Fault Diagnosis of Rolling Bearing Based on Probability box Theory and GA-SVM

HONG TANG1,∗, ZHENGXING YUAN2,∗, HONGLIANG DAI1, AND YI DU2
1 State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha 410083, China
2 City College, Kunming University of Science and Technology, Kunming 650051, China
Corresponding author: Yi Du (duyi12844059@sina.cn)
∗Hong Tang and Zhengxing Yuan contributed equally to this work.
This work was supported in part by the Creative Research Groups Foundation of China under Grant 51621004, in part by the National Natural Science Foundation of China under Grant 51365020, in part by the Hunan Provincial Civil-Military Integration Industry Development under Grant [2018]23, in part by the Open Foundation of China Ship Scientific Research Center (CSSRC) under Grant 702SKL201705, and in part by the State Key Laboratory of Advanced Design Manufacturing for Vehicle Body under Grant 71865009.

ABSTRACT For an intelligent detection of bearing failure in rotating machinery, this paper proposed a fault diagnosis method based on a probability box (p-box) and support vector machine (SVM) with a genetic algorithm (GA) algorithm. Firstly, based on vibration signals of the bearing, the different p-boxes are obtained and fused using the evidence theory. Then, the different bearing p-boxes can be classified by adopting SVM model; the GA algorithm is considered to optimize key parameters of the SVM model, i.e., GA-SVM. Finally, experimental results show that total recognition rate of this method is better than that of the traditional feature extraction method, which demonstrates the effectiveness of the current method.

INDEX TERMS Probability box theory, feature vector, support vector machine, genetic algorithm, fault diagnosis.

I. INTRODUCTION
Rolled bearings are widely applied in rotating machinery, which running status will directly influent the function of the whole machine. Therefore, it is vitally significant to study the fault diagnosis technology of rolled bearings [1]. The no-cross research model of traditional fault diagnosis has been changed by information fusion, which made it a new hotspot [2], [3]. However, there are two inevitable problems in most methods: (1) The feature extraction of raw data is needed before information fusion, which leads to the loss of statistical information in raw data; (2) most researches are limited to the same class information due to the space-time registration problem.

Tang et al., [4] presented the technical reports of implementing the p-box in fault diagnosis. They reported that the p-box was effective and practical for use in fault diagnosis. The p-box theory provides a new research way to solve the above two problems because of its inclusive of raw data and its great advantage in dealing with uncertainties. The p-box modeling can take subjective and objective uncertainties into account. It can make up the defect of discarding rich statistical information due to the feature extraction process. It can reflect the overall uncertainty of the whole system. The raw data is first converted into the p-box and then the p-boxes are fused. Using the p-box as the space-time registration framework can solve the second problem [5].

SVM is a kind of pattern recognition method based on statistical learning theory and VC dimensional structure risk minimization principle. It can effectively deal with learning problems such as a small sample, high dimension, and non-linear in fault classification [6], [7]. A hybrid evolutionary algorithm featuring a competitive swarm optimizer combined with a local search was proposed by Long et al., and applied to the fault diagnosis of delta 3-D printers using attitude data [8], [9]. Aburomman and Reaz [10] compared several methods for creating a multiclass, SVM based classifier from a set of binary SVM classifiers. Wu et al., [11] suggested a novel approximation for the radius of the minimum enclosing ball in feature space, and then proposed a convex radius-margin-based SVM model for joint learning of feature transformation and the SVM classifier. Chauhan et al., [12] presented a review on evolution of linear support vector machine classification, its solvers, strategies to improve solvers, experimental results, current challenges and research directions. A novel self-training hierarchical prototype-based approach for semi-supervised classification was introduced by Gu [13]. Two machine learning models, i.e., a Gaussian process regression
and SVM model, for predicting the concrete breakout capacity of single anchors in shear were proposed by Olausi and Spyridis [14]. A hybrid SVM method for two-channel interleaved Vienna rectifier was developed by Wang et al., [15] to reduce the harmonics distortion and current ripple. One of the key factors affecting the classification accuracy of SVM is the selection of kernel width and penalty factor [16]. Therefore, the optimization of kernel width and penalty factor has an impact on SVM fault classification. The main methods to optimize kernel width and penalty factors are cross-validations, bilinear search, and grid search. The common disadvantages of these methods are that the amount of calculation is too large and the optimization efficiency is low. However, GA not only avoids the disadvantages but also has good global optimization ability.

Defersha and Movahed [17] developed linear programming assisted GA for solving a flexible jobshop lot streaming problem. A fuzzy clustering method based on multiobjective GA was proposed by Dong et al., [18]. According to the mechanism kinematic chain isomorphism identification criteria, a highly efficient hybrid GA model was proposed by Liu et al., [19] for isomorphism identification. The impact of the most relevant control parameters (as initial population, crossover rate, etc.) on the performances and results of a nature inspired algorithm of GA was analyzed by Piersanti and Orlandi [20]. A new encoding method of GA, adaptive degressive any number encoding, was proposed by Zhang and Liu [21]. To deal with the early maturity of the standard binary-coded genetic algorithm, Zou and Zhang [22] studied two methods that use different operators and observation concepts to solve early maturity problem of the standard binary-coded genetic algorithm, chooses to introduce the diversity and niche technology to improve GA. To overcome the disadvantages of traditional GA, which easily fall to local optima, Li and Lei [23] proposed a hybrid GA based on information entropy and game theory. Hence, GA algorithm is considered to optimize the parameters of SVM model.

