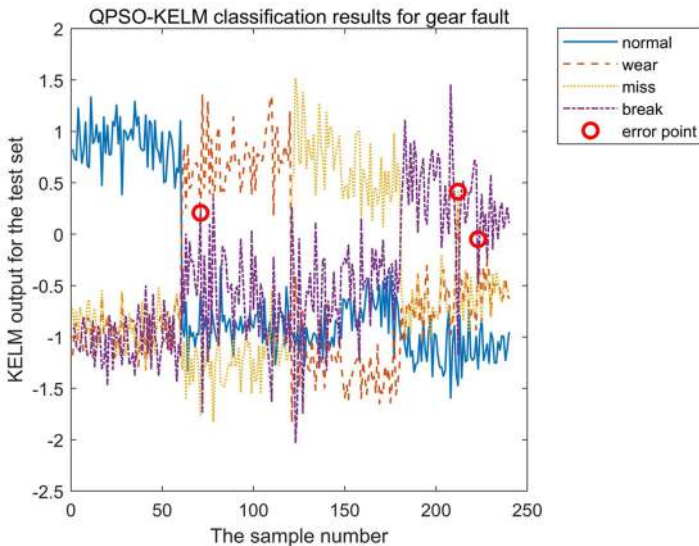
Fig. 8. Iterative process of the two methods

Fig. 9. Classification results of gear test dataset based on QPSO-KELM

Table 5. Gear fault classification accuracy

	SVM	CNN	ELM	PSO-KELM	QPSO-KELM
Training accuracy	83.41 %	98.43 %	80.88 %	98.75 %	99.31 %
Testing accuracy	81.57 %	97.69 %	80.88 %	97.92 %	98.75 %

4.3. Gear-bearing fault diagnosis

At last, the composite faults of gear-bearing are analyzed. The fault parts are the gear and bearing mentioned in the previous section. There are 10 types of gear-bearing data: normal gear-normal bearing data, worn gear – faulty bearing rolling element data, worn gear-faulty bearing inner race data, worn gear-faulty bearing outer race data, gear tooth missing-faulty bearing rolling element data, missing gear tooth-faulty bearing inner race data, missing gear tooth-faulty bearing outer race data, broken gear tooth-faulty bearing rolling element data, broken gear tooth-faulty bearing inner race data, broken gear tooth-faulty bearing inner race data. The mixed faults of gear-bearing is shown in Table 6.

In the experiment, the motor speed is set to 30 r/s, and the signal sampling frequency is 20.48 kHz. The sampling time of each group of data is 2 seconds. The total data is 2472, except the data of the normal state (sample size is 288) and the broken tooth- faulty inner race (sample size is 264), the sample sizes of the other states are all 240 of each group.

Table 6. The mixed faults of gear-bearing

Combinations	Gear fault				Bearing fault			
	Normal	Wear	Miss	Break	Normal	Roller	Inner	Outer
1	+				+			
2		+				+		
3		+					+	
4		+						+
5			+			+		
6			+				+	
7			+					+
8				+		+		
9				+			+	
10				+				+

When diagnosing the gear-bearing fault, follow the procedure shown in Fig. 1. Firstly, the feature parameters of gear-bearing fault in time-domain, frequency-domain and NASA are extracted from each group of gear-bearing data, and 2742 samples will be gotten. Each group of samples is labeled with the following rules: normal-1, wear-roller-2, wear-inner-3, wear-outer-4, miss-roller-5, miss-inner-6, miss-outer-7, break-roller-8, break-inner-9, break-outer-10. Then, the training set and test set are randomly selected from all samples in a ratio of 3:1, and the training dataset is input to KELM for training. In order to maximize the test accuracy, QPSO is used to optimize the parameters of KELM. Finally, in order to verify the effectiveness of the method and the superiority of the method, the ELM and PSO-KELM are also used for gear-bearing fault diagnosis to compare with QPSO-KELM.

