Weighted K-NN Classification Method of Bearings Fault Diagnosis With Multi-Dimensional Sensitive Features

QINGFENG WANG1,2, SHUAI WANG1,2, BINGKUN WEI1,2, WENWU CHEN3, AND YUFEI ZHANG4
1Beijing Key Laboratory of Health Monitoring and Self-Recovery of High-End Machinery Equipment, Beijing University of Chemical Technology, Beijing 100029, China
2Diagnosis and Self-Recovery Engineering Research Center, Beijing University of Chemical Technology, Beijing 100029, China
3State Key Laboratory of Safety and Control for Chemicals, China Petroleum and Chemical Corporation Research Institute of Safety Engineering, Qingdao 266071, China
4SINOMACH Academy of Science and Technology Company Ltd., Beijing 100050, China
Corresponding authors: Qingfeng Wang (wangqf2422@buct.edu.cn) and Yufei Zhang (zhangyufei@synomast.com.cn)

This work was supported by the China Petroleum and Chemical Corporation Ministry of Science and Technology under Grant 320059 and Grant 319022-1.

ABSTRACT Research on the intelligent fault diagnosis method of rolling bearing based on laboratory data has made some achievements. However, due to the change of working conditions and the lack of historical data of the same equipment in the actual diagnosis, some methods mostly have problems such as poor generalization. Model training and verification data are insufficient, and engineering practice still lacks effective intelligent fault diagnosis methods. In this paper, we propose a weighted k-nearest neighbor (WKNN) fault diagnosis model based on multi-dimensional sensitive features, and propose a fault diagnosis method for rolling bearings that adapts to different equipment and different operating conditions. First, we extract time domain, frequency domain, and entropy features of the original signal to form the raw signal feature set. Then, the iterative ReliefF feature screening method is used to evaluate the joint feature set, calculate the weight of each feature, remove insensitive and redundant features, and obtain a high-dimensional sensitive feature set. Finally, the WKNN classification model is used to identify bearing failure modes. The fault diagnosis model was trained using rolling bearing data from the Case Western Reserve University (CWRU), while laboratory data from the Intelligent Maintenance System (IMS), the Society of Mechanical Failure Prevention Technology (MFPT) and the engineering case data were used for testing. The results show that the model proposed in this paper has high fault diagnosis accuracy and can accurately determine the fault type after early warning. Compared with other comparison methods, the fault recognition accuracy rate is higher. And it is suitable for different working conditions and different equipment, and has good engineering application value.

INDEX TERMS Different working conditions, fault diagnosis, multi-dimensional sensitive features, ReliefF, WKNN.

I. INTRODUCTION

With the development of modern industrial Internet, big data analysis, and in particular artificial intelligence, has greatly promoted the development of intelligent fault diagnosis and made great strides towards predictive maintenance (PDM). Rotating machinery is the most common type of mechanical equipment and plays an important role in industrial production. According to statistics, rolling bearing faults account for 30% to 40% of the common rotating equipment faults [1]. Except for a few sudden failures, most failures undergo a gradual process. If incipient fault diagnosis can be achieved and appropriate predictive maintenance measures can be taken in time, planned shutdowns and replacement of parts for high-end equipment can be arranged well before a failure occurs. This has important theoretical significance and
engineering value for realizing the safe and stable operation of mechanical equipment [2].

Mechanism model methods, data-driven methods, and combinations of these approaches are widely used in fault detection. In industrial production, the number of equipment is large and their structure is complicated. Once the equipment fails, serious economic losses or catastrophic failure consequences may be incurred, and there is an urgent need for timely equipment fault diagnosis. However, the formulation of an accurate failure mechanism model is time-consuming, laborious and uneconomical. The corresponding effort required is considerable, so the utilization of this method in practical applications is limited [3]. At the same time, with the rapid development of data mining, signal processing and machine learning [4], data-driven fault diagnosis methods are showing more and more applicability. So the combination of failure mechanism modeling and data-driven methods has become a development trend.

Data-driven fault diagnosis processes generally include four steps: raw data acquisition, data preprocessing, feature extraction, and fault pattern recognition. When the rolling bearing fails, which are key components of rotating equipment, a series of vibration and shock signals will be generated. About 50% of faults are characterized using vibration [5], so collecting vibration signals can directly reflect the operating status of mechanical equipment. Feature extraction is essential for fault diagnosis, and the sensitivity of features determines the accuracy of the final diagnosis. With the continuous development of feature extraction methods, the number of dimensions or categories of extracted feature vectors continues to increase, and irrelevant and redundant feature vectors in high-dimensional feature sets have an increasing trend, which is likely to affect the accuracy of fault diagnosis [6]. In practical applications, due to the different bearing models and operating conditions, different types of features also have different sensitivity [7]. The study the extraction, screening and feature fusion methods of fault characteristic signals to reduce redundant features that have little contribution to fault diagnosis and obtain low-dimensional, high-sensitivity fault feature subsets is of great significance for improving the accuracy of fault diagnosis results and reducing the complexity of corresponding algorithms [8]. Liu et al. [9] used 3-layer wavelet packet decomposition to extract 56-dimensional features, and used kernel principal component analysis (KPCA) to reduce feature dimension and extract the main features, as well as a support vector machine (SVM) for fault identification. Lei et al. [10] proposed a bearing fault classification method, which extracted 11 time-frequency domain features, selected features according to local and non-local preserved projections, and used the k-nearest neighbor (KNN) algorithm to classify bearing health. Tuncer et al. [11] used X-ray images to detect Covid-19. In the feature selection stage, they adopted a novel feature selection method based on iterative ReliefF. Yang et al. [12] proposed a fault classification method based on principal component analysis (PCA) and the KNN algorithm for compressor faults. The method used offline data to construct a dimension reduction model through PCA, and the KNN algorithm to train the classification model for anomaly detection and fault identification in real time. In order to solve the shortcomings of single-feature life prediction algorithms for rolling bearings, Wang et al. [13] proposed a comprehensive rolling bearings life prediction method based on multi-dimensional feature fusion. This method extracts nine time-domain performance degradation features of rolling bearing vibration signals, and uses PCA to fuse multi-dimensional features and characterize the working state of rolling bearings. Zhou et al. [5] first extracted the time-domain features of diagnostic signals. Based on this, they proposed a combination of k-means clustering and k-nearest neighbor for fault diagnosis. Yan et al. [2] proposed a hybrid intelligent fault diagnosis model combining a hierarchical sparse self-encoding network and a KNN classifier. Fei et al. [14] analyzed and diagnosed rolling bearing faults by extracting four information entropies in the time domain, frequency domain, and time-frequency domain. The proposed multi-feature entropy distance method has high diagnostic accuracy and strong robustness. Lin et al. [15] used deep neural networks for frequency domain feature extraction, and proposed a stacked auto-encoder (SAE) deep learning fault diagnosis method to identify and classify different bearing fault modes. Chen and Li [16] used a two-layer stacked auto-encoder neural network for feature fusion, and the fused feature vector was used as a machine health indicator to train a deep belief network (DBN) for fault classification. Generally speaking, as the number of hidden layers of neural networks increases, the depth of learning gradually deepens, and the adaptive learning effect of deep features is better [10].