Designing assembly relationship during operation always involves the analyses of many components and multidiscipline interaction, and seriously influences the reliability and work efficiency of complex machinery [24], [25]. “The angle of the load from the radial plane varies with the position of each rolling element in the bearing, as the ratio of local radial to axial load changes. Thus, each rolling element has a different effective rolling diameter and is trying to roll at a different speed, but the cage limits the deviation of the rolling elements from their mean position, thus causing some random slip. The resulting change in bearing frequencies is typical of the order of 1-2%, both as a deviation from the calculated value and also as a random variation around the mean frequency. This random slip, while small, does give a fundamental change in the character of the signal, and is the reason why the effects of some preprocessing techniques like high-pass, band-pass filtration, envelope detection (demodulation) and wavelet transform of the vibration signals, before feature extraction, are generally studied to remove uncertainty (unknown transmission path) of bearing signals in traditional rolling element bearing diagnostics. However, the uncertainty of bearing signals is the main characterization of bearing signals, because it always exists and changes from the beginning of health signals to the formation of fault signals.”

We proposed firstly a new way for the bearing fault diagnosis based on the p-box theory by collecting uncertainty of the time-domain signals of the bearing, because the p-box theory owns significant advantages in dealing with the uncertainty of bearing signals. The remainder of this study is organized as follows. Basic notions of the p-box are first given in Section II, followed by the p-box modeling method based on the vibration signals and the fusion of the p-boxes. In Section III, feature extraction methods of the p-box are introduced by employing a cumulative uncertainty measurement method. The parameters of SVM model are optimized by using GA in Section IV. Then, the rolling bearing fault diagnosis method based on the p-box and GA-SVM is proposed in Section V. The effectiveness of the proposed method is demonstrated in Section VI. Finally, conclusions are shown in Section VII.

II. PROBABILITY BOX MODELING AND FUSION

A. BASIC NOTIONS OF P-BOX

If the estimated value $\hat{x}$ of any variable $x$ is not a precise point estimate, its cumulative distribution function (CDF) cannot be expressed in a single curve. Define the range of $[\tilde{Y}(x), \bar{Y}(x)]$ as a p-box. $\tilde{Y}(x)$ is the lower bound of p-box. $\bar{Y}(x)$ is the upper bound of p-box. CDF of the variable $x$ can be limited into a p-box, they must satisfy [26], [27]

$$\tilde{Y}(x) < Y(x) < \bar{Y}(x)$$

DSS (The core of the p-box) is composed of a set of multiple focal elements, DSS can be expressed as following [28], [29]

$$\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_n, y_n), m_n\}$$

where $(x_i, y_i), m_i$ is the focal element and must meet the following conditions

$$x_i \leq y_i \quad (2)$$

$$\sum m_i = 1 \quad (3)$$

In Eqs. (2) and (3) $i = 1, 2, \ldots, n$. If $x_i = y_i$, ensure that $y_i \neq y_j$.

The p-box can be plotted by DSS, and the left boundary of p-box can be obtained by

$$\tilde{Y}(z) = \sum_{x_i \leq z} m_i \quad (4)$$

The right boundary of p-box can be obtained by

$$\bar{Y}(z) = \sum_{y_i < z} m_i \quad (5)$$

Fig. 1 shows the relationship between p-box and DSS. The range from the rough straight line is a p-box. The straight is the focal element of DSS, which is composed of two thin
H. Tang et al.: Fault Diagnosis of Rolling Bearing Based on p-box Theory and GA-SVM

lines. Its high value represents the mass value of the focal element. The horizontal coordinate shows the interval. The p-box is obtained by the sum of all DSS focal elements.

FIGURE 1. Discretization of p-box.

The convolution operation between p-boxes must be implemented by DSS. Fig. 2 is the transformation process of p-box to DSS, which is similar to the digital discrete process of an analog signal. Because the distance between every two fine lines is equal, the corresponding mass values of each focal element are equal to each other. Obviously, large discrete will be advantageous for the accurate fitting of the p-box.

FIGURE 2. Equal probability discretization of p-box.

B. P-BOX MODELING METHOD

Three kinds of modeling methods are considered, i.e., p-box modeling method under the condition of known probability distribution type (KPDT-PMM), fault-feature p-box modeling method (FF-PMM), raw-data-direct p-box modeling method (RDD-PMM).

KPDT-PMM method has the following basic ideas. Analyze raw data collected by the experiment. To determine whether they are subject to a certain type of probability distribution. Assume that the data follows a normal distribution to be used as an example. According to the sampling frequency, the data can be divided into several groups. Obtain the mean and variance of each group of data. Obtain the mean value of DSS $[\mu_{\min}, \mu_{\max}, 1]$, where $\mu_{\min}$ and $\mu_{\max}$ represent mean minimum and maximum values [29]. Obtain the variance of DSS $[\sigma_{\min}, \sigma_{\max}, 1]$. Uncertainty interval estimation of mean value and variance is brought into an expert estimation p-box modeling method. The interval estimation of mean value and variance are brought into an expert estimation p-box modeling method. Obtains the DSS according to the given discrete rate. The upper and lower bounds of p-box are plotted respectively by the upper bound and lower bound of the DSS.