The optimization process of PSO-KELM and QPSO-KELM are shown in Fig. 10. It can be seen from the figure that the test progress of QPSO-KELM algorithm reach the highest after six iterations, while the PSO-KELM algorithm has no transformation from the beginning, and the accuracy is low, indicating that the classification ability of PSO-KELM is weaker than that of QPSO-KELM. The initial setting of parameters is the same as that in Section 4.1, and the settings after parameter optimization are as follows: $C = 1000$, $\sigma = 20$.

The final classification accuracy of the three methods for mixed faults of gear-bearing is shown in Table 7. It can be seen from Table 7 that the training accuracy of QPSO is not the highest in the training process, but when comparing the final test accuracy, the classification accuracy of QPSO

is far higher than that of the other two calculations, and the test accuracy is 90.13 %. However, the test accuracy of the other two methods are no higher than 60 %. The QPSO-KELM classification results of the test dataset are shown in Fig. 11. As can be seen from Fig. 11, the error points are mainly concentrated in three states: the worn-rolling element fault state, worn-outer race fault state and broken-outer race fault state. Therefore, more sensitive feature parameters should be extracted for these three states.
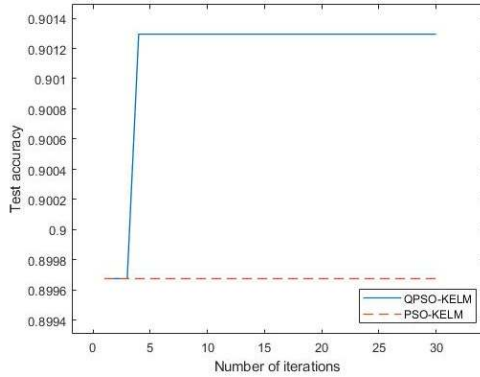

Fig. 10. Iterative process of the two methods

Table 7. Classification accuracy of mixed faults of gear-bearing

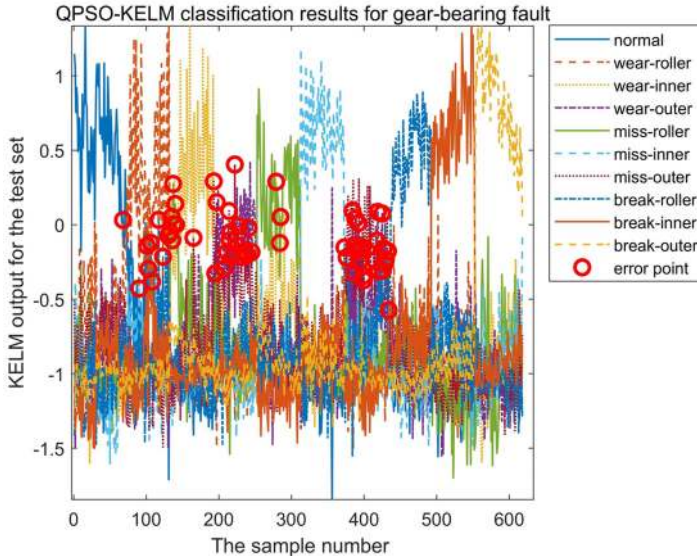|	SVM	CNN	ELM	PSO-KELM	QPSO-KELM
Training accuracy	58.43 %	90.67 %	53.34 %	92.27 %	92.45 %
Testing accuracy	55.52 %	89.49 %	53.34 %	89.64 %	90.13 %

Fig. 11. Classification results of gear-bearing test dataset based on QPSO-KELM

5. Conclusions

A novel QPSO-KELM method is proposed in this paper and applied to gearbox fault diagnosis. The method is implemented by using QPSO algorithm to select the proper KELM parameters, kernel parameter σ and penalty coefficient C. In order to verify the effectiveness and the superiority of the method, the time-domain, frequency-domain and NASA features are collected,

the results of the bearing single fault, gear single fault and gear-bearing fault are analyzed. After comparing with two other methods (ELM and PSO-KELM), and analyzing the test results, the following conclusions can be given: (1) Experimental results show that the proposed method is effective, it can identify the types of faults mentioned above; (2) Compared with SVM, ELM, CNN and PSO-KELM, OPSO-KELM has more advantages. The accuracy of QPSO-KELM is more than 90 % in both training data and test data, which is higher than other methods.