As mentioned above, scholars have conducted a lot of research around feature extraction and fault diagnosis methods, and made significant contributions. However, the above research methods still have the following three problems: First, traditional artificial intelligence technology directly extracts a single fault feature from the original data, which does not fully reflect the fault information, and may need to re-extract features when conducting new diagnostic tasks. This results in poor generalization of models based on single fault characteristics. Second, feature dimensionality reduction mostly uses linear and non-linear methods such as PCA [13] and local tangent space alignment (LTSA) [17]. While new features can be obtained by extracting high-dimensional features using various mathematical methods, but the physical meaning of the new features often cannot be explained, and thus cannot be used to guide equipment maintenance practices. Third, the use of neural networks to adaptively extract deep features leads to problems, such as excessive numbers of hidden layers leading to feature overfitting; in addition, when the number of neurons in one layer differs greatly from the number of neurons in the next layer, feature data may be lost. Moreover, compared with traditional machine learning algorithms, neural networks usually require
more labeled data samples. Existing fault diagnosis methods based on deep learning use vibration data sets obtained in ideal experimental environments to achieve good test results. However, in actual engineering application environments, poor data quality often leads to problems such as poor generalization performance and inability to correctly identify faults.

In order to accurately identify common failures of rolling bearings in an industrial environment and provide technical support for predictive maintenance of rotating equipment, in this paper we formulate a fault diagnosis model that combines data-driven and fault mechanism analysis, and propose a rolling bearing weighted k-nearest neighbor classification fault diagnosis method based on sensitive multi-dimensional features. The fault diagnosis model was trained on the rolling bearing data of Case Western Reserve University (CWRU), while the laboratory data of the Intelligent Maintenance System (IMS), the Society of Mechanical Failure Prevention Technology (MFPT) and engineering case data were used for testing. The results showed that the method proposed in this paper has high fault diagnosis accuracy, can identify failure modes earlier, and shows good generalization when applied to different equipment and different working conditions. This exploration will promote data-driven rolling bearings engineering applications of intelligent fault diagnosis methods. The main contribution of the paper lies in the following aspects.

1) A fault diagnosis model of weighted k-nearest neighbor (WKNN) classification based on multi-dimensional feature parameters is formulated, and a fault diagnosis method based on multi-dimensional sensitive features is proposed and compared with related bearing fault diagnosis methods. The comparative analysis results show that the method proposed in this paper can not only automatically identify the failure category of rolling bearings, but also identify earlier and weaker fault signals.

2) The iterative ReliefF algorithm is used to filter the fault sensitive features. The complementarity of multi-dimensional characteristic parameters is used to characterize the fault information of rolling bearings more comprehensively. A rolling bearing fault identification model with good generalization was established by using the fault data under various working conditions for model training.

3) This paper builds a rolling bearing fault identification model based on multi-dimensional feature signals, and the fault data sets obtained under different speeds and different load working environments were used to train the model. The model identification accuracy of the constructed model was tested using fault data from different equipment and different working conditions in laboratory and industrial installations, and good results were obtained in all tests.

The rest of this paper is organized as follows. In Section II, the basic theory is described. The basic method and system framework are presented in Section III. In Section IV, experimental verification and the results are discussed. Three laboratory bearing datasets and one engineering case are studied, and performance comparisons with other methods are conducted. Finally, the study is summarized in Section V.

II. BASIC THEORIES

A. MULTI-DIMENSIONAL FEATURE FAULT EXTRACTION

1) TIME-DOMAIN FEATURES

Time-domain features can directly characterize signal changes over time and are widely used in different online monitoring systems for rolling bearings. Time-domain features can be divided into dimensional and dimensionless features. Dimensional features include the virtual value, the absolute mean, mean root mean square (RMS) and standard deviation; the dimensionless features include the clearance factor, skewness, kurtosis, the form factor, the peak factor, the pulse factor, and the margin factor [18], etc.

Dimensional features can directly reflect the health status of equipment components. For example, the RMS value of the vibration signal reflects the bearing’s vibration intensity. However, dimensional feature values are not only related to the equipment model, but are also greatly affected by the difference of the operating speed and working load. In addition, the feature values of the same fault can change greatly during different performance degradation stages. On the other hand, dimensionless parameters, such as kurtosis, which reflects the data distribution and can be used to measure the impact characteristics of vibration signals, are less affected by differences in operating conditions such as equipment types, operating speeds and loads.

The bearing operating status information reflected by a single time domain feature is limited, and sometimes the same time-domain feature value may be a manifestation of different underlying fault conditions [19]. The wear of bearing components leads to an increase in the value of shock and vibration when parts rotate. RMS and peak values can be used to quantify the degree of wear of the same bearing at different stages of degradation, but for different faulty bearings, RMS and peak values may also be incomparable.