If the fault signal is not subject to any type of the existing probability distribution types, the above method is not valid. To solve the problem, we can use the FF-PMM method. Its specific algorithm is described as follows. According to the sampling frequency, the data can be divided into several groups. Obtain dimensionless parameters for each set of data. Analyze dimensionless parameters to make sure whether or not it submits to a certain type of probability distribution. Assume that data to be analyzed follows a normal distribution to be used as an example. Determine the DSS of mean value and variance of dimensionless parameters. The intervals of mean value and variance are brought into an expert estimation p-box modeling method to obtain DSS. The upper and lower bounds of p-box are plotted respectively by the upper bound and lower bound of the DSS.

Another RDD-PMM method can be used to solve the above problem too. The specific algorithm is described as follows. Convert the raw data into an array of M rows and n columns based on the sampling frequency, where M and n represent sampling numbers and sampling frequency. Cut redundant data. Obtain a new array according to the order from small to large. Find the minimum and maximum values in each column from M subsample data. Obtain the minimum row vector and the maximum values row vector respectively. The upper and lower bounds of p-box are plotted respectively by the upper bound and lower bound of the DSS.

The p-box boundary is the narrowest by the RDD-PMM method, that is, the compactness of the obtained p-box is the highest. Another advantage of this method is that it does not need to verify whether or not the data is subject to any probability distribution type, so the RDD-PMM method is adopted in this paper.

C. FUSION OF P-BOXES

The measured fault signals can be transformed into p-boxes by the above p-box modeling methods. There is information redundancy in the fault signals obtained from different locations, but there is a piece of complementary information too, which requires the fusion of probability boxes. Whether the collection is similar information or heterogeneous information, all the data ultimately will be transformed into p-boxes which provides a unified framework for the fusion of heterogeneous information fusion. DSS and DS evidence theory has maintained a high degree of consistency and complementarity in the main points of view, so that we use the evidence theory to fuse p-boxes [4].

The method of using evidence theory to fuse p-boxes as follows. Input p-boxes or discretized DSS to be fused, and
then input required discrete rate or discrete points. Set default discrete points. Calculate with the default points, if the discrete points are not raised by the user. Calculate with the discrete points given by the user, if the discrete points are presented by the user. Normalize the input DSS. Use DS evidence synthesis rules to fuse the average discrete DSS pair by pair. A new fused DSS is obtained through DS evidence synthesis rules. Standardize the fused DSS. The upper and lower bounds of p-box are plotted respectively by the upper bound and lower bound of the DSS.

III. FEATURE EXTRACTION OF P-BOX
The cumulative uncertainty measurement method can obtain the whole information about p-boxes from different angles. The form of expression is a single scalar or interval. But each scalar or interval can be understood as a feature of p-box [4]. So, the feature extraction of p-box is realized by the method of cumulative uncertainty measurement in this paper. Several feature extraction methods are introduced in this paper.

Method 1, Obtain cumulative width to use basic probability distribution for all focal element interval of weight, which could be expressed as [4]

$$y = \sum_{i=1}^{r} (m_i \times |x_i - x_j|)$$

(6)

Method 2, Obtain cumulative log width to use basic probability distribution for all focal element interval of weight, which could be expressed as [4]

$$y = \sum_{i=1}^{r} (m_i \times \log_2 |x_i - x_j|)$$

(7)

Method 3, Obtain 1 as cumulative logarithmic width for Radix to use basic probability distribution for all focal element interval of weight, which could be expressed as [4]

$$y = \sum_{i=1}^{r} (m_i \times (1 - \log_2 |x_i - x_j|))$$

(8)

Method 4, Obtain p-box lower and upper bound to use basic probability distribution for cumulative interval boundary value of the weight, which could be expressed as [4]

$$y = \left[ \sum_{i=1}^{r} m_i x_{i1} \right] \left[ \sum_{i=1}^{r} m_i x_{i2} \right]$$

(9)

Method 5, Obtain accumulated boundary value of p-box lower bound and upper bound under the conditional value of cumulative distribution function. Assume that constitute p-box’s DSS = \{[(x_{11}, x_{12}), m_1], [(x_{21}, x_{22}), m_2], \ldots, [(x_{m1}, x_{m2}), m_m] \}, if cumulative distribution function value \( \alpha \) could be expressed as [4]

$$\sum_{i=1}^{k} m_i \leq \alpha \leq \sum_{i=1}^{k+1} m_i$$

(10)

Accumulated boundary value for the lower bound and upper bound of p-box could be expressed as [4]

$$\left[ \sum_{i=1}^{k} x_{i1} \right] \left[ \sum_{i=1}^{k} x_{i2} \right]$$

(11)

Method 6, Obtain the contradiction interval statistics of p-box. Assume that DSS’s form is \{[(x_{11}, x_{12}), m_1], [(x_{21}, x_{22}), m_2], \ldots, [(x_{m1}, x_{m2}), m_m] \}. Cumulative uncertainty measurement results could be expressed as [4]

$$y = - \sum_{i=1}^{n} m_i \times \log_2 (c_2 - c_1')$$

(12)

where \( c_1' \) and \( c_2' \) represent mean probability statistics of the left and right boundaries of DSS.