The study focuses on optimizing the structural parameters of KELM and the application in gearbox fault diagnosis. From the analysis of experimental results, however, it can be found that there is still a space to improve the identification accuracy of gear-bearing hybrid fault. Since the accuracy of fault judgment is closely related to fault feature extraction, in order to further improve the fault identification accuracy of composite faults, the next step of the study is to extract more sensitive and effective fault feature parameters.

References

[1] **Qu Jianling, Yu Lu, Yuan Tao** Adaptive fault diagnosis algorithm for rolling bearing based on one-dimensional convolutional neural network. Chinese Journal of Scientific Instrument, Vol. 39, Issue 7, 2018, p. 134-143.

[2] **Wang Yakun** A diagnosis method for electrical submersible pump based on RSGWPT-MSE and PNN. Mechanical Engineer, Vol. 1, 2018, p. 167-170.

[3] **Yu Yang, Bai Rui, Yang Ping** Application of support vector machine based on genetic algorithm optimization on acoustic emission detection of gear fault. Journal of Mechanical transmission, Vol. 29, Issue 9, 2017, p. 2189-2197.

[4] **Wang X., Yan W.** Fault diagnosis of roller bearings based on the variational mode decomposition and SVM. Journal of Vibration and Shock, Vol. 36, Issue 18, 2017, p. 252-256.

[5] **Saari J., Strombergsson D., Lundberg J., et al.** Detection and identification of windmill bearing faults using a one-class support vector machine (SVM). Measurement, Vol. 137, 2019, p. 287-301.

[6] **Huang G. B., Zhou H., Ding X.** Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics), Vol. 42, Issue 2, 2012, p. 513-529.

[7] **Iosifidis A., Tefas A., Pitas I.** On the kernel extreme learning machine classifier. Pattern Recognition Letters, Vol. 54, Issue 1, 2015, p. 11-17.

[8] **Wen Jiangtao, Yan Changhong, Sun Jiedi** Bearing fault diagnosis method based on compression acquisition and deep learning. Chinese Journal of Scientific Instrument, Vol. 39, Issue 1, 2018, p. 171-179.

[9] **Yang Yu, Lu Peng, Gan Lei, et al.** SADBN and its application in rolling bearing fault identification and classification. Journal of Vibration and Shock, Vol. 38, Issue 15, 2019, p. 11-16.

[10] **Shao H., Jiang H., Wang F., et al.** Rolling bearing fault diagnosis using adaptive deep belief network with dual-tree complex wavelet packet. ISA Transactions, Vol. 3, Issue 69, 2017, p. 187-201.

[11] **He Guangyu, Liu Xinxue, Wei Xuenkai** Applications of genetic algorithm to aeroengine fault diagnosis. Fire Control and Command Control, Vol. 31, Issue 11, 2006, p. 82-84.

[12] **Zhang Ning, Wei Xiuye, Guo Xiaoyong, et al.** Study on fault diagnosis for bearings based on improved fish swarm algorithm to optimize neural network. Bearing, Vol. 474, Issue 5, 2019, p. 49-53.

[13] **Xu J., Wang Y., Ji Z.** Fault diagnosis method of rolling bearing based on WKELM optimized by whale optimization algorithm. Journal of System Simulation, Vol. 29, Issue 9, 2017, p. 2189-2197.

[14] **Duan Haibin, Wang Daobo, Yu Xiufen** Ant colony algorithm: survey and prospect. Engineering Science, Vol. 9, Issue 2, 2007, p. 98-102.

[15] **Wu Husheng, Zhang Fengming, Wu Lushan** New swarm intelligence algorithm – wolf pack algorithm. Systems Engineering and Electronics, Vol. 35, Issue 11, 2013, p. 2430-2438.

[16] **Agrawal S., Panigrahi B. K., Tiwari M. K.** Multiobjective particle swarm algorithm with fuzzy clustering for electrical power dispatch. IEEE Transactions on Evolutionary Computation, Vol. 12, Issue 5, 2008, p. 529-541.