The pulse waveform of a fault has the characteristics of high peak value. For shock faults, the peak factor is more sensitive than other features, but the peak factors of vibration signals caused by poor bearing lubrication are relatively small [20]. The above time-domain signals have different sensitivities to different faults and fault severity levels, and the same eigenvalue of the same fault may also be quite different when identifying faults under different working conditions. To identify the same or different faults in the same or different equipment, a single characteristic value alone is not sufficient to characterize the health and performance degradation of the bearings accurately.

2) FREQUENCY-DOMAIN FEATURES

The health status of key rotating equipment components during operation can be determined by analyzing the frequency domain characteristics of the vibration signal. The time domain signal can be transformed into a frequency domain signal through the Fast Fourier Transform (FFT); when a bearing fails, the frequency and harmonics of the fault
generated in the frequency domain can be analyzed to determine the fault and its severity. Therefore, extracting bearing frequency domain features can also reflect the information carried by the fault.

Frequency domain features generally include: average frequency (AF), center frequency (CF), root mean square frequency (RMSF), and standard deviation frequency (STDF). CF measures the average of the frequency spectrum line weighted by its amplitude. CF and RMSF characterize the fault characteristic frequency and its distribution in the spectrum, while STDF can characterize the extent of energy dispersion for each characteristic spectrum.

The time domain fault features, frequency domain fault features and entropy fault features of rolling bearing vibration signals have been extracted and listed in Table 1. $X(k)$ is the spectral amplitude, $f_k$ is the frequency, and $k$ is the number of spectral lines.

3) ENTROPY FEATURES

“Entropy” generally refers to a measure of the operating state of equipment, and the degree to which certain operating states may appear [21]. Empirical Mode Decomposition, Wavelet packet Decomposition (WPD) [22] and Variational Mode Decomposition (VMD) [23] have been widely used in fault identification. They decompose the vibration signal into different frequency band signals and use entropy to describe the uncertainty and complexity of each frequency band signal quantitatively. Therefore, methods for extracting bearing fault information through various entropy features are commonly used for fault identification and classification.

Different types of features can bear fault information of different parts. A feasible method is to establish a multi-dimensional multi-parameter feature set to characterize bearing fault information to adapt to different working environments. Applied on raw vibration data, time-domain, frequency-domain, and entropy features of each component after WPD and VMD have been extracted. New sensitive feature set can be synthesized on the basis of feature value screening and removal of redundant and insensitive features from the high-dimensional feature set. Characterization of the health status information of bearings from different domain perspectives is one of the research routes of this paper.

B. FEATURE SCREENING

The purpose of feature selection is to remove useless and redundant features without sacrificing diagnostic accuracy. Filter and Wrapper are the methods frequency-used to screen signal features. [24]. At present, there are a variety of filter feature selection methods; among them, the ReliefF algorithm is regarded as one of the most successful pre-processing algorithms due to its advantages of high efficiency and unlimited data types [25].

The core idea of the ReliefF algorithm is to evaluate the sensitivity of eigenvalues based on their ability to distinguish adjacent instances [26]. When the algorithm deals with multiple types of problems, a sample $R$ is randomly taken out from the training sample set. Then $k$ nearest neighbor samples $H$ (near Hits) of $R$ is found from the same sample set as $R$, and $k$ nearest neighbor samples $M$ (near Misses) are found from different sample sets of $R$. The distance between the sample $R$ and nearest neighbor samples is then calculated. The update of the weight $W(A)$ of attribute $A$ is then calculated according to Eq. (1).

$$W(A) = W(A) - \sum_{j=1}^{k} \frac{\text{diff}(A, R, H)}{mk} + \sum_{C \neq \text{class}(R)} \frac{P(C)}{(1 - P(\text{class}(R)))} \sum_{j=1}^{k} \frac{\text{diff}(A, R, M_j(C))}{(mk)}$$

where, $\text{diff} (A, R_1, R_2)$ is the distance between the samples $R_1$ and $R_2$ based on feature $A$; $m$ is the number of iterations; $P(C)$ is the probability of the $C$-th target; $M_j(C)$ represents the $j$-th target in the $C$-th target nearest neighbor sample.

C. WEIGHTED K-NEAREST NEIGHBOR CLASSIFICATION

The KNN is a non-parametric simple supervised learning algorithm based on $k$ nearest neighbors. It is simple and efficient, and has allowed significant progress in pattern recognition applications [27]. It is different from fault sortation algorithms based on SVM and neural network. KNN directly classifies the test samples based on all the stored training samples. The process can be briefly described as follows: when an unlabeled sample $x$ is input given for classification, KNN searches the $n$-dimensional pattern space of the training data, and extracts the $k$ nearest neighbors of $x$ after calculating their distances. The majority vote of the $k$ nearest neighbors determines the category to which $x$ belongs.

![Figure 1. Decision-making process of KNN algorithm.](image-url)
### Table 1. Time-domain, frequency-domain feature and entropy feature formulas.

| Feature       | Equation                                                                 |
|---------------|--------------------------------------------------------------------------|
| Virtual value | \( F_v = \frac{1}{N} \sum_{i=1}^{N} x_i^2 \)                            |
| Absolute mean | \( F_a = \frac{1}{N} \sum_{i=1}^{N} |x_i| \)                       |
| Mean          | \( F_m = \frac{1}{N} \sum_{i=1}^{N} x_i \)                             |
| RMS           | \( F_{\text{RMS}} = \left( \frac{1}{N} \sum_{i=1}^{N} |x_i|^2 \right)^{1/2} \) |
| Standard deviation | \( F_{\text{std}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2} \) |
| Variance      | \( F_{\text{var}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2} \) |
| Peak factor   | \( F_{\text{p}} = \max(|x_i|) \)                                     |
| Pulse factor  | \( F_{\text{p}} = \max(x_i) \)                                       |
| Margin factor | \( F_{\text{m}} = \max(|x_i|) \)                                       |

### Table 2. Multi-dimensional time-domain features of rolling bearing.

| Symbol | F1   | F2   | F3   | F4   | F5   |
|--------|------|------|------|------|------|
| Feature | Virtual value | Absolute mean | Mean | RMS  | Standard deviation |