IV. SVM OPTIMIZED BY GA
The choice of kernel function has a great influence on the classification accuracy of SVM. At present, the commonly used kernel functions are Q order polynomial kernel function, radial basis function (RBF) and Sigmoid kernel function, where the RBF kernel function only needs to determine the width of a kernel function, and the convergence region is wider; it can get desired effect for the classification of the low dimension, high dimension, small samples, and large sample have the desired effect. Therefore, the RBF kernel function is used in this paper. the width \( \sigma^2 \) and the penalty factor \( C \) of the RBF kernel function have a great influence on the SVM classification accuracy, which reflect in the following aspects. \( \sigma^2 \) mainly affects the complexity of the sample data for distribution in high dimensional feature space. \( C \) adjusts the confidence range of the learning machine and the proportion of experience risk in determining the feature space. Therefore, to ensure that the SVM has a high classification accuracy, it is necessary to select the appropriate \( \sigma^2 \) and \( C \). The procedure of using GA to optimize \( \sigma^2 \) and \( C \) can be described by the flow diagram as shown in Fig. 3:

FIGURE 3. The flow diagram of GA-SVM parameters optimization.
Step-1: The search intervals of $\sigma^2$ and $C$ are used as the initial population; the relevant parameters of GA are determined, including population size (generally ranging from 10 to 200), evolution algebra, selection probability, crossover probability, and mutation probability.

Step-2: A group of codes is randomly selected from the initial population as parents, and the fitness value of each individual is calculated; since the actual fault cases and the correct recognition rate of classification results fully reflect the classification accuracy of SVM, the correct classification rate is regarded as the fitness function of GA, and is defined as follows:

$$f(x) = \frac{x_1}{x_2}$$  \hspace{1cm} (13)

where $x_1$ is test samples correctly classified by SVM, $x_2$ is the number of actual test samples.

Step-3: Genetic operations are performed on the parents, including selection, crossover, and mutation; the selection method of roulette is used to carry out the genetic operation on the parents; with a certain selection probability (generally 0.08), individuals with higher fitness value are selected and inherited into the next generation population. The crossover probability ranges from 0.4 to 0.99; the mutation probability ranges from 0.0001 to 0.1.

Step-4: The procedure is stopped if the population’s maximum fitness converges or the evolution algebra reaches the set maximum algebra, and the individual with the highest fitness in the last generation is taken as the optimal solution; then the optimized $\sigma^2$ and $C$ are obtained; otherwise, the best individual in the parent generation is used to replace the worst individual in the generation population, and the generation population is converted into the parent, then steps-2 to 3 are repeated.

V. FAULT DIAGNOSIS BASED ON P-BOX AND GA-SVM

Fig. 4 shows a flow chart of fault diagnosis based on p-box and GA-SVM. Firstly, collect rolling bearing vibration signals of different faults on the mechanical fault simulation experiment platform. Secondly, adopt different p-box modeling methods to deal with the obtained signals. The raw data are firstly converted into p-boxes and then p-boxes are fused. Adopt feature extraction of p-box to conduct feature extractions for fused p-boxes. Obtain p-box feature vector of the N group. Train M groups. Test (N-M)groups. input M groups training feature vectors to the SVM. Use GA algorithm to train SVM parameters and optimizes SVM parameters. Obtain the trained SVM which can reduce the classification error. Input (N-M) groups test feature vector into the trained SVM. Finally, the status of the fault bearings is identified.

VI. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL DATA ACQUISITION

The fault signal of rolling bearing is used as the research object. The experimental platform is the QPZZ-II rotating machinery comprehensive fault simulation test bench. Signal acquisition is the PXI-1042Q NI high-performance sound vibration testing system. Software is the LabVIEW data acquisition software. Sensors are American M603C01 ICP PCB acceleration sensors. Fig. 5 shows the experimental apparatus. Fig. 6 shows the ICP sensor arrangement point.

Experimental bearing is NU205 cylindrical roller bearing. The line cutting grooves (length 15mm, width 0.5mm, deep 0.5mm) are cut out respectively in the outer ring and the inner ring, which are used to simulate the outer ring, inner ring damage fault respectively. The line cutting groove (width 0.5mm, deep 0.5mm) is cut out on one of the rolling elements, which is used to simulate the rolling element local damage. Fig. 7 shows fault bearing entities. Motor speed is 800r/min.

Rolling bearing fault characteristic frequency is obtained by theoretical formula, as shown in Table 1.
The sampling frequency is set to 1024Hz. Fig. 8 shows the time domain signals of the V direction.

Fig. 9 shows the time domain waveforms and the power spectrums of the experimental bearing acceleration signals with sampling frequency 10240Hz.

**B. LOW FREQUENCY’S EFFECT ON P-BOX**

Using the time domain signals with 1024 Hz, the p-boxes of V direction are established as shown in Fig. 10, where UB and LB denote the upper and lower bounds, respectively; N is normal bearing, IR inner ring fault, OR outer ring fault, and RE rolling element fault.

The overlaps between the p-boxes are serious, and the p-boxes cannot be separated from each other obviously in Fig. 10. Therefore, for high-frequency bearing signals, there is a distortion problem between the signals if the signals do not meet the sampling law. The distortion problem is not suitable for spectrum analysis but also for bearing fault diagnosis based on the p-box. Hence, using the p-box for the bearing fault diagnosis should also satisfy the sampling law. Then, the sampling frequency is considered as 10240 Hz as follows.

**C. RESULTS OF P-BOXES MODELING AND FUSION**

1) P-BOXES MODELING OF EXPERIMENTAL DATA

As the KPDT-PMM method for fault signal needs to obey someone type of probability distribution, this method does not be used as the experimental instructions in this paper. The
FF-PMM method and RDD-PMM method are used to p-box modeling of the data. Based on the vertical and horizontal signals with 10240 Hz, the different p-boxes are established by using FF-PMM method, as shown in Fig. 11.