[17] **Wang Tiantian, Wang Yan, Ji Zhicheng** Fault diagnosis of rolling bearing based on improved extreme learning machine. Journal of System Simulation, Vol. 30, Issue 11, 2018, p. 390-397.

[18] **Wu Wei, Gui Boxiang** Research on intelligent diagnosis method of machinery fault based on ant colony algorithm. Mechanical Research and Application, Vol. 6, 2014, p. 15-17.

[19] **Pei Fei, Chen Xuechen, Zhu Yongli, et al.** Transformer fault diagnosis based on particle swarm optimization and kernel-based extreme learning machine. Computer Engineering and Design, Vol. 5, 2015, p. 1327-1331.

[20] **Liu Min, Zhang Yingtang, Li Zhining, et al.** Engine fault diagnosis based on IVMD and modified KELM. Journal of Vibration, Measurement and Diagnosis, Vol. 39, Issue 4, 2019, p. 875-883.

[21] **Coello C. A. C., Pulido G. T., Lechuga M. S.** Handling multiple objectives with particle swarm optimization. IEEE Transactions on Evolutionary Computation, Vol. 8, Issue 3, 2004, p. 256-279.

[22] **Pang S., Yang X., Lin X.** Evolutionary extreme learning machine optimized by quantum-behaved particle swarm optimization. Journal of System Simulation, 2017.

[23] **Rehman O. U., Yang S., Khan S.** A quantum particle swarm optimizer with enhanced strategy for global optimization of electromagnetic devices. IEEE Transactions on Magnetics, Vol. 55, Issue 8, 2019, p. 7000804.

[24] **Luo Deyang** Principle and application of time-domain synchronous averaging. Journal of Vibration, Measurement & Diagnosis, Vol. 19, Issue 3, 1999, p. 202-207.

[25] **Li Jie, Zhao Jianmin** Chaotic characteristic verification of gearbox vibration signal based on the time synchronous averaging and separation technology. Journal of Mechanical Transmission, Vol. 7, 2018, p. 33.

[26] **Del Moral P., Doucet A., Jasra A.** On adaptive resampling strategies for sequential Monte Carlo methods. Bernoulli, Vol. 18, Issue 1, 2012, p. 252-278.

[27] **Sawalhi N., Randall R. B.** Signal pre-whitening using cepstrum editing (liftering) to enhance fault detection in rolling element bearings. Proceedings of the 24th International Congress on Condition Monitoring and Diagnostic Engineering Management, 2011.

[28] **Li Haiping, Zhao Jianmin, Song Wenyuan** Method of planetary gearbox feature extraction and condition recognition based on EMD and EDT. Journal of Vibration and Shock, Vol. 35, Issue 3, 2016, p. 48-54.

Shuo Meng received the B.Sc. degree in Army Engineering University of PLA, Shijiazhuang, China, in 2018. Now, he is a full-time M.S. student of the Army Engineering University of PLA. His current research is focused on mechanical fault detection, intelligent diagnosis, and prognostics.

Jianshe Kang received the Ph.D. degree in mechatronical engineering from Beijing Institute of Technology, Beijing, China. He is a Professor at Army Engineering University of PLA, Shijiazhuang, China. He is a Direct General of China Ordnance Industry Society and selected as the editorial board of Acta Armamentarill. He authored one book in the field of maintenance engineering. He published about 60 journal papers. His current research interests include system reliability analysis, condition-based prognostics and health management of capital assets.

Kuo Chi received the B.Sc. degree in mechanical engineering from Jimei University, Xiamen, China, in 2013 and the M.S. degree in maintenance engineering from Mechanical Engineering College, Shijiazhuang, China, in 2015. Now, he is a full-time Ph.D. student of the Army Engineering University of PLA, Shijiazhuang, China. His current research is focused on mechanical fault detection, diagnostics, and prognostics.

Xupeng Die received the B.Sc. degree in maintenance engineering from Mechanical Engineering College, Shijiazhuang, China, in 2017 and the M.S. degree in Army Engineering University of PLA, Shijiazhuang, China, in 2019. His current research interests include mechanical fault feature extraction and fault diagnosis.