### Table 3. Multi-dimensional frequency-domain features and wavelet entropy of rolling bearing.

| Symbol | F14  | F15  | F16  | F17  | F18  | F19  |
|--------|------|------|------|------|------|------|
| Feature | AF   | CF   | RMSF | STDF | Wavelet Energy Entropy | Wavelet Singular Entropy |

### Table 4. Multi-dimensional variational mode decomposition entropy features of rolling bearing.

| Symbol | F20-22 | F23-25 | F26-28 |
|--------|--------|--------|--------|
| Feature | Sample Entropy | Information Entropy | Permutation Entropy |

The WKNN classification is an improvement of the \( k \)-nearest neighbor classification (KNNC). It not only has the advantage of high calculation efficiency of KNNC, but also can avoid the disadvantage that the recognition accuracy of KNNC is easily affected by the neighborhood size \( k \). Huang et al. [28] proposed a KNN algorithm based on class contribution and feature weighting. A comparative experiment of University of California Irvine dataset showed that the classification recognition accuracy increases with the increase of \( k \).

WKNN classification assigns different weights to the neighbor samples according to the similarity between each neighbor sample and the test sample, so the classification result of the test sample is closer to the training sample with higher similarity. Euclidean distance is used to measure the distance between samples, and the inverse distance square weighting method is applied to calculate the sample weight. This method is a common and simple spatial interpolation method. The nearest neighbor’s votes are weighted based on the calculated distance; as the weight is the reciprocal of the distance square, smaller distances result in higher weights.

### III. METHOD AND SYSTEM FRAMEWORK

#### A. MULTI-DIMENSIONAL FEATURE SET CONSTRUCTION

Before feature extraction, Z-score normalization is carried out for each sample to eliminate the numerical differences of vibration signals caused by different equipment and working conditions.

The time-domain feature extraction is shown in Table 2. These features are commonly used health indicators for fault diagnosis and analysis of rolling bearings.

The FFT algorithm is used to convert the time-domain sample to a frequency domain signal, and 4 common frequency domain features are extracted. The db4 wavelet is utilized to decompose the collected raw signal into a 3-layer wavelet, and 8 frequency bands of the decomposed signal are obtained. Then, the wavelet energy entropy and wavelet singular entropy are extracted (Table 3).

The vibration entropy feature is extracted using the variational mode decomposition method. By calculating the correlation coefficient between the modal components, the presence of frequency aliasing between the modal components is determined. On this basis, the number of \( K \) modes of the variable mode decomposition was determined to be 3. Then, the sample entropy, information entropy, permutation entropy, fuzzy entropy and dispersion entropy of each modal component are calculated. The extracted entropy features are shown in Table 4.

\( h \) is the distance from the unknown sample to the training sample.

In this manner, the recognition accuracy is improved and the sensitivity to the value of \( k \) is weakened, so the accuracy and robustness of the recognition results are improved.
In this study, we selected 13 time-domain features, 4 frequency-domain features and 17 entropy features based on the WPD and VMD decomposition signals to construct a multi-dimensional feature set to characterize the health of the bearing.

B. RELIEFF FEATURE SCREENING

In order to overcome the inefficiency of artificially-selected fault-sensitive features based on qualitative experience without prior knowledge, the ReliefF algorithm is used to calculate the weight of each feature and remove features that are not sensitive to faults. First, based on the training data, the extracted high-dimensional feature set is subject to arctangent normalization; then, the ReliefF algorithm is used for feature selection.

![Flow chart of ReliefF algorithm](image1.png)

The ReliefF algorithm flow is shown in Fig. 2. A sample \( R \) is randomly selected from the training set \( D \), and the weight of each feature is updated according to Eq. (1). The weight \( W \) of each feature is calculated after \( m \) cycles. The larger the feature weight, the more distinguishable the feature is, which means that it is a sensitive feature.

![WKNN classification fault diagnosis model framework schematic diagram](image2.png)

FIGURE 3. WKNN classification fault diagnosis model framework schematic diagram.

C. CONSTRUCTION OF WKNN-BASED FAULT CLASSIFICATION MODEL

It can be seen in Fig. 3 that the WKNN classification of the rolling bearing fault identification model based on multi-dimensional sensitive features operates under two modes: offline training and online monitoring. The offline training mode includes multi-dimensional feature parameter extraction, sensitive feature parameter screening and fault diagnosis rule library creation. The online monitoring mode includes bearing performance degradation detection and early warning, sensitive feature value calculation, and failure mode identification.

1) OFFLINE TRAINING MODE

Step 1: Multi-dimensional feature parameter extraction.

Raw vibration signal data of the bearing in the normal state and degraded operating states are used as the training data set. First, Z-score normalization preprocessing is applied on the training sample data. Then, the FFT, WPD and VMD of the preprocessed signal are calculated. Finally, feature extraction is performed to obtain 34 time-domain, frequency-domain and entropy feature parameters, and their eigenvalues are calculated.

Step 2: Sensitive feature parameter screening.

First, the 34 feature values are normalized using the arc tangent to eliminate differences in the sizes of each eigenvalue. Second, the iterative ReliefF feature screening method is used to calculate the weight coefficient of each feature parameter, and invalid redundant feature parameters are removed based on rules. Finally, an \( m \)-dimensional rolling bearing fault-sensitive characteristic parameter set is formulated.

Step 3: Fault diagnosis rule library construction.

First, the \( m \)-dimensional sensitive feature vector space of \( n \)-type faults is established. Second, the intersection verification method is applied to select the optimal value of \( k \). Then, the Euclidean distance is used to measure the distance between the training and test samples. The distance-square weighting method is used to determine the distance weight coefficient. This is followed by the calculation of the distance between the test sample and the \( k \) nearest samples,
classification of the test sample into the category with the largest population among the \( k \) samples. Then, the accuracy of the sample classification result is verified; if it is low, it is necessary to optimize the sensitive feature parameters and the value of \( k \) using representative training samples. Finally, the verified \( n \)-type fault \( m \)-dimensional sensitive feature vector with good generalization performance, \( k \) value, and distance weight coefficient together constitute the WKNN classification fault diagnosis classification rule base.

2) ONLINE MONITORING WORK MODE

Step 1: Fault determination.