![Figure 11](image1.png)

**FIGURE 11.** Experimental results by FF-PMM method along with the V and H directions.

A comparison of the p-boxes between the V and H directions can be found that the overlap areas are more or less among the different p-boxes (see projection in Fig. 11). Fusing the p-boxes between the H and V directions may improve the overlapping phenomenon.

Using RDD-PMM method, the p-boxes are obtained along with the V and H directions as shown in Fig. 12; to better present each p-box, we cropped each p-box data to the interval values between $-2$ and $2$.

![Figure 12](image2.png)

**FIGURE 12.** Experimental result by RDD-PMM method; (a) V direction; (b) H direction.

Comparison of the p-boxes between the V and H directions can be found that the p-box of N and the p-box of IR are separate in V direction but overlap in H direction; the p-box of OR and the p-box of RE are separate in V and H directions.

Comparing the p-boxes by FF-PMM method (Fig. 11) and the p-boxes by RDD-PMM method (Fig. 12) can be found that the compactness of the latter is higher than that of the former. This is because the process of RDD-PMM method does not exist in the process of feature extraction, which means there is no loss of statistical information for raw data.

It can be seen that there are some similarities and differences for the p-boxes of four conditions in Figs. 11 and 12, such as the compactness of the p-boxes, the overlapping area of the p-boxes, etc. It is demonstrated that the p-boxes in different directions are redundant and complementary to each other. It is necessary to study the fusion problem of p-boxes to realize the complementary utilization of p-boxes information in different directions.

2) P-BOXES FUSION

Fig. 13 shows the fusion results of the p-boxes between the V and H directions from Fig. 11. Similarly, the fusion results of the p-boxes between the V and H directions from Fig. 12 are shown in Fig. 14.

![Figure 13](image3.png)

**FIGURE 13.** Fusion the p-boxes of FF-PMM between the V and H directions.

The overlap of the p-boxes is improved after the fusion in Fig. 13, there is still a partial overlap at the bottom. Based on RDD-PMM method, the p-boxes are better separated after fusion in Fig. 14. It demonstrates that the fusion of the p-boxes is effective to realize the complementary utilization of p-boxes information in different directions.

D. SVM OPTIMIZED BY GA

1) FEATURE EXTRACTION OF P-BOXES

Use the feature extraction of p-box mentioned in Section III, the results of feature extraction for the 4 probability boxes in Fig. 17 are shown in Table 2. 150 groups of p-box feature vectors for each type of bearing are obtained respectively according to the experimental data. 100 groups are used for training, and the rest 50 groups are used for testing. The total sample number is 600 groups for 4 types of bearings. Training
sample set T1 accounts for 2/3 of the whole samples, and the test sample set T2 accounts for the rest 1/3. Table 2 shows the feature vector of a group of p-box. 8 features are obtained from 6 different cumulative uncertainty statistics methods.

2) OPTIMIZED SVM PARAMETERS RESULTS BY GA ALGORITHM
The optimal interval of $\sigma^2$ is set as $[0, 0.500]$ and $C$ as $[1, 2000]$. The parameter settings of the GA algorithm are shown in Table 3.

The error correction coding (ECC) method is considered to solve the multi classification SVM in this case. The patterns are normal bearing state, fault bearing inner ring state, outer ring fault bearing state and rolling element bearing state. Table 4 shows SVM with ECC code.

SVM1~SVM4 is one group, where the SVM1 regards the normal state as 1, and others as 0. SVM5~SVM8 is another group, where the SVM5 regards the normal state and rolling element fault as 1, and others as 0. According to Table 4, we can judge that it is the normal mode if the binary code is 10001001. As the SVM classification is not possible to achieve the correct recognition rate of 100%, the code 10001001 may not be obtained for the normal bearing signal. Hence, 8-bit binary codes obtained from eight SVM are compared with four ideal binary codes such as “10001001” and “Hamming distance minimization method”. The one with the minimum Hamming distance is regarded as the target. The method owns certain fault tolerance.

The correct identification rate is used as the optimization objective function. The optimization objects are the width of the kernel function $\sigma^2$ and the penalty coefficient $C$. The optimized results are shown in Table 5.

E. CLASSIFICATION RESULTS AND CONTRAST
The minimum Hamming distance method is used to compare the 8 SVM classification results, and the results obtained are shown in Table 6.

As can be seen from Table 6, although there are a small number of samples were identified errors in the rolling element fault, the method can fully identify the normal state of the bearing and the fault conditions of the outer ring and inner ring. Additionally, using SVM parameters without optimization, the experimental results showed that the total correct classification rate is 90.7%. Hence, the optimized SVM parameters are used in the following.

Table 7 shows the pattern recognition results obtained by different p-box modeling methods.

As shown in Table 6, the total correct recognition rate of RDD-PMM method is higher than that of FF-PMM method.

To demonstrate the effectiveness of the current method, the traditional statistical features are considered as the 8 feature vectors, i.e., waveform index, pulse index, peak value index and margin index, skewness index and kurtosis index, skewness index, and kurtosis index; then, the SVM is used as the pattern recognition tool. The sample data are still 600 groups. 2/3 are for testing and the remaining 1/3 are for identification. The results obtained are shown in Table 8.

The result shows that the total recognition rate of this method is better than that of the traditional feature extraction method.

“Composite multiscale fuzzy entropy is an effective method to analyze the complexity of time series in bearing fault diagnosis. It can not only reflect the complexity characteristics of time series from multiple scales, but also has the advantages of short data and good robustness. Fig. 15 presents the results of the composite multiscale fuzzy entropy for each bearing condition based on the current data, where the value of largest scale is 20, embedding dimension 2, gradient of exponent function 2, and similarity tolerance 0.15SD (SD denotes standard deviation of raw bearing data), respectively [33].”

In Fig. 15, “the fuzzy entropy of H is larger on the relatively large scale, and changes gently with the increase of the scale values; the curve of composite multiscale fuzzy entropy for other bearing conditions shows the obvious decreasing trend. In this contrastive study, the steps used in this method...
can be described as the following: Firstly, Total 600 samples were used in this study; the feature set was obtained by calculating the values of composite multiscale fuzzy entropy for each bearing condition. Then, 2/3 were for testing and the remaining 1/3 were for identification. Finally, the correct classification of faults can be given by the SVM model; the

| Rolling bearing state | Method 1 | Method 2 | Method 3 | Method 4 | Method 5 | Method 6 |
|-----------------------|----------|----------|----------|----------|----------|----------|
| Normal state          | 0.012    | -6.597   | 0.017    | [-0.053, -0.042] | [0.335, 0.346] | 5.490    |
| Inner ring fault      | 0.058    | -5.357   | 0.051    | [-0.124, -0.066] | [0.630, 0.665] | 5.458    |
| Outer ring fault      | 0.0427   | -5.088   | 0.0578   | [-0.064, -0.021] | [0.907, 0.910] | 5.299    |
| Rolling element fault | 0.023    | -6.059   | 0.031    | [-0.045, -0.020] | [0.491, 0.517] | 5.417    |

| Population size | The number of generation populations | Selection Method | Probability | Crossover Method | Probability | Mutation Method | Probability |
|-----------------|--------------------------------------|------------------|-------------|------------------|-------------|-----------------|-------------|
| 100             | 100                                  | Roulette         | 0.08        | Single-point     | 0.8         | Basic bit       | 0.1         |

| Rolling bearing state | SVM1 | SVM2 | SVM3 | SVM4 | SVM5 | SVM6 | SVM7 | SVM8 |
|-----------------------|------|------|------|------|------|------|------|------|
| Normal state          | 1    | 0    | 0    | 0    | 1    | 0    | 0    | 1    |
| Inner ring fault      | 0    | 1    | 0    | 0    | 0    | 0    | 1    | 0    |
| Outer ring fault      | 0    | 0    | 1    | 0    | 1    | 1    | 0    | 0    |
| Rolling element fault | 0    | 0    | 0    | 1    | 1    | 1    | 0    | 0    |

| SVM | SVM1 | SVM2 | SVM3 | SVM4 | SVM5 | SVM6 | SVM7 | SVM8 |
|-----|------|------|------|------|------|------|------|------|
| $\sigma^2$ | 23.42 | 20.01 | 26.31 | 92.53 | 7.35 | 84.19 | 83.71 | 51.53 |
| C   | 0.92 | 9.61 | 0.03 | 30.3 | 0.31 | 1.02 | 1.37 | 3.23 |

| Conditions          | Test sample number | Number of correct classifications | Number of error classification | Correct classification rate | Total correct classification rate |
|---------------------|--------------------|-----------------------------------|-------------------------------|----------------------------|----------------------------------|
| Normal state        | 50                 | 50                                | 0                             | 100%                        | 99.5%                            |
| Inner ring fault    | 50                 | 50                                | 0                             | 100%                        | 99.5%                            |
| Outer ring fault    | 50                 | 50                                | 0                             | 100%                        | 99.5%                            |
| Rolling element fault | 50                | 49                                | 1                             | 98%                         | 98%                              |

| P-box modeling method | Total sample number | Training sample number | Testing sample number | Total correct recognition rate |
|-----------------------|---------------------|------------------------|-----------------------|-------------------------------|
| FF-PMM method         | 600                 | 400                    | 200                   | 93.33%                        |
| RDD-PMM method        | 600                 | 400                    | 200                   | 99.5%                         |

| Method               | Total sample number | Training sample number | Testing sample number | Total correct recognition rate |
|----------------------|---------------------|------------------------|-----------------------|-------------------------------|
| New method           | 600                 | 400                    | 200                   | 99.5%                         |
| Traditional method   | 600                 | 400                    | 200                   | 86.15%                        |
experimental results showed that the total correct recognition rate is 83.70%, which is smaller than the correct recognition rates 93.33% and 99.5% from the proposed methods in this paper. It is because that the method of composite multiscale fuzzy entropy requires additional empirical effort in the bearing fault diagnosis [33]."

VII. CONCLUSION

The fault modeling method of the rolling bearing based on the p-box has been proposed. The modeling method makes up the defect of the traditional feature extraction method which discards the abundant probability statistics information, and the following conclusions can be obtained.

1. The influence of the sampling frequency on the overlapping degree of the p-box has been discussed. The experimental results show that the data of fault diagnosis based on the p-box must also satisfy the sampling law.

2. Employing the evidence theory, the p-boxes have been fused to reduce the overlapping degree; the p-boxes based on the RDD-PMM method have been better separated after fusion than the p-boxes established by the FF-PMM method.