Online incipient fault detection method based on improved \( l_1 \) trend filtering and support vector data description can be used to detect and confirm the initiation of rolling bearings performance degradation [29]. As shown in Figure 4, using this early warning method, the early warning point can be found online, and fault diagnosis can be started.

Step 2: Sensitive feature vector value calculation.

First, we obtain the raw vibration waveform data of the rolling bearing that has demonstrated performance degradation or incipient fault. Then, \( Z \)-score normalization preprocessing is applied on the raw data, and, finally, the \( m \)-dimensional sensitive feature vector for \( n \) types of faults is calculated.

Step 3: Failure mode identification.

First, the \( m \)-dimensional sensitive feature vector is normalized using the arc tangent. Then, the processed feature vector is input into the WKNN for calculation, and a classification is output based on the classification rule base.

IV. MODEL TRAINING AND VERIFICATION

A. MODEL BUILDING

1) CWRU TESTING BENCH AND DATASET

To verify the validity of the constructed model, the rolling bearing fault dataset provided by the Case Western Reserve University (CWRU) [30] laboratory was used for model training and testing. The bearing parameters are shown in Table 5. It can be seen from Fig. 5 that the CWRU test bench consisted of a three-phase asynchronous motor (left), a torque encoder (middle), a dynamometer (right) and an acceleration sensor. The latter collected the vibration signal of the bearing at the driving end of the test bench at a sampling frequency of 12KHz. The acquired vibration data included bearing data under four states: normal (N), inner ring failure (IF), outer ring failure (OF) and rolling element failure (RF).

Based on the bearing speed and the sampling frequency of the sensor, it can be inferred that about 400 sampling points were collected per bearing revolution. As shown in Table 6, in order to obtain complete and accurate bearing health status information, 2048 sampling points were taken as the length of the test samples. Under each workload operating environment, 50 test samples were obtained for each health state of the bearing. As shown in Table 7, three testing datasets A, B, and C were established under different operating conditions; these could then be used as training and test sets.

2) WKNN CLASSIFICATION MODEL ESTABLISHMENT PROCESS

Step 1: \( Z \)-score normalization

In order to eliminate the impact of the difference in the magnitude of the distribution of the model training or verification sample data caused by the bearing defects, working speed, and operating load differences, the \( Z \)-score method was used to normalize the magnitudes of the different measurements, so that they conform to the standard normal distribution and are dimensionless. This ensures the comparability between different test and verification sample data.

\[
z = \frac{x - \mu}{\sigma}
\]

where \( x \) is a sample vector containing 2048 sampling points, \( \mu \) is the mean of the data, and \( \sigma \) is the standard deviation of the data.

Step 2: Feature extraction

The time-domain vibration signal is subjected to FFT, WPD and VMD, and 13 time-domain features,
4 frequency-domain features, and 17 entropy features are obtained, jointly forming the raw signal feature set. When the eigenvalues are input into the WKNN classification, arctangent normalization is used to eliminate the difference in the size of each eigenvalue so as to avoid the phenomenon that “in a multi-dimensional feature space, when the magnitude of the feature is large, the smaller feature cannot be effectively represented”.

\[ x^* = \frac{\text{atan}(x)^2}{2\pi} \]  

(5)

Step 3: Screening of sensitive feature parameters based on relieff

The iterative Relieff algorithm was used to calculate the respective weight coefficients of the 34-dimensional original feature parameters. Iterative calculations were performed on the training and test sample data. The number of iterations \( m \) of the Relieff algorithm was selected as 100, and the nearest neighbor \( k \) value was 5. The weight of each original feature parameter was calculated using Eq. (1), and feature parameters whose weight coefficient was lower than 0.1 were removed, thus resulting in a 30-dimensional sensitive feature parameter set.

Step 4: WKNN classification \( k \) value optimization

The cross-validation method was used to select the optimal \( k \) value for nearest neighbor classification. As shown in Fig. 7, because the number of data files for a single test sample was 50, the recognition accuracy would decrease significantly as \( k \) increased beyond 50. Therefore, \( k \) was set to 5, the Euclidean distance was used, and the distance weight coefficient was determined using the inverse distance square weighting method.

Step 5: Fault diagnosis rule base establishment

The distance between the test sample and \( k \) recent training samples was calculated, and different weights were assigned according to the distance. Using the voting system, the test samples were classified into the category with the most votes among the \( k \) samples.

3) MODEL TRAINING AND TESTING

To confirm the comprehensiveness of the algorithm more accurately, 6 groups of experiments were carried out. Experiments 1 and 2 used dataset A as the training set, and dataset B and C as test sets. In experiments 3 and 4, B was used as the training set, and A and C as the test sets. Experiments 5 and 6 experiments used C as the training set, and A and B as the test sets.

As shown in Table 8, the proposed fault identification model showed good accuracy under the three operating environments A, B, and C. This demonstrates that the proposed method is not affected by the change of speed and load of between the training and test set significantly.

Then, the three CWRU datasets were combined as the training set for cross-device diagnosis. Each dataset contained four health states, with 50 groups of samples for each health state, for a total of 200 groups of samples, and 600 groups of samples for cross-device training were obtained. The number of neighbors for WKNN classification was set to 51, Euclidean distance was used, and the distance weight is determined by the inverse distance square weighting method. Other experimental platform data and engineering case data were selected for the test set.

B. WKNN CLASSIFICATION MODEL ESTABLISHMENT PROCESS

1) MODEL VERIFICATION BASED ON IMS BEARING DATASET

The IMS bearing experiment platform is shown in Fig. 8. A corresponding dataset [31] was used for model verification, while Dataset 2 is used for model verification and analysis. The bearing vibration signal of the test bench was collected every 10 minutes, until the outer ring of bearing 1 was severely peeled off after 9840 minutes of operation, and the experiment was terminated, at which point 984 data files had been collected.

The fault recognition model now operated in online operation mode. Fault recognition was started after the confirming of the bearing’s performance degradation using the incipient fault detection model. According to the
conclusions of Wang et al. [29], bearing performance degradation occurred in the 533th data file in this dataset. Therefore, we used the confirmed first failure data file as the first test sample, and the 100 subsequent test sample as the test group to verify the fault diagnosis model proposed in this study. The fault identification conclusions are shown in Table 9, were the accurate identification of the outer ring fault is evident in all case.

2) MODEL VERIFICATION BASED ON MFPT LABORATORY DATA

The bearing experiment dataset provided by MFPT was used for model verification [32]. Acceleration sensor data under 8 load environments were obtained, as shown in Table 10. The experimental data of 7 groups of inner and outer ring faults under 7 load environments were selected as model test data. The characteristics of the test data are as follows. There were 7 sets of acceleration test data for rolling bearing outer ring failure, collected under loads of 25lbs, 50lbs, 100lbs, 150lbs, 200lbs, 250lbs and 300lbs, with the same sampling frequency and duration. The model verification results based on the inner and outer ring fault test data are shown in Table 11.

It can be seen from Table 11 that the model can identify the failure of the inner ring and groups O-1, O-3 and O-4 of the MFPT rolling bearing with an accuracy of 100%. For the O-5 group, O-6 group and O-7 groups of test data, the recognition accuracy rates were 44%, 72% and 0%, respectively.

C. MODEL VERIFICATION BASED ON ENGINEERING CASE DATA

The rolling bearing data collected in a factory environment are different from data collected in the laboratory. To validate the applicability of the proposed fault verification model under industrial environments, the P3409A centrifugal pump bearing fault data of a petrochemical hydrocracking unit was used to test the fault identification accuracy of the model. The layout of the bearing vibration monitoring points of the centrifugal pump is shown in Fig. 9. The centrifugal pump operated at 2980r/min, and the bearing designation was 6217, the sampling frequency was 25600Hz, and a set of raw vibration data file was obtained every two hours.

The raw acceleration data of the P3409A centrifugal pump from 24:00 on December 15, 2013 to 16:00, January 12, 2014 were obtained. There was a total of 332 data files, each composed of 16,384 sampling points. The bearing performance degradation detection model of Wang et al. [29] determined that the 142nd data file was the starting point of bearing performance degradation, and confirmed it with the envelope spectrum analysis method. The first fault data file that was confirmed was used as the first test sample, and 100 subsequent test samples were used as test groups to verify the fault diagnosis model proposed in the paper. It is known that the rolling bearing fault was present in the outer ring.
The identification results of the proposed model are shown in Table 12.

### D. MODEL VERIFICATION AND COMPARATIVE ANALYSIS

In the following, we compare and analyze our method’s performance with that of the following commonly used methods: the PCA (Method 1) and KPCA (Method 2) feature dimensionality reduction methods of [12] and [9], respectively; feature screening (Method 3) [11]; deep learning method (Method 4) based on the stacked auto-encoding network of [16]; the SVM (Method 5) and KNN classification (Method 6) methods of [11]; multi-feature entropy extraction method (Method 7) of [14]; the 34 - dimensional features (Method 8) and 30 - dimensional sensitive features selected (Method 9). The detailed description of fault diagnosis methods is shown in Table 13.

In Methods 1 and 2, PCA and KPCA are used for linear and non-linear dimensionality reduction on the 34-dimensional raw signal features; the main feature components (those retaining 95% of the information) are extracted and used as input for WKNN classification. In Method 3, feature selection based on iterative ReliefF is applied, and the top five weighted features are selected for WKNN classification. In Method 4 the frequency domain signals are the input, and a three-layer stacked auto-encoder network is used, with the number of hidden layer nodes in the network being the same as the input data. In Method 5, a one-to-one SVM is used as a multi-classifier. KNN classification is the basis of Method 6, while WKNN is used in Method 9, both using the 30-dimensional fault feature set selected in this paper as input for their respective classifications. Method 7 uses 4 information entropy features as the input of WKNN, and Method 8 uses 34-dimensional original features without processing as the input of WKNN, as the control experiment group of the method (Method 9) in this paper.

#### 1) VERIFICATION AND COMPARISON OF THE SAME EQUIPMENT UNDER DIFFERENT WORKING CONDITIONS

The three sets of CWRU data (A/B/C) under different operating conditions were used as training and test datasets respectively, and the 7 fault diagnosis methods were used for verification and comparison. Comparative analysis shows that the methods based on dimensionality reduction (Method 1 and Method 2) had the worst diagnostic accuracy among the seven methods, while the accuracy of the remaining methods was above 95%. This shows that methods 3-7 can be used for fault diagnosis of this type of equipment with high accuracy. It also shows that the bearing failures in the CWRU bearing experiment are typical, and the different raw failure datasets are also representative. Therefore, it is selected as the failure training set, and other laboratory and engineering case data were used as the verification set to compare and validate the practicality of the method proposed in this paper.

#### 2) VERIFICATION AND COMPARISON UNDER DIFFERENT EQUIPMENT AND DIFFERENT WORKING CONDITIONS

IMS, MFPT laboratory data and P3409A project case data were used as validation sets to compare the advantages and disadvantages of the above methods.

a) Model verification and comparative analysis based on IMS dataset.

The IMS datasets were selected as test data for the fault diagnosis model. The first confirmed fault data file was used as the first test sample and 100 subsequent test samples were used as test groups to verify the aforementioned 7 fault diagnosis models. The test results can be seen in Fig. 10, Method 4 and Method 5 could not identify the fault model at all; the fault recognition accuracy of Methods 1 and 2 was
b) Model verification and comparative analysis on MFPT datasets.

The MFPT experimental dataset includes seven kinds of bearing failure parts under different operating environments; these were used as test samples to verify the seven fault diagnosis methods. As shown in Fig. 11, none of the seven fault diagnosis methods could identify O-7 outer ring faults correctly. Methods 1 and 2 had low accuracy in identifying the inner and outer ring faults of rolling bearings; Method 3 and 7 had high recognition accuracy for inner ring faults, but low accuracy for outer ring faults; Method 4 could not identify any type of rolling bearing faults; Methods 5 and 6 showed high accuracy for outer ring faults, and lower for inner ring faults; Method 8 has lower accuracy for inner and outer ring fault identification; and Method 9 achieved a high fault recognition accuracy rate for both the inner and outer ring faults of the rolling bearing under different operating environments.

c) Model verification and comparative analysis based on P3409A data.