3. The parameters of the SVM model have been optimized to improve the total correct classification rate; using SVM parameters with optimization, the experimental results showed that the total correct classification rate is 99.5%. However, the total correct classification rate is 90.7% under SVM parameters without optimization. Hence, the optimized SVM parameters are effective. Additionally, the experimental results showed that the total correct recognition rates of the traditional statistical feature extraction method and composite multiscale fuzzy entropy are 86.15% and 83.70%, respectively, which is smaller than the correct recognition rates from the proposed methods in this paper; then, the effectiveness of the current method has been demonstrated.

REFERENCES

[1] Q. Gao, H. Tang, J. Xiang, Y. Zhong, S. Ye, and J. Pang, “A Walsh transform-based teager energy operator demodulation method to detect faults in axial piston pumps,” Measurement, vol. 134, pp. 293–306, Feb. 2019, doi: 10.1016/j.measurement.2018.10.085.

[2] L. Wang and J. Xiang, “A two-stage method using spline-kernelled chirplet transform and angle synchronous averaging to detect faults at variable speed,” IEEE Access, vol. 7, pp. 22471–22485, 2019, doi: 10.1109/access.2019.2898884.

[3] S. Wang and J. Xiang, “A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps,” Soft Comput., vol. 24, no. 4, pp. 2983–2997, Feb. 2020, doi: 10.1007/s00500-019-04076-2.

[4] H. Tang, Y. Du, and H.-L. Dai, “Rolling element bearing diagnosis based on probability box theory,” Appl. Math. Model., vol. 80, pp. 944–960, Apr. 2020, doi: 10.1016/j.apm.2019.10.068.

[5] K. Worden, W. J. Staszewski, and J. J. Hensman, “Natural computing for mechanical systems research: A tutorial overview,” Mech. Syst. Signal Process., vol. 25, no. 1, pp. 111–114, Jan. 2011, doi: 10.1016/j.ymssp.2010.07.013.

[6] A. Cury and C. Crémona, “Pattern recognition of structural behaviors based on learning algorithms and symbolic data concepts,” Struct. Control Health. Syst., vol. 19, pp. 161–186, Oct. 2012, doi: 10.1002/stc.412.

[7] H. Witten and F. Eibe, Data Mining: Practical Machine Learning Tools and Techniques, 2nd ed. Berlin, Germany: Morgan Kaufmann, 2005, pp. 26–27.

[8] J. Long, Z. Sun, C. Li, Y. Hong, Y. Bai, and S. Zhang, “A novel sparse echo autoencoder network for data-driven fault diagnosis of delta-3-D printers,” IEEE Trans. Industr. Meas., vol. 69, no. 3, pp. 683–692, Mar. 2020, doi: 10.1109/tim.2019.2905752.

[9] J. Long, S. Zhang, and C. Li, “Evolving deep echo state networks for intelligent fault diagnosis,” IEEE Trans. Ind. Informat., vol. 16, no. 7, pp. 4928–4937, Jul. 2020, doi: 10.1109/tii.2019.2938884.

[10] A. A. Aburomman and M. B. Ione Reaz, “A novel weighted support vector machines multiclass classifier based on differential evolution for intrusion detection systems,” Inf. Sci., vol. 414, pp. 225–246, Nov. 2017, doi: 10.1016/j.ins.2017.08.007.

[11] X. Wu, W. Zuo, L. Lin, W. Jia, and D. Zhang, “F-SVM: Combination of feature transformation and SVM learning via convex relaxation,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 11, pp. 5185–5199, Nov. 2018, doi: 10.1109/tnnls.2018.2791507.

[12] V. K. Chauhan, K. Dahiya, and A. Sharma, “Problem formulations and solvers in linear SVM: A review,” Artif. Intell. Rev., vol. 52, no. 2, pp. 803–855, Aug. 2019, doi: 10.1007/s10462-018-9614-6.

[13] X. Gu, “A self-training hierarchical prototype-based approach for semi-supervised classification,” Inf. Sci., vol. 535, pp. 204–224, Oct. 2020, doi: 10.1016/j.ins.2020.05.018.

[14] O. B. Olalusi and P. Spyridis, “Machine learning-based models for the concrete breakout capacity prediction of single anchors in shear,” Adv. Eng. Sofw., vol. 147, Sep. 2020, Art. no. 102832, doi: 10.1016/j.advengsoft.2020.102832.

[15] T. Wang, C. Chen, P. Liu, T. Liu, Z. Chao, and S. Duan, “A Hybrid Space-Vector Modulation Method for Harmonics and Current Ripple Reduction of Interleaved Vienna Rectifier,” IEEE Trans. Ind. Electron., vol. 67, no. 10, pp. 8088–8099, Dec. 2020, doi: 10.1109/tie.2019.2947810.

[16] C. M. Wen, S. L. Hung, C. S. Huang, and J. C. Jan, “Unsupervised fuzzy neural networks for damage detection of structures,” Struct. Control Hlth. Meas., vol. 14, no. 1, pp. 144–161, 2007, doi: 10.1002/stc.116.

[17] F. M. Defersha and A. Movahed, “Linear programming assisted (not embedded) genetic algorithm for flexible jobshop scheduling with lot streaming,” Comput. Ind. Eng., vol. 117, pp. 319–335, Mar. 2018, doi: 10.1016/j.cie.2018.02.010.

[18] Z. Dong, H. Jia, and M. Liu, “An adaptive multobjective genetic algorithm with fuzzy c-Means for automatic data clustering,” Math. Problems Eng., vol. 2018, pp. 1–13, Apr. 2018, doi: 10.1155/2018/6123874.