The P3409A data were used to test the fault diagnosis models. The first confirmed fault data file was used as the first test sample and the 100 subsequent test samples were used as the test groups to verify the aforementioned 7 methods. As shown in Fig. 12, Method 4 could not identify the failure mode; Methods 1 and 2 had fault recognition accuracy rates of about 40%; Methods 3 and 6 rates ranged from 75% to 100%; Method 5 could not identify early failures of rolling bearings; the accuracy of Method 7 decreases in the later stage; the fault identification accuracy of Method 8 and Method 9 proposed in this paper always reaches 100% accuracy.

3) DISCUSSION

The verification results of the above groups of models are explained as follows. Although the feature sets after the dimensionality reduction of Methods 1 and 2 retain the main information of the bearing operating state, this information may not be fault feature information. It is relatively easy to extract signal features using the stacked self-encoding network deep learning methods, but there is an over-fitting problem, and the extracted signal features only have a high fault recognition accuracy rate on the same device, but are not effective for the identification of the same faults in different devices. Methods 5 and 6 build SVM, KNN and WKNN fault diagnosis models based on the filtered 30-dimensional fault sensitive feature set, and use P3409A engineering case data as the model testing dataset. The test results show that the SVM model’s had poor accuracy in identifying faults, compared with the KNN model, the WKNN model can better eliminate the influence of different sensitivity values K on the accuracy of fault identification. Method 7 is based on the WKNN accuracy for outer ring faults, and lower for inner ring faults; Method 8 has lower accuracy for inner and outer ring fault identification; and Method 9 achieved a high fault recognition accuracy rate for both the inner and outer ring faults of the rolling bearing under different operating environments.
In this paper, a weighted $k$-nearest neighbor classification model based on multi-dimensional sensitive features is proposed as the rolling bearing fault diagnosis method. By extracting time domain, frequency domain and entropy features, a high-dimensional feature parameter set is established, sensitive features are determined via the iterative ReliefF algorithm, and a WKNN classification model is established for fault pattern recognition. This method not only considers the sensitivity differences of different features to faults, but also fully takes into account the complementarity of different feature combinations for fault identification, and can characterize the operating status of mechanical components comprehensively.

The WKNN classification fault diagnosis model constructed based on CWRU laboratory data training was tested and verified using IMS, MFPT laboratory data and P3409A engineering data. The results show that the constructed fault diagnosis model can automatically identify bearing fault types. It can adapt to different experimental conditions of the same equipment, different experimental environments of different equipment, and complex working conditions of industrial applications. Under various test conditions, it has high rolling bearing fault recognition accuracy and good model generalization.

A fault diagnosis model combining data-driven and fault mechanism analysis is established. The input of the model is raw vibration data, and the output of the model is the determination of bearing fault diagnosis. The process of fault detection and classification does not need to rely on external experts and their prior knowledge. In the industry 4.0 environment, this has important engineering application value for realizing predictive maintenance and timely planning of maintenance operations.

REFERENCES

[1] I. Attoui, N. Fergani, N. Boutasseta, B. Oudjani, and A. Deliou, “A new time–frequency method for identification and classification of ball bearing faults,” J. Sound Vibrat., vol. 397, pp. 241–265, Jun. 2017.

[2] Z. Yan, X. Yuan, F. Zhou, Y. Song, Q. Xu, and Y. Shao, “Fault diagnosis based on a stacked sparse auto-encoder network and KNN classifier,” in Proc. Chin. Autom. Cong. (CAC), Nov. 2019, pp. 3129–3134, doi: 10.1109/CAC48633.2019.8996955.

[3] A. Soualhi, H. Razik, and G. Clerc, “Data driven methods for the prediction of failures,” in Proc. 12th Int. Symp. Diagn. Electr. Mach., Power Electron. Drives (SDEMPED), Aug. 2019, pp. 474–480.

[4] S. Yuanyuan, G. Lili, and W. Yongming, “Artificial intelligence and learning techniques in intelligent fault diagnosis,” in Proc. 4th Int. Conf. Comput. Sci. Netw. Technol. (ICCSNT), Dec. 2015, pp. 702–707.

[5] Q. Zhou, Y. Lv, L. Tu, M. Wang, and S. Li, “Study of fault diagnosis for rolling bearing based on clustering algorithms,” in Proc. 5th Int. Conf. Control Robot. Eng. (ICCRE), Apr. 2020, pp. 58–62.

[6] M. Kang, J. Kim, J.-M. Kim, A. C. C. Tan, E. Y. Kim, and B.-K. Choi, “Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis,” IEEE Trans. Power Electron., vol. 30, no. 5, pp. 2786–2797, May 2015.

[7] R. Li, P. Sopon, and D. He, “Fault features extraction for bearing prognostics,” J. Intell. Manuf., vol. 23, no. 2, pp. 313–321, Apr. 2012.

[8] N. Li, G. Zhang, Z. Xiao, and O. P. Malik, “Feature extraction based on adaptive multiwavelets and LTSA for rotating machine fault diagnosis,” Shock Vibrat., vol. 2019, pp. 1–15, Jan. 2019.

[9] Z. Liu, W. Guo, J. Hu, and W. Ma, “A hybrid intelligent multi-fault detection method for rotating machinery based on RSGWPT, KPCA and twin SVM,” ISA Trans., vol. 66, pp. 249–261, Jan. 2017.

[10] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, “An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data,” IEEE Trans. Ind. Electron., vol. 63, no. 5, pp. 3137–3147, May 2016.

[11] T. Tuncer, S. Dogan, and F. Ozurt, “An automated residual exemplar local binary pattern and iterative ReliefF based COVID-19 detection method using chest X-ray image,” Chemometric Intell. Lab. Syst., vol. 203, Aug. 2020, Art. no. 104054, doi: 10.1016/j.chemolab.2020.104054.