[19] H. Liu, S. Shi, P. Yang, and J. Yan, “An improved genetic algorithm approach on mechanism kinematic structure enumeration with intelligent manufacturing,” J. Intell. Robot. Syst., vol. 89, nos. 3–4, pp. 343–350, Mar. 2018, doi: 10.1007/s10846-017-0564-z.

[20] S. Piersanti and A. Orlandi, “Notice of retraction: Impact of genetic algorithm control parameters on chip PDN decoupling capacitors placement,” IEEE Trans. Electromagn. Comput., vol. 60, no. 4, pp. 945–951, Aug. 2018, doi: 10.1109/temc.2017.2773220.

[21] Y. Zhang and M. Liu, “An improved genetic algorithm encoded by adaptive degreessive ary number,” Soft Comput., vol. 22, no. 20, pp. 6861–6875, Oct. 2018, doi: 10.1007/s00500-017-2981-0.

[22] G. X. Zou and F. Z. Zhang, “An improved genetic algorithm for site selection problem,” Comput. Syst. Sci. Eng., vol. 40, no. 4, pp. 712–722, 2018, doi: 10.3969/j.issn.1007-130X.2018.04.019.

[23] L. Jiacheng and L. Lei, “A hybrid genetic algorithm based on information entropy and game theory,” IEEE Access, vol. 8, pp. 36602–36611, 2020, doi: 10.1109/access.2020.2971060.

[24] C.-W. Fei, H. Li, H.-T. Liu, C. Lu, B. Keshetgar, and L.-Q. An, “Multilevel nested reliability-based design optimization with hybrid intelligent regression for operating assembly relationship,” Aerosp. Sci. Technol., vol. 103, Aug. 2020, Art. no. 105906, doi: 10.1016/j.ast.2020.105906.

[25] C. Lu, Y.-W. Feng, C.-W. Fei, and S.-Q. Bu, “Improved decomposed-coordinated kriging modeling strategy for dynamic probabilistic analysis of multicomponent structures,” IEEE Trans. Rel., vol. 69, no. 2, pp. 440–457, Jun. 2020, doi: 10.1109/tr.2019.2954379.

[26] S. Ferson, Constructing Probability Boxes and Dempster-Shafer Structures. Livermore, CA, USA: Sandia National Laboratories, 2003.

[27] G. Dempster, “Upper and lower probabilities induced by a multivalued mapping,” Ann. Math. Statist., vol. 38, pp. 325–339, Oct. 1967.
[28] G. Shafer, “The combination of evidence,” Int. J. Intell. Syst., vol. 1, pp. 155–179, Dec. 1986.
[29] F. Yang, H. Wei, and P. Feng, “A hierarchical dempster-shafer evidence combination framework for urban area land cover classification,” Measurement, vol. 151, Feb. 2020, Art. no. 105916, doi: 10.1016/j.measurement.2018.09.058.
[30] M. Yar and J. K. Hammon, “Parameter estimation for hysteretic systems,” J. Sound. Vib., vol. 117, pp. 161–172, Aug. 1987.
[31] G. Saleheh and M. R. Maheri, “Multi-stage approach for structural damage detection problem using basis pursuit and particle swarm optimization,” J. Sound. Vib., vol. 384, pp. 210–226, Dec. 2016, doi: 10.1016/j.jsv.2016.08.024.
[32] S. Rezaeian-Marjani, S. Galvani, V. Talavat, and M. Farhadi-Kangarlu, “Optimal allocation of D-STATCOM in distribution networks including correlated renewable energy sources,” Int. J. Elec. Power, vol. 122, pp. 3445–3449, Oct. 2020, doi: 10.1016/j.ijepes.2020.106178.
[33] J. Zheng, H. Pan, and J. Cheng, “Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines,” Mech. Syst. Signal Process., vol. 85, pp. 746–759, Feb. 2017, doi: 10.1016/j.ymssp.2016.09.010.

ZHENGXING YUAN is currently pursuing the B.E. degree with the Kunming University of Science and Technology, Kunming, China.

He is also a Senior Manager of the Machinery Association. His research interests include engineering structure optimization design and reliability, structural mechanics, and NVH.

HONGTANG DAI received the Ph.D. degree from Shanghai Jiaotong University, Shanghai, China, in 2005.

He is currently a Professor, a Ph.D. Supervisor, and the Vice President with the School of Mechanical and Vehicle Engineering, Hunan University. He is the Technical Expert of the China New Energy Vehicle Battery Echelon Utilization Alliance. He has published more than 110 SCI papers. His research interests include intelligent materials and structural mechanics, NVH, and engineering structure optimization design and reliability.

Dr. Dai is a Reviewer of more than 20 international journals, such as International Journal of Solids and Structures, Applied Mechanics, European Mechanics, Composite Structures, Thermal Fabric, and International Journal of Pressure Vessels and Piping.

HONG TANG received the B.E. degree in mechanical engineering from the Kunming University of Science and Technology, China. He is currently pursuing the Ph.D. degree with the College of Mechanical and Vehicle Engineering, Hunan University, China.

He is currently a member of the Machinery Association. His research interests include engineering structure optimization design and reliability, intelligent materials and structural mechanics, and NVH.

YI DU received the Ph.D. degree from the Kunming University of Science and Technology, Kunming, China, in 2012.

He is currently a Professor and the Deputy Director of the Mechanical Department, Kunming University of Science and Technology. His research interests include mechanical fault diagnosis, information fusion, and uncertainty analysis.

* * *