[12] Y. Yinghua, S. Guoqiang, and S. Xiang, “Fault monitoring and classification of rotating machine based on PCA and KNN,” in Proc. Chin. Control Decis. Conf. (CCDC), Jun. 2018, pp. 398–403.

[13] H. Wang, G. Ni, J. Chen, and J. Qu, “Research on rolling bearing state health monitoring and life prediction based on PCA and Internet of Things with multi-sensor,” Measurement, vol. 157, Jun. 2020, Art. no. 107657, doi: 10.1016/j.measurement.2020.107657.

[14] C.-W. Fei, Y.-S. Choy, G.-C. Bai, and W.-Z. Tang, “Multi-feature entropy distance approach with vibration and acoustic emission signals for process feature recognition of rolling element bearing faults,” Struct. Health Monit., vol. 17, no. 2, pp. 156–168, 2018.

[15] X. Lin, B. Li, X. Yang, and J. Wang, “Fault diagnosis of aero-engine bearing using a stacked auto-encoder network,” in Proc. IEEE 4th Int. Conf. Mechatronics Eng. Conf. (IOTOE), Dec. 2018, pp. 545–548, doi: 10.1109/IOTOE.2018.8745045.

[16] Z. Chen and W. Li, “Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network,” IEEE Trans. Instrum. Meas., vol. 66, no. 7, pp. 1693–1702, Jul. 2017.

[17] J. S. Jiang, H. Q. Wang, Y. L. Ke, and W. Xiang, “Fault diagnosis based on LTSA and K-nearest neighbor classifier,” J. Vibration. Shock, vol. 36, no. 11, pp. 134–139, 2017.

[18] Y. Li, W. Dai, and W. Zhang, “Bearing fault feature selection method based on weighted multidimensional feature fusion,” IEEE Access, vol. 8, pp. 19008–19025, 2020.

[19] M. S. Ballal, Z. J. Khan, H. M. Suryawanshi, and R. L. Sonolikar, “Adaptive neural fuzzy inference system for detection of inter-turn insulation and bearing wear faults in induction motor,” IEEE Trans. Ind. Electron., vol. 54, no. 1, pp. 250–258, Feb. 2007.

[20] L. Zhen, H. Zhengjia, Z. Yangyan, and C. Xuefeng, “Bearing condition monitoring based on shock pulse method and improved redundant lifting scheme,” Math. Comput. Simul., vol. 79, no. 3, pp. 318–338, Dec. 2008.

[21] J. Li and J. Guo, “A new feature extraction algorithm based on entropy cloud characteristics of communication signals,” Math. Problems Eng., vol. 2015, pp. 1–3, Jun. 2015.

[22] L. Eren and M. I. Devaney, “Bearing damage detection via wavelet packet decomposition of the stator current,” IEEE Trans. Instrum. Meas., vol. 53, no. 2, pp. 431–436, Apr. 2004, doi: 10.1109/TIM.2004.823323.

[23] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” IEEE Signal Process., vol. 62, no. 3, pp. 531–544, Feb. 2014.

[24] Y. Zhang, D. Gong, Y. Hu, and W. Zhang, “Feature selection algorithm based on bare bones particle swarm optimization,” Neurocomputing, vol. 148, pp. 150–157, Jun. 2015.
[25] R. Z. Zhao, K. J. Li, and L. F. Deng, “Fault feature selection method based on ReliefF and genetic algorithm,” J. Vib. Shock, vol. 33, pp. 211–214, Aug. 2014. [Online]. Available: http://d.wanfangdata.com.cn/conference/8428866

[26] I. Kononenko, “Estimating attributes: Analysis and extensions of RELIEF,” in Machine Learning: ECML-94. Berlin, Germany: Springer, 1994, pp. 171–182.

[27] R. S. Gunerkar and A. K. Jalan, “Classification of ball bearing faults using vibro-acoustic sensor data fusion,” Exp. Techn., vol. 43, no. 5, pp. 635–643, Oct. 2019, doi: 10.1007/s40799-019-00324-0.

[28] J. Huang, Y. Wei, J. Yi, and M. Liu, “An improved kNN based on class contribution and feature weighting,” in Proc. 10th Int. Conf. Measuring Technol. Mechatronics Autom. (ICMTMA), Feb. 2018, pp. 313–316, doi: 10.1109/ICMTMA.2018.00083.

[29] Q. Wang, X. Liu, B. Wei, and W. Chen, “Online incipient fault detection method based on improved $\ell_1$ trend filtering and support vector data description,” IEEE Access, vol. 9, pp. 30043–30059, 2021, doi: 10.1109/ACCESS.2021.3058907.

[30] (2013). Bearing Data Center. [Online]. Available: https://csegroups.case.edu/bearingdatacenter/home

[31] (2014). IMS Bearings Dataset. [Online]. Available: http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/

[32] E. Bechhoefer. (2016). A Quick Introduction to Bearing Envelope Analysis, MFPT Data. [Online]. Available: http://www.mfpt.org/FaultData/FaultData.htm.

QINGFENG WANG received the Ph.D. degree from the College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, China, in 2011. He is currently an Associate Researcher with the College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology. He is also a Security Expert of the State Administration of Work Safety, China. His research interests include electromechanical equipment monitoring, diagnosis, and maintenance.

SHUAI WANG received the B.S. degree from the College of Chemical Engineering, China University of Mining and Technology, China, in 2019, where he is currently pursuing the M.S. degree. His current research interest includes fault diagnosis.

BINGKUN WEI received the B.S. degree from the College of Electrical Engineering, Beijing University of Chemical Technology, China, in 2018, where he is currently pursuing the M.S. degree. His current research interests include fault diagnosis and health evaluation.

WENWU CHEN is currently a Professorate Senior Engineer with the SINOPEC Qingdao Research Institute of Safety Engineering, China. His current research interests include integrity management, fault diagnosis, and prediction technology application in oil refining and chemical equipment.

YUFEI ZHANG received the master’s degree from the College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, China, in 2016. He is currently a Head of the Department of SINOMACH Academy of Science and Technology Company. His research interests include electromechanical equipment monitoring, diagnosis, and maintenance.